DNS Stability:
The Effect of New Generic Top Level Domains on the Internet Domain Name System

Introduction: Maintaining Stability through Growth

The Internet consists of a backbone of networks and servers connected through various protocols. The Internet allows the sharing of information at remote sites from institutions such as governmental organizations, research facilities, businesses offering products and services, and individuals who want a presence on the Internet. The loosely coupled technologies that enable sharing of this information include IP addresses and domain names. An important infrastructure for these Identifiers is the Internet's distributed Domain Name System (DNS).

At the very top of the hierarchy of the DNS is the root zone, which contains information pertaining to top-level domains, or TLDs (of which there are currently 281). The root zone is published via a globally distributed system of domain name servers known as the root servers.

The ICANN community has recently completed a policy development process regarding the introduction of new generic top-level domains (gTLDs). This effort took place within ICANN's Generic Names Supporting Organization (GNSO) (see http://gnso.icann.org/) and resulted in a set of recommendations (see http://gnso.icann.org/issues/new-gtlds/pdp-dec05-fr-parta-08aug07.htm#_Toc43798015) to guide ICANN in introducing new gTLDs to the namespace.

In its recommendations, the GNSO concluded that "ICANN must implement a process that allows the introduction of new top-level domains," and called for a procedure respecting the principles of fairness, transparency and non-discrimination. In preparing for the expected implementation of the gTLD policy recommendations, staff is conducting some review and analysis of the technical issues involved in this development. The addition of gTLDs to the namespace is an expansion of the DNS on a potentially large scale, to include many more names at the top level.

Specifically, the policy recommendations developed in relation to the introduction of new gTLDs included the requirement that "Strings must not cause any technical instability." ICANN is publishing this paper to solicit informed input on the technical issues relevant to the addition of new gTLDs, and to provide transparency toward how it will interpret and implement this recommendation. The goal is a clear set of rules that will be available to potential new gTLD applicants, so that it is known from the outset what tests will be applied to each application. The areas discussed in this paper are:

I. TLD strings that might impact stability.

II. Technical issues with expansion of the root zone that might impact stability.

III. Operational issues resulting from expansion of the root zone that might impact stability.

ICANN is seeking feedback on the proposed approach to these areas as a step in its implementation planning for the introduction of new gTLDs.

I TLD Strings and Technical Instability

Syntax Rules

ICANN's first area of consideration in response to the GNSO's recommendation is whether there are certain TLD strings that would tend to result in instability. Conformity to existing standards and syntax rules, as described below, will be a requirement for any new TLD.

RFC 952, “DOD Internet Host Table Specification,” describes certain assumptions about the syntax of host names, including:

- They will consist of characters drawn from the Latin alphabet (A-Z), digits (0-9), the minus sign (-), and period (.). Note that periods are only allowed when they serve to delimit components of “domain style names.”
- No blank or space characters are permitted as part of a name.
- No distinction is made between upper and lower case.
- The first character must be an alpha character (Note modification to this at [1] below).
- The last character must not be a hyphen.
- Single character names or nicknames are not allowed.

RFC 952 was liberalized by RFC 1123, “Requirements for Internet Hosts – Application and Support.” RFC 1123 states that: [1] “One aspect of host name syntax is hereby changed: the restriction on the first character is relaxed to allow either a letter or a digit. Host software MUST support this more liberal syntax.”

RFC 2181, “Clarifications to the DNS Specification,” states that: “The DNS itself places only one restriction on the particular labels that can be used to identify resource records. That one restriction relates to the length of the label and the full name. The length of any one label is limited to between 1 and 63 octets. A full domain name is limited to 255 octets (including the separators).”

RFC 3696, “Application Techniques for Checking and Transformation of Names,” states that: “The LDH rule, as updated, provides that the labels (words or strings separated by periods) that make up a domain name must consist of only the ASCII [ASCII] alphabetic and numeric characters, plus the hyphen. No other symbols or punctuation characters are permitted, nor is blank space. If the hyphen is used, it is not permitted to appear at either the beginning or end of a label. There is an additional rule that essentially requires that top-level domains not be all-numeric.” In the case of an IDN TLD, the Punycode version would be a string comprised of LDH characters and complying with the above rules, with a single label being no longer than 63 octets long.

In order to prevent confusion with IPv4 addresses which are expressed as numeric labels separated by dots (e.g. 192.168.1.1) ICANN expects to disallow proposed TLDs containing only numeric characters. Such a prohibition will avoid a TLD typed into a browser from being resolved as an IP address.

Labels beginning with 0x (the number “0” and the letter “x”) followed by a hexadecimal character [a-f, 0-9] would not be allowed since inet_addr() converts hex strings to IP addresses. For example a "ping 0xa" command will be converted to 0.0.0.10 0xaf.0x7d.0x9f.0xc and would be translated to 175.125.159.12.

Hyphens in both the third and fourth positions of a label are allowable only in a valid Punycode string, where the currently approved IDNA prefix (currently xn) is used. The previously-used IDN .TEST evaluation strings (such as <.xn—zckzah>) will not be allowed in production. As described in the IDN .TEST Evaluation Plan, “Any label that is entered into the root zone of the DNS for the purposes of IDN testing will be categorically barred from subsequent delegation as a production domain.” (See http://www.icann.org/topics/idn/idn-evaluation-plan-v2-9-2-14aug07.pdf.)

Many of these rules were also discussed by the GNSO working group on reserved names (see http://gnso.icann.org/issues/new-gtlds/final-report-rn-wg-23may07.htm).

To summarize, the proposed limitations to be placed on new TLDs for technical stability reasons are:

- Labels must consist only of ASCII LDH characters (letters, digits, and hyphen).
- Labels must be 63 characters or less.
- Labels must not be made up entirely of digits.
- Labels featuring hyphens in the third and fourth character positions must be valid Punycode labels using the approved IDNA prefix.
- Labels must not begin or end with a hyphen.

- Labels must not begin with the characters “0x” followed by a hexadecimal character.

File Extensions

ICANN has also considered the possible impact of allowing commonly-used file extensions as TLDs in the root, examining whether this might result in users or applications confusing URLs with filenames, or create other vulnerabilities for the user. Examples of such strings include:

- .EXE
- .HTML
- .DOC
- .JPG
- .CSV
- .PPT
- .ZIP
- .PDF
- .TXT
- .MP3
- .GIF
- .XLS
- .RTF

Staff explored this issue by consulting the Security and Stability Advisory Committee (see http://www.icann.org/committees/security/), as well as conducting some preliminary outreach to browser and operating system developers on the question of whether TLDs identical to file extensions could cause technical problems when attempting to be resolved in their particular programs. Responses to these inquiries have indicated that if there were problems resulting from the addition of TLD labels coinciding with common file extensions, they would be problems of user confusion rather than breakage in the DNS. It should also be noted that strings that are also file extensions exist as TLDs in the DNS today without adverse effects.

The DNS will perform look-ups of a TLD string without consideration to its perceived meaning in another context. If a user types “mydocument.pdf” into a browser, the auto completion rules of most commonly-used browsers will add the “http://” notation to the front of that string and perform a DNS lookup. Only by expressly indicating that the user is looking for a file with file://mydocument.pdf will the browser assume that a path to a file is indicated rather than a domain name. These are functions of the application, however, not the DNS. Based on this analysis, there does not appear to be a technical need to restrict common file extensions from also being TLDs.

The question remaining to be answered is whether there should be any application-oriented limitations on TLD strings. If ICANN wished to prevent potentially significant problems in the application layer or avoid user confusion issues resulting from a string like .EXE, it would need to do so based on a defined list of extensions that were disallowed, a list maintained outside of ICANN and internationally recognized as the authoritative source for “commonly used file extensions.” Additionally, a mechanism for maintaining and updating such a list would need to be in place, because new file extensions could become prevalent at any time. To date, staff has not been able to locate a list of common file extensions that is generally acknowledged to be authoritative.

Given preliminary feedback that there is not a technical need to prevent file extensions as TLDs, as well as the lack of an authoritative source of common file extensions to draw from, staff determined that it is not workable to prevent common file extensions from being used as TLDs.

To summarize, it is the recommendation of the ICANN technical staff to allow applications for TLD strings that may also be commonly used for file extensions.

Other Issues

ICANN invites feedback on any potential TLD strings or categories of strings that might cause technical instability which have not been considered here.

II Capacity of the Root Zone

In connection with its exploration of DNS stability, ICANN has engaged in an examination of possible technical limitations of the root zone. One way to approach this is by drawing comparisons between the root zone and other zones, while noting the similarities and differences. The difference between the root zone and any other zone lies in the way in which the root zone is found, rather than the characteristics of the zone itself. With the exception of the use of a root hints file to find the DNS servers for its zone, the root is technically similar to the delegations within it.

Currently, the largest top-level zone within the DNS is the .COM zone. As reported to ICANN by VeriSign in September of this year, over sixty million names are registered under .COM (see http://www.icann.org/tlds/monthly-reports/com-net/verisign-200709.pdf). Even if not all of these names are actually propagated to the zone, the size of the .COM zone indicates that it is technically possible to have a zone that has registrations numbering in the tens of millions. Similarly, zones such as .DE and .NET reinforce the technical feasibility of having zone files containing over a million domain names. At a minimum, the DNS should be able to function at its current level with at least 60 million TLDs. This allows significant room for large-scale expansion without concerns about a negative effect on stability.

Some of the technology used to distribute and manage a larger zone would, however, require changes to the current mechanisms. Currently, the standard AXFR (asynchronous full transfer zone) mechanism is used to move the root zone for the hidden master servers to the publicly available root name servers (a.root-servers.net through m.root-servers.net). As the size of the zone increases it will be required to change this method to an Incremental Zone Transfer or IXFR. All of the currently large zones use an incremental method of update to accommodate the larger zone file size. There would be significant administrative planning, work, and testing needed across all of the root server operators as well as the distribution master to make this transition for distribution of the root zone, but there is no technical limitation.

Another factor to be considered in the inclusion of a large number of TLDs in the root zone is increased deployment of the DNSSEC specifications. One exercise conducted recently indicated that the size ratio of the signed zone (with DS records) to the current zone is just under 4-to-1. The inclusion of a large number of signed TLD zones would require more time and effort to generate and publish the root zone, but this would not have any impact on performance or be apparent to the end user.

It could also be expected that with the increase in the number of TLDs, the number of look-ups to the servers would also increase as caching servers need to retrieve a greater number of records from the root servers. There will not be definitive data as to the level of increased look-ups resulting from a larger number of TLDs until these are actually added and become operational. However, the limited number of additions to the root zone over the course of ICANN's existence has not made a measurable impact at the servers. Any increase in traffic to the servers due to the addition of more TLDs is likely to be minimal, as the main source of traffic to root servers is not the number of TLDs but rather the number of end systems initiating queries. Additionally, the root servers provision far above normal capacities for security reasons (such as DDOS mitigation) and thus no technical issues are foreseen due to an increase in the number of queries.

To summarize, there is not currently any evidence to support establishing a limit to how many TLDs can be inserted in the root based on technical stability concerns. The staff has made a distinction between technical instability (that causes direct adverse impact to the DNS) and operational impacts which may not be harmful to the Internet technically, but do impose operational challenges in the management and operation of the DNS. Operational challenges are a limiting factor, and this is discussed below.

III Operational Capacity

Operational capacity is also a key part of any discussion on technical issues. One aspect of the operational impact of having a large number of TLDs is a corresponding increase in the resources needed to manage the root zone, both in creating a new TLD and maintaining the existing TLDs. Staff is currently reviewing the resources required for creating a new top-level domain and evaluating how many new TLDs could be added in a given period of time. In processing change requests, the IANA currently handles less than 25 changes to the root zone in a peak month, based on a zone with less than three hundred child delegations. Rounding

this off, in an extraordinary month 10% of the TLDs may require a change that uses administrative resources. In a typical month statistics show this to be about 5%. Another way to view this is to say that there is roughly one change per TLD per year. This work is currently undertaken by 1.5 FTE. Increases in the size of the zone will likely lead to a proportional increase in the number of changes requested. This would require additional staffing and a higher level of automation. However, many registry organizations have successfully accommodated similar patterns of growth.

Entities inside and outside of ICANN play critical roles in the change management processes of the root and would potentially need to increase resources to handle an increase in the number of TLDs in the root as well. These operational limits should also be a factor in determining what the optimum number of TLDs is. Within ICANN, legal, compliance, finance, and other staff play a role in maintaining and supporting the registry once the TLD exists. ICANN is examining the impact that the new gTLD program will have on the organization operationally. This involves operational readiness and operational impact analyses, including a review of each department's services and processes with a variety of projected volume scenarios. Additionally, ICANN is looking at the potential technical impact of business continuity issues with a large number of TLDs in the root. These are ongoing projects within ICANN and will inform the way that ICANN moves forward with the introduction of new gTLDs.

IV Conclusion

ICANN is issuing this paper to show its proposed approach to implementing the GNSO's recommendation on new gTLD strings and technical instability. As this paper indicates, it is possible to establish a clear set of rules that can be made available to applicants at the outset of the gTLD process. ICANN welcomes comments on any section of this paper. Comments should be submitted by **March 7 2008**.