NEW GTLD COLLISION RISK MITIGATION

Proposals to mitigate the collision risks between new gTLDs and existing private uses of the same strings

INTRODUCTION

ICANN's mission and core values call to preserve and enhance the operational stability, reliability, security, and global interoperability of the Internet's system of unique identifiers (names, IP numbers and protocol parameters). In pursuing these goals and following the direction of its Board of Directors as well as the advice of the Security and Stability Advisory Committee, ICANN commissioned a study on the potential security impacts of the applied-for new-gTLD strings. The study was to consider whether name collisions might occur between applied-for new gTLD strings and non-delegated TLDs that may be in use in private namespaces. The study was also to review the possibility of name collisions arising from the use of X.509 digital certificates.

The name collision study (the Study) identifies categories of strings according to the risk they represent. The Study also identifies options to mitigate the risks, however it does not make specific recommendations for each of the categories. Building upon the Study, this paper describes proposals to mitigate the risk of collisions between new gTLDs and existing private uses of the same strings. The risk mitigation proposals are tailored to the categories of strings identified in the Study. Familiarity with the Study will be helpful to better understand the proposal depicted here.

RISK PROFILES

The Study used as input: 1) samples of DNS requests to root servers from the "Day in the Life of the Internet" initiative from DNS-OARC and 2) information from Certificate Authorities regarding the issuance of internal name certificates (e.g., TLS/SSL certificates for un-delegated names). A full description of the methodology of the Study can be found in section 3.4 of the Study. From the Study the following risk profiles can be identified:

LOW-RISK (~80% OF STRINGS)

As described in the Study, a "*reasonable threshold for "low risk" could be established by reference to the number of queries for existing TLDs that are empty (meaning that their zones contain only the necessary DNS meta-data).*" In other words, applied-for new gTLDs that appear in the query stream at the root less frequently than existing TLDs with "empty zones" in the 2013 Day in The Life of the Internet (DITL) data used for the Study, will be considered to fit in the low-risk profile. This will include almost 80% of the proposed new gTLDs, i.e., the strings with frequency ranks between 282 and 1395, inclusive, as shown in Appendix B of the Study report.

HIGH-RISK (HOME, CORP)

The Study identifies two strings that would likely cause problems (as described in section 6 of the Study report) if delegated given their high frequency of appearance in queries to the root. Both ***home*** and ***corp*** will be considered high-risk strings given that they occur an order of magnitude more often in the 2012 and 2013 DITL data than the next most frequently occurring string. The Study identifies these strings as having a level of queries in the realm of heavily used TLDs. Additionally, in the case of ***corp***, it is identified as the string that has the most internal name certificates as shown in Appendix C of the Study report.

UNCALCULATED-RISK (20% OF STRINGS)

The remaining 20% of the strings, i.e., those between ranks 3 and 281, inclusive, as shown in Appendix B of the Study report will be considered part of the uncalculated-risk category. The Study did not find enough information to properly classify these strings given the short timeline.

PROPOSAL TO MITIGATE RISK

For each of the string risk profiles described in the previous section, a proposal to mitigate risk is presented below. A listing of the applied for strings in each category can be found here.

LOW-RISK

The Study establishes a low-risk profile for 80% of the strings. ICANN proposes to move forward with its established processes and procedures with delegating strings in this category (e.g., resolving objections, addressing GAC advice, etc.) after implementing two measures in an effort to mitigate the residual namespace collision risks.

First, registry operators will implement a period of no less than 120 days from the date that a registry agreement is signed before it may activate any names under the TLD in the DNS[1]. This measure will help mitigate the risks related to the internal name certificates issue as described in the Study report and SSAC Advisory on Internal Name Certificates. Registry operators, if they wish, may allocate names during this period, i.e., accept registrations, but they will not activate them in DNS. If a registry operator were to allocate names during this 120-day period, it would have to clearly inform the registrants about the impossibility to activate names until the period ends.

[1] Impact on TLD launch should be minimum in most cases since there are a set of activities that need to be completed between contracting and launch that account for a good fraction of 120 days.

Second, once a TLD is first delegated within the public DNS root to name servers designated by the registry operator, the registry operator will not activate any names under the TLD in the DNS for a period of no less than 30 days. During this 30-day period, the registry operator will notify the point of contacts of the IP addresses that issue DNS requests for an un-delegated TLD or names under it. The minimum set of requirements for the notification is described in Appendix A of this paper. This measure will help mitigate the namespace collision issues in general. Note that both no-activate-name periods can overlap.

The TLD name servers may see DNS queries for an un-delegated name from recursive resolvers – for example, a resolver operated by a subscriber's ISP or hosting provider, a resolver operated by or for a private (e.g., corporate) network, or a global public name resolution service. These queries will not include the IP address of the original requesting host, i.e., the source IP address that will be visible to the TLD is the source address of the recursive resolver. In the event that the TLD operator sees a request for a non-delegated name, it must request the assistance of these recursive resolver operators in the notification process as described in Appendix A.

HIGH-RISK

ICANN considers that the Study presents sufficient evidence to classify ***home*** and ***corp*** as high-risk strings. Given the risk level presented by these strings, ICANN proposes not to delegate either one until such time that an applicant can demonstrate that its proposed string should be classified as low risk based on the criteria described above. An applicant for one of these strings would have the option to withdraw its application, or work towards resolving the issues that led to its categorization as high risk (i.e., those described in section 7 of the Study report). An applicant for a high-risk string can provide evidence of the results from the steps taken to mitigate the name collision risks to an acceptable level. ICANN may seek independent confirmation of the results before allowing delegation of such string.

UNCALCULATED-RISK

For the remaining 20% of the strings that do not fall into the low or high-risk categories, further study is needed to better assess the risk and understand what mitigation measures may be needed to allow these strings to move forward. The goal of the study will be to classify the strings as either low or high-risk using more data and tests than those currently available. While this study is being conducted, ICANN would not allow delegation of the strings in this category. ICANN expects the further study to take between three and six months. At the same time, an applicant for these strings can work towards resolving the issues that prevented their proposed string from being categorized as low risk (e.g., those described in section 7 of the Study report). An applicant can provide evidence of the results from the steps taken to mitigate the name collision risks to an acceptable level. ICANN may seek independent confirmation of the results before allowing delegation of such string. If and when a string from this category has been reclassified as low-risk, it can proceed as described above for the low-risk category strings.

CONCLUSION

ICANN is fully committed to the delegation of new gTLDs in a secure and stable manner. As with most things on the Internet, it is not possible to eliminate risk entirely. Nevertheless, ICANN would only proceed to delegate a new gTLD when the risk profile of such string had been mitigated to an

acceptable level. We appreciate the community's involvement in the process and look forward to further collaboration on the remaining work.

APPENDIX A – NOTIFICATION REQUIREMENTS

Registry operator will notify the point of contact of each IP address block that issue any type of DNS requests (the Requestors) for names under the TLD or its apex. The point of contact(s) will be derived from the respective Regional Internet Registry (RIR) database. Registry operator will offer customer support for the Requestors or their clients (origin of the queries) in, at least, the same languages and mechanisms the registry plans to offer customer support for registry services. Registry operator will avoid sending unnecessary duplicate notifications (e.g. one notification per point of contact).

The notification should be sent, at least, over email and must include, at least the following elements: 1) the TLD string; 2) why the IP address holder is receiving this email; 3) the potential problems the Requestor or its clients could encounter (e.g., those described in section 6 of the Study report); 4) the date when the gTLD signed the registry agreement with ICANN, and the date of gTLD delegation; 5) when the domain names under the gTLD will first become active in DNS; 6) multiple points of contact (e.g. email address, phone number) should people have questions; 7) will be in English and may be in other languages the point of contact is presumed to know; 8) ask the Requestors to pass the notification to their clients in case the Requestors are not the origin of the queries, e.g., if they are providers of DNS resolution services; 9) a sample of timestamps of DNS request in UTC to help identify the origin of queries; 10) email digitally signed with valid S/MIME certificate from well-known public CA.