

Mitigating the Risk of DNS Namespace Collisions

*A Study on Namespace Collisions in the Global Internet DNS
Namespace and a Framework for Risk Mitigation*

Phase One Report



4 JUNE 2014

TABLE OF CONTENTS

| | | |
|----------|--|-----------|
| 1 | Discussion of Public Comments and Revisions | 1 |
| 2 | Summary and Preface to Phase One Report..... | 2 |
| 2.1 | Summary of Recommendations | 6 |
| 2.2 | Acknowledgements | 8 |
| 3 | Detection and Response | 9 |
| 3.1 | Approach to Delegation | 14 |
| 3.2 | Root Level Data, Monitoring, and Day-In-The-Life (DITL) | 27 |
| 3.3 | Collisions in Existing DNS Namespace | 30 |
| 3.4 | Description of Forthcoming Phase Two Report..... | 31 |



1 Discussion of Public Comments and Revisions

JAS would like to sincerely thank all of the individuals that have participated in the review and comment process. We received significant and valuable feedback from the public draft and numerous other discussions since the initial release of our document in February. We have amended our report accordingly. Specifically, we have addressed the following issues:

- Discussion of IPv6-related issues (see new Section 3.1.3);
- Recognition of emergent data and experience (see new text in Section 2);
- Additional discussion concerning the implementation tradeoffs of using a 127/8 IP vs. an Internet IP (“Honeypot”) for Controlled Interruption (see additions to Section 3.1.7);
- Reduction of Controlled Interruption period to 90 days (see new Section 3.1.2);
- Additional description of our findings regarding probability and severity of possible impacts resulting from name collision occurrences (see new text in Section 3.1.6);
- Discussion of staggered vs. consistent introduction of Controlled Interruption (see new text in Section 3);
- Recommendation to collect additional logs to support long-term measurement of the collisions phenomena (see new text in Section 3.2);
- Description of content that is expected to appear in our Phase Two report (see new section 3.4); and
- Other minor modifications, improvements, and elaborations throughout.

2 Summary and Preface to Phase One Report

Collisions in the global Domain Name System (DNS) namespace have the potential to expose serious security-related issues for users of the DNS. This report dives right into the technical discussion and is targeted at readers who have been following the issue. Those new to the issue should first read the introductory documents located at: <http://www.icann.org/en/help/name-collision>.

We do not find that the addition of new Top Level Domains (TLDs) fundamentally or significantly increases or changes the risks associated with DNS namespace collisions. The modalities, risks, and etiologies of the inevitable DNS namespace collisions in new TLD namespaces will resemble the collisions that already occur routinely in the other parts of the DNS. The addition of multiple new TLDs over the past decade (generic and country code) has not suggested that new failure modalities might exist; rather, the indication is that the failure modalities are similar in all parts of the DNS namespace. Our research has shown that a very few root causes are responsible for nearly all collisions, and these root causes appear in nearly every classification of TLD, albeit in varying proportions.

That said, DNS namespace collisions are a complex and pervasive occurrence that manifests throughout the global Internet DNS namespace. Collisions in all TLDs and at all levels within the global Internet DNS namespace have the ability to expose potentially serious security and availability problems and deserve serious attention. While current efforts to expand the global DNS namespace have collision-related implications, the collision problem is bigger than new TLDs and must be viewed in this context.

In summary, our recommendations describe a comprehensive approach to reducing current and future DNS namespace collisions, alerting operators of potential DNS namespace related issues, and providing emergency response capabilities in the event that critical (e.g., life safety) systems are adversely impacted.

DNS namespace collisions exist outside of, and independently from, the current efforts to expand the DNS namespace. These collisions have almost certainly existed since the emergence of a global public DNS. As early as 2003, multiple researchers have pointed to the existence of queries into undelegated space received at the root.^{1,2,3,4} Our research shows that every TLD that has been added to the root since

¹ *Understanding DNS Evolution*, Castro, Zhang, John, Wessels, claffy, 2010, http://www.caida.org/publications/papers/2010/understanding_dns_evolution/understanding_dns_evolution.pdf

² *Is Your Caching Resolver Polluting the Internet?*, Wessels, 2004, <http://dns.measurement-factory.com/writings/wessels-netts2004-paper.pdf>

³ *RFC 4697: Observed DNS Resolution Misbehavior*, Larson, Barber, 2006, <http://tools.ietf.org/html/rfc4697>



consistent data collection has occurred (2007) has exhibited some symptoms of collision activity prior to delegation.

The issue of collisions is not specific to TLDs; rather, risk exists wherever a collision crosses an administrative control boundary in the DNS. Said differently, the most dangerous DNS namespace collisions occur when *the resulting DNS query is resolved by a different administrative party than expected by the querier*. This makes intuitive sense. Because of the hierarchical nature of the DNS, the vast majority of administrative control separations occur at the TLD and Second Level Domain (2LD) levels.

Over the course of the study, JAS found no evidence to suggest that the security and stability of the global Internet DNS itself is at risk. This finding confirms the results of the *DNS Stability String Review* performed on each string during Initial Evaluation pursuant to Section 2.2.1.3.1 of the Applicant Guidebook (AGB).^{5,6} The remainder of our research is focused on issues from the perspective of end-systems as consumers of the global DNS.

When faced with a range of unknowns and hypotheticals, it is important not to overlook emergent facts and experience. As we write this update, 275 New gTLDs have been delegated and over 835,000 second level registrations have been added. TLDs representative of the complete range of the taxonomy JAS developed (see Section 3.4) are represented. .berlin – a geographic term that our research suggests is heavily present in DNS search paths – has the third largest number of registrations of all new TLDs. .email and .link – short, technology-oriented generic terms that our research suggests are present in a number of hardcoded configurations – rank 6th and 7th respectively, each with over 30,000 2LD registrations. .company, .solutions, and .agency – terms that our research suggests are commonly hardcoded into small business-oriented configurations – are also delegated and have thousands of registrations each. Neither JAS nor ICANN is aware of even a single instance of a problematic collision. Of course, this fact certainly doesn't "prove the negative" but it also can't be ignored at this point.

Certainly the nature of the string impacts the drivers behind colliding behavior. As we presented at Verisign's *Workshop and Prize on Root Causes and Mitigation of*

⁴ *Wow, that's a lot of packets*, Wessels, Fomenkov, 2003, <http://www.caida.org/publications/papers/2003/dnspackets/wessels-pam2003.pdf>

⁵ *gTLD Applicant Guidebook*, ICANN, 2012, <http://newgtlds.icann.org/en/applicants/agb>

⁶ The process followed by ICANN's vendor for this review, Interisle Consulting Group, process is documented at <http://newgtlds.icann.org/en/program-status/evaluation-panels/dns-stability-process-07jun13-en.pdf>



Name Collisions (WPNC)⁷ in London, strings with the potential to introduce new failure etiologies have been introduced into the TLD in the past. .post (delegated in 2012) saw the most collision activity prior to delegation of any of the nine TLDs added since 2007. .post is interesting because “post” is also an HTTP method and a not insignificant proportion of the collisions appeared to be related to erroneous DNS lookups of text intended to be transmitted to an HTTP server. History provides lessons and data regarding the introduction of a variety of strings at the TLD.

We believe the introduction of new TLDs offers an opportunity to educate operators regarding DNS namespace collisions and help find and remedy potential collision-related issues that may be present in their systems. As such, we recommend implementation of a 90-day “controlled interruption” period for all approved new TLDs with the exception of .corp, .home, and .mail. Registries that have not yet been delegated to the root zone shall implement controlled interruption via wildcard records; registries that have elected the “alternative path to delegation” shall implement controlled interruption by adding appropriate resource records for the labels appearing in their respective block lists. Following the 90-day controlled interruption period, registries will not be subject to further collision-related restrictions. Like the Certificate Authority (CA) revocation approach, which may be partially implemented in parallel, we believe the 90-day controlled interruption period offers a conservative buffer between potential legacy usage of a TLD and the new usage.

Lacking clear RFC 1918-like guidance directing operators to DNS namespaces safe for internal use, several such namespaces have been “appropriated” for this purpose over the years. While the etiology is subtly different, the .corp and .home TLDs are clear outliers in this respect; the use of .corp and .home for internal namespaces/networks is so overwhelming that the inertia created by such a large “installed base” and prevalent use is not likely reversible. We also note that RFC 6762 suggests that .corp and .home are safe for use on internal networks.⁸

Given that the Internet has demonstrated a need for RFC 1918-like DNS namespaces, we recommend that .corp and .home be referred to the Internet Engineering Task Force (IETF) for potential RFC 1918-like protection/treatment.⁹

⁷ <http://namecollisions.net>

⁸ *RFC 6762: Multicast DNS* (appendix G), Cheshire, Krochmal, 2013, <http://tools.ietf.org/html/rfc6762>

⁹ [RFC 6761](#) may be the appropriate vehicle for implementing a permanent reservation.



Like .corp and .home, the TLD .mail also exhibits prevalent, widespread use at a level materially greater than all other applied-for TLDs. Our research found that .mail has been hardcoded into a number of installations, provided in a number of example configuration scripts/defaults, and has a large global “installed base” that is likely to have significant inertia comparable to .corp and .home. As such, we believe .mail’s prevalent internal use is also likely irreversible and recommend reservation similar to .corp and .home and similarly recommend ICANN not delegate that TLD at this time.

RECOMMENDATION 1: The TLDs .corp, .home, and .mail be referred to the Internet Engineering Task Force (IETF) for potential RFC 1918-like protection/treatment.

JAS uncovered a vulnerability not directly related to ICANN’s New gTLD Program nor to new TLDs in general that has the potential to impact end-systems. Pursuant to ICANN’s Coordinated Vulnerability Disclosure Process,¹⁰ ICANN shall: “...privately disclose information relating to a discovered vulnerability to a product vendor or service provider (“affected party”) and allow the affected party time to investigate the claim, and identify and test a remedy or recourse before coordinating the release of a public disclosure of the vulnerability with the reporter.” Furthermore, ICANN’s process states: “All parties to the disclosure generally agree to refrain from disclosing the vulnerability to the public until a remedy is identified and tested or until the threat is considered contained.”

After extensive discussions with impacted vendors and ICANN executives, JAS is concerned that publication of the experimental methods and data contained in the complete JAS report may accelerate discovery of the vulnerability and/or serve to facilitate exploitation of the vulnerability after it is discovered. As such, pursuant to ICANN’s process and out of an abundance of caution, JAS has recommended against publication of a complete report at this time.

A description of our expected Phase Two report appears in a section 3.4; the Phase Two report will be published as soon as it is prudent.

¹⁰ *Coordinated Vulnerability Disclosure Reporting at ICANN*, ICANN, 2013, <https://www.icann.org/en/about/staff/security/vulnerability-disclosure-05aug13-en.pdf>



2.1 Summary of Recommendations

RECOMMENDATION 1: The TLDs .corp, .home, and .mail be referred to the Internet Engineering Task Force (IETF) for potential RFC 1918-like protection/treatment.

RECOMMENDATION 2: ICANN continue efforts to make technical information available in fora frequented by system operators (e.g., network operations groups, system administration-related conferences, etc.) regarding the introduction of new gTLDs and the issues surrounding DNS namespace collisions.

RECOMMENDATION 3: Emergency response options are limited to situations where there is a reasonable belief that the DNS namespace collision presents a clear and present danger to human life.

RECOMMENDATION 4: Root-level de-delegation of a production TLD is not considered as an emergency response mechanism under any circumstances.

RECOMMENDATION 5: ICANN leverage the EBERO mechanisms and functionality to respond to DNS namespace-related issues. ICANN must have the following capabilities on a 24x7x365, emergency basis: 1). Analyze a specific report/incident to confirm a reasonable clear and present danger to human life; 2). Direct the registry on an emergency basis to alter, revert, or suspend the problematic registrations as required by the specific situation; 3). Ensure that the registry complies in a timely manner; and 4). Evaluate and monitor the specific situation for additional required actions. Furthermore, we recommend that ICANN develop policies and procedures for emergency transition to an EBERO provider in the event the registry is unable and/or unwilling to comply. We recommend ICANN maintain this capability indefinitely.

RECOMMENDATION 6: ICANN require new TLD registries to publish the controlled interruption zone immediately upon delegation in the root zone. After the 90-day period, there shall be no further collision-related restrictions on the registry.

RECOMMENDATION 7: ICANN require registries that have elected the “alternative path to delegation” rather than a wildcard, instead publish appropriate A and SRV resource records for the labels in the ICANN 2LD Block List to the TLD’s zone with the 127.0.53.53 address for a period of 90 days. After the 90-day period, there shall be no further collision-related restrictions on the registry.

RECOMMENDATION 8: ICANN relieve the prohibition on wildcard records during the controlled interruption period.

RECOMMENDATION 9: ICANN monitor the implementation of controlled interruption by each registry to ensure proper implementation and compliance.



RECOMMENDATION 10: ICANN work with the IETF to identify a mechanism for IPv6 that provides similar functionality to that available in IPv4's "localhost" reserved prefix.

RECOMMENDATION 11: ICANN, DNS-OARC, and the root operators explore a medium-latency, aggregated summary feed describing queries reaching the DNS root.

RECOMMENDATION 12: ICANN, DNS-OARC, and the root operators explore establishment of a single, authoritative, and publicly available archive for historical data related to the root.

RECOMMENDATION 13: ICANN explore collecting NXDOMAIN entries in DNS query logs from registry operators and contribute them to an independent data repository such as DNS-OARC for further analysis. To limit the potential for commercial gaming or use by malicious parties, we recommend that logs be provided six months in arrears.

RECOMMENDATION 14: ICANN request that the appropriate bodies further explore issues relating to collisions in existing DNS namespace, the practice of "domain drop catching," and the associated data feeds that may be leveraged by attackers when attempting to exploit collisions.



2.2 Acknowledgements

JAS is grateful for the constructive engagement by numerous members of the community. We specifically want to recognize and thank:

- the Security and Stability Advisory Committee (SSAC) for thoughtful and valuable interaction while we drafted this report;
- Burt Kaliski and his team at [Verisign Labs](#), for extensive and insightful public comments, valuable interaction with the JAS team throughout our study, and for their overall leadership on this issue including hosting the *Workshop and Prize on Root Causes and Mitigation of Name Collisions* (WPNC) in London;
- [Farsight Security](#) for contributing valuable data;
- [OpenRegistry](#) for contributing valuable data; and
- our longtime partner [simMachines](#) for their analytical contributions.



3 Detection and Response

Since risk cannot be totally eliminated, a comprehensive approach to risk management contains some level of *a priori* risk mitigation combined with investment in detection and response capabilities. Consider fire protection; most major cities have *a priori* protection in the form of building codes, detection in the form of smoke/fire alarms, and response in the form of 9-1-1, sprinklers, and the fire department.

In terms of detecting problematic DNS namespace collisions, the initial symptoms will almost certainly appear through various IT support mechanisms, namely corporate IT departments and the support channels offered by hardware/software/service vendors and Internet Service Providers. When presented with a new and non-obvious problem, professional and non-professional IT practitioners alike frequently turn to Internet search engines for answers. This suggests that a good detection/response investment would be to “seed” support vendors/fora with information/documentation about this issue in advance and in a way that will surface via search engines when IT folks begin troubleshooting. We collectively refer to such documentation as “self-help” information. ICANN has already begun developing documentation designed to assist IT support professionals with namespace-related issues.¹¹

RECOMMENDATION 2: ICANN continue efforts to make technical information available in fora frequented by system operators (e.g., network operations groups, system administration-related conferences, etc.) regarding the introduction of new gTLDs and the issues surrounding DNS namespace collisions.

One valuable suggestion from Google in the public comment period¹² is to stagger introduction of the Controlled Interruption periods such that impacted parties have a reprieve between the detection and mitigation phases of their response. However, staggered Controlled Interruption periods will have the side effect of causing intermittent failures, which are maddening and hard to diagnose from a system administrator perspective. Moreover, we found that systems configured in a way to create collision-related effects in the existing DNS namespaces routinely experience and tolerate intermittent failures (for example, when using a different DNS resolver) so intermittent failures are likely to resemble the status quo for impacted systems, and not communicate a problem. We believe a sustained and consistent Controlled Interruption period is the best opportunity to communicate with administrators.

¹¹ *Name Collision Resources & Information*, ICANN, retrieved January 2014, <http://www.icann.org/en/help/name-collision>

¹² <http://forum.icann.org/lists/comments-name-collision-26feb14/pdfBGWsaf8Vuk.pdf>



However, providing advice to system administrators regarding technical mechanisms they may deploy to temporarily gain reprieve during Controlled Interruption is valuable. Such advice may include the use of Response Policy Zones to temporarily rewrite query responses to something non-problematic (presumably NXDOMAIN), temporarily becoming authoritative for certain zones, etc. We recommend ICANN augment the existing technical advice to system administrators with such temporary remediation information and techniques.

It is likely that in the vast majority of expected cases, the IT professional “detectors” will also be the “responders” and any issues detected will be resolved without involving other parties.¹³ However, situations in which other parties may be expected to have a role in response must be considered.

For the sake of this discussion, assume that an Internet user is experiencing a problem related to a DNS namespace collision. The term “Internet user” is intended broadly as any application, system, or device that is a consumer of the global Internet DNS. At this point in the thought experiment, disregard the severity of the problem. The affected party (or parties) will likely exercise the full range of typical IT support options available to them – vendors, professional support, IT-savvy friends and family, and Internet search. If any of these support avenues are aware of ICANN, they may choose to contact ICANN at some point. Let’s further assume the affected party is unable and/or unwilling to correct the technical problem themselves and ICANN is contacted – directly or indirectly.

There is a critical fork in the road here: Is the expectation that ICANN will provide technical “self-help” information or that ICANN will go further and “do something” to technically remedy the issue for the user? We consider the options below in an escalating progression:

Option 1: ICANN provides technical support above and beyond “self-help” information to the impacted parties directly, including the provision of services/experts. Stated differently, ICANN becomes an extension of the impacted party’s IT support structure and provides customized/specific troubleshooting and assistance. *We rule out this option as inappropriate and out-of-scope for ICANN.*

Option 2: At ICANN’s request, referral, or direction, the registry provides technical support above and beyond “self-help” information to the impacted parties directly, including the provision of services/experts. Stated differently, the registry becomes an extension of the impacted party’s IT support structure and provides customized/specific troubleshooting and assistance. *We rule out this option as inappropriate and out-of-scope for a registry.*

¹³ Availability issues are typically detected internally whereas security issues are often detected by third parties and reported to the system operators.



Option 3: ICANN forwards the issue to the registry with a specific request to remedy. In this option, assuming all attempts to provide “self-help” are not successful, ICANN would request that the registry make changes to their zone to technically remedy the issue. This could include temporary or permanent removal of second level names and/or other technical measures that constitute a “registry-level rollback” to a “last known good” configuration. *We consider this option feasible but undesirable as it creates considerable opportunity for operational complexities and unintended consequences. This option should only to be used in excessively serious circumstances.*

Option 4: ICANN initiates a “root-level rollback” procedure to revert the state of the root zone to a “last known good” configuration, thus (presumably) de-delegating the impacted TLD. In this case, ICANN would attempt – on an emergency basis – to revert the root zone to a state that is not causing harm to the impacted party/parties. *We consider this option feasible but even more undesirable as it creates considerable opportunity for operational complexities and unintended consequences. This option should only to be used in excessively serious circumstances after all previous mitigation attempts have failed.*

We note that ICANN’s New gTLD Collision Occurrence Management Plan and SAC062 contemplate some of these emergency response options in a broad sense.

In any theater of operations – not just the global Internet DNS – emergency responders must be mindful of “cure is worse than the disease” scenarios wherein the response actually creates additional risks, harms, and significant potential for unintended consequences. Because of the potential operational impacts to the global Internet DNS, changes to the root zone are not to be taken lightly.

From a practical perspective, we conclude that the de-delegation of a TLD in the root would effectively be a permanent death for that TLD regardless of whether the TLD reappeared in the future.¹⁴ This is a steep price for a registry to pay for anything but the most egregious and flagrant disregard for a serious harm.

Obviously, the severity of the harm is a critical variable. In risk analysis, severity is almost always measured economically and from multiple points of view. Any party expected to “do something” will be forced to choose between two or more economically motivated actors: users, registrants, registrars, and/or registries experiencing harm. We must also consider that just as there may be users negatively impacted by new DNS behavior, there may also be users that are

¹⁴ While we note that there has always been some degree of churn in the root zone, the commercial pressures on the current new gTLDs significantly elevate the impact of a de-delegation, no matter how short.

dependent upon on the new DNS behavior. Unfortunately, we cannot give equal consideration to actors that are following the technical standards vs. those depending on technical happenstance or poorly implemented software for proper functionality.

Even attempting to weigh economic harm or “national security” on a global basis creates a slippery slope and forces registries and ICANN to arbitrate impossible scenarios. Concepts like “national security,” “law and order,” and “key economic processes” do not translate well on a global basis and risk another “Morality and Public Order” debate – which is exactly what happened when similar terms were introduced into the ICANN landscape previously. Unfortunately, there will not be time for such a debate in real-time, leaving emergency responders forced to make rapid decisions concerning extremely serious issues – like root-level changes – in a non-deterministic state.

Moreover, an emergency response threshold that is not well defined risks weaponization of the mechanism by commercial or government interests. Sadly, recent history has shown some governments will use a full range of tools to silence distribution of certain viewpoints over Internet channels. It is also reasonable to assume that commercial interests will attempt to “game” any mechanism for competitive advantage.

As such, we recommend that emergency response be limited to scenarios where there is a reasonable belief that the DNS namespace collision presents a clear and present danger to human life. While admittedly a high bar, we believe it is the only deterministic and non-debatable option. We feel creating a path to emergency response (including root-level changes) based on lesser factors is unwise.

Despite the previous recommendation, ICANN must prepare for the worst-case scenario. Fortunately, ICANN has already developed an emergency response mechanism as a part of the Emergency Back-End Registry Operator (EBERO) Program. The EBERO Program is designed to quickly respond to a variety of registry-level technical SLA failures; response options include an emergency (and potentially involuntary) transition of an entire registry to a new operator using a robust process that is highly scripted and exercised.

We recommend that, if necessary (in the event of an unresponsive or non-cooperative registry), a “root-level rollback” be implemented via EBERO as opposed to simply removing a TLD from the root. Shifting a registry to EBERO and making subsequent surgical changes is a superior approach to wholesale removal of an entire production TLD – including potentially many 2LD registrations that are not causing harm.



RECOMMENDATION 3: Emergency response options are limited to situations where there is a reasonable belief that the DNS namespace collision presents a clear and present danger to human life.

RECOMMENDATION 4: Root-level de-delegation of a production TLD is not considered as an emergency response mechanism under any circumstances.

In the case of severe harm being exposed by a DNS namespace collision where the registry is unable or unwilling to take action (by altering or suspending a second level registration), ICANN could transfer the registry to an EBERO on an emergency basis and instruct the EBERO to make the required second level change to remedy the harm. While we recognize any “root-level rollback” is highly undesirable, ICANN should maintain the capability, thus ensuring that timely action can be taken in all circumstances.

RECOMMENDATION 5: ICANN leverage the EBERO mechanisms and functionality to respond to DNS namespace-related issues. ICANN must have the following capabilities on a 24x7x365, emergency basis: 1). Analyze a specific report/incident to confirm a reasonable clear and present danger to human life; 2). Direct the registry on an emergency basis to alter, revert, or suspend the problematic registrations as required by the specific situation; 3). Ensure that the registry complies in a timely manner; and 4). Evaluate and monitor the specific situation for additional required actions. Furthermore, we recommend that ICANN develop policies and procedures for emergency transition to an EBERO provider in the event the registry is unable and/or unwilling to comply. We recommend ICANN maintain this capability indefinitely.



3.1 Approach to Delegation

The delegation of new TLDs presents a unique opportunity to raise awareness of the DNS namespace collision issue and help system operators identify and mitigate potential issues. Therefore, we recommend a “controlled interruption” approach as described below. The idea for controlled interruption springs from past DNS-related experiences and is conceptually similar to a “trial delegation” as proposed in SAC062.

3.1.1 Controlled Interruption

The infamous Microsoft Hotmail domain expiration in 1999¹⁵ and other similar domain expirations led to the implementation of ICANN’s Expired Registration Recovery Policy.

More recently, Regions Bank made news¹⁶ when their domains expired, and undoubtedly countless other similar events go unreported. In the case of Regions Bank, the Expired Registration Recovery Policy seemed to work exactly as intended – the interruption inspired immediate action and the problem was solved, resulting in only a bit of embarrassment. Importantly, there was no opportunity for malicious activity.

For the most part, the Expired Registration Recovery Policy is effective at preventing unintended expirations due to the application of “controlled interruption.” The Expired Registration Recovery Policy calls for extensive notification before the expiration, then a period when “the existing DNS resolution path specified by the Registrant at Expiration (“RAE”) must be interrupted” – as a last-ditch effort to inspire the registrant to take action.

Nothing inspires urgent action more effectively than service interruption.

But critically, in the case of the Expired Registration Recovery Policy, the interruption is immediately corrected if the registrant takes the required action – renewing the registration. It’s nothing more than another notification mechanism – just a more aggressive round after all of the passive notifications failed. In the case of a registration in active use, the interruption will be recognized immediately, inspiring urgent action.

¹⁵ *Good Samaritan squashes Hotmail lapse?*, Hansen/CNET, December 27, 1999, retrieved January 2014, <http://news.cnet.com/2100-1023-234907.html>

¹⁶ *Regions Bank website down, domain not renewed?*, Walsh/al.com, April 15, 2013, retrieved January 2014, http://www.al.com/business/index.ssf/2013/04/regions_bank_website_down_do_ma.html



Like unintended expirations, DNS namespace collisions can be viewed as a notification problem. The system administrator utilizing the colliding namespace (either knowingly or unknowingly) must be notified and take action to preserve the security and stability of their systems.

Leveraging a controlled interruption to raise awareness of DNS namespace collisions draws on the effectiveness of the Expired Registration Recovery Policy with the implementation looking like a modified “Application and Service Testing and Notification (Type II)” trial delegation as proposed in SAC62. But instead of responding with pointers to application layer listeners (or “honeypots”), the authoritative nameserver responds with an address inside 127/8 – the range reserved for Loopback. We recommend this approach be applied to A queries directly and MX and SRV queries via an intermediary A record (the vast majority of collision behavior observed in DITL data stems from A and MX queries).¹⁷

Responding with an address inside 127/8 will likely interrupt any application depending on an NXDOMAIN or some other response, but importantly also prevents traffic from leaving the requestor’s host and does not facilitate a malicious actor’s ability to intercede. In the same way as the Expired Registration Recovery Policy calls for “the existing DNS resolution path specified by the RAE [to] be interrupted”, responding with a localhost reserved address should encourage immediate action by the requesting party while not exposing them to new malicious activity.

If legacy/unintended use of a DNS name is present, one could think of controlled interruption as a “buffer” or “cooling-off” period prior to use by a legitimate new registrant. This is similar to the CA Revocation Period as proposed in the New gTLD Collision Occurrence Management Plan that “buffers” the legacy use of certificates in internal namespaces from new use in the global DNS. As we discussed at ICANN Singapore, and Verisign’s *Workshop and Prize on Root Causes and Mitigation of Name Collisions* (WPNC) in London, 30 to 90 day buffer periods are also commonly deployed in other large important namespaces like postal and phone numbering systems to provide feedback when changes occur. Like the CA Revocation Period approach, a set period of controlled interruption is deterministic for all parties. Unfortunately, human nature often requires a hard deadline to inspire urgent action.

Moreover, instead of using the typical 127.0.0.1 address for localhost, we recommend using a unique “flag” IPv4 address: 127.0.53.53. Because the primary objective is to communicate with system administrators through their logs, this unique and strange IP should stand out in log files, be noticed, and result in the administrator searching the Internet for assistance (we note that as of today, using Google to search for “127.0.53.53,” the top 5 results are relevant). Making it known

¹⁷ AAAA query load suggests that collisions related to IPv6 space are far less pervasive.



that new TLDs will behave in this fashion and publicizing the flag IP (along with self-help materials) will help administrators isolate the problem more quickly than just using the common 127.0.0.1. As hosts often have listening sockets bound to 127.0.0.1, this approach also reduces the probability of creating issues related to those servers. We also suggest that system administrators proactively search their logs for this flag IP address as a possible indicator of problems. Enterprise-wide sensors in the form of DNS query log analysis or Network Intrusion Detection Systems (NIDS) such as SNORT provide an enterprise perspective.

Numerous experiments performed by JAS confirmed that a wide range of application layer software logs something resembling a “failed connection attempt to 127.0.53.53” which is the desired behavior. We also confirmed that all modern Microsoft, Linux, Apple, and BSD-derived operating systems correctly implement RFC 1122 (albeit with variations¹⁸) and keep the traffic within the host system, not transmitted over the network. This includes Linux and Windows-derived embedded operating systems. Of particular importance is Windows XP because our research has indicated that Windows XP is used extensively in industrial control and other embedded systems.

Additionally, we encourage ICANN and the IETF to work with software vendors eventually incorporate functionality and tools to notice DNS queries that respond with this flag IP address and provide meaningful assistance. One could imagine a meaningful event in the Windows Event Log describing the situation if a DNS query returns the flag IP, browsers displaying helpful diagnostic information instead of simply stating “Connection Timeout,” etc.

The ability to “schedule” the controlled interruption serves to further mitigate possible effects. One concern in dealing with collisions is the reality that a potentially harmful collision may not be identified until months or years after a TLD goes live – when a particular second level string is registered. A key advantage to applying controlled interruption to all second level strings in a given TLD in advance and at once via wildcard is that most failure modes will be identified during a scheduled time and before a registration takes place. This has many positive features, including easier troubleshooting and the ability to execute a far less intrusive rollback if a problem does occur. From a practical perspective, avoiding a complex string-by-string approach is also valuable.

The Expired Registration Recovery Policy mandates that the disruption may be for as few as eight days. However, our experiments indicate that the disruptions

¹⁸ Some implementations route the entire /8 to localhost whereas other implementations use a host route resulting in only a /32 being dedicated to localhost. The resulting behavior during a connection attempt is slightly different, but indicative of failure in both cases.

associated with controlled interruption as proposed may be more subtle, justifying a longer disruption period.

We believe the 90-day CA Revocation Period is sufficiently conservative (recall, we characterized our initial recommendation – 120 days – as “exceedingly conservative”). Given the potential seriousness of DNS namespace collisions and the immense value of detecting a harmful collision prior to a registry entering General Availability (GA), we believe the conservative approach is also warranted and recommend a 90-day controlled interruption period.

If there were to be a catastrophic impact, a surgical reversal of a 2LD registration could be implemented relatively quickly, easily, and with low risk while the impacted parties worked on a long-term solution. A new registrant and associated new dependencies would likely not be adding complexity at this point. Our recommended 90-day controlled interruption period is an ample and conservative detection and cure period for impacted parties.

Implementation of controlled interruption achieves these objectives:

- Helps notify system administrators of possible improper use of the global DNS;
- Protects these systems from malicious actors during a cure period;
- Doesn't direct potentially sensitive traffic to registries, registrars, Internet hosts/honeypots, or other third parties;
- Inspires urgent remediation action;
- Is low risk with limited opportunity for unintended consequences; and
- Is easy to implement and deterministic for all parties.

We therefore recommend controlled interruption be implemented by each new TLD registry by publishing a zone similar to the following:

```
$ORIGIN TLD
$TTL 1H
@      IN      MX 10 your-dns-needs-immediate-attention
*      IN      MX 10 your-dns-needs-immediate-attention
@      IN      SRV 10 10 0 your-dns-needs-immediate-attention
*      IN      SRV 10 10 0 your-dns-needs-immediate-attention
@      IN      TXT "Your DNS configuration needs immediate attention see URL"
*      IN      TXT "Your DNS configuration needs immediate attention see URL"
@      IN      A 127.0.53.53
*      IN      A 127.0.53.53
```

We note that some versions of popular DNS servers (notably BIND¹⁹) do not properly validate DNSSEC signed query responses to wildcards in all cases.

¹⁹ Bug 390 - NSD does not return closest provable enclosure NSEC3 on wildcard queries, NLnet Labs, May 26, 2011, retrieved January 2014, https://www.nlnetlabs.nl/bugs-script/show_bug.cgi?id=390; also note ISC RT ticket #26200



However, we also note the potential difficulties and confusion that could arise when treating the controlled interruption zones differently than production zones from an operational perspective. We have considered the tradeoffs and recommend that registries DNSSEC sign the controlled interruption zone using the same policies and procedures they intend to use when the zone is in production. A client downstream of a flawed DNS server may in some situations be “interrupted” due to the DNS server’s inability to validate the signature as opposed to an interruption due directly to controlled interruption.

We recommend that the registry implement the controlled interruption period immediately upon delegation in the root zone and the prohibition on wildcard records be temporarily suspended during this period. Given the objective of controlled interruption and the reality that no registrant data will be in the zone at this point, we believe that temporarily permitting wildcard records for this purpose is not counter to established ICANN prohibitions on wildcard records and does not raise the concerns that lead ICANN to establish these prohibitions.²⁰

RECOMMENDATION 6: ICANN require new TLD registries to publish the controlled interruption zone immediately upon delegation in the root zone. After the 90-day period, there shall be no further collision-related restrictions on the registry.

However, implementing a wildcard record is not prudent for a registry in GA. As such, we recommend publishing A and SRV resource records for labels in the ICANN 2LD Block List for the 90-day controlled interruption period. While arguably not an exhaustive list of queries, the 2LD block lists as currently constructed provide an adequate inventory^{21,22} of queries sent by long-lived systems, which are the ones of most concern. The alternative – wildcard records in production zones – is less attractive and counter to established ICANN prohibitions.²³

²⁰ *SSAC Report: Redirection in the com and net Domains*, ICANN Security and Stability Advisory Committee (SSAC), July 9, 2004, retrieved January 2014, <http://www.icann.org/en/groups/ssac/report-redirection-com-net-09jul04-en.pdf>

²¹ *Public Comments on Proposal to Mitigate Name Collision Risks by Google Inc.*, Google Inc., September 17, 2013, retrieved January 2014, <http://forum.icann.org/lists/comments-name-collision-05aug13/pdfkwCALijJOp.pdf>

²² *Is Your Caching Resolver Polluting the Internet?*, Wessels, 2004, <http://dns.measurement-factory.com/writings/wessels-netts2004-paper.pdf>

²³ *SSAC Report: Redirection in the com and net Domains*

With the exception of .corp, .home, and .mail, this approach would apply to all registries, including the registries not eligible for the “alternative path to delegation.” ICANN will make 2LD Block Lists available as required.

RECOMMENDATION 7: ICANN require registries that have elected the “alternative path to delegation,” rather than a wildcard, instead publish appropriate A and SRV resource records for the labels in the ICANN 2LD Block List to the TLD’s zone with the 127.0.53.53 address for a period of 90 days. After the 90-day period, there shall be no further collision-related restrictions on the registry.



RECOMMENDATION 8: ICANN relieve the prohibition on wildcard records during the controlled interruption period.

RECOMMENDATION 9: ICANN monitor the implementation of controlled interruption by each registry to ensure proper implementation and compliance.

3.1.2 Why 90 days?

By far the most prevalent public comments to our draft report were related to the 120-day Controlled Interruption period. We reviewed these comments carefully and subsequently modified our thinking.

A portion of the public comment from .Club Domains, LLC sums up the issue nicely:

The comments of the NTAG, Donuts, Rightside/United TLD, and Ari Registry Services have thoroughly and competently explained why the 120 day interruption period of Recommendation 7 is excessively conservative. A merely conservative interruption period of 60 days is more than adequate for registries that have already been delegated, because the detrimental effects on public interest must be balanced against the security interest of a longer interruption period. A lengthened interruption period is significantly detrimental to the public interest because it would cause confusion for commercial registrants.²⁴

We like this comment because it speaks to the trade-offs between potential risks/harms and actual risks/harms. In New TLD space, Controlled Interruption is a conservative mitigation against a theoretical harm. Despite a concentrated effort by a number of researchers (JAS included!) for the better part of the past two years to find actual incidences of collision-induced harms related to New TLDs, the reality is that none have been found. As of today, 275 New gTLDs have been delegated and over 835,000 2LD registrations have been added with no indication of issues. As we stated earlier, while this certainly doesn't "prove the negative," the data must be taken into consideration. Based on everything we know now, the harms remain theoretical. Given no indication of actual harms, is it justifiable for JAS to recommend an "excessively conservative" and atypical duration, or is a "merely conservative" and more typical duration more appropriate? What is the tradeoff – what *actual* harms could we be causing with an "excessively conservative" approach to a theoretical harm?

After reviewing this issue, we have changed our recommendation to indicate a 90-day Controlled Interruption period.

²⁴ <http://forum.icann.org/lists/comments-name-collision-26feb14/pdfEVFexxB8GK.pdf>

3.1.3 What about IPv6?

Since IPv6 does not support a range of addresses for localhost like IPv4, there is not a straightforward analog of our Controlled Interruption recommendation in v6 space. So the discussion becomes twofold: (1) is a v6 response necessary, and if so, (2) what address would be returned?

Addressing the first, we do not believe v6 responses are necessary at this time. The data we analyzed revealed a miniscule number of resolvers seeking v6-only responses (less than 1%) where the resolver doesn't appear to be dual-stacked. As of this writing, Google reports that roughly 3.5% of their users access Google over v6.²⁵ So while v6 adoption is certainly important and growing, v6-only hosts experiencing a DNS namespace collision does not appear to be a real problem today.

Regarding the second item, an address that is not a direct conceptual equivalent to 127.0.53.53 in v4 space would need to be selected (or "appropriated") for the purpose of Controlled Interruption. While experts can certainly debate this topic (we considered ::1, ::53, IP addresses within fd00::/8, fe80::/10, and ::ffff:127.0.53.53) at the end of the day each approach has plusses, minuses, and importantly the potential for unintended consequences. It's critical to remember that v6 implementations are comparatively young when compared to v4 implementations; the behavior of the vast majority of v4 stacks when presented 127.0.53.53 is well understood whereas the behavior of v6 implementations and their associated infrastructure when presented with ::53, fd00::53, or ::ffff:127.0.53.53 is certainly less deterministic.

So we're left with a tradeoff: do we risk potential unintended consequences of experimenting in the "fringes" of v6 for what is very likely a small benefit? Do we risk causing new problems to address what is fairly clearly a corner case? At the end of the day, we are left with no strong rationale for a v6 response and numerous reasons to be cautious of the potential for unintended consequences.

That being said, v6 support is certainly desirable in the long-term. One possible solution is working with the IETF to extend the definition of localhost to ::0/64 instead of ::1/128 to create a direct equivalent of the 127/8 space in IPv4. We recommend that ICANN work with the IETF to identify a workable long-term solution for IPv6.

²⁵ <https://www.google.com/intl/en/ipv6/statistics.html>



RECOMMENDATION 10: ICANN work with the IETF to identify a mechanism for IPv6 that provides similar functionality to that available in IPv4's "localhost" reserved prefix.

3.1.4 Controlled Interruption Trial

In January, JAS deployed the controlled interruption zone in multiple 2LD namespaces that exhibited evidence of significant collision and collision-like behavior.

As we had previously established bi-directional communication with multiple parties querying these names, we gave our contacts advance notice that we were making changes to the zone and asked them to observe and report the behavior of their systems during the controlled interruption windows.

Despite publishing phone numbers and email addresses via http and Whois, in the event the controlled interruption caused harm, not a single call or email was received. Additional details of this trial will be available in a future report.

3.1.5 Alternatives to Controlled Interruption

We considered several alternatives to controlled interruption as described above, including several honeypot approaches, use of DNAME, and various 2LD string-by-string and TLD-by-TLD approaches. While we eventually concluded that controlled interruption approach offers the most value and presents the least risk, discussion of alternatives is worthwhile.

3.1.6 String-by-String Approaches (TLD and 2LD)

While the occurrence and risk associated with DNS namespace collisions is not uniform across all TLDs and 2LDs, our analysis concluded that any collision and any harm could – at least in theory – occur anywhere in the global DNS namespace. We found ample evidence supporting this conclusion, and found that it would be a quixotic undertaking to determine the root cause of every incidence of a DNS namespace collision.²⁶ With the exception of .corp, .home, and .mail, which are clear outliers for the reasons mentioned earlier, the several root causes we found are not limited to particular strings, new or existing TLDs, or even specific levels of the DNS.

JAS' assessment is, with the exception of .corp, .home, and .mail, that the risk of a collision in the newly applied-for TLD namespaces causing more than a highly localized disruption is low after the recommended mitigation technique is applied. String-by-string and TLD-by-TLD approaches add significant complexity and

²⁶ *Focused Analysis on Applied-For gTLDs - .cba*, Verisign Inc., September 15, 2013, retrieved January 2014, <http://forum.icann.org/lists/comments-name-collision-05aug13/msg00039.html>

potential for unintended consequences while adding little if any security value. Not a good tradeoff. As such, we recommend an approach that addresses the root causes and does not delineate between specific strings unnecessarily.

3.1.7 Honeypot Approaches

Significant discussion has occurred in several fora regarding various implementations of a trial delegation that directs traffic to an Internet-based honeypot. The honeypot, run by ICANN or some trusted third party, could serve two functions: 1) Present helpful information for operators reaching the site over http and potentially other protocols; and 2) Collect logs to help identify volume, sources, and potential severity of collision and collision-like activity. Some ideas describe a honeypot that runs for a deterministic time period while others continue the honeypot until some threshold is achieved indicating risk has been mitigated to an (undefined) acceptable level.

Because collisions are largely a notification problem, we like the concept of honeypot approaches. However, there are some critical traits of honeypot approaches that make them undesirable.

- Whenever logs are collected, the question “for what purpose” must be asked. How much collision activity is “OK” – what is the acceptable risk? Is the threshold the same for all TLDs? Are all query sources to be treated equally – that is, do we look differently upon log entries that *appear* to be from a nuclear power plant vs. a residential broadband network? These questions, being subjective in nature, may not have answers that can achieve consensus.
- Whenever logs are collected, we must also be vigilant for gaming opportunities. Because there are many interested parties and significant commercial pressures, we assume that competing interests will attempt to exploit any activity that may create an argument for slowing or halting valuable registrations in a TLD. Even the possibility (perceived or actual) of such gaming will virtually assure that gaming occurs.
- There are collision scenarios where returning an Internet IP address will cause traffic to be sent over the Internet that was never previously sent. Ever conscious of “cure being worse than the disease” concerns, we certainly do not want to open these hosts to new risks while we try to help them. Additionally, we are informed by the vulnerability we discovered on this matter; for machines impacted by the issue, honeypotting a popular port will *assure* that sensitive information is transmitted in the clear over the LAN and the Internet to the honeypot. Absent the honeypot, transmission of this sensitive information is not assured. Controlled interruption should not *decrease* the security posture of a system, even temporarily. Or, as Verisign cleverly said in their public comment,



we don't want to risk turning "Controlled Interruption" into "Controlled Exfiltration!"²⁷

- As security researchers have long known, a lot of potentially sensitive information appears in logs. Usernames and passwords regularly appear in http logs. Other protocols raise similar concerns. Our experience confirms that any advertised honeypot IP will receive a host of sensitive information. Managing this information – and convincing the global Internet community that the data is being handled responsibly – is another hurdle with any honeypot approach.
- Different global legal jurisdictions place restrictions on data collected after it was "solicited." As advertising a honeypot IP could be argued as "soliciting traffic," the resulting data may have legal protections, further adding to the complexity.
- Very limited experience exists related to large-scale honeypotting of the service discovery protocols and corporate directory protocols that dominate colliding DNS queries. SAC06²⁸ contains a lengthy discussion of the unintended consequences of these sorts of interactions with non-HTTP protocols. There is sufficient risk of causing collateral damage and unintended consequences.

The final four bullets describe our rationale for a 127/8 IP address that does not cause traffic to leave the host, thereby avoiding those pitfalls.

We also considered a variation wherein the honeypot would be an RFC 1918 IP address as opposed to an Internet address – thereby allowing private network operators to monitor and capture the resulting traffic. However, we ruled out this variation due to the potential for unintended consequences if the RFC 1918 IP happened to be in use in the network where the affected party resides, and because of the potential for causing general confusion. An operator with the requisite sophistication to redirect or capture RFC 1918 traffic likely also has the requisite sophistication to react appropriately to 127/8 responses.

²⁷ <http://forum.icann.org/lists/comments-name-collision-26feb14/pdf/jLkllhcj4.pdf>

²⁸ <https://www.icann.org/en/system/files/files/report-redirectation-com-net-09jul04-en.pdf>



3.1.8 DNAME Approaches

We considered multiple schemes using DNAME records in an attempt to emulate similar controlled interruption behavior. While we eventually concluded that these schemes are not feasible and less effective than localhost-based ideas, discussion is worthwhile.

One option could be implemented via DNAME records in the root. We quickly considered this option infeasible due to the difficulties, unknowns, and potential for unintended consequences surrounding the placement of DNAME records in the root; furthermore, such an approach is very likely not compatible with the IANA/Verisign/NTIA root zone management system as currently implemented and may require modifications to the IANA Functions contract.

However, using wildcards in the delegated zone is a more viable option and emulates most of the desired behavior.

Consider a wildcard DNAME record within the origin of the TLD zone pointing to some identifiable target (e.g., "you-need-to-change-your-dns-config-see-collisions-dot-icann-dot-org."). The target should not be resolvable in order to force an NXDOMAIN response (note that this assumes the specific DNAME implementation returns an NXDOMAIN instead of SERVFAIL or something else – given the relative newness of DNAME in the DNS protocol suite and its lack of significant exercise in implementations, unusual implementation decisions and/or behavior can't be ruled out).

When considering DNAME approaches, client support is a paramount concern. While the experiments²⁹ conducted by Geoff Huston and George Michaelson are valuable and informative, they are biased to heavy clients and human browsing (running Flash and receiving ads). The situation before us is far less biased to these types of clients, so client support is in question at best. Proper support of DNAME (RFC 2672 circa 2000) in legacy, possibly misconfigured, devices is probably less likely than proper localhost support (RFC 1122 circa 1989).

DNAME-based approaches do offer additional flexibility when compared to localhost redirection approaches, specifically in the ability of sophisticated operators to observe, control, and redirect the responses. But again, an IT operation sophisticated enough to control DNAME queries certainly has plenty of other options available to manage DNS namespace collisions. Catering to sophisticated IT

²⁹ *draft-jabley-dnsop-as112-dname-01: AS112 Redirection using DNAME*, Abley, Dickson, Kumari, Michaelson, October 12, 2013, retrieved January 2014, <http://tools.ietf.org/html/draft-jabley-dnsop-as112-dname-01> (see *Appendix A: Assessing Support for DNAME in the Real World*)

operators by providing flexibility and options seems to come at the expense of simplicity, predictability, and widespread client support.

Finally, DNAME-based approaches don't necessarily interrupt, negating the whole purpose of controlled interruption. The DNAME redirect to return NXDOMAIN means folks can continue on as they're currently doing. They won't notice anything so they won't fix it, defeating the purpose of the interruption.

As such, we consider DNAME-based approaches inferior to localhost-based approaches.



3.2 Root Level Data, Monitoring, and Day-In-The-Life (DITL)

We blogged³⁰ about our experiences using the DNS-OARC-maintained “DITL” datasets; these datasets are truly invaluable albeit limited for researchers looking into global Internet DNS traffic. Conscious of the calls for additional datasets and monitoring at the root level, we want to discuss the objectives of monitoring and logging systems at a meta level.

When considering monitoring and logging systems, one must always start with the “for what purpose” questions. Different data consumers have different requirements. For example, operators interested in emergency response demand a low-latency, actionable, “ticket” type of monitoring. They want the “this hard drive is dead” ticket as soon as possible after it dies. Capacity planners want intermediate-latency data with some ad hoc aggregation and trending capabilities to answer questions like “how much data do we have and what is the growth rate?” Product managers want high-latency, highly detailed data repositories that can answer a full range of complex ad hoc queries to observe behaviors, trial new product ideas, etc.

Obviously, these very different consumers have very different requirements driving very different technical implementations.

We observe that from an availability standpoint, low-latency ticket/availability data is already available for the root. Albeit in a highly decentralized fashion, the DNS root is probably one of the most highly monitored systems on Earth in that regard.

Conversely, DITL datasets are at the other end of the spectrum: extremely high latency (one 50 hour period annually), voluminous and unstructured data suitable only for compute-intensive ad hoc analysis by expert researchers.

While individual root operators certainly have a full range of data available to them, there is nothing in the middle available to researchers or the Internet at large.

Looking from a slightly different angle, the *availability* and *content* of the root is exceptionally well monitored with low latency but the *queries* to the root are much less visible.

We believe there is a need for a medium-latency, aggregated, and more “consumable” data stream from the root operators containing aggregated summary data describing the queries seen by the root. This new feed should be in a reasonably accessible and well-documented format like CSV, XML, or YAML and

³⁰ *Demystifying DITL Data [Guest Post]*, Kevin White, JAS Global Advisors LLC, November 16, 2013, retrieved December 2013, <http://domainincite.com/15068-demystifying-ditl-data-guest-post>



ideally have latency on the order of a few days. Mindful of the numerous issues surrounding such an undertaking, we recommend that ICANN, DNS-OARC, and the root operators explore such a mechanism.

We note ongoing efforts by the Root Server System Advisory Committee (“RSSAC”) to address monitoring, and the forthcoming publication of RSSAC 002: *Recommendations on Measurements of the Root Server System*. We applaud the proactive efforts of some root operators to increase the fidelity of root server monitoring.

RECOMMENDATION 11: ICANN, DNS-OARC, and the root operators explore a medium-latency, aggregated summary feed describing queries reaching the DNS root.

Over the course of our research, we were also surprised to find that authoritative historical information regarding the contents of the root zone is not always available. A significant proportion of historical information is only captured informally in email threads and in the heads of various luminaries. As such, we also recommend that a single, authoritative archive for root data be established.

RECOMMENDATION 12: ICANN, DNS-OARC, and the root operators explore establishment of a single, authoritative, and publicly available archive for historical data related to the root.

We recognize that data and measurement regarding the DNS namespace collision phenomenon is important. One of the attractive features of a honeypot approach is a new, high fidelity, and low latency data stream describing this behavior. In lieu of the honeypot, we recommend ICANN explore collecting NXDOMAIN and Controlled Interruption (127.0.53.53 query response) entries in DNS query logs from registry operators and contribute them to an independent data repository such as DNS-OARC where they may be analyzed by the research community. To limit the potential for commercial gaming or use by malicious parties, we recommend that logs be provided six months in arrears. If such logs are collected commencing with delegation, the long-term effectiveness of Controlled Interruption may be measured; we believe it important to be informed by these metrics when considering future mitigation techniques in delegated and un-delegated DNS namespace.

RECOMMENDATION 13: ICANN explore collecting NXDOMAIN and Controlled Interruption (127.0.53.53 query response) entries in DNS query logs from registry operators and contribute them to an independent data repository such as DNS-OARC for further analysis. To limit the potential for commercial gaming or use by malicious parties, we recommend that logs be provided six months in arrears.



3.3 Collisions in Existing DNS Namespace

Because of the popularity of .com, typical software behavior, and common DNS search path configurations/practices, collisions at the 2LD level within .com likely occur at a higher frequency than collisions at any other location in the DNS (2LD and TLD). Because of the sheer size and prevalence of .com, this is not unexpected. With respect to collisions, .com is victim of its own success. Recently, other researchers have quantified the order of magnitude of collisions within .com using different datasets.³¹ Noted security researcher Robert Stucke spoke at DEFCON 21 about vulnerabilities he discovered by leveraging DNS namespace collisions within .com.³²

Researching collisions in existing TLD namespaces was a part of our engagement. Over the course of this study, JAS registered several 2LDs to enhance our understanding of this phenomenon and collect additional data. Based on behaviors uncovered during our research, we made educated guesses as to where problematic collisions may occur. These registrations immediately generated a surprising amount of traffic.

It is worth noting that while selecting 2LDs to register for our research, we made use of publically available tools designed to facilitate “domain drop catching” and various “squatting” activities. One such tool offers to the public the ability to find 2LDs within .com that are “available with traffic” – the very definition of a DNS namespace collision – at the second level within the Internet’s most popular TLD.

While we understand the commercial value of this service, as security practitioners we are deeply concerned about this type of functionality. As such, we recommend that ICANN request that the appropriate bodies (GNSO, SSAC, etc.) further explore issues relating to collisions in existing DNS namespace, the practice of “domain drop catching,” and the associated data feeds that may be leveraged by attackers when attempting to exploit collisions.

RECOMMENDATION 14: ICANN request that the appropriate bodies further explore issues relating to collisions in existing DNS namespace, the practice of “domain drop catching,” and the associated data feeds that may be leveraged by attackers when attempting to exploit collisions.

³¹ <http://forum.icann.org/lists/comments-name-collision-05aug13/pdf056yDnxGje.pdf>

³² <http://www.youtube.com/watch?v=ZPbyDSvGasw>

3.4 Description of Forthcoming Phase Two Report

JAS uncovered a vulnerability not directly related to ICANN's New gTLD Program nor to new TLDs in general that has the potential to impact end-systems. In fact, the vulnerability manifested while researching collisions within .com while analyzing collisions in existing TLDs (please see below). Pursuant to ICANN's Coordinated Vulnerability Disclosure Process,³³ ICANN shall: "...privately disclose information relating to a discovered vulnerability to a product vendor or service provider ("affected party") and allow the affected party time to investigate the claim, and identify and test a remedy or recourse before coordinating the release of a public disclosure of the vulnerability with the reporter." Furthermore, ICANN's process states: "All parties to the disclosure generally agree to refrain from disclosing the vulnerability to the public until a remedy is identified and tested or until the threat is considered contained." As such, pursuant to ICANN's process and out of an abundance of caution, JAS has recommended against publication of a complete report at this time.

The Phase Two report is expected to contain the following information:

Impact of malware/adware/clickfraud tools: We found that malware, adware, and clickfraud tools generated a significant proportion of the random and pseudo-random string queries. We also identified other potential sources of algorithmic queries. In the Phase Two report, we will describe this occurrence and provide specific details and supporting data. It is worth noting that we previously discussed this finding in presentations in ICANN Buenos Aires, ICANN Singapore, and Verisign's *Workshop and Prize on Root Causes and Mitigation of Name Collisions* (WPNC) in London.³⁴ Queries related to malware/adware/clickfraud tools explain in excess of 20% of the colliding queries revealed in DITL datasets.

Analysis of Collisions in previous TLD delegations: We found that collisions have occurred prior to delegation of every TLD since (at least) 2007 and presented high-level data to this effect at the WPNC in London and ICANN Singapore. Said differently, collisions in the DNS namespace are certainly not a new phenomenon (please also see Section 2 above). In the Phase Two report, we will further describe this occurrence and provide specific details and supporting data.

Analysis of Collisions in existing TLDs: As stated in Section 3.3 above, JAS is very concerned about collisions in existing DNS namespace and the tools that facilitate discovery of colliding names. In the Phase Two report, we will describe these

³³ *Coordinated Vulnerability Disclosure Reporting at ICANN*, ICANN, 2013, <https://www.icann.org/en/about/staff/security/vulnerability-disclosure-05aug13-en.pdf>

³⁴ <http://namecollisions.net>



phenomena, including the methods we used to discover the vulnerabilities, locate vulnerable hosts, and datasets where appropriate.

A Taxonomy of Queries and TLDs: We classified behavior leading to collisions and created a high-level description of each applied-for TLD based on the colliding queries present in DITL datasets. Organizing the classification into a taxonomy leads to an understanding that: (1) a very few root causes seem to explain the vast majority of colliding behavior, and (2) nearly all root causes appear in all TLDs in differing proportions. Only .corp, .home, and .mail are clear outliers.

The classification was based on: (1) the diversity of querying source IP addresses and Autonomous Systems; (2) the diversity of labels queried; (3) applying sophisticated “randomness detection” to strings and substrings; (4) presence of linguistic terms and colloquialisms in strings and substrings; (5) temporal patterns; and (6) analysis of the Regular Expressions of the labels queried within each TLD and across all TLDs. Aside from improving our understanding of the behavior within .corp, .home, and .mail, we eventually found that the taxonomy does not directly translate to mitigation techniques. Mitigation techniques addressing the small number of root causes are applicable to all TLDs. Dr. Arnoldo Muller-Molina, Founder and Chief Data Scientist of our partner simMachines presented some information about the classification research we performed at the closed DNS-OARC workshop in Warsaw, Poland in May 2014.³⁵ We also note that the JAS public comment submission included an analysis of colliding queries over a few of the aforementioned dimensions.³⁶

Sources, methods, and experimental results: Over the course of this study, we performed a number of experiments and collected a significant amount of data. We talked to a number of vendors, consultants, and end-users experiencing collisions in existing namespaces today. We purchased a number of names and collected data. Several sources contributed data. The types of analysis we performed have been alluded to in this paper and in our presentations at ICANN Singapore, the WPNC in London, and the DNS-OARC workshop in Warsaw. Phase Two of this report will contain the full detail of these activities and relevant datasets.

³⁵ <https://indico.dns-oarc.net//contributionDisplay.py?sessionId=1&contribId=30&confId=19>

³⁶ <http://forum.icann.org/lists/comments-name-collision-05aug13/pdf3WmZlrH3fo.pdf>

