Managing the Risks of Top-Level Domain Name Collisions

Findings for the Name Collision Analysis Project (NCAP) Study 1

Karen Scarfone, Scarfone Cybersecurity

May 7, 2020

Table of Contents

Acknowledgments

Scarfone Cybersecurity thanks The Internet Corporation for Assigned Names and Numbers (ICANN), the ICANN Office of the Chief Technology Officer (OCTO), and ICANN's Name Collision Analysis Project Discussion Group (NCAP DG) for their support and insights during the development of this draft report. Scarfone Cybersecurity also thanks the individuals and groups who submitted feedback during the first public comment period.

Executive Summary

This report presents the findings for Study 1 of the Name Collisions Analysis Project (NCAP). A name collision occurs when a single domain name, like .EXAMPLE, is used in contradictory ways at the same time, so it is unclear which resource is being requested. There are many forms of name collisions for top-level domains (TLDs), and Study 1 examined four types: duplicate, shortened, search list, and re-registered.

Study 1 had three goals, which can be summarized as documenting prior work on name collisions, assessing name collision datasets, and recommending whether or not the proposed follow-on Studies 2 and 3 should be performed. This report provides a thorough account of all relevant prior work on name collisions and the datasets used for that work. The report's major findings from that survey of prior work and datasets are as follows:

1. Name collisions have been a known problem for decades, possibly as early as the late 1980s. Reports, papers, and other work regarding name collisions were sparse and sporadic until 2012, at which point many organizations and individuals began publishing extensively on the topic. Workshops were held in 2013 and 2014. Since ICANN approved the Name Collision Occurrence Management Framework in 2014, which instituted controlled interruption as the mitigation strategy for new TLDs, the volume of work on name collisions by academic institutions, the security industry, IT product and service vendors, and others has greatly decreased. The only known work on name collisions during the past few years has been from ICANN by the NCAP Discussion Group (DG) and the New gTLD Subsequent Procedures (SubPro) Working Group. Since mid-2017, there has not been any published research into the causes of name collisions or new name collision mitigation strategies.
2. Since controlled interruption was instituted, there have been few instances of name collision problems being reported to ICANN or reported publicly through technical support forums, mailing lists, and other means. Most problems occurred during 2014, 2015, or 2016, with only a single problem reported to ICANN during the three-year period from 2017 through 2019, as well as a sharp dropoff in public reports during the same period. Only one of the reports to ICANN necessitated action by a registry, and none of the public reports surveyed mentioned major harm to individuals or organizations.
3. Prior work and name collision reports have indicated there are several types of root causes of name collisions, perhaps a dozen or more. These root causes have typically been found by individuals researching a particular leaked TLD to find its origin, not by examining datasets. There is unlikely to be any dataset that would contain root causes; identifying root causes is generally going to require research of each TLD involved in name collisions on a case-by-case basis.
4. No gaps or other issues have been identified in accessing the datasets that would be needed for Studies 2 and 3.

Recent discussions among NCAP DG members indicate differences of opinion as to whether controlled interruption has been "successful." It does not appear that criteria for success are formally defined, and until such criteria are defined, disagreements are likely to continue.

That being said, however, there have been minimal name collision problems reported since controlled interruption was instituted, given the number of new TLDs it has been used for in the past six years. Research conducted for this report included extensive searches for evidence, and NCAP DG members were repeatedly asked to provide information on any evidence they were aware of. The counterargument to this has been the old saying, "Absence of evidence is not evidence of absence." Although that saying has merit, over time the continued absence of evidence that controlled interruption has not been successful makes it less likely to be true. The lack of interest in alternatives to controlled interruption outside a few groups within ICANN further supports the likelihood that controlled interruption has been successful.

Given these findings, the recommendation is that Studies 2 and 3 should not be performed as currently designed. Regarding Study 2, analyzing datasets is unlikely to identify significant root causes for name collisions that have not already been identified. New causes for name collisions are far more likely to be found by investigating TLD candidates for potential delegation on a case by case basis. Regarding Study 3, controlled interruption has already proven an effective mitigation strategy, and there does not appear to be a need to identify, analyze, and test alternatives for the vast majority of TLD candidates.

All of that being said, this does not mean further study should not be conducted into name collision risks and the feasibility of potentially delegating additional domains that are likely to cause name collisions. Most notably, the Study 3 question of how to mitigate name collisions for potential delegation of the corp, home, and mail TLDs is still unresolved. However, the proposals for Studies 2 and 3, which were developed years ago, do not seem to be effective ways of achieving the intended goals.

1 Study Overview

This report presents the findings for Study 1 of the Name Collision Analysis Project (NCAP) [1]. The purpose and scope of Study 1 were defined in a July 2019 Request for Proposal (RFP) [2]. The draft of this report addresses all three goals of Study 1, as stated in the RFP:

> "1. Production of a summary report on the topic of name collision that brings forth important knowledge from prior work in the area. The report will be a primer for those new to the subject. The report will be based on an examination of all relevant prior work on the issue of name collisions.
>
> 2. Creation of a list of datasets used in past name collision studies; an identification of gaps, if any; and creation of a list of additional data sets that would be required to successfully complete Studies 2 and 3.
>
> 3. A recommendation if Studies 2 and 3 should be performed based on the results of the survey of prior work and the availability of data sets." [2]

For the purposes of Study 1, the term *name collision* "refers to the situation where a name that is defined and used in one namespace may also appear in another. Users and applications intending to use a name in one namespace may attempt to use it in a different one, and unexpected behavior may result where the intended use of the name is not the same in both namespaces. The circumstances that lead to a name collision could be accidental or malicious.

Study 1 concerns name collisions in the context of top-level domains (TLDs), where the conflicting namespaces are:

- the global Internet Domain Name System (DNS) namespace reflected in the root zone overseen by the Internet Assigned Numbers Authority (IANA) Function; and
- any other namespace, regardless of whether that other namespace is intended for use with the DNS or any other protocol." [2]

Also from the RFP:

> "Name collision refers to the situation in which a name that is used in one namespace may be used in a different namespace, where users, software, or other functions in that domain may misinterpret it. In the context of top level domains, the term 'name collision' refers to the situation in which a name that is used in the global Domain Name System (DNS) namespace defined in the root zone as published by the root zone management (RZM) partners ICANN and VeriSign (the RZM namespace) may be used in a different namespace (non-RZM), where users, software, or other functions in that domain may misinterpret it." [2]

The report contains the following sections addressing tasks from the RFP:

- Section 2 contains a name collision primer (task 2a).

- Section 3 provides a review of pertinent previous work (tasks 1 and 2c).
- Section 4 details evidence of harm caused by name collisions (task 2b) and discusses the technical impact of name collision mitigation techniques employed to date (task 2d).
- Section 5 assesses datasets used in past name collision studies, identifies additional datasets that would be needed for Studies 2 and 3, and discusses the availability of those additional datasets (tasks 3, 4, and 5, respectively).
- Section 6 makes recommendations on performing Studies 2 and 3 (task 8).

All sources referenced in this report are cited in Section 7, Bibliography.

2 Name Collision Primer

This section explains the basics of name collisions. Readers who are already well-versed on the topic of name collisions should still read this section because it defines new terms for the purposes of this report and establishes the scope for the report. Other concepts in this section are based on material from the ICANN Acronyms and Terms tool. [3]

2.1 Domains

A *domain name* maps to a piece of data, like an IP address. For example, icann.org is a domain name for the ICANN organization. You can use "icann.org" to reach ICANN's computing resources, like websites and email servers, instead of typing in icann.org's IP address every time you want to access an ICANN website.

Every domain name consists of one or more labels, and the labels go from most specific on the left to least specific on the right. The icann.org domain name has two labels. "org" is the label for the *top-level domain (TLD)*. "icann" is the label for the *second-level domain (SLD)*. The SLD is a domain name that is associated with a TLD—in other words, an SLD is *registered* to a TLD. There are usually many SLDs registered to a single TLD. Many domain names have three or more levels, such as "www.icann.org", but for the purposes of this explanation, we will focus on the highest two levels (TLDs and SLDs) only.

In the past, there were a small number of *generic TLDs (gTLDs)* like com and org. [4] A few more were added in 2000 and in 2004. [5] Efforts began in 2005 to consider adding many more gTLDs, and in October 2013 the first of these new gTLDs was made available for usage on the Internet, a process better known as *delegation*. The gTLDs are frequently referred to as "names" or "strings", so when you see a term like "delegated strings", it just means that a new gTLD was made available on the Internet. For more information on gTLDs, see ICANN's resources [6] [7] [8].

In addition to gTLDs, there are also TLDs specific to country names—*country code TLDs (ccTLDs)*. The original ccTLDs were all two letters long, such as fr and us, taken from the two-letter country codes in International Organization for Standardization (ISO) 3166, *Codes for the representation of names of countries and their subdivisions*. [9] The two-letter ccTLDs have since become known as *ASCII ccTLDs*. In 2009, ICANN approved an effort to delegate new internationalized domain name ccTLDs (IDN ccTLDs) through what is called the IDN ccTLD Fast Track Process. [10] IDN ccTLDs use non-Latin characters, such as the alphabet of the primary language spoken in a particular country. For more information on ccTLDs, see ICANN's resources [11] [12] [13].

In this report, usage of "TLD" refers to both gTLDs and ccTLDs.

2.2 Name Collisions

There are many forms of name collisions. To understand what name collisions are and which types are in scope for this report, let's look at four examples, which are based on Section 2.3.3 of the study RFP [2]. Each example maps to one or more of the situations described in the RFP.

Explanation of Name Collision Type	Mapping to RFP
1. Suppose that Alice uses .EXAMPLE internally only as her top-level domain, which works without ambiguity because .EXAMPLE is not a TLD delegated on the Internet. If Alice types "www.example" in a web browser, it would take her to her own website. The next year, .EXAMPLE is delegated as a new TLD. Now when Alice tries to access "www.example", it's no longer clear whether she is trying to access her own website or the new public domain on the Internet. The .EXAMPLE used internally by Alice and the .EXAMPLE used publicly by someone else *collide*. This report will refer to these as *duplicate name collisions*—the collision is caused by the same TLD being used in two places at the same time.	**A.a:** User Alice intentionally uses .EXAMPLE in a non-RZM context and .EXAMPLE is now delegated in the public DNS. User Alice suffers adverse impact as a result. **A.b:** User Alice unintentionally uses .EXAMPLE in a non-RZM context (for example as the result of a software behavior) and .EXAMPLE is now delegated in the public DNS. User Alice suffers adverse impact as a result.
2. Suppose that Alice uses shortened forms of domain names—for example, she might type "dashboard.example" instead of "dashboard.example.com"—and there's a list of domain suffixes like ".com" that automatically get appended to what she typed in order to find the desired domain. This is known as *search list processing*, and this works as long as there's no TLD for .example. However, the next year, .EXAMPLE is delegated as a new TLD. When Alice wants to go to "dashboard.example.com" and types "dashboard.example", she'll be taken to the latter instead of the former. This report will refer to these as *shortened name collisions*—the collision is caused by someone using a shortened name that matches a TLD being used elsewhere at the same time.	**A.c:** Registrant Alice uses EXAMPLE as a label anywhere except as a non-RZM TLD, and relies on search list processing where the label EXAMPLE is the terminal label, as an intermediate step in that search list processing. *(e.g. User searches for dashboard.example.com by typing in dashboard.example)* .EXAMPLE is now registered in the public DNS and the search list processing behavior of Alice now changes.

Explanation of Name Collision Type	Mapping to RFP
3. Suppose that there is a public domain EXAMPLE.COM, and Alice uses it as her domain. The next year, .EXAMPLE is delegated as a new TLD. Some external users might have search list processing that automatically appends the ".com" domain suffix to requests, so some queries for .EXAMPLE domains may mistakenly go to .EXAMPLE.COM instead. Alice will be receiving traffic that she was not intended to receive. This report will refer to these as *search list name collisions*—the collision is caused by the search list not recognizing .EXAMPLE as a new TLD and instead going through its search list to try to find the domain.	**B.a:** Registrant Alice uses EXAMPLE.COM (or EXAMPLE.TLD where TLD is any current TLD in the public DNS) and .EXAMPLE is now registered in the public DNS. Registrant Alice now receives multiple queries as a result of search list processing of users of domains under .EXAMPLE
4. Suppose that Alice registers a TLD or SLD and uses it for some time, then lets it expire. Subsequently someone else registers the same domain and delegates it. Now queries looking for the old domain (for Alice) will go to the new domain (for someone else). This report will refer to these as *re-registered name collisions*—the collision is caused by someone registering a domain that was previously registered by someone else.	**B.b:** Registrant Alice uses .EXAMPLE as a TLD in the public DNS and then lets the registration expire. Registrant Bob then registers and delegates .EXAMPLE. Traffic intended for Alice's use of .EXAMPLE is now received by Bob's use of .EXAMPLE **B.c:** Registrant Alice uses EXAMPLE.COM and then lets the registration expire. Registrant Bob then registers and delegates EXAMPLE.COM. Traffic intended for Alice's use of EXAMPLE.COM is now received by Bob's use of EXAMPLE.COM

All four of these types of name collisions are in scope for Study 1. Only duplicate name collisions and shortened name collisions (types A.a, A.b, and A.c from the RFP) are in scope for Section 5 of this report (on data sets for Studies 2 and 3). No other types of name collisions are in scope for any parts of Study 1.

For more information on name collisions, see ICANN's resources [14].

3 Review of Previous Work

This section provides a review of previous work on name collisions. All reviewed work meets at least one of the following criteria from the Study 1 RFP:

> "i. Peer reviewed paper
> ii. Report/Analysis based on data
> iii. Qualitative research on name collision experience
> iv. Proposed or agreed technical standards" [2]

The search for previous work was rigorous. Some previous work was cited in the Study 1 RFP, and members of ICANN's Name Collision Analysis Project Discussion Group (NCAP DG) also submitted lists of previous work. The author of this report also conducted extensive searches online for previous work, and then used the references or other sources cited in all the identified documents to identify additional previous work, and so on. During the first public comment period, no suggestions for additional previous work to include were made. There is no way to be sure that every previous work on name collisions has been identified, but there is reasonable confidence that all relevant online documents in English have been found.

The review is broken into several sections based on timeframe and topic area. It is largely chronological, but some items are intentionally out of sequence—for example, it may have taken a few years to finalize a standard on a particular topic, so that standard is included in the topic area's section, where it fits thematically, instead of a later section where it would fit chronologically.

In cases where the previous work includes correspondence on a particular document, such as public comments on a draft report, the review points to the archived correspondence as a whole and does not list or mention each piece of correspondence. In some cases, particular pieces of correspondence are mentioned and discussed. This does not imply that only the cited correspondence is relevant; often several parties made similar points, so one or a few instances are cited as examples, and readers are encouraged to read the others if they desire.

Each subsection within this section indicates which type or types of name collisions are applicable to its contents, if any.

3.1 DNS Wildcard Address Records: 2003 – 2009

Applicability: No name collision types (background material)

In September 2003, Verisign launched what they called Site Finder. Site Finder changed how requests for nonexistent domain names were handled by adding a DNS wildcard address record that matched to every com and net address that didn't otherwise have a match. People and services were used to the previous behavior and were unaware it was changing, so the sudden deployment of DNS wildcard address records inadvertently caused a lot of problems. Then the workarounds for the problems caused even more problems. [15] [16] [17] [18]

Note that while Site Finder's launch was the event that brought a great deal of attention to the subject of DNS wildcard address records, the possibility of domain name requests being resolved in unintended ways was not a new one, with formal treatments of the subject going back to at least 1993. [19] What made Site Finder so noteworthy was that it affected many people and services at one time. Site Finder in particular, and the use of DNS wildcard address records more broadly, did not cause name collisions; however, they are relevant to this study because there are obvious parallels between wildcard address records and name collisions. Both involve domain name queries being resolved in unexpected ways that can disrupt Internet usage for affected parties. So reviewing the recommendations for avoiding another Site Finder-like incident helps indicate potential ways of avoiding negative impacts from name collisions as well.

The ICANN Security and Stability Advisory Committee (SSAC) conducted a review of Site Finder and DNS wildcard address records, and issued their findings and recommendations in report SAC 006 in July 2004. [16] Most pertinent to this study is recommendation 4 from SAC 006:

> "Changes in registry services should take place only after a substantial period of notice, comment and consensus involving both the technical community and the larger user community. This process must (i) consider issues of security and stability, (ii) afford ample time for testing and refinement and (iii) allow for adequate notice and coordination with affected and potentially affected system managers and end users. Thirty years of experience show that this strategy ensures robust engineering and engenders trust in the systems and the processes surrounding their maintenance and development."

Also of particular relevance for this study is the Reserved Names portion of the "Problems encountered in recent experiences with wildcards" section of [15], which is also duplicated by [16]:

> "This sort of wildcard usage is incompatible with any use of DNS which relies on reserving names in a registry with the express intent of not adding them to the DNS zone itself. An example of such a use is the JET-derived IDN approach of 'registry restrictions' and 'reserved names', which depends on the existence of names that are reserved and can be registered only by the holder of some related name, but which do not appear in the DNS. By some readings of the current ICANN IDN policy, support for that 'reserved name' approach is required. To accomplish the goal of reduced consumer confusion, the reserved names must not be resolvable at all. This reserved name approach appears to be completely incompatible with this sort of wildcard usage: since the wildcard will always cause a result to be returned, even for a reserved name which does not appear in the zone, one can support either one or the other, but not both."

In November 2006, the SSAC posted SAC 015 [17], an advisory explaining why wildcards should not be used for TLDs. It explained how wildcards work and gave examples of problems that resulted in the past from wildcard resource records in TLDs. SAC 015 referenced previous SSAC work and also cited a report from the ICANN Registry Services Technical Evaluation Panel [20] that reached the same conclusion as SAC 006 did: wildcards were too risky to be used in TLDs.

SAC 032 [18] was released in June 2008. It contained a broader discussion of DNS response modification, with wildcards part of that discussion. SAC 032 provided preliminary recommendations for addressing DNS response modification, including this: "SSAC concurs with

the IAB and recommends that entrusted agents should not use DNS wildcards in a zone without informing the domain registrant of the risks identified in this Report and elsewhere, that entrusted agents should not generate wildcards and synthesized responses without the informed consent of the registrant, and that entrusted agents should provide opt-out mechanisms that allows clients to receive the original DNS answers to their queries."

SAC 041, published in June 2009, summarized the SSAC's study of DNS wildcarding since 2004 and advised "ICANN that new TLDs, including both new gTLDs and new ccTLDs, should not use DNS redirection and synthesized DNS responses. […] The redirection and synthesizing of DNS responses by TLDs poses a clear and significant danger to the security and stability of the domain name system." [21] The references to redirection and synthesizing included the use of DNS wildcard address records.

For additional information on Site Finder, see the ICANN Archives for Verisign's Wildcard Service Deployment. [22]

3.2 Collisions from Failure to Renew a Domain: 2006

Applicability: Re-registered name collisions

Re-registered name collisions can occur when someone fails to renew a domain and someone else subsequently acquires the same domain. This topic was extensively discussed in SAC 010 [23] and SAC 011 [24], both from June 2006.

SAC 010 provided information and guidance for registrants. From SAC 010: "…registrants may not appreciate that expired domain names are commonly registered to another registrant within a few weeks or months of the date of expiry of the domain name registration agreement. The new registrant may not use the domain name for the same purposes as a former registrant. Incidents show that previously registered domain names may be exploited, at the expense of the reputation of a former registrant. In this Advisory we refer to this form unexpected consequence as *reputational harm*."

SAC 011 provided more technical information on the situation, with examples of the disruptions that could be caused by a domain name not being renewed and a malicious party subsequently renewing it. In this situation, the attacker could receive traffic that was intended for the organization that originally had the domain.

Both SAC 010 and SAC 011 made it clear that it is ultimately the registrant's responsibility to ensure they renew their domains in a timely fashion. The importance of keeping contact information up to date was emphasized.

3.3 Initial TLD Delegation Concerns: 2008 – 2013

3.3.1 Invalid TLD Queries Reaching Root Servers

Applicability: Duplicate name collisions, re-registered name collisions

ICANN started work in mid-2008 to figure out processes for parties to apply for new gTLDs and for ICANN to evaluate the applications. [25] In 2009, the SSAC was asked to look at the issue of invalid TLD queries reaching root servers, stemming from someone registering a TLD that others had already been using and the root server had been responding to. At that time, an estimated 26% of all query load at root servers was invalid TLDs. One of the earliest mentions of this problem was in a June 2009 blog posting. [26] The SSAC studied the issue and released their SAC 045 report in November 2010. [27] SAC 045 cited the possibility of someone applying for a TLD that had appeared in queries before or had been issued before, and the problems with queries that had been failing suddenly succeeding once the TLD was delegated.

SAC 045 acknowledged how difficult it would be to eliminate inadvertent instances of such problems: "It is likely that many of the same conditions that cause the current set of invalid TLD queries to appear at the root level of the DNS will persist despite efforts to encourage end users, private networks, software and equipment manufacturers to correct configuration and programming errors."

SAC 045 had recommendations for reducing other instances of query ambiguity. Recommendation 2 said, "Prohibit the delegation of certain TLD strings. Internet Engineering Task Force (IETF) Request for Comments (RFC) 2606, 'Reserved Top Level Domain Names,' currently prohibits a list of strings, including test, example, invalid, and localhost." Section 2.2.1.2 of ICANN's *gTLD Applicant Guidebook* [25] released in June 2012 specified names that could not be gTLDs; these included the names from RFC 2606 [28], plus a few dozen more, in what was termed the Top-Level Reserved Names List.

3.3.2 Certificates for Internal Domains That May Also Become gTLDs

Applicability: Duplicate name collisions

In November 2012, the SSAC became aware of a problem with certificate issuance that could negatively affect new gTLD delegation. At that time, the application period for new gTLDs had already closed but no new gTLDs had yet been delegated. In March 2013, the SSAC released SAC 057. [29] SAC 057 was an advisory cautioning the ICANN Board about certificate authorities (CAs) issuing certificates for internal domains that are also TLDs. One type of certificate is called an Internal Name certificate and is meant for an organization's internal use only, so the CA cannot resolve the name or look up the owner. Certificates can also contain Subject Alternative Names, which are supposed to be additional names for the same domain. At that time, the typical practice was for CAs to confirm that the internal domains were not the same as an already-delegated TLD; CAs were not checking the list of applied-for new gTLDs to see if there was a match. This could allow someone to get an Internal Name certificate matching a domain name that would soon be a TLD. SAC 057 presented a case study showing how an SSAC member was able to get a certificate issued for the www.site domain.

The most important findings from SAC 057 related to name collisions were the following:

- From Finding 3: "There are at least 37,000 internal name certificates used in thousands of enterprises. […] with the introduction of new gTLDs, namespace collisions and other man-

in-the-middle attacks (see Finding 4) will become more apparent. In addition, because many of the applied for TLDs are common, generic terms the risk of collisions increases."

- Finding 4: "The practice for issuing internal name certificates allows a person, not related to an applied for TLD, to obtain a certificate for the TLD with little or no validation, and launch a man-in-the-middle attack more effectively."

SAC 057's recommendations included "requesting that they [CAs] treat applied for new gTLDs as if they were delegated TLDs as soon as possible…."

3.3.3 Verisign Labs Report on New gTLD Security and Stability

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions

Verisign Labs issued two similar versions of a technical report, "New gTLD Security and Stability Considerations" in March 2013 (version 2.1 [30] and version 2.2 [31]). The report noted the lack of data and metrics on queries for TLDs, and how this meant the impact of delegating new gTLDs would not be quantifiable. The report emphasized the technical and logistical complexity of new gTLD adoption for registry operators, and it pushed back on ICANN's timelines for the new gTLD program, saying it did not give registry operators enough time to prepare.

Section V of the Verisign Labs report discussed name collisions specifically:

- Subsection A referenced the SAC 057 advisory and acknowledged the benefits of the advisory's recommendation to have CAs "treat applied for new gTLDs as if they were delegated TLDs" when issuing Internal Name certificates. However, it also criticized the changes to the proposed gTLD delegation processes as still providing a window of opportunity for attackers.

- Subsection C cited the lack of studies to identify the complex, subtle issues of name collisions. One particularly noteworthy statement was, "...the introduction of .info over a decade ago highlighted just what sort of obvious and nuanced interdependencies may exist as new gTLDs are delegated and made available on the Internet while applications and other systems are ill-prepared."

3.3.4 PayPal Concerns about Delegating Certain gTLDs

Applicability: Duplicate name collisions

In mid-March 2013, between the release of versions 2.1 and 2.2 of the Verisign Labs report, PayPal sent a letter to ICANN [32] regarding SAC 045 [27] and RFC 6762 [33]. The PayPal letter warned ICANN of issues with delegating certain gTLDs:

> "ICANN should consider not just the potential costs and unwanted network traffic sent to applicants for these names, but the substantial and severe costs imposed on the general Internet community arising from delegation of names that have been common *de facto* private network suffixes for nearly two decades. At minimum, the top ten observed invalid

TLDs plus those recommended for use by RFC 6762 should be permanently reserved for private use to prevent large scale disruption and damage to the millions of users and systems that rely upon them today. A more prudent approach would be to consider the negative externalities for each of the applied for new gTLDs."

The 13 names that PayPal recommended be permanently reserved were: invalid, wpad, home, belkin, corp, lan, domain, localdomain, localhost, local, intranet, internal, and private.

3.3.5 Internet-Draft on TLD Delegation Procedures

Applicability: Duplicate name collisions

On May 2, 2013, an Internet-Draft proposing procedures for TLD delegation was released. [34] It was authored by representatives of NLnet Labs, Dyn, and Google. It was updated twice in the following three months, with the last draft published on August 1. [35] The Internet-Draft specifically addressed situations where queries for a never-delegated TLD had already been seen on the Internet, and the TLD was subsequently going to be delegated.

The Internet-Draft emphasized the potential consequences if commonly used internal TLDs were delegated as public TLDs, citing SAC 045 and RFC 6762. These consequences included security issues. Section 2 stated, "Responsible administration of the public namespace therefore requires great care in permitting public delegation of any name where there is good reason to suppose it is in widespread use as a private namespace…." Section 2.1 gave a hypothetical example of a name collision caused by an organization using an internal subdomain, corp, that subsequently was delegated as a TLD, so it was no longer clear how queries for corp names should be resolved.

Section 3 recommended that zone operators monitor the frequency of queries for nonexistent domains. Such domains that receive the most queries should not be delegated as public TLDs. Section 3.2 sketched out parts of a methodology zone operators could use to determine which names are most likely to be problematic. However, the authors did not update the Internet-Draft further, and it expired.

3.4 gTLD Risk Profiles: 2013 – 2014

3.4.1 ICANN Report from Interisle Consulting Group

Applicability: Duplicate name collisions

The next major milestone was the August 2, 2013 release of an ICANN report from Interisle Consulting Group titled, "Name Collision in the DNS: A study of the likelihood and potential consequences of collision between new public gTLD labels and existing private uses of the same strings, version 1.5." [36] It was studying duplicate name collisions for gTLDs only: an internally used domain is subsequently delegated as a new gTLD. The Interisle study reviewed much of the same body of work as this NCAP Study 1 report (through mid-2013), but its most significant contribution was to analyze data sets collected from several root servers of all DNS requests they

received (totaling 94 billion, from the DNS-OARC Day in the Life of the Internet data [37]), plus a global DNS resolver service providing 53 billion requests it saw.

The findings of the Interisle study that are most noteworthy within the context of this report were:

- "**The potential for name collision with proposed new gTLDs is substantial.** Based on the data analyzed for this study, strings that have been proposed as new gTLDs appeared in 3% of the requests received at the root servers in 2013. Among all syntactically valid TLD labels (existing and proposed) in requests to the root in 2013, the proposed TLD string `home` ranked 4th, and the proposed `corp` ranked 21st. DNS traffic to the root for these and other proposed TLDs already exceeds that for well-established and heavily-used existing TLDs."

- "**The delegation of almost any of the applied-for strings as a new TLD label would carry some risk of collision.** Of the 1,409 distinct applied-for strings, only 64 never appear in the TLD position in the request stream captured during the 2012 'Day in the Life of the Internet' (DITL) measurement exercise, and only 18 never appear in any position. In the 2013 DITL stream, 42 never appear in the TLD position, and 14 never appear in any position."

- "**The designation of any applied-for string as 'high risk' or 'low risk' with respect to delegation as a new gTLD depends on both policy and analysis.** This study provides quantitative data and analysis that demonstrate the likelihood of name collision for each of the applied-for strings in the current new gTLD application round and qualitative assessments of some of the potential consequences. Whether or not a particular string represents a delegation risk that is 'high' or 'low' depends on policy decisions that relate those data and assessments to the values and priorities of ICANN and its community; and as Internet behavior and practice change over time, a string that is 'high risk' today may be 'low risk' next year (or vice versa)."

- "**For a broad range of potential policy decisions, a cluster of proposed TLDs at either end of the delegation risk spectrum are likely to be recognizable as 'high risk' and 'low risk.'** At the high end, the cluster includes the proposed TLDs that occur with at least order-of-magnitude greater frequency than any others (corp and home) and those that occur most frequently in internal X.509 public key certificates (mail and exchange in addition to corp). At the low end, the cluster includes all of the proposed TLDs that appear in queries to the root with lower frequency than the least-frequently queried existing TLD; using 2013 data, that would include 1114 of the 1395 proposed TLDs."

In summary, the Interisle study concluded that there was a risk of name collision with practically any new gTLD, but that most gTLDs would be low risk because there were few queries already being seen by the root servers for those domain names. A small number of TLDs were already mistakenly requested so often that it would surely cause significant disruptions to delegate them as public gTLDs.

3.4.2 ICANN Proposal on New gTLD Collision Risk Mitigation

Applicability: Duplicate name collisions

A few days after the Interisle study [36] was released, ICANN posted a proposal called "New gTLD Collision Risk Mitigation." [38] It proposed how the risks identified in the Interisle study [36] could be mitigated. The proposal was based on the data sets used by Interisle and information provided by CAs on the domains specified within Internal Name certificates they had issued. The proposal defined three risk profiles for applied-for new gTLDs:

- **Low-risk:** there were fewer queries being received for the not-yet-delegated TLD at the root servers as there were for other delegated TLDs that were "empty". The low-risk profile fit roughly 80% of the applied-for new gTLDs.

- **High-risk**: the number of queries being received for the not-yet-delegated TLD at the root servers was an order of magnitude higher than other such queries. The high-risk profile would fit two names, home and corp. Also, corp was the string most often seen in Internal Name certificates.

- **Uncalculated-risk**: there was not enough information to determine if these were low or high risk. This was roughly 20% of all applied-for new gTLDs.

The proposal included recommendations for mitigating the risk for each of the three risk profiles. Low-risk TLDs could be delegated, with a mandatory 120-day waiting period between signing an agreement and activating names. Also, for at least the first 30 days after first delegating a TLD, the registry operator would not activate names under the TLD, and during that time would notify the appropriate contacts for any IP address that requested a name under the TLD. High-risk TLDs were not to be delegated for the time being. Uncalculated-risk TLDs were not to be delegated until further study was completed, and applicants would also be expected to "provide evidence of the results from the steps taken to mitigate the name collision risks to an acceptable level."

ICANN also posted an announcement that gave an overview of the Interisle report, the mitigation proposal, and other information ICANN was making available related to the topic. [39]

3.4.3 Public Comments on ICANN Proposal

Applicability: Duplicate name collisions

There were dozens of responses to the ICANN proposal [38] over the next few months. Some of these also commented on the findings of the Interisle report. The entire archive of approximately 80 public comments is available online. [40] There was also a report from ICANN summarizing the public comments. [41]

There were public comments from over 15 companies about the ICANN proposal and the Interisle report overstating the risk from new gTLDs, especially those in the uncalculated-risk profile. The public comment summary [41] listed several of these concerns, including the following:

- "…all applied for new TLDs other than .corp and .home represent a combined 0.016% of the total query rate in the 2012 DITL data provided by Interisle. This figure and the potential reasons that these queries are taking place simply do not warrant mitigation through a 3-6 month delay."

- "Merely counting the number of requests for each string is completely insufficient when judging risk. The true origin of the 'collision' must be taken into account. The vast majority of requests provided in Table 12 either posed no potential risks or risks that could be handled with simple mitigations."

- "Basing risk measurement on total query counts is fundamentally flawed, especially when using data collected after the new TLD applications were posted. The Interisle report makes no mention of investigating the possibility that some of the requests were issued intentionally."

- "Risks listed by Interisle or Verisign already exist and many are prevalent in existing gTLDs such as .com. Future studies would gain credibility if the listed risks were compared against the situation in current gTLDs."

There were also numerous comments criticizing the methodology used in the Interisle study and questioning the findings of that study. The comment from Donuts [42] indicated that their own analysis found a lower rate of requests for applied-for gTLDs than Interisle's analysis did, because Donuts accounted for time to live (TTL) for DNS answers, and that Interisle had admitted they had insufficient time to perform their analysis. Donuts downplayed the risk of name collision, which included stating that, "In order to make a fair comparison of the relative risk regarding collision, it's critical to point out that Verisign, as manager of the .COM registry, experiences collision at a rate of at least 2,000 names per day for the studied period in 2013, and at least 16,000 names per day for the study period in 2012…." Donuts also provided an explanation for the prevalence of .home requests: 92% of them were from Google Chrome querying for random SLDs within the .home TLD, seeking replies that there was no such domain. Donuts provided its own set of recommendations for addressing name collisions, and stated that "no applied-for TLDs need mitigation, with the possible exception of a very few."

There were dozens of comments questioning the uncalculated-risk profile, with most asserting that only a few TLDs should be high-risk and all others should be considered low-risk and allowed to proceed with applications. One of these was from DigiCert. [43] DigiCert performed its own analysis of the Interisle data combined with "additional data on certificates, SLD information, and total number of domains." DigiCert looked at potential collisions at all levels (not just the top level). Their conclusion was that six TLDs should be considered high-risk: corp, home, mail, ice, global, and ads. All other applied-for TLDs should be considered low-risk.

One of the submitted comments was a study titled "Namespace Expansion" from JAS Global Advisors and simMachines. [44] This study used the same data that was the basis for the Interisle study, and it analyzed it to look for queries for applied-for gTLDs with a focus on the SLD names in the queries and the IP addresses making these queries. The study provided statistics, not conclusions. It was based on the assumption that "there is risk inherent in interacting on the Internet", so this study was trying to help differentiate unusual risks from typical risks.

Another comment in support of treating all but a few applied-for TLDs as low risk came from the New gTLD Applicant Group (NTAG). [45] Although the NTAG agreed that the two high-risk profile strings should not immediately proceed with delegation, the NTAG did not find justification for delaying any others. They stated, "A Verisign analysis using data from January 2006, prior to the launch of several active TLDs, found that .xxx received more queries before delegation than any other new TLD. Despite having more queries than of all of the TLDs currently under consideration in the 'Uncategorized Risk' category, .xxx was delegated in 2011. This TLD launched without incident, and no public complaints or technical issues have been identified since."

Verisign Labs submitted a report analyzing the risk for three applied-for TLDs: website, coffee, and club. [46] The website and coffee domains were initially classified as low risk, while the club domain was considered uncalculated risk. Their analysis indicated greater levels of risk for all three domains than originally estimated, and their report criticized the original analysis methodology as being inadequate both in terms of the length of time data was collected (two days) and in the importance given to the number of queries.

Verisign Labs also did its own interdisciplinary study on the risk that gTLD delegation could cause to end users. It published this study in late August 2013 and submitted it as a comment. [47] Verisign Labs proposed a methodology for measuring risk for applied-for TLDs. Based on their analysis, they discovered several cases where a particular string or strings meant for internal use was reaching root servers because of proxies, Internal Name certificates, and other reasons, and they believed delegating applied-for gTLDs would put the users in these cases at immediate risk from man-in-the-middle attacks. The study was very cautious about delegating more gTLDs, and it recommended more study and more implementation of existing recommendations for mitigating the risks.

A final example of a public comment on the ICANN proposal proposing a different risk analysis methodology was Neustar's report, *A Methodology for Assessing Collision Risk and New gTLDs*. [48] It said that "ICANN's mitigation strategy rests entirely on the possibility of collision, not the consequences." Also, "ICANN already has all the data and research necessary to calculate the risk and develop mitigation strategies that are carefully tailored to the specific risk associated with each TLD." Neustar proposed its own methodology for assessing impact based on "(i) TLD query volume; (ii) query source IP address volume; (iii) queried second-level domain volume; and (iv) volume of SSL certificates." By far the highest scoring TLDs were corp and home, with mail in third and all others far behind mail. Accordingly, they proposed having those three TLDs as high-risk and all others as low-risk.

3.4.4 ICANN Proposal on New gTLD Collision Occurrence Management

Applicability: Duplicate name collisions

ICANN released a second proposal on October 4, 2013. [49] The first proposal was on collision risk mitigation; this subsequent proposal was on managing collisions that occurred. The proposal stated that ICANN would have a name collision occurrence management framework developed. The framework would be used for each applied-for TLD to assess the likelihood of collisions and their potential impact, and to help create a name collision occurrence assessment for the TLD.

Each assessment would include suggested mitigations for SLDs within that TLD, such as blocking particular SLDs (temporarily or indefinitely). The proposal also stated ICANN would perform outreach to raise public awareness of name collisions and to educate network operators and software and equipment manufacturers about name collisions and how they can mitigate them.

Appendix I of the October proposal provided ICANN's response to the public comments on the August proposal. Responses of particular interest are as follows:

- "ICANN agrees that other parameters, besides request frequency, should be considered in assessing the threat, particularly the potential for harm caused by name collisions. ICANN will adopt the advice regarding the use of the other proposed parameters when developing a collision occurrence management framework."

- "ICANN will adopt the idea by NTAG and others to block Second Level Domain names (SLDs) that are being queried."

- "ICANN will enable an affected party to report and request the suspension of a domain name that by virtue of name collisions is causing severe harm."

- "DotGreen requested that strings in the uncalculated-risk category be allowed to proceed to contracting. Similarly, other commenters complained about ICANN not allowing these strings to proceed to contracting when the public comment period for the proposal is still open. ICANN understands the interest of applicants to see their strings move as fast as possible through the new gTLD process and will remove that restriction. The adoption of the blocking of SLDs makes this restriction unnecessary."

On October 7, 2013, the New gTLD Program Committee (NGPC) passed a resolution to have the proposal implemented. [50]

3.4.5 DNS-OARC Workshop Session on High-Risk Strings Collisions

Applicability: Duplicate name collisions

The Domain Name System Operations Analysis and Research Center (DNS-OARC) held a workshop session on high-risk strings collisions on October 5, 2013. There were four presentations in the session:

- Jim Reid from Interisle [51] spoke on the data analysis Interisle performed for their study for ICANN. He explained many of the logistical issues they experienced while attempting to analyze terabytes of data in a matter of weeks.

- Roy Hooper from Demand Media [52] discussed numerous challenges encountered when attempting to perform additional analysis of the same data Interisle had analyzed.

- Andrew Simpson from Verisign [53] presented the results of research he and his colleagues had performed on queries for applied-for gTLDs to see if there was any significance to their origins (e.g., a disproportionate number coming from a particular country). This could help identify countries at greater risk from a particular gTLD being delegated.

- Andrew Sullivan from Dyn [54] spoke about an Internet-Draft he was co-authoring with Olaf Kolkman from NLnet Labs and Warren Kumari from Google. [55] The idea was that "test delegations be used to enable empirical research on the extent of the possible disruption prior to actual allocation and delegation of any label in the root zone." The Internet-Draft proposed a methodology for doing the test delegations and collecting the necessary data. (Note that the Internet-Draft was updated twice in the following few months, but the authors eventually let it expire.)

3.4.6 SSAC Advisory SAC 062 on Mitigating Name Collision Risk

Applicability: Duplicate name collisions

On November 7, 2013, the SSAC released its SAC 062 advisory on mitigating name collision risk. [56] This advisory was based on the Interisle study [36], the August ICANN proposal [38], the October ICANN proposal [49], and SSAC's own analysis of the subject. SAC 062 stated that the SSAC generally agreed with the October proposal, and it made a few additional recommendations:

- The first was for ICANN to work with the Internet Architecture Board (IAB), IETF, and potentially others to determine which domain names should be reserved, both TLDs and lower-level names.

- The second involved trial delegation. The concept was to delegate a TLD with a short time to live, then collect data on queries for that TLD. The trial could cause name collisions for a short time, which might be temporarily disruptive but would also allow issues to be identified and addressed before permanent delegation occurred.

- The third was having policies and processes in place to roll back delegation of a TLD, if the TLD was causing security or stability issues that couldn't be immediately mitigated through other means.

3.4.7 SLD Blocking List Effectiveness

Applicability: Duplicate name collisions

As mentioned in Section 3.4.4, one of the outcomes of the public comments on ICANN's August proposal [38] was ICANN's decision to adopt SLD blocking, as announced in ICANN's October 4 proposal. [49]

Verisign Labs published a preliminary analysis of SLD blocking list effectiveness on November 5, 2013. [57] This was based on a group of gTLDs that had SLD blocking lists released on October 29. The gTLDs were: camera, clothing, equipment, guru, holdings, lighting, singles, ventures, and voyage. The initial results of the analysis indicated that the SLD blocking lists were "ineffective." It also raised questions about how the SLDs on the blocking lists were selected.

Another Verisign Labs report was released on November 15 on SLD blocking effectiveness. [58] This report continued the analysis in the November 5 report, expanding it to include newly published SLD blocking lists for 16 more gTLDs released on November 6. The report asserted that

"a fundamental reason that SLD blocking based on DITL datasets is ineffective is that the set of SLDs in queries evolves. New SLDs appear in queries all the time." Verisign Labs' analysis indicated that "the number of SLDs observed in the DITL data for the first time each year is on a significant upward trend."

On November 17, 2013, ICANN posted an announcement about SLD blocking. [59] The announcement indicated that for some gTLDs, SLD blocking was not sufficiently effective:

> "The gTLDs that were considered ineligible were those for which the growth of the number of SLDs queried year over year significantly exceeded the average growth rate for all applied for gTLDs in at least two of the DITL years (2006-2012), and for which one of the years in which this was observed was the most recent year, 2012. The analysis of this data showed that for some strings, the variance of SLDs queried varied so significantly from year to year that the mechanism of blocking SLDs might not be an effective way of addressing the name collision issue."

The announcement then listed 25 applied-for gTLDs that would not be delegated for these reasons, plus a mention that home and corp would also not be delegated.

On March 8-10, 2014, the Workshop and Prize on Root Causes and Mitigation of Name Collisions (WPNC) was held. [60] All the talks at this workshop were related in some way to name collisions. The talks listed below pertained to SLD blocking lists. Note that this report also summarizes talks with other material in their subject areas, such as the creation of the Name Collision Occurrence Management Framework (see Section 3.6). See RFC 8023, *Report from the Workshop and Prize on Root Causes and Mitigation of Name Collisions* for a summary of the workshop. [61]

- Verisign Labs personnel gave a talk and released a paper on using block lists to prevent collisions. [62] Their work was almost entirely focused on data analysis to attempt to quantify the effectiveness of SLD blocking. Their conclusion was that SLD queries change so often that SLD blocking would not be effective in mitigating name collisions.

- Paul Hoffman from the VPN Consortium [63] commented during his talk on the uncertainty about the effectiveness of SLD blocking.

- Another came from RTFM. [64] This work was also focused on data analysis, but with the purpose of determining if SLD blocking would cause harm to "naïve DNS clients" like "stub resolvers and forwarding-only devices. If these query the root servers, they can receive referral responses that they are unable to process and that would result in undefined behavior." The conclusion of the work was that these types of clients were unlikely to be harmed by SLD blocking.

3.5 Research on Name Collision Causes: 2013 – 2016

3.5.1 Search List Processing and FQDN Usage

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions

The Réseaux IP Européens Network Coordination Centre (RIPE NCC) posted an article on November 12, 2013 about new gTLDs and search list processing [65]. Search list processing is when a person specifies only a portion of a domain name instead of a fully qualified domain name (FQDN), and the search list adds to that domain name and attempts to resolve it as an FQDN, trying again with another addition if the first one fails, and so on. Although search list processing was not a novel topic at the time—it had been discussed in numerous documents before—this was when it got greater attention on its own. The RIPE NCC article argued that the risks from name collisions caused by searches for internal domains leaking out to the Internet were much lower than had been claimed. However, it also appealed for operating systems and web browsers to handle searches for names in standard ways so that the leaking of queries for internal domains would stop.

ICANN released a blog posting on December 6, 2013 on managing name collision occurrences, focusing on issues with search list processing [66]. The blog posting referenced version 1.0 of the *Guide to Name Collision Identification and Mitigation for IT Professionals*, which ICANN had released the day before the blog posting. [67] Both the blog posting and the guide had the same motivation: to stop the leaking of queries for internal domains, which would prevent name collisions from occurring, by encouraging organizations to migrate from internal-only, shortened domain names to public FQDNs.

In February 2014, the SSAC released its SAC 064 advisory on DNS search list processing. [68] This advisory discussed the inconsistencies in how search list processing was performed by operating systems, web browsers, email clients, and other software. Although there were RFCs (1123 [69], 1535 [19], and 1536 [70]) with search list guidelines, these RFCs were informational and had not been widely adopted, and there were also concerns that the RFCs were not as clear and specific as they needed to be. The SAC 064 advisory documented some of the differences in search list processing among commonly used client operating systems. Most importantly, Section 4 of SAC 064 proposed improvements to search list processing that would reduce the likelihood of name collisions.

In March 2014, there were several talks at the Name Collision Workshop pertaining to search list processing:

- Warren Kumari [71] spoke about the need to educate developers on not using shortened names instead of FQDNs, and on the value of reserving .alt as a local-only domain name.

- Paul Hoffman from the VPN Consortium [63] talked about what organizations could do to help mitigate name collisions. He encouraged organizations to stop using shortened names and to use FQDNs instead. He also mentioned that "determining the so-called 'potential for collisions' for a private namespace is nearly impossible."

- Colin Strutt from Interisle [72] presented on the corp.com domain. This domain had been registered in 1994 by Mikey O'Connor, but no SLDs within it were registered, so any queries for this domain were likely to be queries for organizations' internal .corp names that, because of search list processing, had leaked onto the Internet. Mikey O'Connor and NetChoice sponsored a small study of this query data by Interisle. The corp.com domain was receiving approximately 2 million queries a day from a wide variety of IP addresses, domains, and countries. The study also attempted to contact some of the organizations and

internet service providers where the queries were coming from in order to get them to stop issuing these queries, but this met with little success.

- Casey Deccio from Verisign Labs [73] spoke on quantifying risk from name collisions by creating a name collision model that takes search lists into account.

- Andrew Simpson from Verisign [74] spoke about detecting search lists. He experimented with real-world systems to observe their name resolution behavior and compare this with DITL data on queries for nonexistent domains. He also gave a second talk [75] on DNS query analysis techniques that might be of use in name collision discovery.

3.5.2 Causes of Internal Domain Leakage

Applicability: Duplicate name collisions

Verisign Labs published a paper in November 2014 on leakage of the .onion domain, which is a non-delegated TLD meant for Tor usage only. [76] The nature of Tor is such that .onion queries should not leak onto the Internet, but analysis of root server data on queries indicated leakage was definitely happening. The reasons for the leakage were unknown, but there were several possibilities cited:

- User error
- Client software misconfiguration
- Browser prefetching
- Third-party applications or plug-ins
- Search list processing
- Web crawlers
- Malware

Another paper on .onion domain leakage was submitted in March 2016 and published in October 2017. [77] Written by university researchers, one of whom was also a co-author on the Verisign Labs 2014 paper [76], this paper used data from DITL and other sources, and it provided a more rigorous analysis of the data for .onion domain leakage than the 2014 paper. It also looked at some potential causes of the leakage in greater detail:

- User error and misconceptions: the researchers surveyed graduate students in a computer security class about the .onion domain, and only half the students who considered themselves “very knowledgeable” about Tor knew the special function of the .onion domain.

- Browser prefetching and web crawlers: some web browsers would try resolving the links on a webpage in advance of anyone clicking on those links, so that could cause .onion links

to try to be resolved when Tor isn't running. To evaluate this, the researchers did a website crawl and looked for strings ending in ".onion", and they found that 17% of the instances of .onion queries seen in the DITL data corresponded to strings seen during their crawl.

- Malware: the researchers looked for .onion queries for SLDs known to host malware, but there was not a clear correlation.

RFC 7686, *The ".onion" Special-Use Domain Name*, was published in October 2015. [78] It explained the unique role of the onion domain name, and it defined how queries for onion names should be resolved, which would prevent further leakage.

In May 2017, there was a presentation about the Operational Research Data from Internet Namespace Logs (ORDINAL) dataset. [79] This presentation gave numerous examples of protocols and applications that misused DNS by using it for authentication, not identification—essentially, they trusted whatever result they got from DNS as being sufficient confirmation of the legitimacy of the destination, instead of subsequently performing authentication with the destination to verify it.

3.5.3 Detection of Leaking Clients

Applicability: Duplicate name collisions

As discussed in Section 3.6.3, the JAS Global Advisors report mentioned observations that indicated some attackers were purposely choosing domains with collisions so they could take advantage of those collisions. There have been works published since that timeframe regarding how attackers could utilize name collisions by detecting internal queries leaking from clients onto the Internet and registering the searched-for names.

Two of these papers focused on vulnerabilities in the Web Proxy Auto-Discovery (WPAD) protocol. Both papers were published in May 2016; one was authored by Verisign Labs personnel only [80] and the other was co-authored by Verisign Labs personnel and University of Michigan researchers [81]. Also published in May 2016 was an alert from the National Cyber Awareness System on the WPAD name collision vulnerability. [82]

The WPAD issue involved internal-only domain names not being found when laptops using WPAD were used on external networks, so the laptops were sending DNS queries to the Internet. Attackers aware of this behavior could register the domains the laptops were erroneously trying to reach and perform man-in-the-middle attacks on the laptops. The problem was first found on Microsoft laptops, but it was soon confirmed that Apple and Linux laptops had the same problem. The recommended mitigation was to disable WPAD if not needed, otherwise to hard-code proxy addresses instead of using WPAD to acquire them. The [80] and [81] papers both highlighted that transient devices like laptops might encounter name collisions more frequently than other devices because transient devices go from one network to another.

The same researchers who authored [81] plus two additional University of Michigan researchers wrote a conference paper published in November 2017 on client-side name collision vulnerabilities. [83] This paper covered a broader range of vulnerabilities than just WPAD. The authors created a general name collision threat model for clients querying internal-only names that

were being leaked onto public networks. The authors then analyzed DITL query data for 2011 through 2016 to find evidence of internal services (WPAD and many others) being exposed. They found that 115 registered services and an undetermined (but large) number of unregistered services were exposed, and they chose 48 of the most commonly seen services for further analysis. The researchers then looked for vulnerabilities in those services and determined that 93.8% of them were vulnerable, for reasons such as lack of server authentication or accepting a different server certificate than the one expected without notifying the user. Further discussion of the contents of the paper is outside the scope of this document, because the service vulnerabilities were not name collision related; the relevance of the paper is that leakage of internal names associated with services puts those services at high risk of exploitation if an attacker registers a particular name collision domain.

Four of the co-authors of the papers mentioned above filed a patent application on March 24, 2017. [84] This patent proposed ways to detect internal names leaking, especially for the WPAD vulnerability, and remediate the problems causing the leaks. As of this writing, the patent application is still pending.

3.6 Name Collision Occurrence Management Framework: 2014 – 2015

3.6.1 JAS Global Advisors Phase One Report Draft

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

The Name Collision Occurrence Management Framework was to be developed so it could be applied to any newly requested gTLD to assess risk and identify mitigations before the gTLD was delegated. JAS Global Advisors was selected by ICANN in November 2013 to create the Name Collision Occurrence Management Framework. The phase one report draft for their work was released for public comment on February 26, 2014. [85] JAS Global Advisors also presented on this work at the Name Collision Workshop in March 2014 [86] and at the March ICANN meeting in Singapore [87] [88]. Among their findings at that time were the following:

- DNS name collisions happened frequently and have happened before the delegation of each new TLD since at least 2007. Issues caused by collisions date back to approximately 1987. There was "no evidence to suggest that the security and stability of the global Internet DNS itself is at risk."

- There were several causes of these collisions, including shortened name usage, search list processing differences, misunderstandings about DNS, expired registrations, human error, and intentional acquisition of colliding names.

- Other types of namespaces have had collisions, with numerous examples from phone numbers and mailing addresses. These were handled through advance notification of the transitions, and in having a grace period of some sort when feasible.

- The corp, home, and mail TLDs should not be delegated because of their existing widespread internal use by organizations. RFC 6762 [33] "suggests that .corp and .home are safe for use on internal networks." Also, ".mail has been hardcoded into a number of installations…and has a large global 'installed base' that is likely to have significant inertia comparable to .corp and .home."

- There should be processes in place to act if a TLD delegation presents "clear and present danger to human life."

The phase one report draft compared the disruption caused by a name collision to the disruption caused by failing to renew a domain. "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."

The report discussed at length how new gTLDs could be delegated using a method called "controlled interruption." Instead of simply delegating a new gTLD and allowing traffic that previously would have gone elsewhere to inadvertently reach the newly delegated gTLD instead, queries for the gTLD would receive a response that directs them to a different address that essentially indicates an error has occurred. The report discussed two options for this "different address"—a honeypot or a loopback address—and recommended loopback addresses because this "prevents traffic from leaving the requestor's network and blocks a malicious actor's ability to intercede."

JAS Global Advisors recommended that a standard loopback address should be used for all controlled interruptions: 127.0.53.53 (with 53 chosen because it is the port number associated with DNS). Having the same unusual IP address returned in all controlled interruption replies should help system administrators to identify the problem. They recommended having a 120-day controlled interruption period:

> "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 120-day controlled interruption period, registries will not be subject to further collision-related restrictions. …we believe the 120-day controlled interruption period offers a conservative buffer between potential legacy usage of a TLD and the new usage."

In other words, for a new TLD on "alternative path," reply with the loopback address for SLDs on the blocking list only. For all other new TLDs, reply with the loopback address for every SLD. That is the equivalent of a wildcard, and the draft report recommended that wildcard records be permitted for the purpose of controlled interruption for TLDs not on "alternative path," since the TLD would not have any registrant data during that period.

Section 2.1.2 of the report draft briefly discussed a trial JAS Global Advisors had performed of controlled interruption, and "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."

The report draft also mentioned alternatives to controlled interruption, "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." See Section 3.6.3 for further discussion of the alternatives to controlled interruption and other contents of the phase one report draft, as presented in the final version of the report.

3.6.2 Public Comments on Phase One Report Draft

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

Over 25 public comments were submitted on the phase one report draft. The complete set of public comments is available at [89]. An ICANN report summarizing these comments was also posted. [90]

Most of the comments received largely agreed with the report draft and its recommendations, except for the topic of 120-day controlled interruption. Many felt it should be shorter, perhaps 38, 45, or 60 days, although one felt 120 days was too short to allow organizations to remediate problems.

There was some disagreement about whether using the 127.0.53.53 loopback address or a honeypot would be better. The majority felt the address would be better than a honeypot. An alternate solution proposed was to use a public IP address and website that could provide information on the nature of the problem to end users.

On the topic of not delegating the corp, home, and mail TLDs, there was no consensus. Some felt they should all be permanently reserved, while others thought more discussion and evaluation was needed, and yet others disagreed with permanently reserving any of them.

Several commenters mentioned that collisions happened all the time in .com and nothing was done about that, but great scrutiny was being given to a much smaller problem with collisions involving new gTLDs.

Verisign released preliminary public comments on the phase one report draft in late February 2014 and updated those comments on March 31. [91] [92] Verisign released an additional set of comments on the phase one report draft on April 21, 2014, in part to clarify the preliminary comments. [93] [94] Topics of the Verisign comments included the following:

- Verisign did not find the expected elements of the Framework in the draft report; instead, the draft report focused on using controlled interruption when delegating new gTLDs, as if an undefined Framework were already being applied.

- Controlled interruption had not been tested. There were two scenarios where it might not succeed in notifying organizations of an impending change in name resolution. The first scenario was an "alternative path" delegation where someone queries for an SLD that isn't on the blocking list. The second scenario involved use of WPAD. Implementations of

WPAD vary, and a non-recommended WPAD implementation might receive a controlled interruption reply, ignore it, and continue with its search list processing.

- "There is therefore a reasonable case to be made, at least for some new gTLDs and SLDs, that the controlled interruption should be done more selectively." Verisign termed this "selective interruption" and said it "requires careful qualitative analysis" to determine when it can be used instead of the broader controlled interruption. There were also some drawbacks to the selective interruption approach. Another alternative approach from other public commenters (United TLD, the NTAG, and the China Internet Network Information Center [CNNIC]) was to allow all SLDs in a new gTLD to be interrupted except for those SLDs that have already been delegated.

- Verisign was opposed to using external honeypots because of the likelihood of sensitive data inadvertently being leaked over unsecured networks (e.g., the Internet) to the honeypot.

- Expired registration name collisions (the term used in this NCAP Study 1 report, not the JAS Global Advisors report or Verisign comments) were not actually a form of name collision.

- "If controlled interruption is adopted, the only way to get a better understanding of the appropriate period is by qualitative analysis of the effectiveness of the mitigation measure in practice."

3.6.3 JAS Global Advisors Final Phase One Report

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

The final phase one report was released on June 4, 2014 [95]. Its assertions of no significant problems caused by name collisions are noteworthy:

> "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."
>
> "As we write this update, 275 New gTLDs have been delegated and over 835,000 second level registrations have been added." … "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."

Significant changes from the draft report included the following:

- The length of the controlled interruption period was dropped from 120 days to 90 days.

- A discussion of controlled interruption for IPv6 addresses was added. The assertion was that there was not a significant problem with IPv6-only hosts having name collisions, and there was not an IPv6 counterpart to 127.0.53.53 that could be used for controlled interruption.

- A discussion of having staggered controlled interruption periods instead of continuous controlled interruptions was added. This idea was suggested by Google in the public comments. The report did not favor staggered controlled interruptions because it would cause "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…."

Section 3.1 of the final report discussed alternatives to using controlled interruption with loopback addresses:

- **String-by-string approaches (TLD and SLD)** (as detailed in [96]): "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 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."

- **Honeypots**: Honeypots would be more useful from a notification standpoint than loopback addresses, but honeypots have several drawbacks. In addition to the potential exposure of sensitive data across networks and to the honeypot itself, honeypot use would also make it possible for someone to "game" things by sending queries to make it look like the new gTLD's delegation should be delayed. Also, the data on the honeypot could be subject to privacy laws and regulations from numerous jurisdictions.

- **Use of DNAME records**: While wildcard DNAME records could point to something like "you-need-to-change-your-dns-config-see-collisions-dot-icann-dot-org.", DNAME has only been well supported since around the year 2000, whereas loopback addresses have been well supported since around 1989. Also, "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."

The final report also added a new Section 3.3, "Collisions in Existing DNS Namespace". To measure collisions within existing TLDs, JAS Global Advisors registered some SLDs and found that "these registrations immediately generated a surprising amount of traffic." They noted that they used tools meant to aid people with "domain drop catching" and "squatting". *Domain drop catching* is, in the parlance of NCAP Study 1, a re-registered name collision that is performed immediately after an expired domain becomes available, typically for malicious purposes. *Squatting*, which refers to someone registering a TLD to prevent someone else from registering it, is not a form of name collision and is outside the scope of the NCAP study. The volume of queries received immediately after registration indicated the tools may have had access to feeds with data on queries to nonexistent domains, and took advantage of that data to intentionally cause name collisions and benefit from them.

3.6.4 SSAC Response to the Final Phase One Report

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

The SSAC released SAC 066, their comment on the final phase one report, on June 6, 2014. [97] The Recommendations from SAC 066 that differed from the recommendations in the final phase one report and are applicable to NCAP Study 1 are as follows (with recommendation text bolded and supporting text not bolded):

- "**Instead of a single controlled interruption period, ICANN should introduce rolling interruption periods, broken by periods of normal operation, to allow affected end-user systems to continue to function during the 120-day test period with less risk of catastrophic business impact.** Controlled interruption periods starting at 24 hours and eventually lengthening to 30 days would be separated by periods of at least 3 days, to allow users or system administrators to identify or develop and put in place solutions or workarounds."

- "**ICANN should perform an evaluation of potential notification approaches against at least the requirements provided by the SSAC prior to implementing any notification approach.**" This was due to SSAC's concerns that use of the 127.0.53.53 loopback address would only effectively notify some system administrators, but not typical end users. The SSAC felt that there was "a wealth of operational expertise" in handling sensitive data sent to honeypots and that privacy concerns should not preclude the use of honeypots instead of loopback addresses.

- "**ICANN should implement a notification approach that accommodates IPv6-only hosts as well as IPv4-only or dual-stack hosts.**"

- "**ICANN should consider not taking any actions solely based on the JAS Phase One Report.**"

- "**ICANN should seek to provide stronger justification for extrapolating findings based on one kind of measurement or data gathering to other situations.**" This was in response to this assumption from the Phase One report: "The modalities, risks, and etiologies of the inevitable DNS namespace collisions in the new TLD namespaces will resemble the collisions that already occur routinely in other parts of the DNS." The SSAC questioned whether this assumption was valid.

Appendix A of SAC 066 discussed four alternative notification approaches:

1. "Do nothing. Users…will experience failures or misconnections and come to realize their configurations are problematic only after the new gTLD and domains within that gTLD are delegated and elicit operational impacts to their systems." This approach was deemed unacceptable.

2. "Perform qualitative analysis of query sources as measured at root and TLD servers and provide proactive user notification." This approach would require the root servers to have measurement capabilities they did not yet possess, so it was not an option in the short term.

3. "Implement structured, short-term test periods ('controlled interruption'), in which end users utilizing a proposed gTLD will experience a failure, and then be given time (after each short-term test period) for planning and effectuating remediation efforts specific to their environment. This approach triggers the errors in a more controlled environment, and can be used as an early warning system to notify potentially impacted parties. There are two variations to notification in this approach:"

    a. Loopback address usage (127.0.53.53)

    b. Redirection to honeypot

3.6.5 Approval of the Name Collision Occurrence Management Framework

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

ICANN approved the Name Collision Occurrence Management Framework on July 30, 2014. [98] ICANN stated in the introduction to the Framework that they took into consideration the JAS Global Advisors final Phase 1 report [95], the public comments submitted on the draft Phase 1 report [90], and the SAC 062 [56] and SAC 066 [97] documents.

The Framework required registry operators to do continuous controlled interruption for each new gTLD for a minimum of 90 days. It was stated that "there is already a mechanism in place (name collision reporting) for affected parties to find temporary relief from name collision harm, if needed, making the intermittent approach an unnecessary burden" for registries. There was not yet an IPv6 counterpart to the IPv4 loopback address, but ICANN was to collaborate with other organizations in finding a suitable mechanism for IPv6-only hosts.

The Framework also stated that ICANN would defer the corp, home, and mail TLDs indefinitely and would work with other organizations to determine how to handle them long-term.

A few days after approval of the Framework, ICANN released version 1.1 of the *Guide to Name Collision Identification and Mitigation for IT Professionals*. [99] This new version included the requirements from the Framework, such as the 90-day continuous controlled interruption period. A few days after that, ICANN released a Name Collision Occurrence Assessment document for gTLD applicants and registry operators. [100] This included details of implementing continuous controlled interruption and responding to requests from ICANN regarding name collision report handling. In November 2014, ICANN released Addendum to Name Collision Occurrence Assessment, which pertained to trademark claims. [101] ICANN also released a Frequently Asked Questions (FAQ) on the Framework for registries [102] and a FAQ on name collisions for IT professionals [103].

3.6.6 Controlled Interruption for New ccTLDs

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

In February 2014, while the Name Collision Occurrence Management Framework [98] was under development, ICANN released a name collision briefing document that summarized ICANN's ongoing efforts to address name collisions for new gTLDs. [104] Section 4 of the briefing document pointed out the potential relevance of the gTLD name collision work for ccTLDs: "The issue of name collision is not unique of new gTLDs and could present in new ccTLDs too, both ASCII and IDN. ICANN is requesting the ccNSO to review the name collision issue and its implications for new ccTLDs." The briefing document also stated the following:

> "Until advice is received from the ccNSO, ICANN plans to send each new ccTLD manager the same kind of interim report that new gTLDs received for the alternate path to delegation. It will remain the responsibility of the local Internet Community and the ccTLD manager to either: 1) proceed to delegation while temporarily blocking the SLDs identified in the report; 2) temporarily defer delegation until receipt of their full collision occurrence assessment and implementation of the measures described; or 3) some other course of action determined by the local Internet Community and the ccTLD manager."

Shortly after the Name Collision Occurrence Management Framework [98] was approved for new gTLDs, ICANN recommended it also be used for each new ccTLD. This included having a continuous controlled interruption period of at least 90 days. The same resources ICANN had already made available for name collisions for new gTLDs (as described in Section 3.6.5) were also relevant for new ccTLDs, and people were pointed to those resources for more information. [105]

3.6.7 JAS Global Advisors Phase Two Report

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

Release of JAS Global Advisors' phase two report was delayed because it would have disclosed a vendor security vulnerability. The final report wasn't publicly released until October 2015 [106]. The first few sections were duplicates of the JAS Global Advisors final phase one report [95], with a few notable changes and additions that reflected events occurring during the delay:

- "…several vendors have…included detection and messaging around the 127.0.53.53 response. For example, recent builds of Google's Chrome browser now include the new error 'ERR_ICANN_NAME_COLLISION' which provides specific and richer error messaging to the user over a general connection timeout."

- There was a new Section 3.1.5, "Effectiveness of Controlled Interruption" that discussed gTLD delegations and real-world name collisions.

    - There had been more than 650 gTLD delegations, and ICANN had "received fewer than 30 reports of disruptive collisions since the first delegation in October of 2013.

None of these reports have reached the threshold of presenting a danger to human life."

- "As expected, controlled interruption caused some instances of limited operational issues as collision circumstances were encountered with new gTLD delegations. While some system administrators expressed frustration at the difficulties, overall it appears that controlled interruption in many cases is having the hoped-for outcome. … JAS would characterize the overall response as 'annoyed but understanding and generally positive.'"

- "JAS also is aware of specific examples where controlled interruption, for whatever reason, did not cause underlying DNS issues to be remedied." In regard to one example: "JAS suspects that in this specific instance, controlled interruption was probably not disruptive enough to get the attention of operators; or if it did get the attention of operators, the issue was not viewed as important enough to cause action. Based on JAS' knowledge of the specific circumstances surrounding this operator, it is unlikely that a longer controlled interruption period or an entirely different approach to controlled interruption would have made a difference."

- JAS Global Advisors tested HTTP honeypots in SLDs known to have high volumes of collisions; reaching a honeypot would return a web page with contact information for JAS and a request to contact JAS. They received no replies. "Reviewing our HTTP logs, less than 8% of DNS resolutions ultimately led to the retrieval of one of our HTTP honeypot pages. Reviewing the HTTP logs further, less than 12% of those 8% reported an HTTP user-agent that could be considered a user-facing application (i.e. a Browser)."

The substantive new material in the phase two report started in Section 4.1, which said that JAS analysis showed many of the queries to nonexistent domains were generated by malware. It also mentioned the queries generated by Google Chrome, which queried for random SLDs within the .home TLD to try to get replies that the domains did not exist. Taken together, queries to nonexistent domains automatically generated by malware and Google Chrome represented "nearly 80% of the random and pseudo-random labels we detected in DITL datasets and in excess of 41% of the total NXDOMAIN traffic described in the DITL datasets. This is consistent with the observation that the 'Alternate Path to Delegation' Second Level Domain (SLD) Collision Block Lists published by ICANN are comprised largely of these seemingly random, pseudo-random, machine-generated or otherwise linguistically nonsensical labels."

Section 5 of the phase two report elaborated on the material briefly discussed in Section 3.3 of the phase one report, where JAS Global Advisors had registered some SLDs in order to measure collisions within existing TLDs. The phase two report indicated that they registered over 50 SLDs. Through their research and analysis, they eventually discovered a vulnerability in Microsoft products, which was the disclosure-related issue that caused the delay in releasing the phase two report.

3.7 Potential Changes to Existing gTLD Processes: 2016 – present

3.7.1 ICANN New gTLD Subsequent Procedures (SubPro) Working Group

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions, re-registered name collisions

The purpose of the New gTLD Subsequent Procedures (SubPro) Working Group is to use "the community's collective experiences from the 2012 New gTLD Program round to determine what, if any changes may need to be made to the existing Introduction of New Generic Top-Level Domains policy recommendations from 8 August 2007." [107]

A July 2018 initial report from the SubPro Working Group [107] indicated that their Work Track 4 would address name collisions. Section 2.7.8, "Name Collisions," of the initial report (pages 156-164) discussed the changes that had occurred regarding name collisions since 2012. The SSAC had previously provided input to the SubPro Working Group, as documented in SAC 094 (May 22, 2017). [108]

The SubPro Working Group's initial report included a set of preliminary recommendations for name collisions:

- "2.7.8.c.1: Include a mechanism to evaluate the risk of name collisions in the TLD evaluation process as well during the transition to delegation phase.
- 2.7.8.c.2: Use data-driven methodologies using trusted research-accessible data sources like *Day in the Life of the Internet* (DITL) and *Operational Research Data from Internet Namespace Logs* (ORDINAL).
- 2.7.8.c.3: Efforts should be undertaken to create a 'Do Not Apply' list of TLD strings that pose a substantial name collision risk whereby application for such strings would not be allowed to be submitted.
- 2.7.8.c.4: In addition, a second list of TLDs should be created (if possible) of strings that may not pose as high of a name collision risk as the 'Do Not Apply' list, but for which there would be a strong presumption that a specific mitigation framework would be required.
- 2.7.8.c.5: Allow every application, other than those on the 'do not apply' list, to file a name collision mitigation framework with their application.
- 2.7.8.c.6: During the evaluation period, a test should be developed to evaluate the name collision risk for every applied-for string, putting them into 3 baskets: high risk, aggravated risk, and low risk. Provide clear guidance to applicants in advance for what constitutes high risk, aggravated risk, and low risk.

- 2.7.8.c.7: High risk strings would not be allowed to proceed and would be eligible for some form of a refund.

- 2.7.8.c.8: Aggravated risk strings would require a non-standard mitigation framework to move forward in the process; the proposed framework would be evaluated by an RSTEP panel.

- 2.7.8.c.9: Low risk strings would start controlled interruption as soon as such finding is reached, recommended to be done by ICANN org for a minimum period of 90 days (but likely more considering the typical timeline for evaluation, contracting and delegation).

- 2.7.8.c.10: If controlled interruption (CI) for a specific label is found to cause disruption, ICANN org could decide to disable CI for that label while the disruption is fixed, provided that the minimum CI period still applied to that string."

The SSAC provided feedback on the SubPro Working Group's initial report in SAC 103, posted October 3, 2018. [109]

As of this writing, the final report from the SubPro Working Group is not yet available.

3.7.2 Requests to Delegate corp, home, and mail

Applicability: Duplicate name collisions, shortened name collisions, search list name collisions

Background on corp, home, and mail being reserved

As previously discussed in Section 3.3.1, the ICANN SSAC's SAC 045 report in 2010 [27] recommended prohibiting the delegation of certain domain names as TLDs. ICANN's *gTLD Applicant Guidebook* [25] released in 2012 specified prohibited names that included the reserved TLD names from RFC 2606 [28]—test, example, invalid, and localhost—plus a few dozen more, in what was termed the Top-Level Reserved Names List. Most of the additional names were specific to internet infrastructure, like apnic, iab, iana, icann, ietf, and ssac, while a few were more general, such as local.

In February 2013, RFC 6761, *Special-Use Domain Names* [110] defined how the RFC 2606 names should be treated, with RFC 6762, *Multicast DNS* [33] providing additional guidance on handling usage of the reserved names from RFC 6761.

Section 3.4 of this report discussed in detail the evaluation of various TLD names, and Section 3.6 covered the recommendations from the JAS Global Advisors Phase One final report [95] and the ICANN approval of the Name Collision Occurrence Management Framework [98], which prevented delegation of the corp, home, and mail TLDs for the time being.

Since at least 2013, perhaps earlier, parties have been asking for the corp, home, and mail TLDs to be delegated. In August 2016, a group of applicants for those three TLDs sent a letter to ICANN asking for the names to be released because the risks that were present some years ago have been mitigated. [111] The letter included the following:

> ".HOME, .CORP, and .MAIL were originally put on the high-risk list due to an anticipated combined effect of conflict with internal name certificate authority use and the number of queries to the root where no name existed (sometimes referred to as 'name collisions'). The unreliability of self-assigned certificates, however, was mitigated last year with the reassignment of certificates to internal names and private IP addresses (i.e., for internal networks). This effective mitigation, coupled with the completion of controlled interruption of new gTLDs without incident, presents evidence that risks anticipated by the JAS report were grossly overstated.
>
> These results, at a minimum, call for a new examination to determine whether the basis for the Board's earlier decision to stymie .HOME, .CORP, and .MAIL remains valid, and whether the original assumptions and recommendations continue to hold, given current experience. Just as the name collision issues were mitigated in all other gTLDs, the same likely is true for these three gTLDs."

RFC 8244, *Special-Use Domain Names Problem Statement* was released in October 2017. [112] Among the challenges it discussed were those involving reserving additional domain names so they are not publicly delegated as TLDs. RFC 8244 referenced an Internet-Draft from 2015, *Additional Reserved Top Level Domains*, that was not finalized and expired. [113] That Internet-Draft proposed classifying the corp, home, and mail domain names as reserved in compliance with RFC 6761. [110] RFC 8244 also referenced RFC 7788, *Home Networking Control Protocol*, [114] which specified in Section 8 the use of ".home" as the default "network-wide zone" for name resolution on a home network.

In response to the August 2016 letter, the ICANN Board approved resolutions on November 2, 2017 regarding the corp, home, and mail strings. [115] The resolutions indicated that "the effect of name collisions on interoperability, resilience, security and/or stability of the DNS is not fully understood" and "the Board has made no determination as to the efficacy or feasibility of potential mitigation mechanisms for Name Collision, and remains focused on minimizing or avoiding risk to the security and stability of the DNS."

Consequently, the Board asked the ICANN SSAC "to conduct a study… to present data, analysis and points of view, and provide advice to the Board regarding the risks posed to users and end systems if .CORP, .HOME, .MAIL strings were to be delegated in the root, as well as possible courses of action that might mitigate the identified risks," as well as a study on several questions related to name collisions in general. That was the driver for this NCAP Study 1 and report.

> Note that while there has been continued interest in delegating corp, home, and mail, there has also been continued interest in not delegating them and in reserving additional names. For example, there was a 2017 Internet-Draft proposing reservation of ".internal" as a TLD. [116] There is another Internet-Draft, started in 2014 and still in progress as of this writing, proposing ".alt" as a reserved domain name not to be used for DNS. [117]

In April 2020, the ICANN Office of the Chief Technology Officer (OCTO) published *Study of the Prevalence of DNS Queries for CORP, HOME, and MAIL*. [118] Conducted in 2017, the study analyzed a representative sampling of root server traffic over an extended period (19 months for one server and 9 months for a second server). The study examined the queries for nonexistent domain names, and compared the relative prevalence of the most commonly queried names with

1263 those from the Interisle report [36], which used DITL data from 2012 and 2013. The OCTO study
1264 found that the corp and home strings were still the most requested nonexistent domain names, and
1265 the ranking of the mail string had not substantially changed either.

1266

4 The Known Harm of Name Collisions and the Technical Impact of Controlled Interruption

The study RFP [2] specified that this report must include the following:

- Study task 2b: "summarizes the known (evidenced) harm of name collisions"
- Study task 2d: "documents any mitigations/actions taken so far, specifically including controlled interruption, and the technical impact of those mitigations only (no examination to be undertaken of the non-technical impacts such as resourcing or costs)" (note: the first part of this was already documented in Section 3.6)

Much of the publicly available information on the known harm of name collisions is not relevant for evaluating current and future risks because it occurred before controlled interruption usage began in 2014. The 90-day controlled interruption periods became mandatory for new gTLD delegation starting in August 2014 and were recommended for new ccTLD delegation in October 2014 [119].

Before then, what happened is name collisions occurred, and at some point there was increased awareness of a particular cause of name collisions, so that cause was addressed and future harm was avoided. An example is re-registered name collisions, as discussed in Section 3.2. There were definitely organizations harmed by their domains expiring and subsequently being registered and misused by others, but this has been a known issue for many years, and organizations have full control over and responsibility for preventing this form of name collision. Similarly, the duplicate name collision risks from Internal Name certificates (see Section 3.3.2) were addressed by CAs changing their processes.

Accordingly, this section of the report summarizes the known harm of name collisions for TLDs since controlled interruption for new TLD delegation began. This section of the report also describes, documents, and analyzes the technical impact of controlled interruption. Controlled interruption is intended to reduce harm—for example, by preventing an organization's network traffic from inadvertently leaking to another organization—but it can still cause harm, such as by causing that network traffic to be routed to the special loopback address. Any discussion of harm from name collisions will be closely tied with a discussion of the technical impact of controlled interruption, so both topics are discussed jointly in this section.

Note that this report does not attempt to define "harm" and that it recognizes a definition of "harm" is needed. For the purposes of this report, a broad interpretation of "harm" is taken to mean anything that negatively affects anyone or any entity using DNS.

4.1 Preparation

As described in Section 3.6.1, controlled interruption was proposed for use to help mitigate name collision risks for new gTLDs. Section 3.6.5 explained that the Name Collision Occurrence Management Framework [98] was approved on July 30, 2014, and it required registry operators to do continuous controlled interruption for each new gTLD for a minimum of 90 days. The same

controlled interruption measures were recommended for each new ccTLD on October 2, 2014. [119]

Attempts to query a new TLD during the controlled interruption period for an "A" record (an IP address) would result in a reply utilizing the loopback address 127.0.53.53. The idea was that this address would be unexpected and unusual, with the repeated "53" values implying the relationship to DNS. DNS queries looking for text records ("TXT") would return the following: "Your DNS configuration needs immediate attention see https://icann.org/namecollision". Other types of DNS queries would return an answer containing the string "your-dns-needs-immediate-attention." as part of the domain name. Doing a subsequent query for that domain name would return the 127.0.53.53 address. [99]

ICANN also increased awareness of controlled interruption through other means. This ranged from creating online technical resources like webpages [14] and the *Guide to Name Collision Identification and Mitigation for IT Professionals* [99] to having social media [120], articles [121], and even Google ads [122] that referenced the 127.0.53.53 address, controlled interruption, and ICANN's name collision resources website [14].

Finally, ICANN provided a webform so any parties adversely affected by a name collision (including a controlled interruption) could report it. [123] The page currently says, in part, "If you believe your name collision meets the criteria above (i.e. your system is suffering demonstrably severe harm as a consequence of name collision or you have a reasonable belief that the name collision presents a clear and present danger to human life), please use the form below to submit your report to ICANN."

4.2 Name Collision Reports

There is no way to quantify the number of name collisions encountered during controlled interruption periods, let alone the nature of the collisions, such as severity, length of time, or name collision cause. Any study of actual name collisions during controlled interruption will be largely anecdotal. To get a somewhat broader view of name collisions, this section looks at reports made both to ICANN and to others (e.g., system administrator sites, user forums, bug tracking systems).

Note that, as mentioned in Section 3.6.7 of this report, JAS Global Advisors stated in their phase two report [106] that there were no significant problems from the delegation of new gTLDs (as of that writing, approximately 650 gTLDs).

4.2.1 Reports to ICANN

The lower (blue) line in Figure 1 shows the number of name collision reports ICANN received by half-year. [124] The upper (orange) line shows the number of new gTLDs delegated during the same half-year periods. [8] Note that as of this writing, there have been a total of 57 IDN ccTLDs delegated through the IDN Fast Track Program since the first new IDN ccTLD requests were submitted in 2009. [125] Because there have been so few new ccTLDs compared to new gTLDs, and even compared to new collision reports, a line on the graph for the new ccTLDs would not be distinguishable from zero values, so it has been omitted.

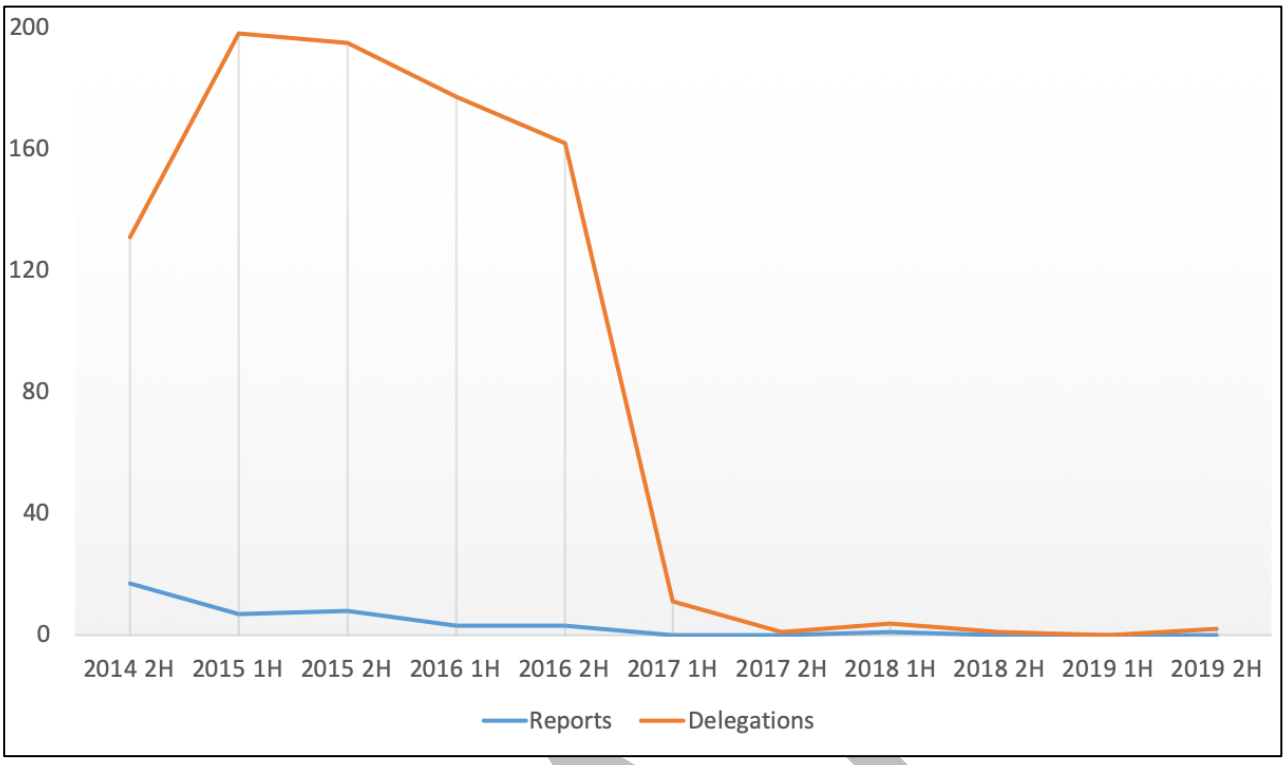

Figure 1: Name Collision Reports to ICANN by Half Year

The following statements are based on the data used for Figure 1:

- The vast majority of new TLDs delegated since July 2014 have not been the subject of any name collision reports to ICANN.
- For every one report in the second half of 2014, there were approximately eight TLDs delegated. During 2015, the ratio was roughly one report to 26 TLDs, and in 2016 it was one report to 57 TLDs.
- During the three-year period from 2017 through 2019, there was only one report to ICANN.

Additional analysis was performed on the name collision reports ICANN received. A few of those reports were incomplete, so the following statements are based on analysis of the complete reports only:

- Each report specified how many days after the new TLD's delegation the problem began. As a reminder, controlled interruption was to last at least 90 days after initial delegation.
    - The range was from 1 day to 991 days (roughly 2.7 years).
    - The median value was 23 days.
    - About one-fourth of the reported problems were detected within seven days of delegation. Just over half the problems were detected within 30 days of delegation.

- Nearly 30% of the problems were not detected until after 90 days of delegation. However, 80% of the problems not detected at 90 days were still not detected after 180 days, and half of those were not yet detected even a year after delegation.

- Around 60% of the reporting organizations said their corporate network was affected. Just over 25% said individual computers were affected, and over 10% cited applications or application development.

Of all the reports to ICANN, only one led to action by a registry. In that case, a large organization had reported disruption of its services on the first day after new TLD delegation. The registry operator for the new TLD voluntarily chose to temporarily stop controlled interruption for that TLD. After the affected organization updated its systems to correct the problem, the registry operator was able to resume controlled interruption for the TLD.

4.2.2 Reports to Others

For the purposes of this study, a member of the ICANN NCAP Discussion Group created and provided a list of URLs for 50 publicly reported instances of name collisions identified through controlled interruption. These accounts were found in technical support forums, mailing lists, and other places where people encountering signs of controlled interruption like 127.0.53.53 asked for help. Each instance was reviewed, and the 33 instances where the nature of the problem could be determined based on the available information were further evaluated. Note that these reports are strictly anecdotal, so while some insights can be gleaned from analyzing them, the accuracy of each report cannot readily be verified, and thus drawing specific conclusions from individual reports is unwise.

Most of the 33 evaluated instances involved duplicate name collisions, where there was internal-only use of a domain that was subsequently publicly delegated. In nearly half of those cases, dev was the TLD in question, with the prod, bar, and box TLDs each also cited in multiple cases, and several other TLDs cited once. The rest of the 33 instances involved shortened name collisions.

Several of the evaluated instances affected an individual, typically someone using a domain on a personally owned computer or home network until public delegation of that domain caused the home configuration to stop working.

None of the evaluated instances mentioned major harm to individuals or organizations—the reactions were curiosity, annoyance, or minor disruption.

To look for additional publicly known instances of name collisions besides those on the list of 50, searches were conducted using terms such as “127.0.53.53”, “name collision”, “controlled interruption”, and “outage” to identify news articles, blog postings, forum discussions, and other accounts of the technical impact of name collisions and controlled interruption. No significant new information was found other than additional instances of name collisions found through controlled interruption, similar to those on the list of 50. The total number of all such postings since 2014 appeared to be in the hundreds, and the volume of new postings of name collision-related problems has dropped sharply over the past few years, with only a handful of such postings made during all of 2019. None of the reviewed postings mentioned major harm to individuals or organizations.

5 Datasets for Name Collision Studies

The study RFP [2] specified that this report must include the following:

- Study task 3: "Identify datasets used in past studies and determine if those datasets are still available and any constraints there may be regarding access."
- Study task 4: "Identify gaps in the datasets used by previous studies, resulting in a list of additional datasets or data providers that would be necessary to successfully complete Studies 2 and 3."
- Study task 5: "Assess the potential availability of these additional datasets."

Section 5.1 discusses the first item (datasets from past studies), and Section 5.2 covers the other two (identify gaps in datasets from past studies, list what is needed to fill those gaps, and assess the availability of items on the list).

5.1 Datasets Used in Past Studies

Most past studies of name collisions have used data from DNS-OARC Day in the Life of the Internet (DITL) [37]. Authors of work cited in this report that used data from DITL include Demand Media [52], DigiCert [43], Donuts [42], ICANN [38], Interisle [36] and [51], JAS Global Advisors [95] and [106], JAS Global Advisors and simMachines [44], Verisign [74], Verisign Labs [58], and Verisign Labs and University of Michigan [77] and [83].

According to [37], DITL data is currently available for every year from 2006 through 2018, and "access to this data requires a current OARC paying membership, or in lieu of payment…a mutually beneficial form of in-kind membership." The lowest-priced paid membership as of this writing is $1100 per year, which allows two people to participate. [126] Note that [37] states that OARC members have access to OARC analysis machines, and that OARC requires "that the data may not be copied off OARC servers to any other host or network beyond OARC's access and control."

Another dataset mentioned by a previous name collision report is the Operational Research Data from Internet Namespace Logs (ORDINAL) dataset [79]. ORDINAL is housed by the Information Marketplace for Policy and Analysis of Cyber-Risk & Trust (IMPACT). It appears that access to IMPACT is free, and researchers in the United States and several other countries approved by the US Department of Homeland Security (DHS) are eligible for IMPACT access. [127] ORDINAL data is being provided on an ongoing basis by JAS Global Advisors, and ORDINAL "contains robust DNS protocol layer data, select application layer data, standard activity logs, received select transmissions, and packet captures of associated activity originally intended to study the impact of DNS namespace collisions. The dataset is generated via Internet activity to sensor nodes which are linked to high activity Domain Names." [128]

In addition to the DITL and ORDINAL datasets, there are also ICANN name collision reports with pertinent information on actual name collisions, their characteristics, and their outcomes. Sanitized summaries of all name collision reports received to date were provided for the purposes of this

report, and it is assumed up-to-date summaries could be provided for the authors of Studies 2 and 3 as needed and appropriate. [124]

The ICANN OCTO's *Study of the Prevalence of DNS Queries for CORP, HOME, and MAIL* [118] analyzed a representative sampling of traffic for two root servers. The data was collected between late 2015 and early 2017 in order to analyze queries for nonexistent domain names. The continued availability of these particular data sources is unknown, but it is reasonable to assume the same data sources or similar data sources could be used.

Finally, there was a dataset particular to the corp.com domain, as mentioned in 2014 in [72] and discussed in Section 3.5.1 of this report. The corp.com domain was receiving many queries that were believed to be leaking from internal .corp domains. The current availability of data for the corp.com domain and any constraints on its access are unknown and would need to be assessed early during the performance of Study 2. See [129] for more information on the status of corp.com as of this writing. A recent update indicated that new corp.com data is no longer being provided to ORDINAL. [130]

5.2 Additional Datasets Needed for Studies 2 and 3

The plans for Studies 2 and 3 are outlined in Section 3.3 of the *SSAC Proposal for the Name Collision Analysis Project* from February 2019. [1] Study 2, "Name Collision Root Cause and Impact Analysis, and Data Repository" would involve gathering datasets in a data repository and conducting an analysis of that data to understand the root cause of most name collisions. Study 3, "Analysis of Mitigation Options," would be analysis and testing of mitigation strategies, with specific guidance to be produced on the potential delegation of the corp, home, and mail TLDs, as well as other TLDs likely to cause name collisions.

In April 2020, the NCAP Discussion Group published the NCAP Gap Analysis Brief to help inform the design of Studies 2 and 3. [131] It stated the following regarding datasets:

> "Since the new gTLD program, various new data sets have become available that may provide additional telemetry to better understand and assess name collision risks. The new gTLD name collision risk assessment was conducted against a few years of DITL DNS traffic data. Unfortunately, the DITL data set has several limitations, as it only provides a few days per year of authoritative root server DNS traffic, is contributed by root server operators on a voluntary basis, may be anonymized due to privacy concerns, and […] may require a different method of analysis. Since the last TLD round, the collection of DITL data has continued and may provide better longitudinal measurements pre/post the new TLD delegations. Other entities have also started to retain high fidelity root DNS traffic that may provide better insights. The emergence of popular open recursive resolvers has also transpired and dramatically shaped the DNS ecosystem since the new gTLD delegations. These recursive services may provide a richer and more complete understanding of name collisions if they can be utilized for analysis. Other potential data repositories of interest would also include the ORDINAL DNS data as well as Certificate Transparency records, neither of which existed during the previous assessment."

No gaps or other issues have been identified in accessing the datasets that would be needed for Studies 2 and 3. Information on previous and recent leakage of corp, home, and mail should already be captured in the DITL and ORDINAL datasets. A current dataset for corp.com could be valuable for comparing current leakage of the corp domain to 2014-era leakage. Similar datasets for the home and mail counterparts to corp.com (e.g., home.com and mail.com) might also be valuable, although much of the same information might be available through the DITL and ORDINAL datasets. A current dataset similar to what was collected for the 2017 ICANN OCTO study would provide information on current corp, home, and mail leakage. Additional data from sources like the recursive services mentioned by the NCAP DG could also be beneficial.

6 Recommendation for Studies 2 and 3

This section addresses the third goal of Study 1, as stated in the RFP [2]: "a recommendation if Studies 2 and 3 should be performed based on the results of the survey of prior work and the availability of data sets." As Section 5.2 already mentioned, Study 2 would involve gathering datasets in a data repository and conducting an analysis of that data to understand the root cause of most name collisions. Study 3 would be analysis and testing of mitigation strategies, with specific guidance to be produced on the potential delegation of the corp, home, and mail TLDs, as well as other TLDs likely to cause name collisions. [1]

Major findings from the survey of prior work and datasets are as follows:

1. Name collisions have been a known problem for decades, possibly as early as the late 1980s. Reports, papers, and other work regarding name collisions were sparse and sporadic until 2012, at which point many organizations and individuals began publishing extensively on the topic. Workshops were held in 2013 and 2014. Since ICANN approved the Name Collision Occurrence Management Framework in 2014 [98], which instituted controlled interruption as the mitigation strategy for new gTLDs and ccTLDs, the volume of work on name collisions by academic institutions, the security industry, IT product and service vendors, and others has greatly decreased. The only known work on name collisions during the past few years has been from ICANN by the NCAP DG and the New gTLD SubPro Working Group. Since mid-2017, there has not been any published research into the causes of name collisions or name collision mitigation strategies. [Section 3]
2. Since controlled interruption was instituted, there have been few instances of name collision problems being reported to ICANN or reported publicly through technical support forums, mailing lists, and other means. Most problems occurred during 2014, 2015, or 2016, with only a single problem reported to ICANN during the three-year period from 2017 through 2019, as well as a sharp dropoff in public reports during the same period. Only one of the reports to ICANN necessitated action by a registry, and none of the public reports surveyed mentioned major harm to individuals or organizations. [Sections 4.1 and 4.2]
3. Prior work, such as [76], and name collision reports have indicated there are several types of root causes of name collisions, perhaps a dozen or more. These root causes have typically been found by individuals researching a particular leaked TLD to find its origin, not by examining datasets. There is unlikely to be any dataset that would contain root causes; identifying root causes is generally going to require research of each TLD involved in name collisions on a case-by-case basis. [Sections 3 and 4.2]
4. No gaps or other issues have been identified in accessing the datasets that would be needed for Studies 2 and 3. [Section 5]

Recent discussions among NCAP DG members (see the threads beginning with [132] and [133]) indicate differences of opinion as to whether controlled interruption has been "successful." It does not appear that criteria for success are formally defined, and until such criteria are defined, disagreements are likely to continue.

That being said, however, there have been minimal name collision problems reported since controlled interruption was instituted, given the number of new TLDs it has been used for in the past six years. Research conducted for this report included extensive searches for evidence, and NCAP DG members were repeatedly asked to provide information on any evidence they were aware of. The counterargument to this has been the old saying, "Absence of evidence is not evidence of absence." Although that saying has merit, over time the continued absence of evidence that controlled interruption has not been successful makes it less likely to be true. The lack of interest in alternatives to controlled interruption outside a few groups within ICANN further supports the likelihood that controlled interruption has been successful.

Given these findings, the recommendation is that Studies 2 and 3 should not be performed as currently designed. Regarding Study 2, analyzing datasets is unlikely to identify significant root causes for name collisions that have not already been identified. New causes for name collisions are far more likely to be found by investigating TLD candidates for potential delegation on a case by case basis. Regarding Study 3, controlled interruption has already proven an effective mitigation strategy, and there does not appear to be a need to identify, analyze, and test alternatives for the vast majority of TLD candidates.

All of that being said, this does not mean further study should not be conducted into name collision risks and the feasibility of potentially delegating additional domains that are likely to cause name collisions. Most notably, the Study 3 question of how to mitigate name collisions for potential delegation of the corp, home, and mail TLDs is still unresolved. However, the proposals for Studies 2 and 3, which were developed years ago, do not seem to be effective ways of achieving the intended goals.

7 Bibliography

[1] ICANN SSAC, "SSAC Proposal for the Name Collision Analysis Project," February 2019. [Online]. Available: https://community.icann.org/download/attachments/79437474/NCAP%20Proposal%20for%20Board%20%28revised%20by%20OCTO%20based%20on%20V2.5BTClean%29%20REDACTED.pdf?api=v2.

[2] ICANN, "Project Overview for the Name Collision Analysis Project (NCAP) Study 1: Request for Proposal," 9 July 2019. [Online]. Available: https://www.icann.org/en/system/files/files/rfp-ncap-study-1-09jul19-en.pdf.

[3] ICANN, "ICANN Acronyms and Terms," [Online]. Available: https://www.icann.org/icann-acronyms-and-terms/icann-acronyms-and-terms/en/nav/A.

[4] J. Postel, "RFC 1591, Domain Name System Structure and Delegation," March 1994. [Online]. Available: https://www.ietf.org/rfc/rfc1591.txt.

[5] ICANN, "New gTLD Program," October 2009. [Online]. Available: https://archive.icann.org/en/topics/new-gtlds/factsheet-new-gtld-program-oct09-en.pdf.

[6] ICANN, "About the Program," [Online]. Available: https://newgtlds.icann.org/en/about/program.

[7] ICANN, "New Generic Top-Level Domains," [Online]. Available: https://newgtlds.icann.org/en/.

[8] ICANN, "Delegated Strings," [Online]. Available: https://newgtlds.icann.org/en/program-status/delegated-strings.

[9] International Organization for Standardization (ISO), "ISO 3166 Country Codes," [Online]. Available: https://www.iso.org/iso-3166-country-codes.html.

[10] ICANN , "IDN ccTLD Fast Track Process," [Online]. Available: https://www.icann.org/resources/pages/fast-track-2012-02-25-en.

[11] ICANN, "Resources for Country Code Managers," 25 February 2012. [Online]. Available: https://www.icann.org/resources/pages/cctlds-21-2012-02-25-en.

[12] ICANN, "IDN ccTLD Fast Track String Evaluation Completion," 19 February 2014. [Online]. Available: https://www.icann.org/resources/pages/string-evaluation-completion-2014-02-19-en.

[13] ICANN, "Frequently Asked Questions: IDN ccTLDs by Country," 25 February 2012. [Online]. Available: https://www.icann.org/resources/pages/faqs-5b-2012-02-25-en.

[14] ICANN, "Name Collision Resources & Information," [Online]. Available: https://www.icann.org/resources/pages/name-collision-2013-12-06-en.

[15] Internet Architecture Board, "IAB Commentary: Architectural Concerns on the Use of DNS Wildcards, September 2003," 19 September 2003. [Online]. Available: https://www.iab.org/documents/correspondence-reports-%20documents/docs2003/2003-09-20-dns-wildcards/.

[16] ICANN SSAC, "Redirection in the com and net Domains," 9 July 2004. [Online]. Available: http://www.icann.org/committees/security/ssac-report-09jul04.pdf.

[17] ICANN SSAC, "SAC 015: Why Top Level Domains Should Not Use Wildcard Resource Records," 10 November 2006. [Online]. Available: https://www.icann.org/groups/ssac/documents/sac-015-en.

[18] ICANN SSAC, "SAC 032: Preliminary Report on DNS Response Modification," June 2008. [Online]. Available: https://www.icann.org/en/system/files/files/sac-032-en.pdf.

[19] E. Gavron, "RFC 1535, A Security Problem and Proposed Correction With Widely Deployed DNS Software," October 1993. [Online]. Available: https://tools.ietf.org/rfc/rfc1535.txt.

[20] ICANN Registry Services Technical Evaluation Panel (RSTEP), "Report on Internet Security and Stability Implications of the Tralliance Corporation search.travel Wildcard Proposal," 2 November 2006. [Online]. Available: https://www.icann.org/en/system/files/files/tralliance-report-09nov06-en.pdf.

[21] ICANN SSAC, "SAC 041: Recommendation to prohibit use of redirection and synthesized responses by new TLDs," 10 June 2009. [Online]. Available: https://www.icann.org/en/system/files/files/sac-041-en.pdf.

[22] ICANN, "ICANN Archives, Verisign's Wildcard Service Deployment," [Online]. Available: https://archive.icann.org/en/topics/wildcard-history.html.

[23] ICANN SSAC, "SAC 010: Renewal Considerations for Domain Name Registrants," June 2006. [Online]. Available: https://www.icann.org/en/system/files/files/renewal-advisory-29jun06-en.pdf.

[24] ICANN SSAC, "SAC 011: Problems caused by the non-renewal of a domain name associated with a DNS Name Server," June 2006. [Online]. Available: https://www.icann.org/en/system/files/files/renewal-nameserver-07jul06-en.pdf.

[25] ICANN, "New gTLD Application Guidebook," 4 June 2014. [Online]. Available: https://newgtlds.icann.org/en/applicants/agb/guidebook-full-04jun12-en.pdf.

[26] G. Kirikos, "Most Popular Invalid TLDs Should Be Reserved," 18 June 2009. [Online]. Available: http://www.circleid.com/posts/20090618_most_popular_invalid_tlds_should_be_reserved/.

[27] ICANN SSAC, "SAC 045: Invalid Top Level Domain Queries at the Root Level of the Domain System," 15 November 2010. [Online]. Available: https://www.icann.org/en/system/files/files/sac-045-en.pdf.

[28] D. Eastlake and A. Panitz, "RFC 2606, Reserved Top Level DNS Names," June 1999. [Online]. Available: https://tools.ietf.org/html/rfc2606.

[29] ICANN SSAC, "SAC 057: SSAC Advisory on Internal Name Certificates," 15 March 2013. [Online]. Available: https://www.icann.org/en/system/files/files/sac-057-en.pdf.

[30] Verisign Labs, "Verisign Labs Technical Report #1130007 version 2.1: New gTLD Security and Stability Considerations," March 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdfY5loOoWatX.pdf.

[31] Verisign Labs, "Verisign Labs Technical Report #1130007 version 2.2: New gTLD Security and Stability Considerations," 28 March 2013. [Online]. Available: https://www.verisign.com/assets/gtld-ssr-v2.1-final.pdf.

[32] B. Hill and B. Smith, "Re: Proposed delegation of invalid names from SAC 045 and RFC 6762," 15 March 2013. [Online]. Available: https://www.icann.org/en/system/files/correspondence/hill-smith-to-chehade-crocker-15mar13-en.pdf.

[33] S. Cheshire and M. Krochmal, "RFC 6762, Multicast DNS," February 2013. [Online]. Available: https://tools.ietf.org/html/rfc6762.

[34] O. Kolkman, A. Sullivan and W. Kumari, "Internet-Draft draft-kolkman-cautious-delegation-00, A Procedure for Cautious Delegation of a DNS Name," 2 May 2013. [Online]. Available: https://www.ietf.org/archive/id/draft-kolkman-cautious-delegation-00.txt.

[35] O. Kolkman, A. Sullivan and W. Kumari, "Internet-Draft draft-kolkman-cautious-delegation-02, A Procedure for Cautious Delegation of a DNS Name," 1 August 2013. [Online]. Available: https://www.ietf.org/archive/id/draft-kolkman-cautious-delegation-02.txt.

[36] Interisle Consulting Group, "Name Collision in the DNS: A study of the likelihood and potential consequences of collision between new public gTLD labels and existing private

uses of the same strings, version 1.5," 2 August 2013. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-02aug13-en.pdf.

[37] DNS-OARC, "Day In The Life of the Internet (DITL)," [Online]. Available: https://www.dns-oarc.net/oarc/data/ditl.

[38] ICANN, "New gTLD Collision Risk Mitigation: Proposals to mitigate the collision risks between new gTLDs and existing private uses of the same strings," 5 August 2013. [Online]. Available: https://www.icann.org/en/system/files/files/new-gtld-collision-mitigation-05aug13-en.pdf.

[39] ICANN, "Addressing the Consequences of Name Collisions," 5 August 2013. [Online]. Available: https://www.icann.org/news/announcement-3-2013-08-05-en.

[40] ICANN, "[comments-name-collision-05aug13] Chronological Index," [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/index.html.

[41] ICANN, "Report of Public Comments: Proposal to Mitigate Name Collision Risks," 5 August 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdf3wmJxwMJoR.pdf.

[42] Donuts, "Donuts' Comments Regarding Proposal to Mitigate Name Collision Risks," 5 August 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdfuanEVzPqbD.pdf.

[43] DigiCert, Inc., "Letter from DigiCert to the ICANN Board," 27 August 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdfUNz0liz2VL.pdf.

[44] J. Schmidt, K. White, D. Conrad and A. Muller-Molina, "Namespace Expansion," 17 September 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdf0r8YJwS4iG.pdf.

[45] ICANN New gTLD Applicant Group (NTAG), "NTAG Comments on ICANN Name Collision Report," 14 August 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/msg00001.html.

[46] E. Osterweil, "Illustrating the Need to Undertake Qualitative Impact Assessments for Applied-For Strings: .WEBSITE, .COFFEE, and .CLUB," 17 September 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdf5H5Sqf0igA.pdf.

[47] Verisign Labs, "Verisign Labs Technical Report #1130008 Version 1.1: New gTLD Security, Stability Resiliency Update: Exploratory Consumer Impact Analysis," 22

August 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdfu6z5kKHEV5.pdf.

[48] Neustar, "A Methodology for Assessing Collision Risk and New gTLDs," 17 September 2013. [Online]. Available: https://www.home.neustar/resources/whitepapers/new-tlds-dns-collision.

[49] ICANN, "New gTLD Collision Occurrence Management: Proposal to manage the collision occurrences between new gTLDs and existing private uses of the same strings," 4 October 2013. [Online]. Available: https://www.icann.org/en/system/files/files/resolutions-new-gtld-annex-1-07oct13-en.pdf.

[50] ICANN New gTLD Program Committee, "Approved Resolutions | Meeting of the New gTLD Program Committee," 7 October 2013. [Online]. Available: https://www.icann.org/resources/board-material/resolutions-new-gtld-2013-10-07-en.

[51] J. Reid, "DITL Crunching for gTLD Name Collision Study," 5 October 2013. [Online]. Available: https://indico.dns-oarc.net/event/1/contributions/46/attachments/38/166/Reid-Crunching.pdf.

[52] R. Hooper, "Abusing Resources to Process 7TB of PCAP Data...Or how not to fork-bomb yourself," 5 October 2013. [Online]. Available: https://indico.dns-oarc.net/event/1/contributions/49/attachments/41/169/DNS-OARC-Abusing-Resources.pdf.

[53] A. Simpson, D. McPherson, E. Osterweil, M. Thomas and D. Wessels, "Regional Affinity for Applied for gTLD Strings," 5 October 2013. [Online]. Available: https://indico.dns-oarc.net/event/1/contributions/37/attachments/44/174/gTLD_Regional_Affinity.pdf.

[54] A. Sullivan, "Using Test Delegations from the Root Prior to Full Allocation and Delegation," 5 October 2013. [Online]. Available: https://indico.dns-oarc.net/event/1/contributions/42/attachments/49/180/Sullivan-Test_delegations.pdf.

[55] O. Kolkman, A. Sullivan and W. Kumari, "Internet-Draft draft-kolkman-root-test-delegation-00, Using Test Delegations from the Root Prior to Full Allocation and Delegation," 20 September 2013. [Online]. Available: https://tools.ietf.org/html/draft-kolkman-root-test-delegation-00.

[56] ICANN SSAC, "SAC 062: SSAC Advisory Concerning the Mitigation of Name Collision Risk," 7 November 2013. [Online]. Available: https://www.icann.org/en/system/files/files/sac-062-en.pdf.

[57] Verisign Labs, "Preliminary Analysis of SLD Blocking Effectiveness," 5 November 2013. [Online]. Available: https://www.icann.org/en/system/files/correspondence/kaliski-to-atallah-crain-05nov13-en.pdf.

[58] Verisign Labs, "Continued Analysis of SLD Blocking Effectiveness," 15 November 2013. [Online]. Available: https://www.icann.org/en/system/files/correspondence/kaliski-to-atallah-crain-15nov13-en.pdf.

[59] ICANN, "Reports for Alternate Path to Delegation Published," 17 November 2013. [Online]. Available: https://newgtlds.icann.org/en/announcements-and-media/announcement-2-17nov13-en.

[60] "Workshop and Prize on Root Causes and Mitigation of Name Collisions (WPNC)," [Online]. Available: http://namecollisions.net/program/index.html.

[61] M. Thomas, A. Mankin and L. Zhang, "RFC 8023, Report from the Workshop and Prize on Root Causes and Mitigation of Name Collisions," November 2016. [Online]. Available: https://www.rfc-editor.org/in-notes/rfc8023.html.

[62] M. Thomas, Y. Labrou and A. Simpson, "The Effectiveness of Block Lists in Preventing Collisions," 9 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc2014_paper_effectiveness_block_lists.pdf.

[63] P. Hoffman, "Name Collision Mitigation for Enterprise Networks," 10 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc14_slides_hoffman_name_collision_mitigation.pdf.

[64] J. Reid, "Analysing the use of the RA and RD Bits in Queries to Root Servers," 9 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc2014_paper_reid.pdf.

[65] G. Huston, "New gTLD Concerns: Dotless Names and Name Collisions," 12 November 2013. [Online]. Available: https://labs.ripe.net/Members/gih/dotless-names.

[66] D. Piscitello, "Managing Name Collision Occurrences," 6 December 2013. [Online]. Available: https://www.icann.org/news/blog/managing-name-collision-occurrences.

[67] ICANN, "Guide to Name Collision Identification and Mitigation for IT Professionals, Version 1.0," 5 December 2013. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-mitigation-05dec13-en.pdf.

[68] ICANN SSAC, "SAC 064: SSAC Advisory on DNS 'Search List' Processing," 13 February 2014. [Online]. Available: https://www.icann.org/en/system/files/files/sac-064-en.pdf.

[69] R. Braden, "RFC 1123, Requirements for Internet Hosts -- Application and Support," October 1989. [Online]. Available: https://tools.ietf.org/html/rfc1123.

[70] A. Kumar, J. Postel, C. Neuman, P. Danzig and S. Miller, "RFC 1536, Common DNS Implementation Errors and Suggested Fixes," October 1993. [Online]. Available: https://tools.ietf.org/html/rfc1536.

[71] W. Kumari, "ALT Special Use TLD," 9 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc14_slides_internet_engineeringpanel_kumari.pdf.

[72] C. Strutt, "Looking at corp.com as a Proxy for .corp," 9 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc14_slides_strutt_looking_at_corpcom.pdf.

[73] C. Deccio and D. Wessels, "What's in a Name (Collision): Modeling and Quantifying Collision Potential," 10 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc2014_paper_deccio.pdf.

[74] A. Simpson, "Detecting Search Lists in Authoritative DNS," 10 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc2014_paper_simpson.pdf.

[75] M. Thomas and A. Simpson, "Analysis Techniques for Determining Cause and Ownership of DNS Queries," 9 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc14_paper_simpson_thomas.pdf.

[76] M. Thomas and A. Mohaisen, "Measuring the Leakage of Onion at the Root: A measurement of Tor's .onion pseudo-top-level domain in the global domain name system," November 2014. [Online]. Available: https://www.verisign.com/assets/labs/Measuring-the-Leakage-of-Onion-at-the-Root.pdf.

[77] A. Mohaisen and K. Ren, "Leakage of .onion at the DNS Root: Measurements, Causes, and Countermeasures," October 2017. [Online]. Available: http://seal.cs.ucf.edu/doc/17-tnet.pdf.

[78] J. Appelbaum and A. Muffett, "RFC 7686, The ".onion" Special-Use Domain Name," October 2015. [Online]. Available: https://tools.ietf.org/html/rfc7686.

[79] "Introducing: The ORDINAL Dataset," 13 May 2017. [Online]. Available: https://www.icann.org/en/system/files/files/presentation-ordinal-datasets-colliding-domains-13may17-en.pdf.

[80] Verisign, "Enterprise Remediation for WPAD Name Collision Vulnerability," 23 May 2016. [Online]. Available: https://www.verisign.com/assets/Enterprise_Remediation_for_WPAD_Name_Collision_Vulnerability.pdf.

[81] Q. Chen, E. Osterweil, M. Thomas and Z. Mao, "MitM Attack by Name Collision: Cause Analysis and Vulnerability Assessment in the New gTLD Era," May 2016. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7546529.

[82] Cybersecurity and Infrastructure Security Agency (CISA), "Alert (TA16-144A), WPAD Name Collision Vulnerability," 23 May 2016. [Online]. Available: https://www.us-cert.gov/ncas/alerts/TA16-144A.

[83] Q. Chen, M. Thomas, E. Osterweil, Y. Cao, J. You and Z. Mao, "Client-Side Name Collision Vulnerability in the New gTLD Era: A Systematic Study," November 2017. [Online]. Available: https://www.ics.uci.edu/~alfchen/alfred_ccs17.pdf.

[84] E. Osterweil, D. McPherson, M. Thomas and Q. Chen, "Detecting and Remediating Highly Vulnerable Domain Names Using Passive DNS Measurements". United States of America Patent US20170279846A1, 24 March 2017.

[85] JAS Global Advisors, "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," 24 February 2014. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-mitigation-26feb14-en.pdf.

[86] J. Schmidt, "Mitigating the Risk of DNS Namespace Collisions," 10 March 2014. [Online]. Available: http://namecollisions.net/downloads/wpnc14_slides_jas_framework_session.pdf.

[87] J. Schmidt, "Name Collision Mitigation Update," 24 March 2014. [Online]. Available: https://archive.icann.org/meetings/singapore2014/en/schedule/mon-name-collision/presentation-name-collision-24mar14-en.pdf.

[88] "Transcript of ICANN Name Collision Mitigation meeting in Singapore," 24 March 2014. [Online]. Available: https://archive.icann.org/meetings/singapore2014/en/schedule/mon-name-collision/transcript-name-collision-24mar14-en.pdf.

[89] ICANN, "[comments-name-collision-26feb14] Chronological Index," [Online]. Available: https://forum.icann.org/lists/comments-name-collision-26feb14/index.html.

[90] ICANN, "Report of Public Comments: Mitigating the Risk of DNS Namespace Collisions," 10 June 2014. [Online]. Available: https://www.icann.org/en/system/files/files/report-comments-name-collision-10jun14-en.pdf.

[91] Verisign, "Preliminary Comments on 'Mitigating the Risk of DNS Namespace Collisions' Phase One Report," 24 February 2014. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-26feb14/pdfNPWfDHk1pu.pdf.

[92] B. Kaliski, "Name Collisions in the Domain Name System," 17 April 2014. [Online]. Available: http://www.verisign.com/assets/Verisign-Kaliski-Collisions-US-Telecom-04162014.pptx.

[93] Verisign, "Additional Comments on “Mitigating the Risk of DNS Namespace Collisions” Phase One Report," 21 April 2014. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-26feb14/pdfTWUAZM3gBN.pdf.

[94] B. Kaliski, "The Real Uneven Playing Field of Name Collisions," 16 May 2014. [Online]. Available: https://blog.verisign.com/security/the-real-uneven-playing-field-of-name-collisions/.

[95] JAS Global Advisors, "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. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-mitigation-study-06jun14-en.pdf.

[96] Verisign, "Re: ICANN's Proposal to Mitigate Name Collision Risks - .CBA Case Study," 15 September 2013. [Online]. Available: https://forum.icann.org/lists/comments-name-collision-05aug13/pdfJUJTT9vS7d.pdf.

[97] ICANN SSAC, "SAC 066: SSAC Comment Concerning JAS Phase One Report on Mitigating the Risk of DNS Namespace Collisions," 6 June 2014. [Online]. Available: https://www.icann.org/en/system/files/files/sac-066-en.pdf.

[98] ICANN, "Name Collision Occurrence Management Framework," 30 July 2014. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.pdf.

[99] ICANN, "Guide to Name Collision Identification and Mitigation for IT Professionals, Version 1.1," 1 August 2014. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-mitigation-01aug14-en.pdf.

[100] ICANN Global Domains Division, "Name Collision Occurrence Assessment," 4 August 2014. [Online]. Available: https://newgtlds.icann.org/sites/default/files/agreements/name-collision-assessment-04aug14-en.htm.

[101] ICANN Global Domains Division, "Addendum to Name Collision Occurrence Assessment," 14 November 2014. [Online]. Available: https://newgtlds.icann.org/sites/default/files/agreements/name-collision-assessment-addendum-14nov14-en.htm.

[102] ICANN, "Frequently Asked Questions: Name Collision Occurrence Management Framework for Registries," [Online]. Available: https://www.icann.org/resources/pages/name-collision-ro-faqs-2014-08-01-en.

[103] ICANN, "Frequently Asked Questions: Name Collision for IT Professionals," [Online]. Available: https://www.icann.org/resources/pages/name-collision-it-pros-faqs-2014-08-01-en.

[104] ICANN, "Briefing on Name Collision Risk for New TLDs," 18 February 2014. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-risk-18feb14-en.pdf.

[105] ICANN, "Name Collision Occurrence Mitigation for New ccTLDs," 2 October 2014. [Online]. Available: https://www.icann.org/resources/pages/cctld-mitigation-2014-10-02-en.

[106] JAS Global Advisors, "Mitigating the Risk of DNS Namespace Collisions: A Study on Namespace Collisions in the Global Internet DNS Namespace and a Framework for Risk Mitigation, Final Report," 28 October 2015. [Online]. Available: https://www.icann.org/en/system/files/files/name-collision-mitigation-final-28oct15-en.pdf.

[107] ICANN GNSO New gTLD Subsequent Procedures Working Group, "Initial Report on the new gTLD Subsequent Procedures Policy Development Process (Overarching Issues & Work Tracks 1-4)," 3 July 2018. [Online]. Available: https://gnso.icann.org/sites/default/files/file/field-file-attach/subsequent-procedures-initial-overarching-issues-work-tracks-1-4-03jul18-en.pdf.

[108] ICANN SSAC, "SAC 094: SSAC Response to the New gTLD Subsequent Procedures Policy Development Process (PDP) Working Group Community Comment 2," 22 May 2017. [Online]. Available: https://www.icann.org/en/system/files/files/sac-094-en.pdf.

[109] ICANN SSAC, "SAC 103: SSAC Response to the new gTLD Subsequent Procedures Policy Development Process Working Group Initial Report," 3 October 2018. [Online]. Available: https://www.icann.org/en/system/files/files/sac-103-en.pdf.

[110] S. Cheshire and M. Krochmal, "RFC 6761, Special-Use Domain Names," February 2013. [Online]. Available: https://tools.ietf.org/html/rfc6761.

[111] .Home Registry Inc. et al, "Letter to Members of the ICANN Board," 24 August 2016. [Online]. Available: https://www.icann.org/en/system/files/correspondence/home-registry-inc-et-al-to-icann-board-24aug16-en.pdf.

[112] T. Lemon, R. Droms and W. Kumari, "RFC 8244, Special-Use Domain Names Problem Statement," October 2017. [Online]. Available: https://tools.ietf.org/html/rfc8244.

[113] L. Chapin and M. McFadden, "Internet-Draft draft-chapin-additional-reserved-tlds-02, Additional Reserved Top Level Domains," 2 March 2015. [Online]. Available: https://www.ietf.org/archive/id/draft-chapin-additional-reserved-tlds-02.txt.

[114] M. Stenberg, S. Barth and P. Pfister, "RFC 7788, Home Networking Control Protocol," April 2016. [Online]. Available: https://tools.ietf.org/html/rfc7788.

[115] ICANN, "Approved Board Resolutions for 02 Nov 2017, section 2a, Consideration of .CORP, .HOME, and .MAIL and other Collision Strings, Rationale for Resolutions 2017.11.02.29-2017.11.02.31," 2 November 2017. [Online]. Available: https://www.icann.org/resources/board-material/resolutions-2017-11-02-en#2.a.

[116] W. Kumari, "Internet-Draft draft-wkumari-dnsop-internal-00, The .internal TLD," 2 July 2017. [Online]. Available: https://www.ietf.org/archive/id/draft-wkumari-dnsop-internal-00.txt.

[117] W. Kumari and A. Sullivan, "Internet-Draft draft-ietf-dnsop-alt-tld-12, The ALT Special Use Top Level Domain," 23 August 2019. [Online]. Available: https://datatracker.ietf.org/doc/draft-ietf-dnsop-alt-tld/.

[118] R. Arends, "OCTO-007, Study of the Prevalence of DNS Queries for CORP, HOME, and MAIL," 14 April 2020. [Online]. Available: https://www.icann.org/en/system/files/files/octo-007-en.pdf.

[119] ICANN, "Name Collision Occurrence Mitigation for New ccTLDs," 2 October 2014. [Online]. Available: https://www.icann.org/resources/pages/cctld-mitigation-2014-10-02-en.

[120] ICANN, 27 September 2014. [Online]. Available: https://twitter.com/icann/status/515976677734629377.

[121] S. Vaughan-Nichols, "ICANN offers fix for domain name collisions," 15 August 2014. [Online]. Available: https://www.zdnet.com/article/icann-offers-fix-for-domain-name-collisions/.

[122] djchuang, "What does 127.0.53.53 mean? It's a system alert notification thing.," 10 September 2014. [Online]. Available: https://djchuang.com/127-0-53-53-mean-system-alert-notification-thing/.

[123] ICANN, "Report a Name Collision," [Online]. Available: https://forms.icann.org/en/help/name-collision/report-problems.

[124] ICANN, *Summary of sanitized name collision reports,* January, 2020.

[125] ICANN, "ICANN Announces Successful String Evaluation for the European Commission and Laos IDN ccTLDs," 5 June 2019. [Online]. Available: https://www.icann.org/news/announcement-2019-06-05-en.

[126] DNS-OARC, "DNS-OARC Participation Agreement," January 2020. [Online]. Available: https://www.dns-oarc.net/files/agreements/oarc-participation.pdf.

[127] IMPACT, "Join Impact," [Online]. Available: https://www.impactcybertrust.org/joinus.

[128] IMPACT, "Dataset Details for DS-0794," [Online]. Available: https://impactcybertrust.org/dataset_view?idDataset=794.

[129] M. Larson, "[NCAP-Discuss] [Ext] Revised draft of NCAP Study 1 report," 5 February 2020. [Online]. Available: https://mm.icann.org/pipermail/ncap-discuss/2020-February/000202.html.

[130] J. Schmidt, "[NCAP-Discuss] Latest draft of Study 1," 12 February 2020. [Online]. Available: https://mm.icann.org/pipermail/ncap-discuss/2020-February/000233.html.

[131] NCAP Discussion Group, "NCAP Gap Analysis Brief," 21 April 2020. [Online]. Available: https://mm.icann.org/pipermail/ncap-discuss/2020-April/000269.html.

[132] M. Larson, "[NCAP-Discuss] Draft final Study 1 report," 24 April 2020. [Online]. Available: https://mm.icann.org/pipermail/ncap-discuss/2020-April/000275.html.

[133] J. Schmidt, "[NCAP-Discuss] Additional comments on the comments to the Scarfone Draft," 6 May 2020. [Online]. Available: https://mm.icann.org/pipermail/ncap-discuss/2020-May/000348.html.

8 Acronyms

Acronym	Definition
APNIC	Asia-Pacific Network Information Centre
CA	Certificate Authority
ccNSO	Country Code Names Supporting Organization
ccTLD	Country Code Top-Level Domain
CI	Controlled Interruption
CISA	Cybersecurity and Infrastructure Security Agency
CNNIC	China Internet Network Information Center
DHS	Department of Homeland Security
DITL	Day in the Life of the Internet
DNS	Domain Name System
DNS-OARC	Domain Name System Operations Analysis and Research Center
FAQ	Frequently Asked Questions
FQDN	Fully Qualified Domain Name
GNSO	Generic Names Supporting Organization
gTLD	Generic Top-Level Domain
HTTP	Hypertext Transfer Protocol

Acronym	Definition
IAB	Internet Architecture Board
IANA	Internet Assigned Numbers Authority
ICANN	Internet Corporation for Assigned Names and Numbers
IDN	Internationalized Domain Name
IETF	Internet Engineering Task Force
IMPACT	Information Marketplace for Policy and Analysis of Cyber-Risk & Trust
IP	Internet Protocol
ISO	International Organization for Standardization
IT	Information Technology
MITM	Man in the Middle
NCAP	Name Collision Analysis Project
NCAP DG	Name Collision Analysis Project Discussion Group
NGPC	New gTLD Program Committee
NTAG	New gTLD Applicant Group
OCTO	Office of the Chief Technology Officer
ORDINAL	Operational Research Data from Internet Namespace Logs
PDP	Policy Development Process

Acronym	Definition
RFC	Request for Comments
RFP	Request for Proposal
RIPE NCC	Réseaux IP Européens Network Coordination Centre
RSTEP	Registry Services Technical Evaluation Panel
RZM	Root Zone Management
SLD	Second-Level Domain
SSAC	Security and Stability Advisory Committee
SSL	Secure Sockets Layer
SubPro	Subsequent Procedures (Working Group)
TLD	Top-Level Domain
TTL	Time to Live
URL	Uniform Resource Locator
WPAD	Web Proxy Auto-Discovery
WPNC	Workshop and Prize on Root Causes and Mitigation of Name Collisions
2LD	Second-Level Domain