# 1 Introduction

The Name Collision Analysis Project (NCAP) Study Two final report brings together the research and analysis of several studies and years of NCAP Discussion Group (DG) presentations and meetings that touch on the critical issues surrounding name collisions. This report takes the reader through the methodology and findings of the three research studies and the analysis of the DG’s work activities. The conclusions from those studies provide guidance for the topics regarding name collisions that the ICANN Board laid out in the ICANN Board resolution [2017.11.02.30](https://www.icann.org/resources/board-material/resolutions-2017-11-02-en#2.a.rationale).

The Internet has evolved since the last round of new gTLD delegations began in 2012. Changes include the use of new DNS transports (such as DNS-over-TLS, DNS-over-HTTPS, and DNS-over-QUIC), additional DNS privacy extensions (such as QNAME minimization and Oblivious DNS), and features that address both privacy and query volume, such as aggressive negative caching and local root instances. Additionally, the rise of global public DNS services has resulted in the increased consolidation of query traffic seen at authoritative servers, including the root servers. This changing landscape, in combination with the research done since 2012 (see “Background and Related Work”) and community feedback, resulted in the Board’s resolutions requesting the ICANN Security and Stability Advisory Committee (SSAC) provide more definitive guidance as to what should be the next steps for the applications requesting delegation of .CORP, .HOME, and .MAIL, three of the top collision strings identified in the 2012 round of gTLD delegations.

Since 2014, Controlled Interruption has been ICANN’s sole mechanism to alert users and system administrators to potential name collision issues. Several reports, from the "Mitigating the Risk of DNS Namespace Collisions Final Report," a commissioned piece by JAS Global Advisors (the “JAS Report”) to the Root Cause Analysis as commissioned through NCAP Study Two, have found Controlled Interruption to be effective in disrupting systems that might be impacted by the general availability of a new gTLD as a preemptive alert to the issues posed by that delegation. However, this very disruption has had an impact ranging from mild to severe on affected systems. These side effects have caused investigators to reevaluate the use of Controlled Interruption and to explore additional techniques for identifying and mitigating the risks of name collision. Furthermore, the DG also evaluated gaps in the availability and completeness of data used to identify name collisions. The result of these evaluations is a workflow that offers guidance to ICANN org and gTLD applicants on how to identify some of the risks of name collision before granting the delegation of a proposed gTLD to a Registry Operator and thus providing some mitigation against consequences experienced by affected systems.

The first section of this report describes the background of the NCAP and the mandate set forth by the ICANN Board in 2017. It goes on to describe the background that informed the direction of Study Two; the methodology of the study group as a whole, including the timeline of research, community outreach, study group consensus; and the terminology necessary to have a common understanding of how these terms are used in this report.

Section 2 of this report summarizes the three studies commissioned by Study Two. While additional research may provide more clarity on the root cause and challenges of identifying name collisions, the results of these studies provide information not previously understood and inform the findings and recommendations in sections 4 and 5.

Section 3 captures the years of discussion held by the DG. The expertise within that group provided necessary background and lived experiences that informed the findings in section 4 and the recommendations in section 5.

Sections 4 and 5 take the research described in Section 2 and offer detailed explanations and guidance, as well as a proposed workflow that will support a clear, deterministic process that provides clarity and transparency to new gTLD applicants and the ability to establish a sustainable and repeatable process for the ICANN Board when handling name collisions.

Section 6 presents a sample Technical Review Team report that could be used as a starting point for what the TRT will do as it conducts its analysis of the Collision Assessments the workflow proposes. Specifically, the sample report addresses the Board resolution that specifically asks for guidance with respect to evaluating the status of .CORP, .HOME, and .MAIL.

While this report is primarily intended as input to the ICANN Board, all parties interested in the future expansion of the gTLD space, from applicants to community groups, may find the material relevant to their efforts.

## 1.1 Scope of Study 2

The SSAC was tasked by the ICANN Board in resolutions 2017.11.02.29 – 2017.11.02.31 to address a set of questions related to name collision. To fulfill the Board’s request, the SSAC chartered the Name Collision Analysis Project and developed three studies to answer the Board’s questions. Study One was authorized by the ICANN Board in March 2019 and was completed in July 2020.

On 17 June 2020, the final draft of the Study One report was published for public comment. The report on this public comment recommended that Studies Two and Three should “not be performed as currently designed.” The DGagreed with this assessment and proposed four alterations to Study Two that would address these concerns:

* Removal of the original Study Two Goal of “Building a data repository.”.
* Removal of the Study Two Tasks to “Build a test system which can be used for impact analysis and to test possible mitigation strategies.”.
* Expansion of the Study Two Task “Conduct an impact analysis.” to detail the activities this Task involves.
* Having the DG undertake most of the work which was slated for paid contractors in the original version of the Study Two proposal.

The results of these modifications dramatically reduced the scope, level of effort, total costs, and resources to execute Study Two. The revised Study Two proposal therefore was limited to the following goals:

1. Understand the root cause of most name collisions
2. Understand the impact of name collisions

And the final tasks included:

* Study of ICANN Collision Reports
	+ Perform an analysis of ICANN Collision Reports to determine the underlying cause of these collisions.
	+ Produce a report on the results of the analysis.
	+ Performed by: Technical Investigator
* Impact and Data Sensitivity Analyses
	+ Research the impact of collisions with regards to Root servers and Resolvers for .CORP, .HOME and .MAIL.
	+ Research the impact of collisions with regards to Root servers and Resolvers for other selected strings.
	+ Based on the above research, evaluate the effectiveness of using multiple sources of collision data with regards to assessing the impact of collisions.
	+ Undertake a public consultation on the findings relative to .CORP, .HOME and .MAIL.
	+ Produce a report on the results of this work.
	+ Performed by: Discussion Group (DG) and Technical Investigator (in the capacity guided by the DG / Admin team).
* Response to Board Questions Relating to Study 2
	+ Respond to Board questions based on the results of the Study of ICANN Collision Reports and Impact and Data Sensitivity Analyses.
	+ Produce a report on the responses to Board questions.
	+ Performed by: DG
* Final Report
	+ Produce the final report for Study Two
	+ Undertake a public consultation on the draft version of this report

It was noted by the DG that an item was erroneously included in the “In scope but not intended to be the subject of data studies” Name Collision definition used for Study One and Study and was appropriately corrected per Appendix X.

## 1.2 Background and Related Work

With over a decade’s worth of discussion regarding the issue of DNS name collision, there is a wealth of background material to draw from on the topic. The diagram below shows a timeline view of all the events and publications described in this background.



Much of that material is captured in the NCAP Study One report, the ICANN Community Wiki, and the ICANN website. NCAP Study One provides an extensive, annotated bibliography of prior work related to name collisions, which we refer to in more detail below. The ICANNWiki has a community-sourced page [dedicated to name collisions](https://icannwiki.org/Name_Collision) that includes some history and enumeration of various events, as well as some references to notable material. ICANN maintains a resource on its website called “[Name Collision Resources & Information](https://www.icann.org/resources/pages/name-collision-2013-12-06-en)” with a broad set of materials applicable to the ICANN community, including a definition of name collisions.

A name collision occurs when an attempt to resolve a name used in a private name space (e.g. under a non-delegated Top-Level Domain, or a short, unqualified name) results in a query to the public Domain Name System (DNS). When the administrative boundaries of private and public namespaces overlap, name resolution may yield unintended or harmful results.

We highlight some of the materials from these sources that significantly influenced this report.

As the launch of the New gTLD Program was beginning, SSAC became aware of an issue with how “internal names” (which today we would compare to private use TLD strings) were being used in certificates and issued [SAC057: SSAC Advisory on Internal Name Certificates](https://www.icann.org/en/system/files/files/sac-057-en.pdf). This report included the first use of the term “name collision,” though it was not formally defined in that document. On 18 May 2013, the ICANN Board adopted Resolutions [2013.05.18.09 – 2013.05.18.11](https://www.icann.org/resources/board-material/minutes-2013-05-18-en#2.a.rationale) in response to SAC057, commissioning a study on the use of undelegated TLDs in enterprises. This initial investigation into the risks and harms of name collisions occurred after the application period ended in April 2012. From there, the ICANN community continued to evolve the work as their understanding of the depth and breadth of the issue grew; ICANN org, in turn, continuously evolved the application evaluation workflow to account for the potential of name collisions.[[1]](#footnote-1)

The first publication within the ICANN context to directly address name collisions was an ICANN-commissioned report by Interisle Consulting Group, LLC, published on 2 August 2013.[[2]](#footnote-2) Entitled “Name Collision in the DNS,” the “Interisle Report” was a study of the likelihood and potential consequences of a collision between new public gTLD labels and existing private uses of the same strings. This report established the first documented definition of a name collision:

Name collision: two names that are represented by syntactically identical strings but belong to different semantic domains are said to “collide” when one of them appears in the other’s semantic domain and is (mis)interpreted as if it belonged there.

The Interisle Report is used in this report as the baseline for comparison to all other work. The findings of the Interisle study were primarily defined by the information that can be derived either directly or through analysis from the DNS request stream at the root servers that participated in the “Day in the Life of the Internet” (DITL) exercises organized by the DNS Operations, Analysis, and Research Center (DNS-OARC) in 2012 and 2013.

Among its many important insights are the following.

* The potential for name collisions is substantial and often arises from well-established policies and practices in private network environments.
* The delegation of almost any new TLD label would carry some risk of collision. The risk arises from the potentially harmful consequences of name collision, not the name collision itself.
* 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.
* The absence of evidence is not evidence of absence, i.e., even proposed new gTLD strings that appear to be “low risk” may be in widespread use on private networks.

Building on this study, ICANN published its “New gTLD Collision Risk Mitigation” on 5 August 2013.[[3]](#footnote-3) It included proposals to mitigate the collision risks between new gTLDs and existing private uses of the same strings. The proposals require the strings to be categorized according to their risk profile using the methodology described in the Interisle Report. The three proposals can be characterized as follows.

* For strings with a low-risk profile, the registry operator would deploy an authoritative name server for the TLD with an empty zone. For a period of not less than 30 days, the registry operator would be required to investigate all DNS queries received, contacting the source of the query and notifying that source of the imminent name collision that may result. The report noted the existence of recursive resolvers that would prevent the registry operator from seeing the actual source of the query; the mitigation proposal, therefore, included the requirement that registry operators obtain the cooperation of those recursive resolvers to identify the actual source of the query.
* For strings with a high-risk profile, the registry operator would need to demonstrate that the name collision could be mitigated such that the risk profile could be reduced to a low-risk profile. The low-risk profile mitigation proposal would then apply.
* For strings with an uncalculated-risk profile, ICANN would conduct an additional study to assess the risk and understand what mitigation measures may be needed to allow these strings to move forward.

On 7 November 2013, SSAC published SAC062, “SSAC Advisory Concerning the Mitigation of Name Collision Risk,” establishing its first definition of a name collision.

In the context of top level domains, the term “name collision” refers to the situation in which a name that is properly defined in the global Domain Name System (DNS) namespace (defined in the root zone as published by the root management partners - ICANN, U.S. Dept. of Commerce National Telecommunication Information Administration (NTIA), and VeriSign) may appear in a privately defined namespace (in which it is also syntactically valid), where users, software, or other functions in that domain may misinterpret it.[[4]](#footnote-4)

SAC062 presented advice based on SSAC’s review of the issues identified in the Interisle report and ICANN’s proposals to mitigate potential collision risks. SSAC’s recommendation at the time was that high-risk strings should be considered for permanent reservation for internal or private use, suggesting that high-risk should include strings with documented evidence of broad and significant private usage. That definition could reasonably be expected to include .HOME and .CORP, and perhaps .MAIL, since the volume of DNS query data did suggest significant private usage.

The SAC062 report defines an action called “trial delegation,” which is similar to the Controlled Interruption that was ultimately deployed with a few critical differences.

* SAC062 defines two types of trial delegation: “DNS Infrastructure Testing” and “Application and Service Testing and Notification”.

“DNS Infrastructure Testing” was characterized by the delegation of the prospective TLD string with an empty zone for the purpose of collecting data on the DNS queries received at the authoritative server for the TLD.

“Application and Service Testing and Notification” was characterized by the delegation of the prospective string with a wildcard resource and having it respond with synthesized responses for the purpose of causing a name collision and providing an opportunity to alert the client of the issue in a manner appropriate for the protocol (i.e., not just the DNS protocol) in use.

* The report further notes that if ICANN operated the trial delegation, “it would presumably be easier to quickly reverse the delegation if a significant consequence is discovered that required immediate mitigation.”

SAC062 was followed by ICANN’s publication of the “New gTLD Collision Occurrence Management, a proposal to manage the collision occurrences between new gTLDs and existing private uses of the same strings.”[[5]](#footnote-5) The Board approved this proposal for implementation and outreach [via Resolutions 2013.10.07.NG01 and 2013.10.07.NG02](https://www.icann.org/resources/board-material/resolutions-new-gtld-2013-10-07-en#1.a). It includes the following definition of a name collision:

“A name collision occurs when users unknowingly access a name that has been delegated in the public DNS when the user’s intent was to access a resource identified by the same name in a private network.”

Among the recommendations presented are the following.

* The Board deferred the delegation of .HOME and .CORP indefinitely and directed ICANN org to collaborate with the technical and security community to continue to study the issues presented by these strings.
* The Board further directed ICANN org to commission a study to develop a name collision occurrence management framework. The framework would specify a set of name collision occurrence assessments and corresponding mitigation measures[[6]](#footnote-6), if any, that ICANN or TLD applicants may need to implement per second level domain name (SLD) seen in the DITL and other relevant datasets. The proposed name collision management framework will be made available for public comment.
* The proposal defined a “Collision Occurrence Assessment” that ICANN would conduct and deliver to each applicant and make available to the community. This assessment would include suggested mitigation methods, among which was the option to implement a trial delegation of some form. Details of the proposed methods can be found in Section 3.2 of the proposal.
* Section 3.3 of the proposal defined a mitigation measure called “Alternate Path to Delegation.” This required registry operators to “block” the use of an extensive set of potential second-level domain names (SLDs). This was done to ensure that a client attempting to use the domain name that would result in a name collision would continue to receive a DNS response indicating the name did not exist. Understanding that requirement is critical to the NCAP Study Two report.
* Section 3.4 empowered ICANN to develop an outreach campaign to raise general awareness and provide advice to minimize the potential for unintended consequences or harm.

ICANN completed the “Collision Occurrence Assessment”, using DITL and other relevant data

as an input, for all applied-for strings on 17 November 2013 and published them as “Reports for Alternate Path to Delegation Published”.[[7]](#footnote-7) This assessment found 25 strings ineligible for the Alternate Path to Delegation, .MAIL among them. These strings would have to wait for the name collision management framework to be developed. All other strings, excluding .HOME and .CORP that the Board had indefinitely deferred, could proceed with the Alternate Path to Delegation implemented if they were approved for delegation and the corresponding registry operator chose to do so. According to ICANN’s Delegated String page, 370 TLDs were delegated via the Alternate Path to Delegation.[[8]](#footnote-8)

On 4 June 2014, ICANN published the Phase One Report, "Mitigating the Risk of DNS Namespace Collisions,"[[9]](#footnote-9) a commissioned report by JAS Global Advisors; the final report was published in 2015.[[10]](#footnote-10) ICANN used the JAS Report, which primarily relied upon DITL data analysis, to develop the “Name Collision Occurrence Management Framework,” a guide for ICANN and the new gTLD registry operators on how to handle name collisions. The report includes several recommendations immediately relevant to the Study Two report; we refer the reader to the JAS Report for the supporting analysis associated with each recommendation.

* Recommendation 1: The TLDs .CORP, .HOME, and .MAIL be referred to the Internet Engineering Task Force (IETF) for potential RFC 1918-like protection/treatment.
* Recommendation 3: Emergency response options are limited to situations where there is a reasonable belief that the DNS namespace collision presents a clear and present danger to human life.
* Recommendation 4: Root-level de-delegation of a production TLD is not considered as an emergency response mechanism under any circumstances.
* Recommendation 5: ICANN leverage the EBERO mechanisms and functionality to respond to DNS namespace-related issues.
* Recommendation 6: ICANN require new TLD registries to publish the controlled interruption zone immediately upon delegation in the root zone. After the 90-day period, there shall be no further collision-related restrictions on the registry.
* Recommendation 10: ICANN work with the IETF to identify a mechanism for IPv6 that provides similar functionality to that available in IPv4’s “localhost” reserved prefix.
* Recommendation 14: ICANN request that the appropriate bodies further explore issues relating to collisions in existing DNS namespace, the practice of “domain drop catching,” and the associated data feeds that may be leveraged by attackers when attempting to exploit collisions.

On 6 June 2014, SSAC published SAC066, “SSAC Comment Concerning JAS Phase One Report on Mitigating the Risk of DNS Namespace Collisions.”[[11]](#footnote-11) In that document, SSAC reviewed the Phase One Report by JAS Global Advisors noted in the previous paragraph. SAC066 used the following definition of a name collision in its report.

“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 actually 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.”

SSAC identified eight issues with the Phase One JAS Report and made a recommendation about each of them. These include:

* ICANN should perform an evaluation of potential notification approaches against at least the requirements provided by the SSAC prior to implementing any notification approach.
* ICANN should implement a notification approach that accommodates Internet Protocol Version 6 (IPv6)-only hosts as well as IP Version 4 (IPv4)-only or dual-stack hosts.
* ICANN should seek to provide stronger justification for extrapolating findings based on one kind of measurement or data gathering to other situations.

Finally, we have the current “Name Collision Occurrence Management Framework,” originally published on 30 July 2014 and approved and directed for implementation by the ICANN Board with [Resolution 2014.07.30.NG01](https://www.icann.org/resources/board-material/resolutions-new-gtld-2014-07-30-en#1.a).[[12]](#footnote-12) This framework has remained in force since it was published and is the current mechanism through which ICANN assesses name collisions. ICANN considered the recommendations in the JAS Report and the advice in SAC062 and SAC066. The Framework begins with the following definition of a name collision.

“A name collision occurs when a user unknowingly accesses a name that has been delegated in the public DNS when the user's intent is to access a resource identified by the same name in a private network. Circumstances like these, where the administrative boundaries of private and public namespaces overlap and name resolution yields unintended results, present concerns and should be avoided if possible.”[[13]](#footnote-13)

Key elements of the Framework’s methodology include:

* Registry operators are required to act on name collision reports forwarded by ICANN within two hours of receipt.[[14]](#footnote-14)
* Controlled Interruption, as described by the JAS Report, is required of all new gTLDs, notably because it was decided its good notification features combined with its superior privacy protection were preferred to the use of a honeypot as defined by the SSAC.[[15]](#footnote-15)
* The lack of IPv6 support was accepted as a tolerable risk; while recognized as a gap, it was not described as a blocking concern. The Framework instead suggested that ICANN “will work within the IETF and with other relevant technical communities to identify a mechanism for IPv6 that provides similar functionality to that available in IPv4’s “Loopback” reserved prefix.[[16]](#footnote-16)
* Registry operators agree that ICANN may designate an Emergency Back-End Registry Operator (EBERO) if the Registry Operator is unable or unwilling to comply with a measure to avoid harm from name collision in a timely manner.[[17]](#footnote-17)
* The recommendation in the JAS Report to treat .MAIL the same as .HOME and .CORP was accepted by ICANN, i.e., the delegation of .MAIL was deferred indefinitely.[[18]](#footnote-18)
* ICANN will produce information materials as needed regarding name collision.[[19]](#footnote-19)
* ICANN will limit emergency response for name collision reports to situations where there is a reasonable belief that the name collision presents a clear and present danger to human life.[[20]](#footnote-20)

Moving ahead to 2017, the ICANN Board requested that SSAC conduct studies to present a data analysis on available information and provide advice to the Board on the topics around DNS name collision.[[21]](#footnote-21) The details of the resolutions and the embedded questions are covered later in this report. Two key elements from those resolutions are that SSAC was asked to propose a proper definition of a name collision and that the Board defined a new term, Collision String, as a category for undelegated strings that should be considered strings that manifest name collisions.

In response, the SSAC proposed the “[Name Collision Analysis Project (NCAP)](https://community.icann.org/display/NCAP/SSAC%2BName%2BCollision%2BAnalysis%2BProject%2B%28NCAP%29%2BHome),” which was quite broad and consistent with SSAC’s prior advice on the issue of name collisions. The final [SSAC NCAP Proposal](https://community.icann.org/display/NCAP/NCAP%2BWorking%2BDocuments?preview=%2F79437474%2F105385595%2FFINAL+NCAP+Proposal+September+2018_Redacted.pdf), published in September 2018, was organized into three studies. In broad terms, the purposes were:

Study 1: To establish a shared understanding of what we know about name collisions and a data repository for studying them.

Study 2: To conduct an analysis with the goals of understanding the source of name collisions and developing a sustainable framework for evaluating the risk of the manifestation of a name collision.

Study 3: To study and propose mitigation and remediation strategies for responding to name collisions.

The ICANN Board accepted SSAC’s suggestion for professional project management, and ultimately the project was assigned to ICANN’s Office of the Chief Technology Officer (OCTO) to manage. OCTO reviewed SSAC’s project proposal and, in collaboration with the SSAC, made minor revisions to the project and developed a budget. The ICANN Board approved moving forward with the [Revised Study One](https://community.icann.org/display/NCAP/NCAP%2BWorking%2BDocuments?preview=%2F79437474%2F105390062%2FNCAP+Proposal+for+Board+%28revised+by+OCTO+based+on+V2.5BTClean%29+REDACTED.pdf) on 14 March 2019 with Resolutions [2019.03.14.20 – 2019.03.14.23](https://www.icann.org/resources/board-material/minutes-2019-03-14-en#2.h.1).

The revised proposal reduced the scope of Study One by removing the creation of the data repository and deferring that work until Study Two, thus reducing the duration and cost of the study. The proposal noted the following definition of a name collision as baseline input for the NCAP Project.

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 actually 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. In the context of top-level domains (TLDs), the conflicting namespaces are the global Internet Domain Name System (DNS) namespace reflected in the root zone as published by the Root Zone Management Partners and any other namespace, regardless of whether that other namespace is intended for use with the DNS or any other protocol.

The formation of the DG was announced on 17 April 2019, inviting anyone in the ICANN Community to join the DG.[[22]](#footnote-22) The initial tasks of the DG were to define the term ‘name collisions’ to scope the material to be researched and review the [Request For Proposals](https://www.icann.org/en/system/files/files/rfp-ncap-study-1-09jul19-en.pdf) developed by OCTO seeking a contractor to complete the work. Ultimately, the goals of Study One were three-fold.

1. To produce a summary report on the topic of name collision that brings forth important knowledge from prior work in the area.
2. To create a list of datasets used in past name collision studies; identify gaps[[23]](#footnote-23), if any; and make a list of any additional datasets required to complete Studies 2 and 3 successfully.
3. To offer a recommendation on whether Studies 2 and 3 should be performed based on the results of the survey of prior work and the availability of datasets.

The final [Study One Report](https://community.icann.org/display/NCAP/NCAP%2BDocuments%2Band%2BCorrespondence?preview=/79437474/153519703/ncap-study-1-report-19jun20-en.pdf) was published on 19 June 2020 and included four significant findings, excerpted here from the Executive Summary.

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 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.

The final report also made a significant recommendation regarding the execution of NCAP Studies 2 and 3, that Studies 2 and 3 should *not* be performed as currently designed. The Study One Report Executive Summary continued as follows.

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 Two, 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 necessarily 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.

SSAC agreed with the assessment regarding Studies Two and Three as currently designed and set to work reframing Study Two and working with OCTO, as the Project Manager, to prepare a budget; Study Three would be reconsidered after Study Two completed. On 5 February 2021, the SSAC submitted a Revised Proposal for Study Two[[24]](#footnote-24) to the ICANN Board. On 25 March 2021, the ICANN Board accepted the Study One final report, approved the Revised Study Two, and directed the DG to proceed with the Revised Study Two with [Resolutions 2021.03.25.11 – 2021.03.25.14](https://www.icann.org/resources/board-material/resolutions-2021-03-25-en#2.b). Readers are referred to the revised proposal for a discussion of the detailed changes from the original proposal. The revised Study Two, for which this report is the final work product, stated four objectives.

* Perform a study of ICANN Collision Reports.
* Perform Impact and Data Sensitivity Analyses with respect to name collisions.
* Respond to Board Questions Relating to Study Two.
* Produce a final report on Study Two.

Overlapping the efforts of Study One and Study Two is the output of the ICANN Subsequent Procedures (SubPro) Working Group, which published its final report on 1 February 2021, “[New Generic Top Level Domain (gTLD) Subsequent Procedures Policy Development Process Final Report](https://gnso.icann.org/sites/default/files/file/field-file-attach/final-report-newgtld-subsequent-procedures-pdp-02feb21-en.pdf).” In Topic 29 of that report, the working group focused entirely on the issue of name collisions. They offered a recommendation, several affirmations, and implementation guidance to ICANN org on how to identify and mitigate name collisions before the next round of gTLDs. Readers of this report are encouraged to review the detailed rationale and support for the recommendation, affirmations, and implementation guidance in the final SubPro report. As these are both relevant and important to the NCAP work, their summary is excerpted here for easy reference.

Recommendation 29.1: ICANN must have ready prior to the opening of the Application Submission Period a mechanism to evaluate the risk of name collisions in the New gTLD evaluation process as well as during the transition to delegation phase.

Affirmation 29.2: The Working Group affirms continued use of the New gTLD Collision Occurrence Management framework unless and until the ICANN Board adopts a new mitigation framework. This includes not changing the controlled interruption duration and the required readiness for human-life threatening conditions for currently delegated gTLDs and future new gTLDs.

Implementation Guidance 29.3: To the extent possible, ICANN should seek to identify high-risk strings in advance of opening the Application Submission Period, which should constitute a “Do Not Apply” list. ICANN should also seek to identify aggravated risk strings in advance of the next application window opening and whether it would require a specific name collision mitigation framework.

Implementation Guidance 29.4: To the extent possible, all applied-for strings should be subject to a DNS Stability evaluation to determine whether they represent a name collision risk.

Implementation Guidance 29.5: The ICANN community should develop name collision risk criteria and a test to provide information to an applicant for any given string after the application window closes so that the applicant can determine if they should move forward with evaluation.

Implementation Guidance 29.6: If controlled interruption (CI) for a specific label (usually a 2nd-level domain) is found to cause disruption, ICANN may decide to allow CI to be disabled for that label while the disruption is fixed, provided that the minimum CI period is still applied to that label.

## 1.3 Methodology

With the acceptance of the revised Study Two proposal, the DG kicked off the study reports as described in “Section 2. Study Reports” and settled into a regular meeting cadence. While the DG considered the questions assigned by the ICANN Board, the researchers collected and analyzed available data relevant to understanding how to observe and measure the impact of name collisions; each report describes its specific methodology. Every line of work listed in the proposal was kept at the forefront of the DG’s thoughts to ensure the findings of the study reports informed the responses to the topics in Board resolution [2017.11.02.30](https://www.icann.org/resources/board-material/resolutions-2017-11-02-en#2.a.rationale) and that the study reports stayed in scope with the overall mandate for the group.

The DG chairs called for consensus on the responses to the Board questions, the study reports, and any special terminology after the discussion on each item was concluded during the regular conference calls. Two of the study reports went out for public consultation prior to their being used in this report to finalize the findings and recommendations to the ICANN Board. The NCAP project was also presented at ICANN74[[25]](#footnote-25) and ICANN75[[26]](#footnote-26) to ensure the broader community was aware of the work, findings, and pending recommendations.

|  |  |  |
| --- | --- | --- |
| **Item** | **Final Consensus Call Date** | **Result** |
| Case Study of Collision Strings | July 13th, 2022 | [Full|Partial|None] |
| Perspective Study of DNS Queries for Non-Existent Top-Level Domains (was Data Sensitivity Analysis) | July 13th, 2022 | [Full|Partial|None] |
| Root Cause Analysis |  | [Full|Partial|None] |
| Final Report |  | [Full|Partial|None] |

##  1.4 Terminology

* Active Collision Assessment - Controlled Interruption that returns a routable IP address of a system that would then provide a protocol-appropriate response.
* Allocation - The process by which the Board decides whether to allow an applied-for TLD to be granted to the applicant.
* Collision Strings - (from the [Proposed Definition of Name Collisions and Scope of Inquiry for the Name Collisions Analysis Project](https://www.icann.org/en/public-comment/proceeding/proposed-definition-of-name-collisions-and-scope-of-inquiry-for-the-name-collisions-analysis-project-02-07-2019)) a string that manifests name collisions
	+ Collision String Registry -a registry of names not to be allocated nor delegated (on the collision string list). Note that multiple organizations (i.e., IETF, ICANN) can add names to the reserved list.
* Controlled Interruption - (From [FAQ](https://www.icann.org/resources/pages/name-collision-ro-faqs-2014-08-01-en)) “Controlled interruption is a method of notifying system administrators who have configured their networks incorrectly (knowingly or unknowingly) of the namespace collision issue, and helping them mitigate potential issues.”
* Critical Diagnostic Measurement - properties that help determine the scope, impact, and potential harm of name collisions
* Day-In-The-Life (DITL) - a large-scale data collection project run by DNS-OARC undertaken every year since 2006. This data has historically been the primary measurement asset for name collision studies.
* Delegation - the DNS protocol technical process of adding a name to the DNS root zone[[27]](#footnote-27). This should be explicitly distinct from the process of granting the gTLD to an applicant.
* Harm - may include numerous things, from cybersecurity risks to reputational damage to physical impacts, making it difficult to appropriately apply scale and context to this otherwise broad term within the scope of name collisions.
* Name Collision - (used in Study One and RFP) 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.”
	+ Domain Name Collision - A name collision in the single resolution protocol of the DNS.
	+ Name Collision Assessment - Controlled Interruption is a mechanism for Name Collision Assessment
	+ Namespace Collision - A potential source of name collision involving multiple namespaces, such as the DNS root zone and a blockchain service.
* Name Collision Occurrence Assessment - formal output of the Technical Review Team
* Query Volume - The number of DNS requests received for a string.
* Passive Collision Assessment - a data collection and analysis method in which the applied-for TLD is delegated in the RSS and an authoritative name server for the TLD is configured with an empty zone and all DNS queries are captured.
* Root Server Identity (RSI) - thirteen identities, each of which is named with the letters ‘a’ to ‘m’, collectively administered by twelve root server operators. They are named in the ‘root-server.net’ domain.
* Search List Processing - “A Domain Name System (DNS) “search list” (hereafter, simply “search list”) is conceptually implemented as an ordered list of domain names. When the user enters a name, the domain names in the search list are used as suffixes to the user-supplied name, one by one, until a domain name with the desired associated data is found or the search list is exhausted.” [[28]](#footnote-28)
* Source Diversity - The number of distinct source IP addresses, distinct /24 or /48 IP blocks, and/or distinct number of ASNs requesting a string. This results in three different measurements/numbers used in DNS query analysis
* Static Assessment - The review of longitudinal data to identify any burst in name collision activity against a given string.

### 1.3.1 Impact and Harm

The JAS report described several of the challenges of enumerating harm when it comes to name collisions. Arguments around concepts of national security, economic hardship, and adherence to the law are impossible to manage in a diverse global context. Their final recommendation on the topic was:

“As such, we recommend that emergency response be limited to scenarios where

there is a reasonable belief that the DNS namespace collision presents a clear and

present danger to human life.”[[29]](#footnote-29)

The NCAP DG felt it necessary to extend the discussion of harm to include its potential. As noted in response to the Board questions, the DG approached harm as follows:

“To address the Board’s question, the discussion group focused on two aspects of harm: potential harm and reported harm. Potential harm is a set of circumstances that might lead users and systems to be negatively impacted by name collisions, with their possible levels of impact. Reported harm is based on actual experience disclosed by organizations and individuals impacted by name collisions.”[[30]](#footnote-30)

When considering the risk of name collisions, the potential for harm must be part of the risk assessment. Ultimately, the goal is to prevent reported harm by evaluating the potential and reacting accordingly.

# 2 Précis of NCAP Study Two Reports

As described in its revised scope, the NCAP Study Group 2 conducted three studies:

* Case Study of Collision Strings
* A Perspective Study of DNS Queries for Non-Existent Top-Level Domains[[31]](#footnote-31)
* Root Cause Analysis reports

Each study offered several insights into how to look for and understand the impact of name collisions.

The first study report, the Case Study of Collision Strings, helped define all the Critical Diagnostic Measurements (CDMs) required to identify name collisions and, further, how to assess the impact of a name collision.

The second study report, A Perspective Study of DNS Queries for Non-Existent Top-Level Domains, considered if and how the available data sets from both individual root servers and global public resolvers were representative or not of the overall picture of the DNS queries that would help identify name collisions.

The third report comprises a root cause analysis of name collisions experienced. The reports collectively assess the name collision reports submitted to ICANN via their web submission form and analyze the incidence of name collisions more generally. They provide both a quantitative analysis, using historical DNS query data, as well as a qualitative analysis, using submitted name collision reports and results from a name collision survey. They include assessments of the pervasiveness of private use of newly-delegated TLDs in DNS suffixes, the effectiveness of controlled interruption in notification and root cause identification, the severity of impact felt by affected parties, and anecdotal configurations that were common causes of name collisions.

The following sections describe the results of those studies in greater detail; the full report for each is available in Annexes A through C at the conclusion of this report.

## 2.1 Case Study of Collision Strings

The DG met over the course of approximately two years to evaluate and consider topics posed by the ICANN Board on the delegation of currently reserved TLDs such as .CORP, .HOME, and .MAIL. The group undertook a review of past studies and literature and conducted its own analysis from two root server identities. The result of that review is a modern picture of the impact and potential harm due to name collisions with the undelegated names under study. The analysis provides a sufficient basis from which to draw a number of important findings. Among these include the observation that queries for these undelegated names are both increasing in volume and diversity. These facts suggest that challenges relating to impact and risk are also increasing. The group also identified a number of Critical Diagnostic Measurements that help determine the scope, impact, and potential harm of name collisions.

## 2.2 A Perspective Study of DNS Queries for Non-Existent Top-Level Domains

The report’s analysis shows that no view at a single root server is comprehensive. However, when considering DNS clients that meet a defined query rate, a single root server observes query traffic from about two-thirds of resolvers that are observed across the entire system. Additionally, there are notable differences in DNS traffic observed by recursive resolvers and at the root server system. These findings are significant in terms of how future guidance and advice may be applied to name collision risk assessments. Specifically, these perspective differences affect the effectiveness of top-N lists, particularly when they are generated from a single source.

The publication of top-N lists of non-existent TLDs can make applicants aware of strings that exhibit some risk associated with name collisions. However, the effectiveness of such lists is limited. The very fact that these lists contain only the top N, ranked by some criteria, is constraining. This is particularly so when they are generated only from a single data source (e.g., root server queries or a single recursive resolver). Because there are multiple perspectives in the DNS ecosystem, the absence of a string on a top-N list does not provide any assurance the string is void or absent of name collision risks, nor does the magnitude or ranking of a string that does show up in the list. For example, this analysis shows that non-existent TLDs observed at high volumes by some recursive resolvers are not seen in the same rankings by root servers.

## 2.3 Root Cause Analysis Reports

The motivation for the root cause analysis reports was to investigate the name collisions reports submitted to ICANN to better understand what caused the name collisions, their severity, and the effectiveness of controlled interruption. Beginning with those name collision reports, they systematically begin a more comprehensive study of name collisions associated with the delegation of new TLDs since the introduction of controlled interruption. Their study incorporates four data sets:

* the 47 name collision reports submitted via ICANN’s name collisions Web submission form;
* historical DNS query data extracted from passive DNS observation from the time of delegation of each of the 885 TLDs delegated since August 2014.
* root DNS query data from the 48-hour once-yearly day-in-the-life (DITL) collection from 2014 to 2021;
* results from a Web search for “127.0.53.53”; and
* responses from a name collisions survey sent to both a general technical audience and those inferred to have been affected by name collisions.

Key findings from the research and analysis of available data include:

* The private use of DNS suffixes is widespread.
* The name collision reports are supported strongly by measured data.
* The usage of private DNS suffixes colliding with newly-delegated TLDs has decreased over time.
* Controlled interruption is effective at disruption but not at root cause identification.
* Configuring DNS resolvers as authoritative for DNS suffixes is not a panacea.
* The impact of TLD delegation ranged from no impact to severe impact.
* The respondents' response to controlled interruption was overall neutral.
* Name collisions were diverse, both in terms of the application involved and their root causes.

Seven of the reports submitted via ICANN’s name collisions report form were related to the interception of user Web traffic due to the combination of systems that use the Web Proxy Auto-Discovery protocol (WPAD), inadvertent usage of the domain name ‘domain.name’ in home router software, and the delegation of wpad.domain.name in the public DNS. While these issues do not fit in the same category as name collisions at the TLD level, the largest constituency of reports submitted to ICANN were associated with this issue. Thus, the DG agreed to additional research in a root cause analysis specific to .WPAD. This research contains a full delegation and resolution history of wpad.domain.name, an analysis of related queries observed at the root servers, and a behavioral analysis of the services operated by wpad.domain.name, i.e., what privacy and operability concerns might have been encountered by affected users.

For more detail on these findings, please review the Root Cause Analysis reports.

# 3 Summary of NCAP Discussion Group Activities

The study reports described above combined with a review of the materials gathered in Study One and a review of the evolution of the DNS and Internet infrastructure since the last round of new gTLDs, provided a foundation for consideration of name collisions today as compared to the last round and the opportunity to reconsider how to examine the risk they present to the security and stability of the DNS. In addition, while the prior reports focused on available data, the discussions of the DG worked to put that information, and more, in context.

## 3.1 The NCAP Gap Analysis

NCAP Study One offered an in-depth review of prior work around identifying and handling name collisions. Between the publication of the NCAP Study One [report](https://community.icann.org/download/attachments/79437474/ncap-study-1-report-19jun20-en.pdf?version=1&modificationDate=1607964281000&api=v2) and the Board r[esolutions *2021.03.25.11 – 2021.03.25.14*](https://www.icann.org/resources/board-material/resolutions-2021-03-25-en#2.b) that approved the revised proposal for Study Two, members of the group focused their efforts on identifying the gaps between the technology that uses the DNS and the mechanisms used to identify name collisions. That effort informed the Revised Study Two Proposal.

The NCAP Gap analysis offered both hypotheses to be tested and baseline assertions to inform the direction of work for Study Two and were included in Appendix 2 of the Revised Study Two Proposal.[[32]](#footnote-32) The substantive text is included here for ease of reference.

**1.) Changes in the DNS infrastructure and Protocol:** DNS usage monitoring provides insight into time-resolved traffic evolution patterns useful in the quantification of system stability and performance as well as detecting aberrant events. Longitudinal measurements and usage trends, however, are increasingly difficult to leverage as the underlying system evolves or as bifurcation within the system occurs. These system changes may result in non-symmetric system usage, partial or even total impairments in DNS measurements, and ultimately confound the interpretability of the system’s usage metrics. Since the 2012 round of TLD delegations, several new technologies and recommended best practices within the DNS ecosystem now have a significant impact on the volume and fidelity of DNS queries observed at name servers in the DNS hierarchy. These technologies include running Root on Loopback (RFC 7706), Aggressive Use of DNSSEC-Validated Cache (RFC 8198), DNS Query Name Minimization (RFC 7816), and DNS Queries over HTTPS (RFC 8484). It is in the DNS community’s best interest to develop a better understanding of how these standards and technology changes will influence data collection capabilities as well as their impacts to data analysis of DNS traffic in an ever evolving, technologically fragmented, and highly distributed system (Board Questions 2 and 7 - Study Two).

**2.) Controlled Interruption Efficacy and Data Analysis**: While the NCAP Study One Report highlights some reports around the efficacy of Controlled Interruption, we believe a more thorough assessment of the framework should be commenced. The collected reports (which comprises a data set previously unavailable for study) should at a minimum be analyzed to better understand any trends, commonalities, assumptions, and success attributes (Board Question 4 - NCAP Study Two). Understanding the nature of these reports with a re-examination of previous Day In The Life (DITL) of the internet data may help identify key signals in the DNS that could better inform name collision risk assessments moving forward (Board Questions 5, 6, and 8 - NCAP Study Two). Some applications, including popular browsers, have implemented specific DNS controls to signal when Controlled Interruption events occur. To that end, efforts should be made to identify and contact such vendors to see if instrumentation data is available. Finally, a study should be made to provide additional evidence that Controlled Interruption was a successful mitigation model, which may include creating and running simulation test beds (Board Questions 4, 5, and - NCAP Study 3).

**3.) Vulnerability Understanding and Mitigation Strategies:** Since the 2012 delegation of TLDs, various peer reviewed academic and industry papers have been published that elucidate some of the more detailed nuances of name collisions, specifically as they relate to various potential risks and vulnerabilities (Board Questions 3, 7, and 8 - Study Two). Specifically, many of these publications directly identify known DNS query patterns, typically associated with zero-configuration protocols such as DNS-SD, that we believe may be weaponized and exploited in a name collision environment. If true, this new knowledge should be applied to future TLD delegation risk assessments as it builds upon a foundational understanding of the intent of the DNS queries as opposed to the volume of queries that was originally used in the new gTLD risk assessment (Board Questions 4, 5, and 6 - Study 3).

**4.) Data Sets**: 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 as noted in Item 1 above may require a different method of analysis. Since the 2012 TLD round of delegations, 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 reshaped 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.

## 3.2 Review of Available Datasets

As part of the effort to build a workflow for evaluating name collision risk, the DG explored what DNS data is available for review. In addition to the DITL data and information from two recursive resolvers discussed in the perspective study, two additional areas were explored as possible sources for developing the necessary CDMs to evaluate name collision risk: Identifier Technology Health Indicators (ITHI) metrics and L-root DNS Magnitude data.

* Quality and availability of DITL data has been reduced (less contributing sources and anonymization)
* Recursive resolver data is not openly available, although future availability to ICANN for name collision analysis could be explored
* Then note the additional datasets that have been identified for consideration - need to note the NXDOMAIN data as responsive to SubPro rec (static do not apply list)
* At least a future pointer to the fact that PCA provides some alternate for the declining DITL data

On 4 August 2021, Alain Durand from the Office of the Chief Technology Officer at ICANN and Christian Huitema from Private Octopus presented the ITHI project (started in 2017) monitoring the health of the registered identifiers ecosystem, through a set of ITHI metrics. There are eight detailed metrics for which data can be seen on the site dedicated to the ITHI project.

The metrics are computed using data captured from various sources including data collected by ICANN projects and traces obtained from participating root DNS servers, authoritative DNS servers, and recursive DNS resolvers. Recently, ICANN Office of the Chief Technology Office has published the [OCTO-25](https://www.icann.org/en/system/files/files/octo-025-08jul21-en.pdf) document regarding the ITHI project, which includes an entire section dedicated to name collisions.

In addition to the ITHI data, the DG considered the data available from the ICANN Managed Root Server (IMRS) during the 3 November 2021 DG call (23:36 in the [recording](https://icann.zoom.us/rec/play/q_sQBiDJFQmNLxrala7bGNd2zHBCpLgxQbMndTbdj6FFAXjO2JLHN8VqUzO0yHGgBFGAa_-6Gte-itfk.gVQmFPkJCDlZ5l4i?continueMode=true&_x_zm_rtaid=-ogRgxjzQjuYlgz7OP-hWg.1659537978260.87e6fbcb4027d9be5b84c717c5fde600&_x_zm_rhtaid=833)) , specifically as part of the ICANN DNS Magnitude project.[[33]](#footnote-33) The ICANN DNS Magnitude project assumes that the number of unique networks that send DNS requests reflects the overall popularity of the domain’s services. This DNS-based metric “DNS Magnitude” can be used for estimating the popularity of a domain. As per their website, they apply this ranking and classify top-level domains by their delegation status, and offer the advice that non-existent domains that are heavily queried for by a large number of networks have a high collision risk.

Both datasets are noted as possible sources of information that the Technical Review Team (TRT) (See Section 3.5.1) might use for information prior to root delegation.

## 3.3 The Issue of Manipulation

One area of concern for the DG involves third-party manipulation of the CDMs used to evaluate the risks associated with name collisions. Discussed during ICANN 74 and on the 25 May 2022 call, there are a variety of ways a third party could fabricate the appearance of name collisions in the DNS RSI and resolver logs. At this time, there is no way to predict or prevent this type of manipulation, and identifying the data to differentiate between legitimate name collisions and fabricated ones requires longitudinal data analysis by the TRT. Moreover, a determined attacker with enough lead time could hide the manipulation such that it would be challenging for the TRT to identify itt. There is also a risk here that with the knowledge that the TRT, prospective registrants, or other parties will use the manipulated data creates an unintended incentive for this manipulation, which could result in very large numbers of unnecessary CDM queries, and thus requiring investigation that might delay Name Collision Occurrence Assessment by the TRT.

The DG agreed that reviewing the data and making this judgment call must be part of the responsibilities for the TRT. This is a difficult problem that will likely require bespoke data analysis efforts that may or may not succeed in identifying manipulation. The issue of manipulation is a residual risk that must be accounted for by TRT analysis.

## 3.4 Critical Diagnostic Measurements

As highlighted in the Case Study report, recommendations regarding any course of action in handling name collisions is based on a set of CDMs and no single class of measurement is sufficient to assess the full scale of name collision risks.[[34]](#footnote-37) The different measurements must be taken as a whole to understand how their interactions inform any technical analysis. For example,

“query volume--one of the four major classes of measurements--is an important factor, but a single source that could be easily mitigated with a simple configuration may be responsible for high query of a name. Conversely, if not only query volume was high, but query origin diversity (i.e., from many networks and many systems) and query type diversity were also extremely high, this would suggest collision impact may be greater. This is because the expectation of negative responses is high, and the mitigation across multiple services, networks, and users is increasingly complex to perform.”[[35]](#footnote-38)

The four major classes of measurement that should help the scope, impact, and potential harm of name collisions include, in no particular order:

* Query Origin Diversity - the number of unique query source IP addresses (resolvers)
* Label Diversity - diversity of labels under a name collision string
* Query Volume - the number of queries each RSI receives
* Query Type Diversity - the type of query (i.e., resource record type) being requested
* Open-source Intelligence (OSINT) and qualitative assessments[[36]](#footnote-39)
* Number of unique DNS suffixes identified

## 3.5 Comparison of Proposed Alerting and Data Collection Techniques

A critical component of the NCAP work involves the introduction of passive and active collision assessment as methods to collect name collision data. This introduction is likely to bring up a variety of questions regarding how these differ from each other and from controlled interruption. This short document offers a comparison of the different methods from the perspectives of alerting effectiveness, operational continuity, security and privacy, user experience, root cause identification, public response, and telemetry.

We first present a high-level overview of each technique. Subsequent sections describe the mechanisms in more detail as part of their comparison. Throughout the text, Phase 1 refers to the resolution of colliding names in the public DNS, while Phase 2 refers to the transport- and application-level communications that are dependent on and follow that resolution.

### 3.5.1 Static Assessment

Included in discussions around data manipulation and the contextual information when identifying any CDM as being “high” or “low” was the concept of a collection of historical data that ICANN could use as part of the risk assessment for name collisions.

A “static assessment” is the ability for the TRT (and applicant) to look at any data sets that are available prior to going into PCA, which requires a delegation and carries some amount of risk. While there are limitations to such types of data, due to either collection biases, impairments, or even manipulation, the data is intended to provide some risk mitigating input before delegating the string.

There are “legacy” data sets that were previously used in 2012 that could be used again, .i.e., DITL. ICANN is also currently capturing and summarizing name collision data via their L-root traffic and making available such data in their ITHI and Magnitude collections. Other data sets may be used for static assessment may include passive DNS data or XXX

* ICANN can foster relationships with key DNS providers (e.g., recursive resolvers) to help provide “oracle” functions.
* ICANN can work with DNS-OARC DITL to make the yearly data more accessible, understandable, and actionable for name collision purposes.

[Somewhen we discussed that name collisions could happen in a short burst, possibly as a result of a temporary misconfiguration. Having longitudinal data would help track those spikes and could be used as part of the overall assessment. Need to find where/when the DG covered that topic.]

### 3.5.2 Passive Collision Assessment

With passive collision assessment, an application attempts to resolve a domain name (Phase 1), which results in a DNS query to the client’s DNS resolver and—depending on configuration—to one or more public authoritative servers. The ultimate response from public authoritative servers is a negative response (e.g., name error or NXDOMAIN).

Client

DNS Resolver

Public DNS Authoritative Server

DNS Query

NXDOMAIN

NXDOMAIN

DNS Query

**Phase 1:**

Name Resolution

### 3.5.3. Controlled Interruption

With controlled interruption, if an application’s attempt to resolve a domain name (Phase 1) results in a DNS query to public authoritative servers, the ultimate response from the authoritative servers is 127.0.53.53 for queries of type A (IPv4 address); queries of type AAAA (IPv6 address) result in a negative response. Any attempts by the application to communicate with that destination (Phase 2) will stay local to the client.

Client

DNS Resolver

Public DNS Authoritative Server

DNS Query

TCP or UDP

(IPv4 only; all ports)

**X**

127.0.53.53

127.0.53.53

DNS Query

**Phase 1:**

Name Resolution

**Phase 2:**

Transport- and Application-Level Communication

### 3.5.4. Active Collision Assessment

With active collision assessment, if an application’s attempt to resolve a domain name (Phase 1) results in a DNS query to public authoritative servers, the ultimate response from the authoritative servers is the IPv4 or IPv6 address (for A and AAAA queries, respectively) of the server designated to handle application-layer requests from clients on select ports. Any attempts by the application to communicate with that destination (Phase 2) will leave the client.

Active Collision Assessment Server

192.0.2.1

2001:db8::1

Client

DNS Resolver

Public DNS Authoritative Server

DNS Query

**X**

192.0.2.1

2001:db8::1

DNS Query

Application

(IPv4 or IPv6; select ports)

192.0.2.1

2001:db8::1

TCP or UDP

(IPv4 or IPv6; all other ports)

**Phase 1:**

Name Resolution

**Phase 2:**

Transport- and Application-Level Communication

## 3.6 Alerting Effectiveness and Coverage

*What population of potentially affected users, systems, and applications are expected to be reached by the alerting mechanism?*

With passive collision assessment, there is no intent to alert end-user systems and applications. Because only negative responses are returned from authoritative servers, applications are not expected to behave differently. Thus, this section mostly applies to controlled interruption and active collision assessment.

**DNS Resolution of Queried Names.** One question involves the resolution of colliding names—whether or not they resolve to addresses in the public DNS (Phase 1). With passive collision assessment, there are no positive answers for queried names, so the names do not resolve. However, with both controlled interruption and active collision assessment, *all* colliding queries reaching public authoritative servers are answered with the address appropriate for the mechanism. Thus, alert success is based on the likelihood of queries reaching the public authoritative servers. Two extreme cases are the following:

* No *queries reach public authoritative servers.* This might be the case, for example, if systems are on a corporate network for which the recursive DNS resolvers are configured to answer authoritatively for the colliding namespace *and* those systems never leave the corporate network.
* All *queries reach public authoritative servers.* This might be the case where the offending query is one of the intermediate queries issued in the course of search list processing and, prior to delegation, results in negative response.

The former case constitutes name collision *potential*, which would only alert only if one of the configuration requisites changes while the mechanism (controlled interruption or active collision assessment) is still deployed. The latter would affect *all* end systems using the colliding namespace. Other configurations, including variants of these, might result in alerting affecting some subset of systems some fraction of the time.

|  |  |
| --- | --- |
|  | **DNS Resolution of Queried Names** |
| **Controlled Interruption** | Resolution of queried names depends on DNS configuration and system mobility |
| **Active Collision Assessment** | Resolution of queried names depends on DNS configuration and system mobility |
| **Passive Collision Assessment** | Queries names do not resolve |

**Application Coverage.** Resolution of a domain name is not only dependent on the query reaching the public authoritative server; it is also dependent on the capabilities of the application that initiated the resolution and the network connectivity of the system on which the application runs. Applications that support IPv4, when run on a system with IPv4 connectivity, will attempt to resolve the domain name to an IPv4 address; similarly, applications that support IPv6, when run on a system with IPv6 connectivity, will attempt to resolve the domain name to an IPv6 address.

Active collision assessment allows applications to resolve affected domain names to both IPv4 and IPv6 addresses. However, with controlled interruption, domain names can only resolve to IPv4 addresses. Because of this, only applications with IPv4 connectivity will be affected by controlled interruption. Applications and systems that are IPv6-only will neither resolve colliding domain names to IP addresses (Phase 1) nor attempt application-level communication (Phase 2) with controlled interruption.

The absence of IPv6 in controlled interruption is discussed in section of 3.1.3 of the JAS report[[37]](#footnote-40). That discussion concentrates on two questions: the need for IPv6; and the IPv6 address that would be used.

With regard to the need for IPv6, the question is the prevalence of IPv6-*only* systems, i.e., where the IPv4 mechanism would not even be a possibility. The JAS report discusses the practicalities, expectations, and even measurements associated with IPv6-only systems. From that standpoint, they conclude that controlled interruption responses for IPv6 addresses are unnecessary. Unfortunately, the data from the report is not helpful in supporting this claim. The measurement of DNS *resolvers* that “appear to be” IPv6-only simply cannot be used to approximate end-user systems, which are the primary candidates for name collisions and controlled interruption; they are simply two very different things. Additionally, the report’s reference to the number of systems accessing Google over IPv6 is not supportive, as that does not imply *IPv6-only* activity (even so, the percentage of IPv6-capable clients has risen from 3% to 40% since the JAS report was written[[38]](#footnote-41)). Despite the lack of self-presented evidence supporting the claims from the JAS report, there is data that suggests that it is impractical to expect a measurable presence of IPv6-only systems. Among that supporting data, only 26% of the top 500 Web sites have AAAA records published[[39]](#footnote-42). NAT64 and related systems are a special consideration. In a NAT64 environment, a client is effectively IPv6-only, relying on special infrastructure to make IPv4-only resources available to IPv6-only clients. A study of how applications on NAT64 systems behave when confronted with name collisions and controlled interruption is desirable but beyond the scope of this work.

With regard to the IPv6 address that might be used for controlled interruption, the JAS report authors considered ::1 (loopback), ::53 (public), addresses within fd00::/8 (site-local), and addresses within fe80::/10 (link-local). However, because of the “the potential for unintended consequences” and the relative immaturity of IPv6 implementations compared to IPv4 implementations, the JAS report recommended avoiding the risk associated with “experimenting in the ‘fringes’ of v6 for what is very likely a small benefit.”[[40]](#footnote-43) More experimentation and analysis can be done to test controlled interruption using addresses from these and possibly other ranges, but it is beyond the scope of this work.

In summary, there is currently no IPv6 address for controlled interruption. Such would require additional considerations and, perhaps most importantly, testing. The need for IPv6 in controlled interruption seems low, but remains unclear without thorough studies, including in environments like NAT64.

|  |  |
| --- | --- |
|  | **Application Coverage** |
| **Controlled Interruption** | Only applications using IPv4 are affected |
| **Active Collision Assessment** | Applications using either IPv4 or IPv6 are affected |
| **Passive Collision Assessment** | No applications are affected |

## 3.7 Operational Continuity, Security, and Privacy

*How might users or systems be negatively impacted by interruption to service or subjected to exploit or privacy violations?*

There are five levels of impact to consider with regard to operational continuity, security, and privacy.

**DNS Query Surveillance.** Colliding DNS queries (Phase 1) are observed by public authoritative servers, as well as Internet service providers, and any other operators of infrastructure on the path between recursive and authoritative servers. In the case of passive collision assessment, only a *fraction* of both name collision queries and query names are leaked to public authoritative servers because of negative caching at the DNS resolver. With controlled interruption and active collision assessment, all *query* *names* will be observed. The *rate* of queries observed for any given query name is a lower bound of the rate observed by the recursive resolver, again because of caching.

**Communication Interruption.** Application communications (Phase 2) are interrupted in connection with the intent to alert. *All* cases of controlled interruption and *select* cases of active collision assessment result in this level of impact. In the case of active collision assessment, communications to ports designated to receive incoming communications result in communication interception (discussed hereafter); all others result in communication interruption and application inference (discussed next). Examples of interruption caused by controlled interruption are well documented in the reports submitted to ICANN via their Web submission form and in the Root Cause Analysis document.

**Application Inference.** Application-layer protocols and sometimes the applications themselves can be inferred by observation of destination (TCP or UDP) port in application-layer communication attempts (Phase 2)—whether or not those attempts are intercepted. This level of impact applies to all cases of active collision assessment.

**Communication Interception.** Application communications (Phase 2) are intercepted. This will *never* be the case with controlled interruption because, by design, communications destined for 127.0.53.53 will never leave the local system, where they might be intercepted by foreign systems. With active collision assessment, select ports are designated to accept incoming communications, and application-layer data is exchanged between client and server.

**Data Exfiltration.** Potentially sensitive application-layer data is sent to the server that intercepted application-layer communications (Phase 2). Thus, clients subjected to active collision assessment are subject to data exfiltration, and only in the case where ports have been configured to accept incoming communications. In some cases the data sent might be innocuous. For example, in an SMTP transaction, the initial communication from the client (after the greeting from the server) is merely a HELO or EHLO message, used to identify the client to the server. HTTP, on the other hand, is a request-response protocol, in which the client sends application-layer data first (the request) before the server sends application-layer data (the response). An HTTP request contains information that might be considered sensitive, including the path being requested, the values of any query string name-value pairs, cookie values, and request data associated with POST request. Thus, with HTTP, by the time the server responds, the potentially sensitive information has already been transmitted.

|  |  |
| --- | --- |
|  | **Operational Continuity, Security, and Privacy** |
| **Controlled Interruption** | **DNS Query Surveillance:** all qnames**Communication Interruption:** all**Application Inference:** none**Communication Interception:** none**Data Exfiltration:** none |
| **Active Collision Assessment** | **DNS Query Surveillance:** all qnames**Communication Interruption:** all**Application Inference:** all**Communication Interception:** select**Data Exfiltration:** select |
| **Passive Collision Assessment** | **DNS Query Surveillance:** all SLDs, fraction of qnames**Communication Interruption:** none**Application Inference:** none**Communication Interception:** none**Data Exfiltration:** none |

## 3.8 User Experience

*What is the experience of the end user, in terms of application behavior, path to resolution, etc?*

With passive collision assessment, only negative responses are returned from authoritative servers (Phase 1). Therefore, applications are not expected to behave differently for the user. One possible exception to this is the case where applications querying for the TLD itself expect an NXDOMAIN (rather than a NODATA) and behave differently because of it. We do not know of any such applications, but we note the possibility for completeness.

That being said, this section mostly applies to controlled interruption and active collision assessment, which involve transport- and application-level communication (Phase 2). The user experience is dependent in part on what is experienced by the application. The application experience, in turn, is dependent on what type of name collision configuration is in place. We consider communication interruption and communication interception as two of those configurations.

**Communication Interruption**

Applications communicating with a system whose intention is to interrupt might result in one of many behaviors, depending on factors such as the transport-layer protocol used (TCP or UDP), kernel-specific routing or access policy, firewall and endpoint protection software, and the application itself. One consideration is the timing of an interruption error, for which we give two primary outcomes:

* *Quick-Response Error.* The application detects the communication error relatively quickly and possibly notifies the user. Such errors are consistent with TCP RST packets, which come from the kernel of the “server” to which the application attempted to connect, as well as ICMP port unreachable packets, which are typically sent by a kernel in response to UDP messages destined to a port that is not listening.
* *Timeout.* A potentially lengthy period of time passes before the application detects the error and possibly notifies the user. This is because neither a TCP RST nor an ICMP port unreachable message are received, so the application must wait for communications to time out.

Quick-response errors are expected almost exclusively when controlled interruption is in use. This is because: 1) hosts subject to controlled interruption only communicate with the loopback interface of the host itself, never leaving the host or the network; 2) firewalls on the loopback interface typically do not make sense, allowing queries to be received by the kernel; and 3) the kernel is the responder, doing so in one the two ways described previously.

For active collision assessment, the timing of the error response experienced by the application depends on the configuration of the network path between the client and the server and the destination port. Here are several scenarios:

* *Stealth Firewall*. If an intervening firewall, anywhere on the path, including the server itself, drops packets associated with communication to a given port, then the application will experience a timeout.
* *Active Firewall*. If an intervening firewall responds with a TCP RST or an ICMP error message, then the application will experience a quick-response error.
* *Server Rejection*. If no firewall intervenes, and the server is not listening on a given port, then the kernel responds with a TCP RST, and the application will likely experience a quick-response error.
* *Communication Interception.* If no firewall intervenes, and the server is listening on a given port, then this constitutes communication interception and is covered in the next section.

While we have attempted to enumerate the errors that might be experienced by users, the list should not be taken as definitive for two reasons. First, while the application will almost certainly experience the timeout or quick-response error associated with transport-layer communication issues, how the application handles that response varies, and what the user sees might be different. Second, in some cases the user experience does not come directly from the application that experiences the communication interruption; rather, their experience is with an application that depends on the application experiencing the interruption. For example, one name collision report submitted to ICANN described clients “freezing” when they encountered controlled interruption. Without additional qualitative data regarding the experiences of users, we can only speculate.

|  |  |
| --- | --- |
|  | **Error Response - Application Experience** |
| **Controlled Interruption** | Quick-Response Error |
| **Active Collision Assessment** | Quick-Response Error or Timeout, depending on network configuration and application port |
| **Passive Collision Assessment** | No Error |

|  |  |
| --- | --- |
|  | **Error Response - User Experience** |
| **Controlled Interruption** | Application Dependent |
| **Active Collision Assessment** | Application Dependent |
| **Passive Collision Assessment** | No Error |

**Communication Interception**

As mentioned previously, communication interception applies only to active collision assessment. We describe the user experience from the perspective of several different clients, protocols, and ports.

* *Web Browser / HTTP.* If the server listens on the standard Web port (80), and the browser requests Web content over HTTP (no TLS), then the server will return whatever content it wishes in response, i.e., about the name collision. The user will almost certainly see different content than they were expecting.
* *Web Browser / HTTPS.* If the server listens on the standard HTTPS port (443), it is expected that browsers will initiate a TLS handshake. At the point in which the browser receives the server’s certificate, most browsers will attempt to validate it, using the domain name in the Uniform Resource Locator (URL). In the case that the certificate is successfully validated by the client, two things can be said: 1) the server can return whatever content it wishes in response, i.e., about the name collision; and 2) the browser and the user will have a greater trust in the server because of the validation. However, in the case that the certificate does *not* validate, the browser will prompt the user with a warning about the certificate, discouraging them from continuing with a button like “I accept the risks and wish to proceed anyway.”

While the details of a proposed active collision assessment implementation are not included in this study, there are some practical limitations that must be considered. First, while there is precedent for a TLS certificate having many domain names, including wildcard domains, having a wildcard domain with the asterisk (\*) immediately below a TLD is rare if even possible. Additional research would need to be done to conduct the feasibility of such a setup. Even so, the asterisk (\*) in any wildcard domain can only be substituted for a single label. Thus, even with a wildcard, it is infeasible for the server to have a certificate that includes all domain names that might resolve to the server’s address with active collision assessment.

|  |  |
| --- | --- |
|  | **User Experience - HTTP / HTTPS Browsers** |
| **Controlled Interruption** | Not applicable |
| **Active Collision Assessment** | **HTTP:** unexpected content received**HTTPS:** TLS certificate errors |
| **Passive Collision Assessment** | Not applicable |

* *Other clients and protocols.* For applications other than Web browsers and protocols other than HTTP, user experience depends on the application. For example:
	+ A non-browser client attempting to access content from (or upload content to) a server over HTTP might not get the desired HTTP response—whether it is status code, header value, or response body—causing it to fail. The user experience depends on how automated the process is and how notifications are configured.
	+ A non-browser client attempting to access content from (or upload content to) a server over HTTPS might abort on validation of the server’s TLS certificate. Again, the user experience depends on the extent of process automation and notification configuration.
	+ An SSH user attempting to connect to a server whose identity is already known will be met with a warning message that “IT IS POSSIBLE THAT SOMEONE IS DOING SOMETHING NASTY!” If the user pushes through anyway, they would fail to login without a legitimate account, but only after having saved the intercepting (false) server’s public key.
	+ A mail user agent attempting to connect to a mail server over IMAP, Outlook Web Access, POP3, or the like, would encounter account issues, due to either a bad TLS validation or a login failure.

|  |  |
| --- | --- |
|  | **User Experience - Other Clients and Protocols** |
| **Controlled Interruption** | Not applicable |
| **Active Collision Assessment** | **Non-browser HTTP:** unexpected content received, other unknown errors**Applications that use TLS:** TLS certificate errors**SSH:** man-in-the-middle attack errors  |
| **Passive Collision Assessment** | Not applicable |

**Local firewall alerts**

Some firewall software raises alerts about anomalous connection attempts and/or prompts the user for permission to allow communications to proceed, even if that communication originates from the local system. This behavior would only be manifest with controlled interruption. This behavior is expected to be rare, but is known to happen. One example was found in the Web search results for 127.0.53.53, documented in the Root Cause Analysis document.

|  |  |
| --- | --- |
|  | **User Experience - Local Firewall Alerts** |
| **Controlled Interruption** | Rare but possible |
| **Active Collision Assessment** | Not applicable |
| **Passive Collision Assessment** | Not applicable |

## 3.9 Root Cause Identification

*How useful is the technique in leading users towards the root cause and a possible resolution?*

Because root cause identification implies that users or systems have experienced some sort of alert by way of interruption or interception (Phase 2), this really only applies to controlled interruption and active collision assessment.

Controlled interruption and active collision assessment have distinct strategies for helping users—and ultimately system administrators—identify the root cause of the problems they have experienced because of name collisions. Active collision assessment presents the application—and, where possible, the user—with an explicit message (i.e., actual text) that they have encountered a name collision, returned as application-layer content in the server response. The most obvious example is a Web browser attempting to retrieve content over HTTP or HTTPS, but the server returns the custom message about name collisions instead. However, as mentioned previously, retrieving this content is fraught with challenges related to the user experience. In particular, most anything other than a Web browser over HTTP suffers from warnings that are deterring at best and possibly impassible. Controlled interruption makes no attempt to intercept client-server communications but instead leaves a “hint” to the user or system administrator by way of a specific IP address (127.0.53.53), which, if the subject of a Web search, *should* result in information about ICANN, name collisions, and controlled interruption.

Both approaches have their advantages and disadvantages. Finding the hints left by controlled interruption requires technical expertise that is well beyond that of a typical user; users in that scenario might simply need to get support for a computer, application, or network that “doesn’t work.” In the case of active collision assessment, even if a user is able to retrieve and read the content prepared by the active collision assessment server server—with or without technical challenges—there remain questions about how effectively that content will be processed and acted on by the user. For example, with the message presented, would the user know which steps to take next, whom to contact, etc? In both scenarios, the problem (hopefully) finds its way to system or network administrators. In the best case, those administrators are knowledgeable and are able to track down the relevant information and resolve the problem—not just for the short term, but also for the long term. In the worst case, the root cause is not properly identified, and the problem not appropriately fixed. This might be because the message isn’t sufficiently understood, or because the problem is difficult to reproduce. The survey results in the Root Cause Analysis document showed that 50% of those that observed 127.0.53.53 successfully associated them with ICANN. The analysis of the Web search results for “127.0.53.53”, also contained in the Root Cause Analysis document, showed a success rate of 76%. However, for active collision assessment, only a properly designed and executed user study could properly measure the effectiveness of both the user experience and the messaging.

While we have no data specifically related to active collection assessment, we do have three collections of data that provide insights into the relative effort and intuitiveness associated with identifying the cause of a name collision with controlled interruption, all from the Root Cause Analysis document: 1) the name collision reports submitted to ICANN via their Web form[[41]](#footnote-44); 2) the results from a Web search for “127.0.53.53”; and 3) responses to the survey issued to network operators as part of the root cause analysis. With the name collision reports submitted to ICANN, the submitters obviously found the Web page for the submission form. However, it is unclear whether finding that page came from observing the controlled interruption IP address (127.0.53.53) or from other troubleshooting measures. We also cannot say anything about cases in which name collisions were experienced but the form was not discovered or cases in which the users found the page but were deterred from making a submission because of the statement on the form indicating that only those experiencing “a clear and present danger to human life” should submit a report. Nevertheless, we can learn several things from the reports with regard to root cause identification.

First, relatively few of those that reported name collisions observed the controlled interruption IP address. Of the 34 reports, only eight (34%) mentioned “controlled interruption” or referred to the IP address “127.0.53.53.” We cannot definitively conclude that reports that did not include such references did not observe the controlled interruption IP address. However, responses to the survey support these low numbers: only 28% of those that experienced name collision problems observed the controlled interruption IP address, and only half of those found it useful in identifying the root cause. The Web search results showed a slightly higher figure of those that made the association between the controlled interruption IP address and ICANN: 76% of those that observed 127.0.53.53 also mentioned ICANN.

Second, an analysis of the reports includes very little evidence to suggest that Web browsers are a common way for name collisions to be manifest. The root cause analysis summarizes the experiences of individuals and organizations that submitted name collision reports and rates the level of impact described in those reports. Seventeen (50%) of the 34 reports were rated as either having “severe” or “significant” impact. Of those 17, 14 (82%) did not mention specific applications but described more general problems. For example: “Network down”; “causing … laptops to crash”; “cannot resolve DNS”; and “Users cant loggon to local domain.” The remaining three (18%) mentioned applications more specifically: mail and network shares. Only two reports—from the remaining 10 reports categorized as having “small-scale” impact—possibly referred to Web browsers: “Internet browsing issues from LAN”; “can't access to some servers.” We are limited to the data that was submitted in the reports and have had no avenue to pursue follow-up information. Nevertheless, it is backed up by the analysis of the Web search results for “127.0.53.53,” which shows that only 20% of results for which an application was identified were related to Web browsers. Based on this data, we expect that alerts that would be noticed in browsers are in the minority, diminishing the effectiveness of root cause identification with active collision assessment.

|  |  |
| --- | --- |
|  | **Root Cause Identification** |
| **Controlled Interruption** | **Low** - hint often not observed (34%) or not understood (24% - 50%) |
| **Active Collision Assessment** | **Low** - name collisions experienced in Web browsers are few (12 - 20%) |
| **Passive Collision Assessment** | Not applicable |

## 3.10 Public Response

*In what ways might the techniques be received in the public, with ICANN and others being accountable for complaints and fallout associated with design and execution of the mechanism?*

This section is intended to provide a comparison of the possible public reception of the techniques, based on the experience of end users, system administrators, or other parties with the deployment of the techniques themselves or similar mechanisms.

As previously mentioned, we do not expect passive collision assessment to affect the user experience, and we expect any impact on security and privacy to be negligible at best. Therefore, we do not expect expressions of sentiment of any kind associated with passive collision assessment; in fact, we hardly expect passive collision assessment to be noticed. Our description hereafter applies to controlled interruption and active collision assessment.

Because we have eight years of deployment experience with controlled interruption, we have some insights into the public reception associated with its deployment. The results of the Web search for “127.0.53.53” were analyzed for sentiment, as documented in the Root Cause Analysis document. In 94% of cases, the public comments surrounding a name collision were neutral in nature—neither positive nor negative. Only one result (6%) conveyed a very negative sentiment, towards both ICANN and the registry of the affected TLD. Based on this data, we suspect that controlled interruption is generally low risk with regard to public reception.

On the other hand, we have no deployment experience with active collision assessment. Therefore, we can only refer to experiences associated with the deployment of similar proposals and deployments from the past, while being careful to distinguish key differences. VeriSign’s Site Finder effectively employed a technique almost identical to that being proposed[[42]](#footnote-45). The major difference is that Site Finder introduced a wildcard A record into the com and net zones, while active collision assessment proposes introducing a wildcard into a TLD not previously delegated. Thus, active collision assessment is expected to affect a much smaller and more targeted population of users—namely those with name collision issues. The public outcry associated with Site Finder was extensive. The same day that it was rolled out, there were calls to submit complaints to VeriSign, ICANN, and the United States Department of Commerce, patches to DNS resolver code to bypass or blacklist the Site Finder mechanism, and even calls inflict a denial-of-service attack on VeriSign using their own framework[[43]](#footnote-46)[[44]](#footnote-47). Thus, the general sentiment of Site Finder was negative.

We emphasize that active collision assessment has a different motivation, comes at a different time, and affects a smaller and more targeted population of systems and users than Site Finder. Nonetheless, it provides useful insights in evaluating potential public reception because of its similarity in technique to active collision assessment.

|  |  |
| --- | --- |
|  | **Public Response** |
| **Controlled Interruption** | **Neutral (94%)**, based on actual deployment experience |
| **Active Collision Assessment** | **Unknown, Possibly negative**, based on experience with Site Finder |
| **Passive Collision Assessment** | **No reactions anticipated** |

## 3.11 Telemetry

*How much data is available to investigative parties, and what type of effort will it take to collect and analyze it?*

Earlier sections describe the reach of each data collection technique as well as their impact on security and privacy. These same attributes can be used to assess their effectiveness in terms of data collection.

**DNS Queries.** With all three techniques, DNS queries (Phase 1) can be collected by the DNS servers that are configured to answer authoritatively for the given TLD namespace. In the case of passive collision assessment, a fraction of DNS queries and query names will be collected, based on the average query rate and the negative cache time-to-live value. In the case of controlled interruption and active collision assessment, *every* name queried will be observed at the authoritative server, even though the rate of queries for given query names by end systems is masked by the caching behavior of DNS resolvers.

Because collection of query data at root servers is already common practice, the value of passive collision assessment might be called into question. In the context of telemetry, we now address some of the advantages of passive collision assessment over simply collecting DNS queries at the root DNS servers.

* *Real-time Availability.* The root servers are operated by 12 different organizations, and each has their own policies with regard to DNS query logging. This includes the specific query components that are recorded, how long data is maintained, whether or not it is anonymized, and how it is shared. Thus, root server data is not generally available to third parties. An exception to this is the annual collection of DNS queries carried out by many root server operators and known as the Day in the Life (DITL). DITL data is made available to members of the DNS Operations, Analysis, and Research Center (DNS-OARC). In contrast, authoritative DNS servers operated in connection with passive collision assessment would facilitate continuous and near real-time analysis—not simply data with year granularity.
* *Consolidated Control.* Considering that 12 distinct operators operate the root servers, getting relatively comprehensive query information—outside of DITL—would require cooperation, coordination, and consent from each operator. However, by delegating a TLD to specific authoritative servers designated for passive collision assessment, no coordination or consent is needed.
* *Increased Query and qname Volume.* The DNS servers associated with passive collision assessment are configured to answer authoritatively for a more specific namespace than the root zone. The impact can be explained as follows. A resolver that learns that a *TLD* does not exist can infer without further queries to the root that *subdomains* of that TLD do not exist. Therefore, the resolver doesn’t return to the root server with a query under that TLD until the negative cache value for the root zone expires. However, when the TLD *does* exist (i.e., with passive collision assessment), the resolver must query the servers authoritative for the TLD for every *second-level domain (SLD)* that it doesn’t know about. Once the resolver learns that the SLD does not exist, it does not return to the TLD authoritative servers again for query names under that SLD until the negative cache value for the TLD expires. Thus, while only a fraction of query names are observed in either case, the root might only observe one query per negative cache value for a given TLD and resolver, while passive collision might observe one query for a given SLD and resolver, in the same time period. Additionally, the negative cache value of the TLD is under control of the TLD operator. Thus, it could be reduced to increase the query rate.
* *Reduced QNAME Minimization Effects.* Many DNS resolvers have implemented a feature known as QNAME minimization, in which only the minimum labels are issued in a query to an authoritative server, reducing the authoritative server’s visibility into full query names being issued. At the root servers, only the TLD would be seen in the strictest cases of QNAME minimization. At the TLD authoritative server, however, the SLD would be observed in such cases.
* *Reduced Aggressive Negative Caching Effects.* Many DNS resolvers have implemented aggressive negative caching for DNS zones signed with DNSSEC. With aggressive negative caching, a resolver can *infer* domain names that do not exist with hints provided gratuitously by authoritative DNS servers in connection with queries for *other* domain names. The result is that queries that might otherwise be asked of authoritative servers can be answered by the resolver without querying the authoritative servers. This yields fewer queries and fewer query names from resolvers that support aggressive negative caching. However, this only applies to DNS zones that are DNSSEC-signed. Thus, it applies to the root zone and root servers, but as long as TLDs to which passive collision assessment is applied are not DNSSEC-signed, aggressive negative caching does not apply.

**IPv4/IPv6.** With active collision assessment, data can be collected with regard to which IP address family is used to attempt application-layer communication (Phase 2). This is not possible with either controlled interruption or passive collision assessment.

**Transport-Layer Protocol and Ports.** With active collision assessment, the transport-layer protocol and destination TCP or UDP port can be collected to infer the application-layer protocol with which communication is being attempted (Phase 2). This is not possible with either controlled interruption or passive collision assessment.

**Application-Layer Data.** With active collision assessment, not only can the destination port be observed, but also some amount of application-layer data, depending on the protocol and the logging configured (Phase 2). For example, for an HTTP/HTTPS request from a Web browser, the whole HTTP request could be logged, including request method, path, user-agent, and more. This is not possible with either controlled interruption or passive collision assessment. Additionally, for TCP-based communications, once a connection has been established, there is reasonable assurance that the client IP address was not spoofed by an off-path entity.

|  |  |
| --- | --- |
|  | **Telemetry** |
| **Controlled Interruption** | **DNS queries:** all qnames; end-system query volume masked by caching**Application:** no telemetry |
| **Active Collision Assessment** | **DNS queries:** all qnames; end-system query volume masked by caching**Application:** IPv4 and IPv6; TCP/UDP usage and destination ports; application-layer data |
| **Passive Collision Assessment** | **DNS queries:** all SLDs, fraction of qnames, end-system query volume masked by caching**Application:** no telemetry |

## 3.12 Generated Measurements of Collision Potential

Deployment of the proposed alerting and data collection techniques will result in telemetry data consistent with their respective capabilities. As colliding namespaces are used by end systems and users, and the queries reach public authoritative DNS servers, the activity is captured in DNS query and application logs. However, as noted in the section on “Alerting Effectiveness and Coverage”, there are network environments in which DNS queries *would* collide—should they be allowed to reach public authoritative DNS servers—but the network configuration of these systems prohibits those queries from reaching public authoritative DNS servers. The proposed data collection techniques currently have no way to measure this name collision potential.

To address this possible gap, two additional measurement techniques have been proposed, which will test for the private use of namespace within a network. Both techniques use the same configuration as passive collision assessment. That is, the TLD is delegated, but any queries associated with the TLD result in NXDOMAIN. Thus, these measurement techniques can be considered in the same light as passive collision assessment.

We now discuss specifics of each measurement technique and assess how each would contribute to name collision telemetry.

**RIPE Atlas probes.** The RIPE Atlas platform has thousands of measurement probes embedded in “host” networks around the world. These probes can be used by researchers to run measurements, which include issuing queries against the DNS resolvers designated for that network—that is, the resolvers used by other systems on the host network. The proposed measurement involves issuing queries under a given TLD to the DNS resolvers used by a RIPE Atlas probe and observing the authoritative DNS server logs for those queries. Networks for which queries are not observed at the authoritative servers can be inferred as using those TLDs as a private namespace.

**Ad-based measurements embedded in Web pages**. Special ads, embedded in Web pages, can cause the browser that renders them to fetch a given resource (e.g., an image) over HTTP/HTTPS. Fetching the resource requires a DNS lookup using the DNS resolvers used by the browser. The ads themselves show up in browsers of users world-wide, based on the algorithms of the ad company. In addition to their primary purpose of promotional marketing, ads can be used to run Internet-related measurements. The proposed measurement involves placing an ad that requires the browser to fetch a resource hosted at a domain name under a given TLD and observing the authoritative DNS server logs for queries associated with requests for that resource. Networks for which queries are not observed at the authoritative servers can be inferred as using those TLDs as a private namespace.

Probe or Web browser

DNS Resolver

Public DNS Authoritative Server

DNS Query

NXDOMAIN

NXDOMAIN

DNS Query

DNS Resolver

Public DNS Authoritative Server

DNS Response

DNS Query

**Indicator that namespace *not* used privately**

**Indicator of private use of namespace**

### 3.12.1 Limitations

The proposed measurement techniques promise to enhance the telemetry associated with name collisions, in particular where name collision potential could not otherwise be identified and quantified. Nevertheless, there are limitations with the techniques, which we now discuss.

* While the measurement techniques would identify *networks* for which TLDs are being answered internally—rather than communicating with the public DNS—the resulting measurement data does not necessarily reflect actual activity by *end users and systems*.
* Although several different configurations and usage models result in name collisions, these measurement techniques only address a subset of those—in particular those that involve private use of namespace.
* Queries observed at authoritative DNS servers—both TLD and root servers—will include queries from both actual end systems *and* the active measurements herein proposed. Without further filtering and processing, the queries from the active measurements will affect the data and metrics associated with “normal” behavior. At the very least, the two types of queries should be made distinguishable from one another to make accurate and meaningful assessments of the data. This is possible at the TLD authoritative servers by using query names whose second label is distinguishable. However, for measurements at root servers, this might not be possible due to a growing percentage of resolvers that use qname minimization and for which the second label will not be visible.
* For the RIPE Atlas measurements, not all probes will point at DNS resolvers that are used by end users and systems.
* For the RIPE Atlas measurements, data will only be gathered for networks that host a RIPE Atlas probe.
* For the ad-based measurements, not all browsers are configured to use the DNS resolvers associated with the network to which they are connected, so their experience may or may not be typical of the network.
* qname minimization has been deployed in many DNS resolvers across the Internet. As such, observing the full query name at the public authoritative DNS servers can only be expected a fraction of the time. Thus, any identifiers associated with query names must be embedded in the second label.

## 3.13 Workflow Development

* TRT
* Neutral Third Party
* Workflow

After considering the variability (i.e., both quantitative and qualitative measures) possible in how to identify name collisions and their potential for harm, the DG considered what the actual workflow might look like in order to evaluate the risks associated with name collisions. Given the goal of a sustainable, repeatable process, the DG iterated on a workflow that ICANN would be able to implement consistently and transparently. The workflow included several functions that were categorized as a Technical Review Team and a Neutral Third Party, as well as a timeline that laid out what everyone might expect from a name collision risk assessment process.

###

### 3.13.1 Technical Review Team Development

As part of the proposed name collision workflow, the DG has recognized the need to have a TRT that will serve four functions: assessing the visibility of name collisions, documenting the results, assessing any mitigation or remediation plans, and providing an emergency response if necessary. Broadly speaking, members of the TRT are expected to be data scientists with significant technical expertise and no conflicts of interest that would impede their neutral evaluation of a delegated string.

Whether these functions are handled by a single group or organization or split up into several teams is an implementation decision for ICANN to make. For ease of discussion, the DG described all these functions are part of a single TRT’s remit.

**Assess the visibility of name collisions**

The main purpose of the TRT is to identify high risk strings. Their evaluation would happen at three points in time during the application and delegation process: prior to delegation, during Passive Collision Assessment, and during Active Collision Assessment. At each point, the TRT is expected to document their results as part of making a recommendation to move onto the next assessment activity (i.e., moving from static to passive to active assessment).

Prior to delegation, the TRT would examine the data available prior to root delegation (e.g., ICANN Managed Root Server (IMRS) logs, ITHI data, DITL data, human-submitted reports, and any other contextual data as may be available) to look for evidence of name collisions. During Passive Collision Assessment, the TRT will collect all available CDMs and any other contextual data as may be available, such as unique strings or labels that might help the TRT understand or identify the root cause of the name collision. The evaluation at this stage is expected to expand over time as the TRT builds a record of previous research. Part of the evaluation would then include comparing the string against a historical baseline to look for known trends. During Active Collision Assessment, the TRT will continue to collect data, including any additional CDMs from protocols other than DNS (e.g., web, email, and others as identified by PCA).

**Document the results**

As noted above, at each point of the evaluation process, the TRT must document their findings to summarize the data seen, measured, and assessed. Any conclusions or recommendations would need to be carefully documented in order to support the goal of transparency.

Part of the documentation effort would include offering reports to the applicant that includes one to two degrees of anonymized, aggregated data. Making this data available allows for an open dialogue with the applicant and should provide insight into any steps needed for mitigation and remediation.

At each point, the TRT will be considering what recommendations to make regarding requesting trial delegation, continuing on to ACA, and ultimately the final disposition regarding whether or not to recommend awarding the collision string to the applicant.

**Assess mitigation and remediation plans**

While the full detail regarding what constitutes an appropriate mitigation and remediation plan is assigned to NCAP Study Three, the TRT is expected to identify when there is a need for such plans. Based on the data they have available from their assessment, they would be in the best position to evaluate how the mitigation and remediation plan offered by the applicant are responsive to the technical issues observed from the CDMs.

**Emergency response**

When necessary, the TRT would indicate if an emergency response is necessary to revert the delegation at any point in the assessment process. While no such process exists today for the emergency removal of a delegation, the DG determined this is a natural and necessary part of the assessment workflow.

The TRT should understand that its role is to identify high-risk strings that are problematic, and otherwise there is an expectation that the string will get delegated.

### 3.13.2 Neutral Third Party Development

While the TRT is expected to assess and report on the data, the DG discussed the additional function of a Neutral Third Party that would handle the mechanics of data collection. Broadly speaking, the Neutral Third Party consists of operators with some expertise in data science or at least analyzing the behavior of the protocols being served. Whether this is treated as a single activity or divided into multiple teams is an implementation detail; for ease of discussion, the DG refers to the Neutral Third Party as a single entity.

Given their access to detailed log data, the need for a clear conflict of interest and data use policy to protect the confidentiality of the data being collected is a critical aspect of this function.

**Operate an authoritative DNS name server for the proposed TLD**

The Neutral Third Party must be capable of developing, configuring, and supporting the root-level name server functionality required to allow for basic support for capturing and logging DNS queries, including timestamps, Qname, Qtype, SrcIP, and Protocol. Additional actions such as enabling synthetic NXDOMAIN response behavior and wildcard answers for ACA are also required.

**Operate Active Collisions Assessment environment**

As described in section 3.4.4 Active Collision Assessment, ACA is a critical component of name collision risk assessment. The Neutral Third Party would develop, configure, and support this service, allowing for the collection of CDMs against protocols beyond just DNS itself (e.g., HTTP/HTTPS, SMTP, SSH).

In managing this service, the Neutral Third Party would be responsible for collecting only the data absolutely necessary (e.g., the domain name in a URL and not the entire URL string) in order to protect the privacy of the end user.

**Log processing**

The Neutral Third Party would be responsible for collecting information both from the name server and ACA server logs, as well as from other passive measurements including Passive Collision Assessment and other passive measurement entities (e.g., Ad measurement, Atlas Probes) where possible.

The goal is to provide the TRT with basic summary statistics (e.g., queries/day, number of distinct IPs, etc.) and CDMs. If a service such as Ad measurements or Atlas Probes is implemented, the Neutral Third Party would process and send data on DNS query identification, extraction, and data sharing to the TRT.

**Emergency response**

As noted in the discussion regarding the TRT, the TRT will indicate to when an emergency response is required to revert a delegation. The Neutral Third Party would coordinate with the TRT implement that process as needed.

### 3.13.3 Workflow Evaluation

Being able to offer the ICANN Board cogent advice on how to assess the risk of name collisions required the DG to consider what the workflow for such an assessment might look like. The DG focused on the need for a more granular ability to collect data than is possible via the Controlled Interruption process as followed for the 2012 gTLD round. Discussing the workflow, what would be in scope, and what is missing from ICANN’s existing policies and procedures took several months (see DG notes from October 2021 through April 2022).

The details of that workflow can be found in Recommendation 6.4. The workflow includes an underlying expectation that the goal of the process is that all names should be delegated unless there is a compelling reason not to that is supported by both the quantitative and qualitative analysis of available data.

Each step in the workflow is a linear progression from the previous step; the DG considered it crucial that both the applicant and the TRT be able to remove the string from consideration at any step in the process. This ‘offramp’ option requires the ability for ICANN to do an emergency change to the root zone to remove a delegation; ICANN has no such process at this time.

**Workflow Timeline**



## 3.14 Tabletop Exercises

Towards the end of the Study Two project [meetings 92 and 93], the DG worked through a set of hypothetical tabletop exercises, acting the role of the TRT in reviewing the data to assess name collision risk. The goal of these exercises was to stress the proposed workflow in different ways and identify any gaps in the workflow or where clarification was needed.

For the sake of this exercise, high-end was described as elevated metrics on all of the CDMs previously defined (i.e., relatively high query volume, query type distribution, etc.). Ultimately, the values of high and low are expected to be subjective and will be defined in context by the TRT.

The DG worked through five fictitious gTLDs: .high-end, .low-end, .pca-spike, .aca-spike, and .pca-ad (the last gTLD designed to work through what the TRT would do with additional information as acquired to the ad network). The .high-end and .low-end gTLDs were evaluated against data collected in the past (static assessment), PCA, and ACA. The other gTLDs were used to demonstrate behavior when measurements showed a spike in numbers that needed investigation.

As a result of this exercise, the DG identified the need to explicitly include the capability to exit the assessment process at any point. This aligned with the SubPro Implementation Guidance 29.5 that stated the need for ICANN to develop a process “​​so that the applicant can determine if they should move forward with evaluation.”[[45]](#footnote-48)

# Appendix 1 - Revised Definition of Name Collision and Scope of Work

The original RFP for Study One also touched on the possibility of name collisions going beyond the DNS; this was noted as out of scope for the NCAP studies:

“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.”

However, post-Study 1, it was noted by the DG that an item was erroneously included in the “In scope but not intended to be the subject of data studies” as it was in direct conflict with the definition above. Item B.c in which “Registrant Alice uses EXAMPLE.COM and the lets the registration expire. Registrant Bob the registers and delegates EXAMPLE.COM. Traffic intended for Alice’s use of EXAMPLE.COM is now received by Bob’s use of EXAMPLE.COM”. By the definition provided, B.c is out of scope because it must be in a different namespace. A re-registration, by the above definition, is not a different namespace. The resolution process for that name depends on the IANA root zone.

 This concern of name collisions is more firmly described in ICANN OCTO’s report “Challenges with Alternative Name Systems”[[46]](#footnote-49):

“The Domain Name System (DNS) is a component of the system of unique identifiers ICANN helps to coordinate. It is the main naming system for the Internet. It is not the only one. Some naming systems predate the DNS, and others have been recently proposed in the wake of the blockchain approach of decentralized systems.

Proposing a new naming system is one thing. Making sure everybody on the Internet can use it is another. Alternative naming systems face a huge deployment challenge. A number of solutions exist to bridge the DNS to those parallel worlds, but they all come with their own drawbacks.

Furthermore, the lack of name space coordination, either between those alternative naming systems and the DNS, or simply among those alternative naming systems, will result in unworkable name collisions. This could lead to completely separate ecosystems, one for each alternative naming system, which would further fragment the Internet.

1. “History of the Name Collision Analysis Project,” [https://community.icann.org/display/NCAP/History+of+the+Name++Collision+Analysis+Project](https://community.icann.org/display/NCAP/History%2Bof%2Bthe%2BName%2B%2BCollision%2BAnalysis%2BProject). [↑](#footnote-ref-1)
2. <https://www.icann.org/en/system/files/files/name-collision-02aug13-en.pdf> [↑](#footnote-ref-2)
3. <https://www.icann.org/en/system/files/files/new-gtld-collision-mitigation-05aug13-en.pdf> [↑](#footnote-ref-3)
4. <https://www.icann.org/en/system/files/files/sac-062-en.pdf> [↑](#footnote-ref-4)
5. <https://www.icann.org/en/system/files/files/resolutions-new-gtld-annex-1-07oct13-en.pdf> [↑](#footnote-ref-5)
6. This footnote is quoted from the proposal: “Note that measures taken by ICANN or TLD applicants are attempts to mitigate unintended consequences or harm by preventing a name collision from occurring. These measures do not mitigate the causes of collision occurrences. Mitigating causes is a matter for users, private network operators, software developers, or equipment manufacturers to address. [↑](#footnote-ref-6)
7. <https://newgtlds.icann.org/en/announcements-and-media/announcement-2-17nov13-en> [↑](#footnote-ref-7)
8. <https://newgtlds.icann.org/en/program-status/delegated-strings> [↑](#footnote-ref-8)
9. <https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.pdf> [↑](#footnote-ref-9)
10. <https://www.icann.org/en/system/files/files/name-collision-mitigation-final-28oct15-en.pdf> [↑](#footnote-ref-10)
11. <https://www.icann.org/en/system/files/files/sac-066-en.pdf> [↑](#footnote-ref-11)
12. [https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.](https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.pdf)pdf [↑](#footnote-ref-12)
13. pg 1, [https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.](https://www.icann.org/en/system/files/files/name-collision-framework-30jul14-en.pdf)pdf [↑](#footnote-ref-13)
14. pg 2, ibid [↑](#footnote-ref-14)
15. pg 2, ibid [↑](#footnote-ref-15)
16. pg 5, ibid [↑](#footnote-ref-16)
17. pg 4, ibid [↑](#footnote-ref-17)
18. pg 4, ibid [↑](#footnote-ref-18)
19. pg 4, ibid [↑](#footnote-ref-19)
20. pg 4, ibid [↑](#footnote-ref-20)
21. via [resolutions](https://www.icann.org/resources/board-material/resolutions-2017-11-02-en#2.a) (2017.11.02.29 - 2017.11.02.31) [↑](#footnote-ref-21)
22. https://www.icann.org/en/announcements/details/invitation-name-collision-analysis-project-ncap-discussion-group-17-4-2019-en [↑](#footnote-ref-22)
23. This footnote is quoted from the RFP: Gaps in the data refers to types, sources, specific events captured, etc., that were not used in prior work but would have been useful or even necessary for the prior work to have been comprehensive. [↑](#footnote-ref-23)
24. <https://www.icann.org/en/system/files/files/ssac2021-02-05feb21-en.pdf> [↑](#footnote-ref-24)
25. <https://75.schedule.icann.org/meetings/WxsCLa9h4NapEaq6n> [↑](#footnote-ref-25)
26. <https://74.schedule.icann.org/meetings/wcin8eB2MQNNRwWP6> [↑](#footnote-ref-26)
27. <http://archive.icann.org/en/committees/gac/gac-cctldprinciples-23feb00.htm> [↑](#footnote-ref-27)
28. <https://www.icann.org/en/system/files/files/sac-064-en.pdf> [↑](#footnote-ref-28)
29. pg 12, JAS report [↑](#footnote-ref-29)
30. link to section on Harm in Board Question responses [↑](#footnote-ref-30)
31. Previously termed “Impact and Data Sensitivity Analysis” [↑](#footnote-ref-31)
32. [https://community.icann.org/display/NCAP/NCAP+Working+Documents?preview=/79437474/158140551/5%20Feb%20NCAP%20Package\_Redacted.pdf](https://community.icann.org/display/NCAP/NCAP%2BWorking%2BDocuments?preview=/79437474/158140551/5%20Feb%20NCAP%20Package_Redacted.pdf) [↑](#footnote-ref-32)
33. Also covered in a session held at ICANN 72 (<https://72.schedule.icann.org/meetings/EpPBA8MefE5dw6Ymm>) [↑](#footnote-ref-33)
34. Case Study, pg 26 [↑](#footnote-ref-37)
35. Case Study, pg 27 [↑](#footnote-ref-38)
36. Open-source Intelligence (OSINT) and qualitative assessments are also mentioned in the Case Study and, but those strings require a qualitative rather than a quantitative assessment. OSINT strings require research to understand the semantic meaning of the string and what that string could be associated with. [↑](#footnote-ref-39)
37. JAS Report (<https://www.icann.org/en/system/files/files/name-collision-mitigation-study-06jun14-en.pdf>) [↑](#footnote-ref-40)
38. <https://www.google.com/intl/en/ipv6/statistics.html> Last visited Aug 26, 2022. [↑](#footnote-ref-41)
39. <http://www.delong.com/ipv6_alexa500.html> Last visited Aug 25, 2022. [↑](#footnote-ref-42)
40. JAS Report. [↑](#footnote-ref-43)
41. <https://www.icann.org/en/forms/report-name-collision> [↑](#footnote-ref-44)
42. [https://web.archive.org/web/20041109202247/http://www.verisign.com/static/002702.pdf](https://web.archive.org/web/20041109202247/http%3A//www.verisign.com/static/002702.pdf) [↑](#footnote-ref-45)
43. <https://slashdot.org/story/03/09/16/0034210/resolving-everything-verisign-adds-wildcards> [↑](#footnote-ref-46)
44. <https://mailman.nanog.org/pipermail/nanog/2003-September/166467.html> [↑](#footnote-ref-47)
45. SubPro pg 135 [↑](#footnote-ref-48)
46. https://www.icann.org/en/system/files/files/octo-034-27apr22-en.pdf [↑](#footnote-ref-49)