New gTLD Security, Stability, Resiliency Update: Exploratory Consumer Impact Analysis

[Verisign Labs Technical Report #1130008 Version 1.1]

August 22, 2013

1 Introduction

Interesting times await those who rely on something, and at once cannot imagine it failing. For example, we rely on the Domain Name System (DNS) [47] for almost everything we do online. We often pay little attention to this seemingly simple system because it mostly "just works," and it has been working for more than 30 years. We count on it to always be highly available, but some recent developments in the DNS ecosystem suggest that we might have begun to mistake its stability and ubiquity for unbounded robustness and flexibility. One might argue that the DNS has fallen victim to a famous curse, pronounced by Mark Weiser,

> *'The most profound technologies are those that disappear."*

These sage words suggest much, but also imply the fate that we often inherently overlook important aspects of those technologies that we depend on the most.

Consider what might happen if overnight, some networked systems inside a healthcare provider in Japan began to suffer undiagnosed system failures. Would it be a concern if some installations of banking software in the islands of the Caribbean became non-responsive? Perhaps pause would be warranted when embarking on a visit to a developing nation, and discovering that the hotels in the region have suffered outages of their reservation systems. What if a rash of major enterprises around the world began suffering from widespread networked system failures of their internal operations (payroll, benefits, VoIP systems, etc.)? What if voice communications for home users became impacted by disruptions? Are specially branded names actually *less* secure than they were under more innocuous naming schemes? All of these scenarios have a measurable dependence on the DNS, and our measurements suggest they might also have a measurable dependence on the *lack* of certain generic Top Level Domain (gTLD) strings being delegated from the DNS root zone.

In 2011, the Board of the Internet Corporation for Assigned Names and Numbers (ICANN) [2] voted to release a document called "The New gTLD Applicant Guidebook." In some ways, this formally heralded ICANN's intention to allow private parties to apply for, and add, multitudes of new gTLDs to the root zone. It is noteworthy that before this, the DNS root zone barely grew at a snail's pace. Additions of new TLDs were infrequent. While this was *not* a mechanical limitation (delegations could certainly be allocated at a much greater pace), this was a matter of maintaining stability. As noted in an open letter to ICANN [45], the mechanical capability to delegate a vast quantity of new gTLDs exists, but using this facility could undermine the stability of the DNS ecosystem. A response to [45] was sent from the U.S. Department of Commerce's National Telecommunications and Information Administration (NTIA) [38], in which the NTIA questioned the distinction between the current *ability* to expedite delegations with the *advisability* of such an action, even though multiple organizations have issued specific advice around this distinction for quite some time. In fact, in 2005, the National Research Council released a report with findings from a study called, "Signposts in Cyberspace," [48] whose goal was to extensively examine a number of issues surrounding the DNS' global ecosystem, including its stability and growth. Among the findings of this report was advice on cautious growth of gTLDs:

> ***Considering technical and operational performance alone, the addition of tens of gTLDs per year for several years poses minimal risk to the stability of the root.*** *However, an abrupt increase (significantly beyond this rate) ... could have technical, operational, economic, and service consequences that could affect domain name registrants, registries, registrars, and Internet users.*
>
> ...
>
> ***If new gTLDs are added, they should be added on a regular schedule that establishes the maximum number of gTLDs (on the order of tens per year)*** *... Addition of gTLDs should be carried out cautiously and predictably, so that*

> *on the one hand, the stability and reliability of the system can be protected, and on the other hand, those considering acquiring a gTLD can do so with a realistic view of future prospects.*
>
> ***A mechanism to suspend the addition of gTLDs in the event that severe technical or operational problems arise should accompany a schedule of additions. It should explicitly specify who has the authority to suspend additions and under what conditions.***

Much of this advice remains unresolved by ICANN [44, 42], despite reiteration of caution from ICANN's own Security and Stability Advisory Committee (SSAC) [19, 40, 41, 43], as well as industry [16, 53]. In a recent technical report [16], we cataloged unresolved issues in the new gTLD program's roll-out, issues upon which we believed the security, stability, and safe introduction of new gTLDs is predicated.

To augment that work, in this study we evaluate the risks that are about to be transferred onto Internet users by the introduction of as many as one thousand new gTLDs (in the first year, alone). To evaluate the "risk," we propose a novel set of measures that represent actual risks to end users, and illustrate their incidence by measuring operational threat vectors that could be used to orchestrate failures and attacks. We present our candidate quantification in the form of a *Risk Matrix*, and illustrate one possible way to interpret its results. What we find is that while some may claim that the relatively abrupt addition of over one thousand new gTLDs is not a concern, there are quantifiable signs that profound disruptions might occur if the current deployment trajectory is followed. This may be especially true if recommendations that have been made are not *fully* resolved. For example, we investigate issues that include Man in the Middle (MitM) attacks, *internal Top Level Domain (iTLD)* collisions with applied for gTLD strings, X.509 certificate ambiguities, and regional affinities that could result in collateral damage to unsuspecting regions. Indeed, our measurements suggest that there may be measurable dependencies for undelegated gTLD strings of `.accountant` in the U.S. Virgin Islands, `.medical` in Japan, `hotel` in Rwanda, and `.corp` across many topologically distributed Autonomous Systems (ASes)[1] in the Internet. We also find evidence that there may exist a dependency between a popular Small Office / Home Office (SOHO) router vendor's SIP boxes and the applied-for gTLD string `box`. What's more, with the intention for some applied-for gTLD strings, such as `secure`, to function as "'secure neighborhoods' on the Net" [39] , our risk matrix suggests that their semantic meaning opens them up to risk factors from current traffic that other, lower profile strings don't start off with. For illustrative purposes, throughout this paper, we consider what specific risk factors can be measured to show that delegations under applied-for gTLD strings like `secure` represent demonstrably riskier profiles than they would have under a gTLD that exists today.

The remainder of this paper is structured as follows: Section 2 describes some of the (potentially surprising) ways in which Internet-based systems interact with the DNS today. Next, in Section 3, we describe the general measures we use to quantify "risk." With that, we discuss our measurements in Section 4, and use those results to motivate our Security Analysis in Section 5. This frames some recommendations in Section 6, before we conclude in Section 7.

2 How We Use the DNS

Part of the DNS' central role in our online lives is that its intricacies and the complex ways that we use it can cause it to slip from the forefront of our attention. We may take for granted that resolving a resource (such as a mail server's service address) for `company.example.com` may require multiple round trips to global resources in order to locate the name servers of several private companies before an email transaction can even be initiated with the mail server itself. As an Internet-scale federated multi-administered database, the DNS is one of a kind. Issues arise, such as *transitive trust* [50, 55] (where resolving a simple DNS domain name could involve DNS resolution of hundreds of *other* DNS domain names), to complications stemming from deploying DNS Security Extensions (DNSSEC) [20, 51], to issues in managing the `DNSKEY` Resource Record set (RRset) [52, 49], and more.

Growth Has Been Slow: It should come as little surprise that modifications to the DNS (whether they be its protocol, its name space, or even the service locations of its root name servers) have always been done at a careful pace. There is a multitude of advice from experts that advocate this conservative approach, including the aforementioned 2005 Nation Research Council report [48] and the subsequent Scaling the Root report [19], in 2009. Indeed, Figure 1 shows some of the relative growth that the DNS root has undergone since late 1999. Note the relatively flat line of TLD growth (representing slow growth in the number of TLDs), especially relative to the larger growth of various Resource Record (RR) types. As we reported in some of our previous work [17], the growth rate of the root zone has been fairly modest over the past 14 years, adding only 66 new TLDs since 1999 (45 of which are internationalized domain names, or IDNs). As part of the Root Zone Maintainer (RZM) role, Verisign maintains the authoritative

[1] Including multiple constituent enterprises, in cases where ASes aggregate or provide connectivity for multiple clients.

TLD	PPM
XXX	4018
ASIA	2708
KP	2588
AX	2369
TEL	1593
UM	836
CW	543
POST	388
SX	331

Table 1: This Table shows a measurement of traffic seen for TLD strings in January 2006. The measurement of traffic is in Parts Per Million (PPM), and that can be interpreted as "the number of queries seen for a given string for every 1,000,000 total queries." One PPM is essentially a very small fraction of a percent. For example, `com` generally gets about 200,000 PPM to the roots.

database containing the root zone data for distribution to all the root servers. Figure 2 illustrates changes in the root zone over the past 61 months. Between June of 2008 and June of 2013 there were 1,446 total changes (about 0.8 changes/day, on average), adding only 37 net new TLDs.

To build on these measurements, we can examine query volumes for some of the TLDs that were added to the DNS root *before* and then *after* they were actually delegated, and try to assess any relative impact. Specifically, Table 1 enumerates traffic volumes for several TLDs (some country code TLDs, ccTLDs, and some gTLDs) before they were delegated from the DNS root. Suffice it to say, after these TLDs were delegated, there was relatively little collateral damage reported to systems throughout the Internet. To contrast this, using data collected in Day In The Life (DITL) of the Internet data [24], some of the query volumes for currently applied-for gTLD strings are shown in Table 2. One can see that many of the currently applied-for strings actually have lower traffic volumes than those in Table 1. This could indicate that there is nothing to worry about when adding new TLDs, because there was no global failure of DNS when this was done before. Alternately, one might conclude that traffic volumes are not the only indicator of risk, and the *semantic meaning* of strings might also play a role. We posit that in some cases, those strings with semantic meanings, and which are in common use (such as in speech, writing, etc.) pose a greater risk for naming collision. In fact, what we will show is that the semantic meanings of strings appear to play a large role in how they are used, and there is evidence that suggests that the traffic volume is not the only indicator of risk.

New gTLD	PPM
HOME	27855
CORP	4085
ICE	511
GLOBAL	355
MED	341
SITE	299
ADS	297
NETWORK	260
GROUP	249
CISCO	238
BOX	222
PROD	187
IINET	167
MAIL	162
DEV	154
HSBC	149

Table 2: This Table shows a measurement of traffic currently being seen for newly applied-for gTLD strings, again in PPM.

Dependence on Ossification: Indeed, while the relative stability, and cautious growth, of the DNS root zone has helped stewards safeguard its stability and operational security, it is not always clearly recognized that this *ossification* has had other effects as well. Specifically, there are ways in which this slow change has resulted in a form of inflexibility. For example, some RFCs [30, 26] expect that certain strings will not be valid DNS TLDs, and suggest that administrators can configure them as iTLDs in their networks. In addition, some system administration manuals [3, 1, 5, 54] also suggest that Local Area Networks (LANs) should configure iTLDs as local DNS TLDs for strings that do not exist in the DNS. Some examples include `corp`, and `dev` (both applied-for strings), and even `novell`. The intuition behind this advice is that locally scoped business-centric domains (like `dev`, `corp`, `mail`, etc.) are all user-friendly mnemonic labels that are easier to reference, and the implication of the advice could be read as they will never exist in the DNS. Moreover, some advice to use these strings as iTLDs is intended to help DNS resolution to continue for internal systems in the event of a disruption of online connectivity. That is, Fully-Qualified Domain Names (FQDNs), tethered to the availability of the global DNS, may be problematic if network connectivity issues exist at a given site or location, or if complete operational autonomy is desired. The intuition for this advice is: If a network or organization becomes partitioned from the Internet, using iTLDs can help preserve networked business operations; else alternative workarounds may need to be considered.

However, many of those iTLD strings are now applied-

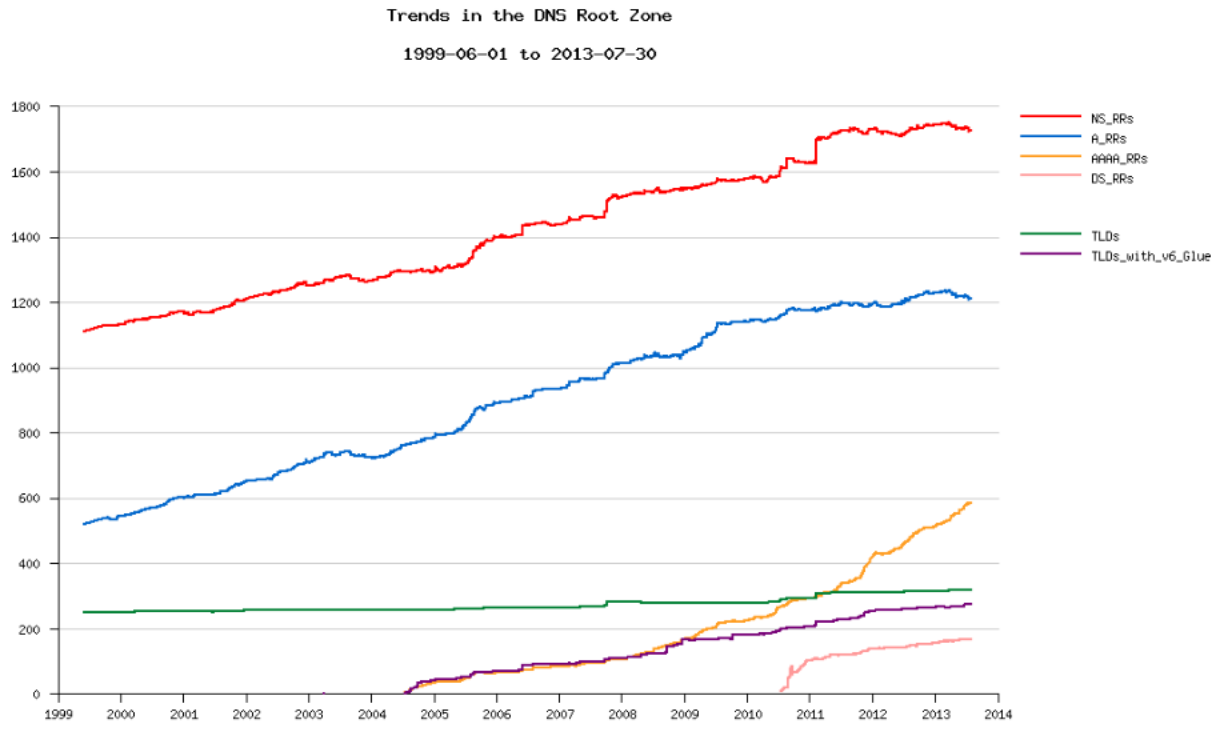

Figure 1: This Figure uses DNS-OARC data to plot the growth of various DNS RR types in the root zone, over time. The green line represents the very modest growth of new TLDs being added over the past 15 years.

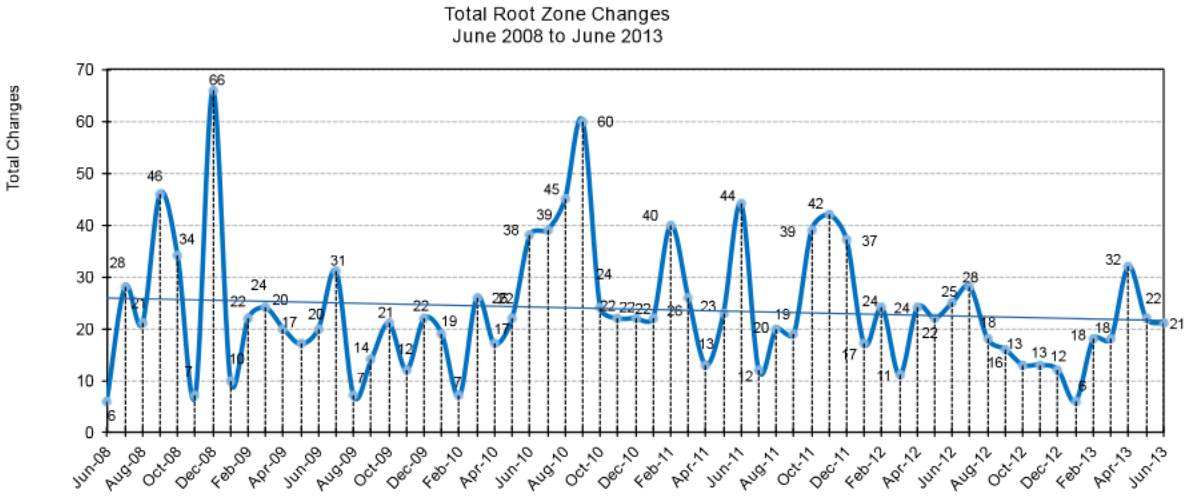

Figure 2: This Figure illustrates the total root zone changes from June 2008 to June 2013.

for gTLDs. This could mean that corporations that are *already depending* on these iTLDs *not* being globally resolvable (as gTLDs) may suffer resolution failures for some of their internal services. Clearly, though, one might ask whether internal DNS queries should even be seen outside a corporation's network. That is, if there is an iTLD `corp`, why would one expect queries to be sent to the global DNS root? The answer is a nuance that even many seasoned system administrators are surprised by (even though it has been a standard behavior of DNS for some time): DNS resolvers may apply search list processing and try the global DNS root first! RFC 1535 [35] outlines this problem, and a recent Internet-Draft [46] illustrated (in its Section 2.1) that DNS search list interactions result in queries being sent to the DNS root *before* being resolved internally.

In fact, this issue was so fundamental to the way DNS works, that ICANN's SSAC provided this advice in SAC045 [40] in 2010:

> ***Recommendation (1):*** *The SSAC recommends that ICANN promote a general awareness of the potential problems that may occur when a query for a TLD string that has historically resulted in a negative response begins to resolve to a new TLD. Specifically, ICANN should:*
>
> - *Study invalid TLD query data at the root level of the DNS and contact hardware and software vendors to fix any programming errors that might have resulted in those invalid TLD queries. ...*
> - *Contact organizations that are associated with strings that are frequently queried at the root. Forewarn organizations who send many invalid queries for TLDs that are about to become valid, ...*
> - *Educate users so that, eventually, private networks and individual hosts do not attempt to resolve local names via the root system of the public DNS.*
>
> ***Recommendation (2):*** *The SSAC recommends that ICANN consider the following in the context of the new gTLD program.*
>
> - *Prohibit the delegation of certain TLD strings. RFC 2606, 'Reserved Top Level Domain Names," currently prohibits a list of strings, including test, example, invalid, and localhost. ICANN should coordinate with the community to identify a more complete set of principles than the amount of traffic observed at the root as invalid queries as the basis for prohibiting the delegation of additional strings to those already identified in RFC 2606.*
> - *Alert the applicant during the string evaluation process about the pre-existence of invalid TLD queries to the applicant s string. ...*
> - *Define circumstances where a previously delegated string may be re-used, or prohibit the practice.*

AFRINIC	IANA-SERVERS	NRO
ALAC	ICANN	RFC-EDITOR
APNIC	IESG	RIPE
ARIN	IETF	ROOT-SERVERS
ASO	INTERNIC	RSSAC
CCNSO	INVALID	SSAC
EXAMPLE*	IRTF	TEST*
GAC	ISTF	TLD
GNSO	LACNIC	WHOIS
GTLD-SERVERS	LOCAL	WWW
IAB	LOCALHOST	
IANA	NIC	

Table 3: *Note that in addition to the above strings, ICANN will reserve translations of the terms "test" and "example" in multiple languages. The remainder of the strings are reserved only in the form included above.

These recommendations were meant to protect two sets of stakeholders. The first and most obvious within the ICANN community was the new gTLD applicants, those who would be associated in some manner with the operations of registry infrastructure supporting new gTLDs. In response to these recommendations, ICANN did reserve a number of strings. Table 3 is taken from Section 2.2.1.2.1 Reserved Names of ICANN's Applicant Guidebook [10], and represents the strings that are prohibited as of June 2012. Table 4 enumerates the amount of query traffic seen for each of these reserved gTLD strings (in PPM). When contrasted with the query rates seen in Tables 1 and 2, this Table suggests that the traffic volume to these reserved strings is relatively negligible. Of note is the fact that the list of reserved names in RFC 2606 [30]: `test`, `example`, `invalid`, and `localhost` (updated by RFC 6761 [27]) all see a reasonably large number of queries at the root, and were included in Table 3. More importantly, while there is no discernible risk-based metric for inclusion in the current reserved names table, there is an abundance of ICANN-associated entities, to which our measurements suggest either very low or no discernible risk exists. Yet, in contrast, there is an obvious absence of potentially problematic strings, such as those discussed in SAC045 [40], and in Appendix G Private DNS Namespaces of RFC 6762 [26]. Furthermore, there seems to be no indication that any of these strings were added based on measurements, as recommended.

PPM	TLD	Reserved By
66963.9	LOCAL	IETF
12023.7	LOCALHOST	IETF
1740.6	INVALID	IETF
432.7	TLD	IETF
137.7	TEST	IETF
45.7	WWW	IETF
30.2	EXAMPLE	IETF
10.1	NIC	ICANN
5.0	GAC	ICANN
1.8	NRO	ICANN
0.7	ASO	ICANN
0.2	WHOIS	ICANN
0.2	IAB	ICANN
0.1	IANA	ICANN
0.0	RIPE	ICANN
0.0	ARIN	ICANN
0.0	ROOT-SERVERS	ICANN
0.0	IESG	ICANN
0.0	IETF	ICANN
0.0	ALAC	ICANN
0.0	SSAC	ICANN
0.0	APNIC	ICANN
0.0	ICANN	ICANN
0.0	GTLD-SERVERS	ICANN
0.0	INTERNIC	ICANN
0.0	GNSO	ICANN
0.0	IRTF	ICANN
0.0	RFC-EDITOR	ICANN
0.0	ISTF	ICANN
0.0	LACNIC	ICANN
0.0	AFRINIC	ICANN

Table 4: This Table shows the amount of traffic (in PPM) for each of the reserved gTLD strings, and which organization reserved the string (the Internet Engineering Task Force, IETF, or ICANN).

Scoping: An additional way in which new gTLD strings can be problematic is in software that has made *implicit* assumptions about which strings are valid TLDs, and the authority structure of the DNS. Consider when a Web browser receives a cookie from a website, such as `www example com`, and then visits `subzone example com`. The browser will protect the scope of the cookies, the X.509 certificates that can be used, etc. This protection is implemented through a global list (maintained at `http //publicsuffix org/`) that details the administrative boundaries of the DNS. It allows Web browsers to determine where various administrative boundaries exist, and discusses issues like "Super Cookies" (described below). Currently, browsers (and other relying party software) download this static file (often at compile time) and then "bake it into" their code. As the DNS delegation structure evolves (administrators subdivide their zones, aggregate their zones, transfer their zones, etc.), the maintainers of this list must struggle to keep its contents current with the state of the global DNS delegation structure. On top of that, as browsers and other software age, their compiled versions of the list becomes more out of date.[2] A number of issues have been reported as a result of this, and we discuss these more in Section 3.1. The suffix list is also used to protect cookies shared between different hosts by not allowing Super Cookies to be set for high level domains, such that cookies can be valid for `example com` but not for all `com` in general.

Now consider, for example, the new gTLD string `secure` [21], which has been called a "'safe neighborhood' on the Net" [39]. The plan for this new gTLD is to offer a branch of the DNS that requires its registrants to have a pronounced security posture through deploying enhanced security precautions, being subject to security scans, and more; all in the vein of conveying greater faith in the security of the domains under this gTLD to end users. Now, suppose a user connects to `www` ⟨`someSite`⟩ `secure`, and then to ⟨`partner-association`⟩ ⟨`someSite`⟩ `secure`. Here, the various parties involved are *all* implicitly trusting that browsers will not allow separate organizations to share cookie information or other cross-administrative data. Moreover, the same infrastructure must ensure that when an HTTPS connection is made to ⟨`partner-association`⟩ ⟨`someSite`⟩ `secure` that any X.509 certificate that may already exist (or even those that might be issued in the future [43]) for the `secure` string (an Internal Name Certificate) cannot be used (similar to a Super Cookie) to impersonate any actual Internet property below `secure`. In such a case, a MitM attack could be successfully launched.

The qualitative liabilities could be seen as a general caution, but rather than debating the possible degree of exposure, we have created a measurement-based approach to codifying one possible example of "risk" to end users and networked systems, which we motivate in Section 3.

3 Gauging the "Risk" Level for New gTLD Strings

Previous studies of the DNS root have noted large amounts of invalid queries [23, 60]. While many queries may not result in positive answers, we contend that this does not necessarily mean they are "invalid." Specifically, we have found indications that many of these queries are likely valid, in some way. For example, when a user clicks on a link that points to a domain name under an applied-for gTLD string (either mistakenly, or because the domain name is meant to resolve internally), the resulting query is "valid" and can pose a direct risk to that user. To illustrate the seriousness of the risks posed to the end user, we begin by detailing a few illus-

[2] A partial list of software that uses this suffix list is maintained at `http://publicsuffix.org/learn/` .

trative examples.

3.1 The Past is Prologue

Constructing hypothetical risks and attacks is a common practice among operational security professionals. However, some have noted that this can be an unbounded exercise that reaches a point of diminishing returns. We, therefore, outline several instances of similar and analogous cases here, and observe that the type of behavior in these examples would likely get *easier* with the inflation of the number delegated gTLDs.

To illustrate what can happen when a namespace that is assumed to be non-delegated goes live, we examine an incident that happened with Apple's iTunes. On September 30^{th}, 2012 Apple released iTunes 10.7 and immediately users started reporting activity on an abnormal domain: `bogusapple com` [11]. In essence, the new version of iTunes began issuing queries to a domain that was *expected* to never exist: `bogusapple com`. Upon seeing this, one person registered that domain and began intercepting traffic, and capturing private information.

Another opportune example was raised at the Security and Stability Session of ICANN 46 [13]. In the transcripts, and as provided in the audio, a participant detailed their experiences of *running* the domain `corp com`. Among other things, this person explained that there was a sustained query load for DNS traffic. At one point, this administrator began servicing email requests and was able to see undisclosed Securities and Exchange Commission (SEC) filings for corporations before they were officially released. This serves as strong caution against the assumption that simply counting query rates is sufficient to measure all aspects of risk.

As an example of iTLD conflicts, [9] reported that Chrome can have difficulty identifying whether the entered text is a domain name or a search term. Problems with an out of date suffix list led to issues where certain TLDs became difficult to access using Chrome. Additionally, [12] noted that the Safari browser was not immune either. In addition, [4, 6] detailed issues with cookie scoping.

3.2 Risks

In order to estimate how much concern might be warranted, we propose a candidate measure to analyze how much *risk* each applied-for gTLD string represents to *Internet users*. To do this, we examine the following set of tangible threats that already exists, and we measure their prevalence on the Internet, today: i) Information leakage and user tracking, ii) Denial of Service (DoS), and iii) MitM attacks. This set of risks was chosen because it covers a range of different concerns to Internet users. While we strive to fully quantify our notion of risk, we acknowledge that this is just one candidate approach to quantifying this general concept, and others may choose alternate approaches, or enhance the approach we have taken with a more comprehensive set of threats considered. In order to measure these risks, we identified specific attack vectors that adversaries could use to orchestrate each of these into actual attacks.

Information Leakage: One result of future delegations of new gTLD strings would be that the private parties that will be servicing the new gTLD strings (*Registry Operators, ROs*) will be *implicitly* observing (and potentially recording) information about DNS queries. Currently, these queries go only to the Root Server Operators (RSOs). While this is still information leakage, the set of observers is poised to grow dramatically (from the current 12 organizations to hundreds), and the restrictions on how the *new* ROs are allowed the use measurements are different than today's RSOs. Moreover, once delegated, the *registrants* under new gTLDs have the ability to register specific domains for targeted collisions. While there are more nuanced differences between the specific attacks that new ROs and registrants can launch, they are beyond the scope of this writing. This form of information leakage can violate privacy of users, provide a competitive advantage between business rivals, expose details of corporate network infrastructures, or even be used to infer details about geographical locations of network assets or users [37]. Another interesting note is that if enterprises follow iTLD provisioning guidance (as discussed above in Section 2), services with naming schemes, like: ⟨`service`⟩ ⟨`location`⟩ `foo-enterprise corp` expose network and business structure to DNS operators. So, an organization that acquires the operation of the new `corp` gTLD could potentially use its collision with *every* company's `corp` iTLD, and (in this example) be able to enumerate the numbers, types, and locations of Foo-Enterprise's internal structure. There is also evidence of similar issues in Novell configurations [8]. Beyond monitoring NXDomain (or `rcode 3`) traffic, new ROs might elect to take a more active role and begin providing positive responses to queries.

Denial of Service: If a Registry Operator (RO) or a domain registrant within that gTLD elected to begin responding to these queries with actual service identifiers (such as IP addresses), this could likely cause the querier to attempt to establish a connection (such as to an IPv4 address, an IPv6 address, mail servers, etc.). Under these conditions, an operator could either DoS the intended service by influencing attempts to resolve a resource's name, or possibly launch a MitM attack and potentially siphon information (such as user credentials, passwords, etc.) from sessions associated with the do-

main name (by returning DNS responses that contain invalid mappings). This concern was recently expressed in a letter from PayPal to ICANN [18]. We note, this predated the disclosure of issues with 'Internal Names Certificates" [43], which we discuss below and which themselves enable even more virulent and stealthy versions of these attacks.

The DoS vector could be intentional, but also inadvertent. If queries that are being issued for any of these applied-for gTLD strings are being serviced by regional or otherwise non-global systems, then any active responses from a newly delegated gTLD could interfere. One of the findings in Section 4 is that some applied-for gTLD strings have a statistically pronounced regional bias. That is, some strings that are seeing query traffic today are heavily queried from specific regions, and this could mean that delegation of those strings would have acute regional effects (even if the global effect seems muted). For example, Estonia shows a very pronounced affinity for the applied-for gTLD string `zone`. Even servicing these requests from a new gTLD (instead of the NXDomain, or `rcode 3`, responses they currently elicit from the root) could disrupt the usage that they may currently have in their regions. However, the threat exists that a RO could also begin answering them, causing a MitM attack.

Man in the Middle: Section 3.1 discusses specific examples of non-existent domain names being registered and then used to launch MitM attacks against domain names that already exist, albeit at much smaller scales than a gTLD. While we believe that these existence proofs actually illustrate a lower bound on the level of concern that is warranted, we leave this judgment to the reader. One common defense against MitM attacks is to use Transport Layer Security (TLS) [29] because it uses end-to-end encryption to help protect sessions from such attacks. Generally, TLS sessions rely on external certificate verification schemes (like an internal list of "trusted" Certification Authorities) to bootstrap authenticity of endpoints during start-up. However, the planned introduction of the new gTLDs has opened even TLS' assurances up to attack as well. With new gTLDs, users may expect that any MitM would be unable to spoof connections, because when a user connects to a TLS service at a domain name, their expectation is that the certificate returned will be checked and its name (either the *CommonName (CN)* or *Subject Alternative Name*) will match that of the DNS domain name being used for the connection. However a recent result (titled, 'The Most Dangerous Code in the World: Validating SSL Certificates in Non-Browser Software") has shown that this verification step is often incorrectly performed or even omitted [36]. One example, cited in Section 9 of this paper, is a code excerpt from PayPal's online shopping cart that left users exposed to exactly this type of vulnerability, when using PayPal services. However, the introduction of new gTLDs may render these checks ineffective anyway, even when checked. A recent report by ICANN's SSAC detailed the dangers posed by "Internal Name Certificates" [43]. This report advised that Certification Authorities (CAs) have had a long-standing practice of issuing certificates for domain names under iTLDs that are not currently delegated gTLDs. The implication of this is that *anyone* can obtain certificates for names that correspond to new gTLD strings. These certificates will have been issued by actual CAs, and pass all TLS verification checks, and *must* be considered a threat not simply until they are revoked, but until they expire [15]. Thus, any TLS connection to any domain name under a new gTLD can be properly established using a certificate that can be easily obtained by anyone.[3]

The dangers posed by this issue include the fact that it would allow an adversary to register domains under new gTLDs and intercept *existing* traffic from unsuspecting users. For example, if a company has provisioned their payroll system on a machine called `foo`,[4] and placed it under their `secure` iTLD, then their internal network would most likely be rife with DNS queries for the name `foo secure`. However, these queries *necessarily* will be issued to the *global* DNS root zone before being serviced internally. Ironically, as a result, the operator of `foo secure` will be ideally positioned to intercept these queries and use their Internet Name Certificate to create *insecure* TLS (or even HTTPS) connections.

In addition to SAC057 [43], without the explicit scoping of authority codified by the rules published on PublicSuffix.org, any internal named certificate under a new gTLD could be applied to arbitrary subdomains throughout that entire branch of the DNS. That is, if a certificate for `* com` were to somehow be minted (which would violate existing CA policies), and if it were *presented* to a web browser, that browser would have a mechanism (offered by PublicSuffix.org) to know to reject it and restrict the information leaks from cookie sharing between unaffiliated zones. This would not be the case with `* secure`, and without rules in PublicSuffix.org, such an Internal Name Certificate would have security scope over *all* Second Level Domains under `secure`.[5] What is, perhaps, more troubling is that Internal Name Certificates for well chosen SLDs (and wildcards below *them*) can certainly be issued throughout the hierarchy of applied for strings. So, for example, names like `* foo secure`, `payroll secure`, `www secure`, `_kerberos _tcp secure`, etc. can all be requested and issued. With well chosen seeds, a pre-

[3] An example of this is illustrated in SAC057.

[4] The label `foo` is just an example, but could just as easily be `payroll`, `Foo-Corp`, `sap`, etc.

[5] It is ironic that this could inherently reduce the innate security of strings like `.secure`.

emptive dictionary attack against `secure` could scorch the earth beneath the entire branch, and (perhaps ironically) render this new "Safe neighborhood," uninhabitable on day one. By contrast, as we discussed in Section 2, today's DNS authority and delegation structure is loosely codified in a suffix list. While this list may be prone to errors and staleness, it may also be viewed as providing some protection. For example, we cannot know how quickly and accurately new gTLDs will be incorporated into that list, or how fast their subzones will be incorporated, or how well the delegation boundaries will be represented, or how quickly *end user* software will pick the new list up.

3.3 Threat Vectors

In order to understand some candidate vectors through which risks might become active threats to users, we examine a few specific instances of online behavior (which are evident in measurements). There are multiple instances of tools and services that attempt to 'help" users overcome connectivity problems through DNS-based discovery. For example, the Web Proxy Autodiscovery Protocol (WPAD) [34] is a technology that attempts to help users automatically discover if their network requires them to configure a Web proxy. Before fetching its first page, a Web browser implementing this method sends the local DHCP server a `DHCPINFORM` query, and uses the URL from the WPAD option in the server's reply. If the DHCP server does not provide the desired information, DNS is used. If, for example, the network name of the user's computer is `pc department branch example com`, the browser will try the following URLs in turn until it finds a proxy configuration file within the domain of the client:

`http://wpad.department.branch.example.com/wpad.dat`
`http://wpad.branch.example.com/wpad.dat`
`http://wpad.example.com/wpad.dat`

While not an Internet standard, we will show evidence in our measurements (in Section 4) that suggests it is indeed in wide use. The danger here is that, while these queries meet with NXDomain responses now, if a new gTLD operator (or a registrant operator under a new gTLD) were to start *answering* these queries with Web proxy information, then that operator could instruct any browser (or any type of WPAD client) to proxy *all future WWW traffic* through the specified proxy.

In addition, the Intra-Site Automatic Tunnel Addressing Protocol (ISATAP) [57] is an automatic transition tunneling technique that discovers endpoints using DNS and may create IP tunnels based on what it finds. ISATAP is meant to transmit IPv6 packets between dual-stack nodes on top of an IPv4 network. The system works by requiring system administrators to maintain a Potential Router List (PRLs) of IPv4 addresses representing ISATAP interfaces of routers. This list is commonly maintained as a FQDN and ISATAP typically builds its PRLs via consulting the DNS and looking up DNS names like `isatap example com`. The danger in this case is similar to the WPAD case: If a new gTLD operator (or registrant operator) were to answer ISATAP queries with PRLs, then clients could begin tunneling *all of their traffic* through the specified routers.

Finally, DNS-Based Service Discovery (DNS-SD) [25] attempts to locate services by using DNS queries. DNS-SD is a protocol that enables network browsing and service discovery using only standard DNS packets and record types (RFC 6763). DNS requests will take the FQDN form of _⟨`Service`⟩ _⟨`Protocol`⟩ _⟨`DnsDomainName`⟩. Some actual examples we observed include:

`_kerberos._tcp.dc._msdcs.HNAH.ADROOT.HSBC.`
`_ldap._tcp.SCAZTM01._sites.dc._msdcs.ent.wfb.bank.corp.`
`_kerberos._tcp.dc._msdcs.sap.corp.`

Of note, LDAP is Lightweight Directory Access Protocol that enables distributed directory access services such as 'single sign-on" where one password for a user is shared across many services. Kerberos is a computer network authentication protocol. However, while these types of services could be considered alarming, in the context of the new gTLDs, some queries are actually specifically configured to query non-existent gTLDs.

Implicit in our discussion of these risks and threat vectors is a need to quantify them. For this, we next examine captures of data from several sources.

4 Measurements

Because our notion of risk includes threats that are beyond just DNS queries and their responses, our measurements cover more than just DNS queries and responses at the DNS root. We, necessarily, included measurements of the World Wide Web, X.509 certificates, and regional trends across all of these measurement modes. In this Section we discuss our measurement apparatuses, and then we broadly break our analysis down into two dimensions: *Spread* and *Impact*. The intuition behind this is to illustrate how broadly measured effects are seen, and to what extent they appear to be having effects.

4.1 Measurement Apparatuses

NXDomain (NXD) Analysis: For our NXDomain (or `rcode 3`) analysis we used a combination of data sets from historical *Day in the Life (DITL) of the Internet* collections [24] and separate traffic captures from the `a` and `j` DNS root servers which Verisign operates. The locations of the root server instances of `a` and `j` are available on the website `http //www root-servers org/`. Our NXDomain analysis allowed us to identify query

patterns, user behavior, and detect some degree of systemic trends. By contrast, our crawl of the Web (using the Internet Profile Service, IPS) allowed us to measure some precursors to DNS queries, and provided additional evidence of X.509 certificate usage.

Internet Profile Service (IPS): The Verisign Internet Profile Service is a platform that is used primarily internally by Verisign to study the health of the overall domain industry. This project crawls a small amount of content from every domain that permits it to do so, and analyzes request traffic from the DNS resolution process. The IPS corpus includes roughly 700 million Web pages.

The crawl process affords Verisign with the opportunity to build detailed statistics about linking relationships between domain names and the certificates that they have installed. The statistics from the resolution process provide insight into which domains are most heavily utilized on the resolution platform and some of the host names that are leveraged beneath the TLD registries that Verisign operates.

SSL Observatory: In addition to the X.509 data gleaned from our IPS crawls, we cross-referenced certificates found with SSL Observatory [31]. While there are some obvious constraints in this data it does provide a lower bound of related certificates for elementary analysis purposes.

4.2 Spread

In this study, we loosely define the term *spread* as representing how widely the effects of queries can be measured. Specifically, the spread of the queries for applied-for gTLD strings can be used to quantify some aspects in which relatively few queries can (with sufficient spread) have large effects, and (alternately) spreading the observation period of measurements out over time can enhance their completeness. For the remainder of the text, when we discuss query measurements, if we do not explicitly mention the source as DNS-OARC DITL data, we implicitly mean the source of measurement is from the `a` and `j` root servers.

Table 5 illustrates the relative percentages of the overall traffic to the DNS root. One of the trends that this data illustrates is that, since at least as far back as 2006, the root system has seen queries for over 90% of the strings that are now being considered for delegation as new gTLDs. In addition, the overall trend is that the query traffic for them is growing. Further, Figure 3 illustrates that one of the longitudinal trends over the last several years is an increasing percentage of the applied-for strings have been seen in queries at the DNS root. Indeed, the most recent DITL collection showed that 98.30% of the currently applied-for strings were queried for within the DITL's 48-hour collection period. Additionally, the lower curve in Figure 3 illustrates the percent of the applied-for strings that were seen year after year, and we can see that in 2013, 96.95% of the applied-for strings were re-observed from previous years. While the historical trend indicates that there has been a longstanding footprint of queries for the current set of applied-for strings, the measurements also suggest that not *all* applied-for strings are immediately visible in 48 hours of query traffic to the DNS root, as provided in the DITL data collection windows throughout this timeframe. Also note that only a single collection of the DITL data (2010) included participation from the full set of root operators. All others were subsets of all root operators.

However, more protracted measurement periods yield broader coverage of the set of applied-for strings. Figure 4 illustrates that almost 6 days of measurements were required from both the `a` and `j` root servers before all applied-for gTLD strings were observed. The queries for these strings were seen from 26,054 distinct Autonomous Systems (ASes). The relative popularity of each of the new gTLD strings (in NXDomain traffic) is plotted as a histogram in Figure 5. This Figure illustrates the heavy-tailed distribution of the query load for applied-for new gTLDs. Figure 6 depicts a cumulative distribution function outlining the number of ASNs requesting an applied-for gTLD string within this collection window. Figure 7 illustrates the applied-for strings with the highest ASN diversity. As RFC6762 suggested, `home` (11,515 ASNs) and `corp` (8,555 ASNs) may in part be the result of private usage of multicast DNS, or general iTLD adoption and use in private networks. However, other measurements (below) suggest that some additional applied-for new gTLD strings may have similar private usage patterns.

In order to begin gauging how richly used any given applied-for gTLD string might be, we investigated the number of unique Second Level Domains (SLDs) that we saw under each applied-for string. Figure 8 illustrates that (by a large margin) `home` has the richest diversity of SLDs. Following this, the breakdown follows a heavy-tailed distribution with `cisco`, `box`, `corp`, and `prod` rounding out the top 5. One possible inference to be drawn from this is that a greater diversity in the namespace of the SLDs may be the result of much more nuanced (and possibly business critical) use by organizations. However, we leave the judgment of this to the reader, as it does not have a direct bearing on our findings. Figure 9 shows the distribution of how many applied-for new gTLD strings appear as links in public Web pages today. This measurement offers evidence of one motivator that users may already be influenced by to direct queries and transactions to proposed new gTLDs. One possible reason for these links could be the intention for the enclosing Web page to direct browsers

When	Valid Queries	NXDomains	Applied-for gTLD NXDomains	Learning Window	Applied-for gTLDs Seen
2006-01-10	60.08%	39.92%	3.70%	48.0hr	90.8%
2007-01-09	67.28%	32.72%	2.55%	46.7hr	91.7%
2008-03-18	72.45%	27.55%	4.80%	47.0hr	93.3%
2009-03-30	72.00%	28.00%	5.52%	71.7hr	94.5%
2010-04-13	72.13%	27.87%	4.62%	47.6hr	96.0%
2011-04-12	65.15%	34.85%	5.22%	49.9hr	96.9%
2012-04-17	59.97%	40.03%	7.24%	47.9hr	97.3%
2013-05-28	56.98%	43.02%	8.99%	47.9hr	98.4%

Table 5: Relative percentages of root system traffic (among DITL participants), percent of all new gTLD strings seen, and amount of time needed to converge on new gTLD set ("Learning Window"). This was measured using DITL data. Note that the Applied-for gTLD NXDomains column represents the percentage of NXDomain traffic.

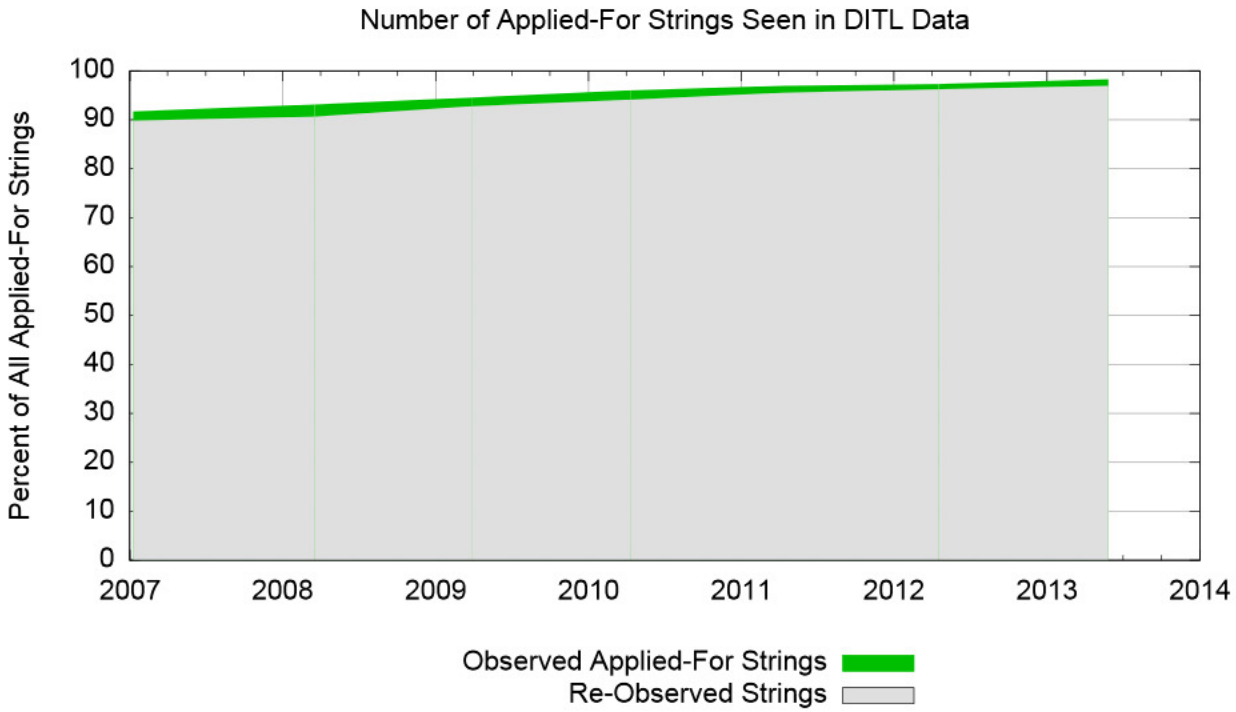

Figure 3: This Figure illustrates the percent of applied-for gTLD strings that were seen in DITL collections, year after year. In addition, the lower curve shows what percent of the seen strings were seen in *previous* years (suggesting more consistent query patterns for strings). Note that this plot begins in 2007 as the 2006 DITL data, while at 90.8% itself, was used for the initial learning set.

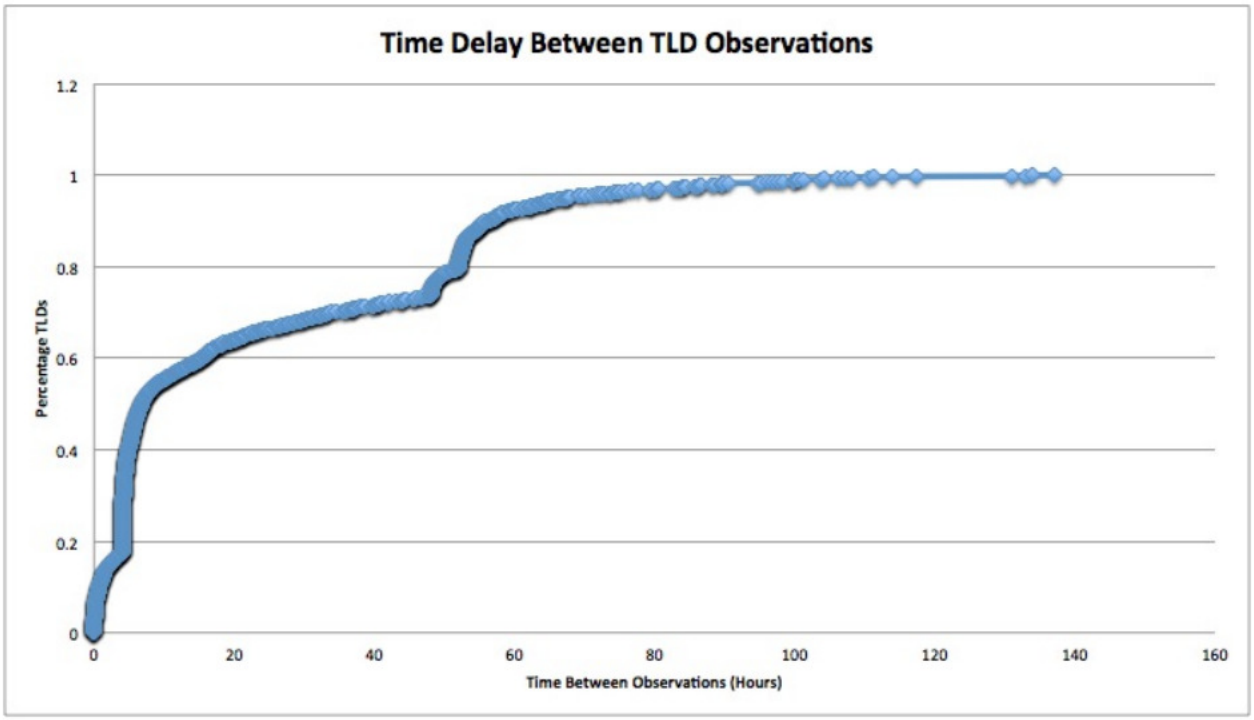

Figure 4: This Figure shows the CDF of the learning rate of new gTLD strings for just the a and j root instances. It shows that it takes just over 5 days to observe all of the new gTLD strings in NXDomain traffic.

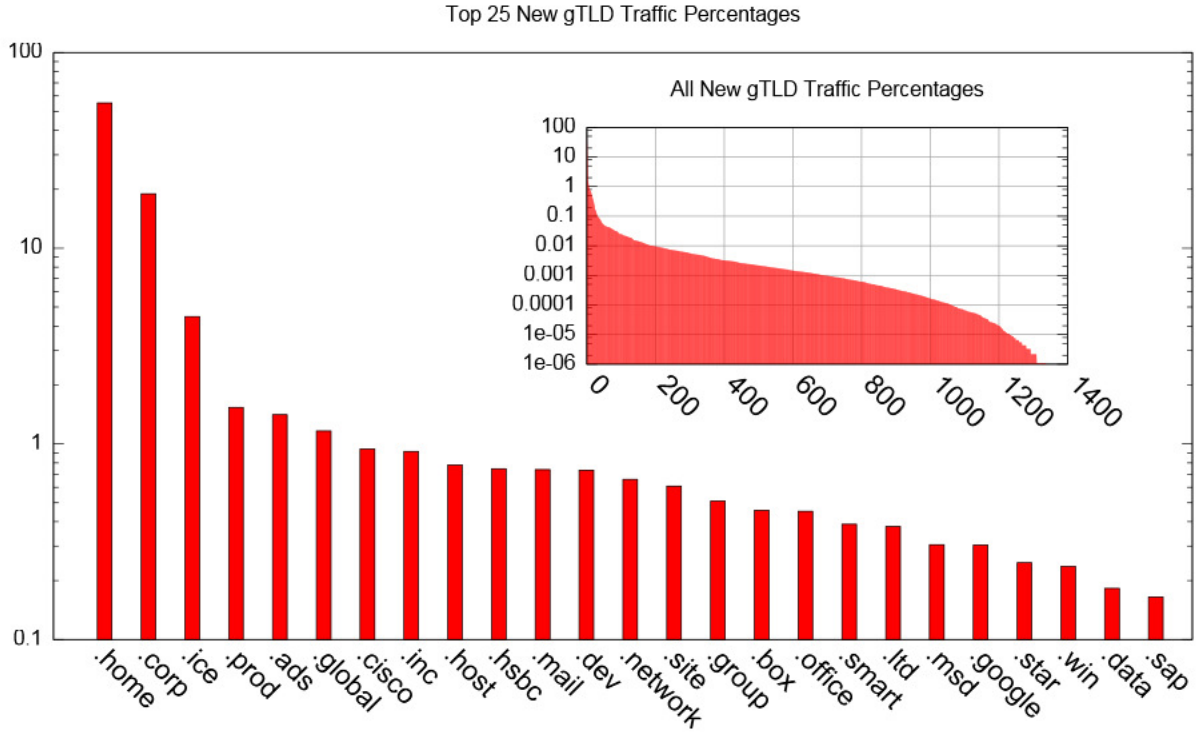

Figure 5: This Figure shows the percentages of all NXDomain traffic that was seen for each of the applied-for gTLD strings (note the logscale).

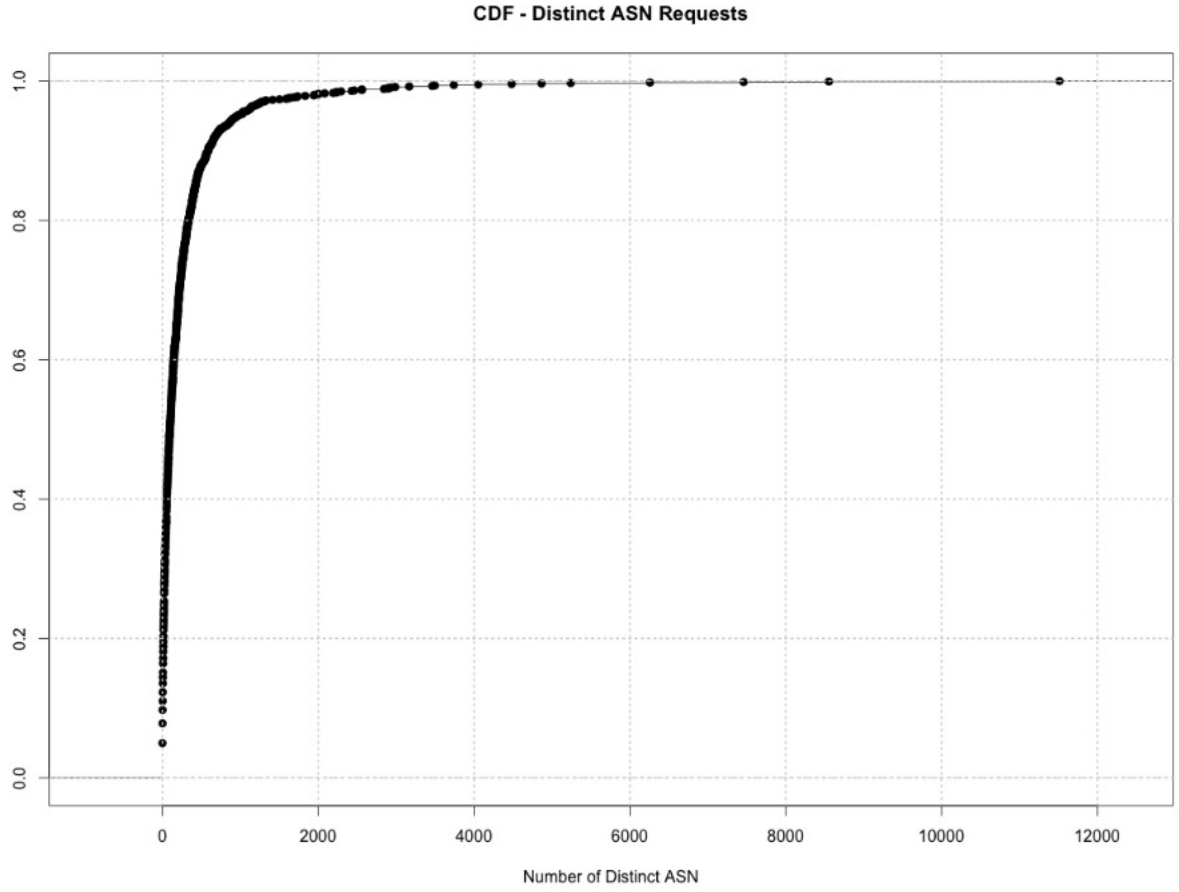

Figure 6: This Figure is a CDF of the number of ASNs that issued queries for each applied-for gTLD string.

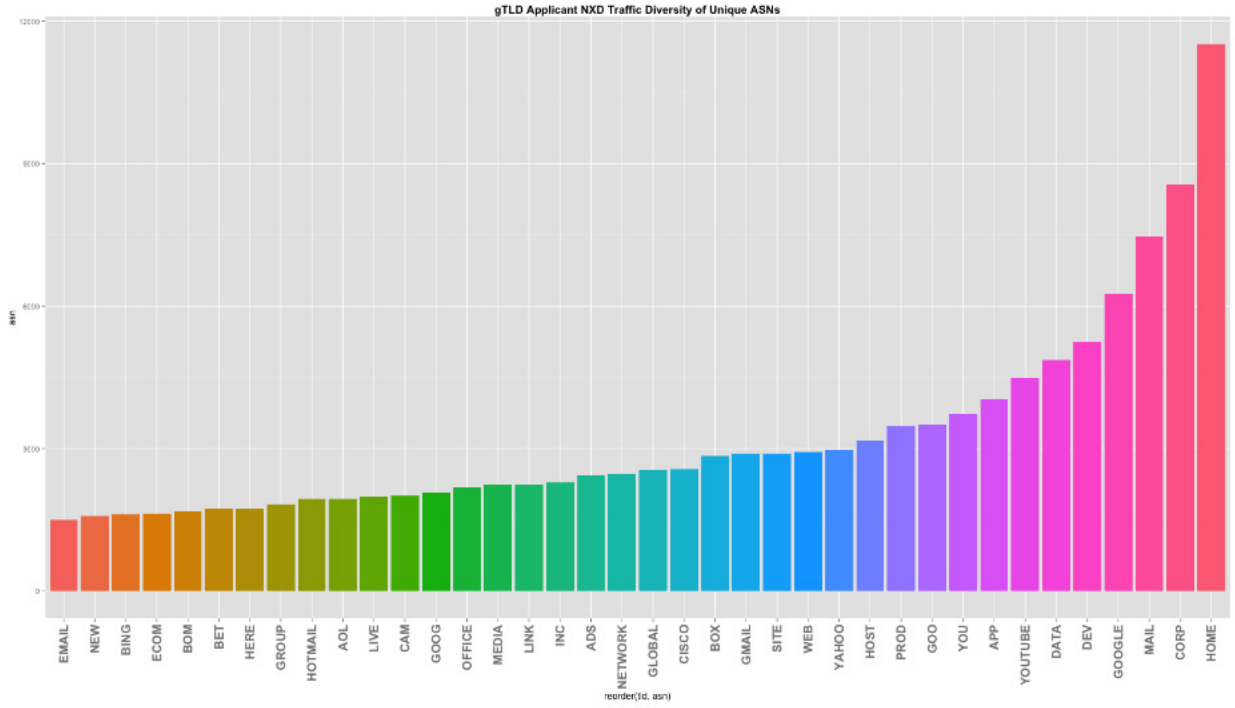

Figure 7: This Figure shows the applied-for gTLD strings with the highest ASN diversity.

to internal sites, another likely explanation for some of them is that they represent typos or information that was not updated when the Web property was moved into a production environment. In any case, these links can be responsible for user traffic, and expose an element of risk.

While the logical converse to wide-spread usage is more narrow usage patterns, this doesn't necessarily mean this type of usage is less important or less critical. Perhaps the opposite is true, in some circumstances. We next consider queries for applied-for new gTLDs that exhibit strong *regional preferences*, and query sources that exhibit marked *periodicity* (i.e., query for applied-for new gTLDs at an abnormally regular interval). The intuition behind these investigations is that new gTLDs that may not be as globally popular as some with broader appeal might actually be very important to certain smaller countries (or regions) and consumers. Our belief is that name conflicts for *those* applied-for gTLDs could have acute impacts on entire regions, without seeming to be as pronounced of a concern in the global query context (i.e., "weak signals").

Regional Preferences: In searching for regional affinities, we develop a candidate metric to determine which regions show disproportionate levels of interest in any of the applied-for gTLD strings. Our metric is just one candidate quantification of this sort of behavior, and others may choose different approaches or enhance our approach. Regardless, our metric offers a concise quantification of this general behavior.

In order to determine if one country (or region) has a pronounced affinity for a given applied-for gTLD string, we begin by mapping query sources to the "region" that they come from, according to ISO 3166 Region Codes [32]. We then establish a baseline affinity for each gTLD for each region across all of the gTLDs that it queries. That is, we determine what each region's "normal" query patterns are for each applied-for gTLD string. Then, we determine if one region has a distinctly different level of interest in any string.

To establish our per-gTLD baseline level of *interest* for a region (i_r^{gTLD}), we first calculate the total number of queries that each region r issues for all new gTLDs Q_r. Table 6 shows some example query counts for some gTLDs, broken out per region. Next, we divide the query count for each new gTLD in region r by the total of all queries seen:

$$i_r^{gTLD} = \frac{q_r^{gTLD}}{Q_r}$$

Where q_r^{gTLD} is the count of queries seen for a given gTLD from a specific region r, and i_r^{gTLD} represents the relative query fraction seen for a specific gTLD in region r. Table 7 illustrates how this normalizes the query rates seen from each region, for each applied-for gTLD.

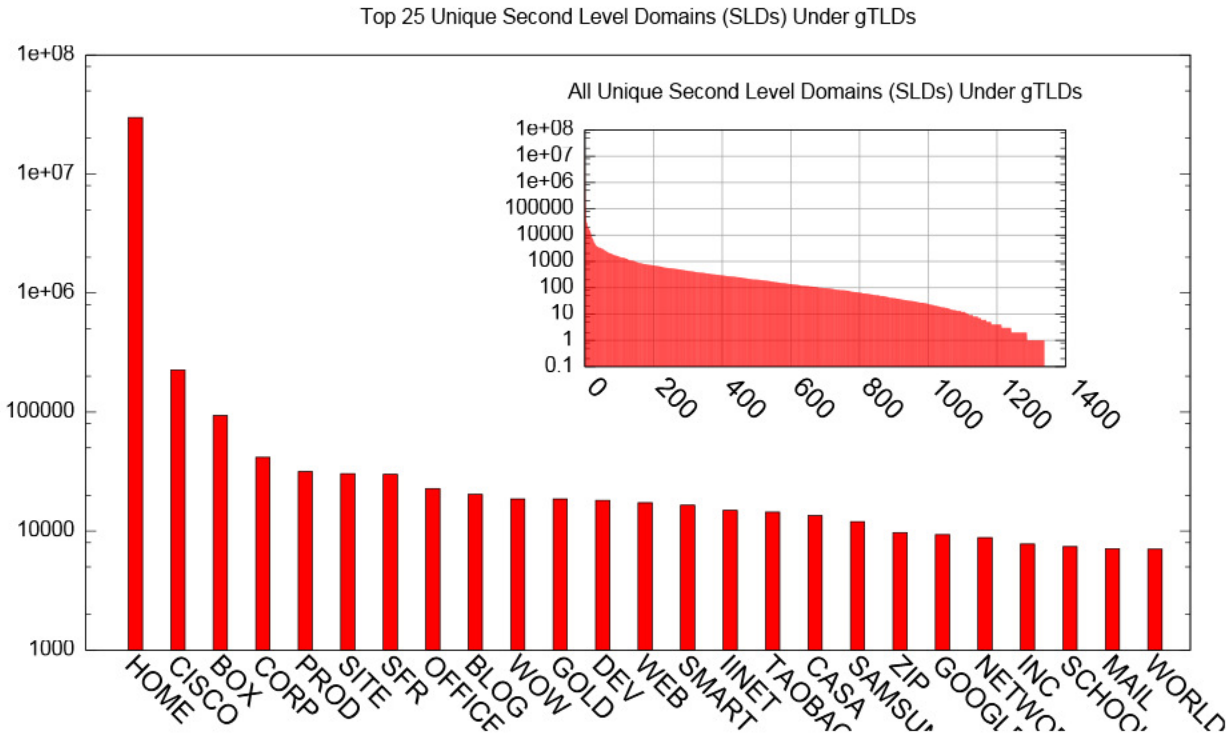

Figure 8: This Figure illustrates (in logscale) the diversity of SLDs under each of the applied-for new gTLD strings. The large plot shows the top 25 applied-for gTLDs, and the smaller plot shows the entire distribution of applied-for strings.

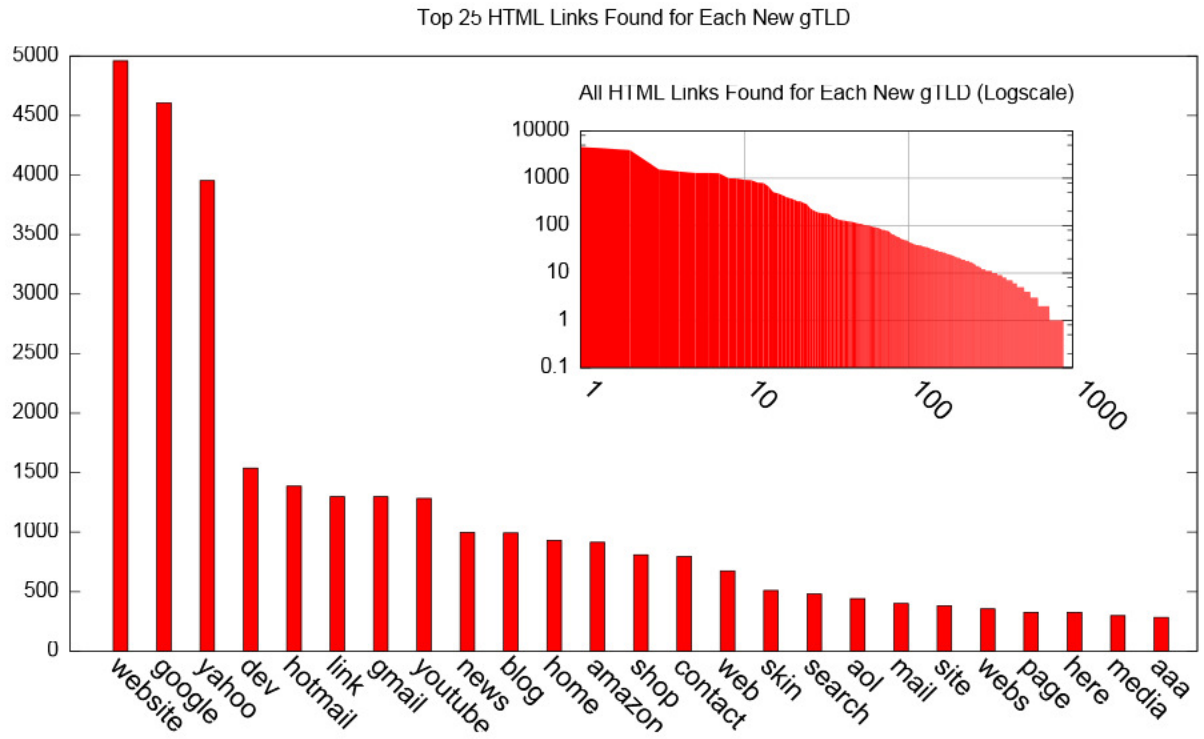

Figure 9: This Figure illustrates the relative counts of new gTLD strings seen in Web pages observed via the IPS Web crawl. The larger Figure illustrates the distribution of the number links (seen in all pages) to the 25 applied-for gTLDs with the greatest link counts, and the smaller (logscale) figure describes the distribution across all applied-for new gTLDs.

gTLD	US	CA	GB	DO	FR
home	14.67M	18.50M	3.05M	3.03M	0.22M
corp	13.76M	0.39M	0.11M	0.03M	0.79M
ice	4.08M	0.00M	0.00M	0.00M	0.00M
prod	1.04M	0.01M	0.00M	0.00M	0.33M

Table 6: This Table shows sample query counts (in units of millions of queries), broken out per region.

gTLD	US	CA	GB	DO	FR
home	0.32	0.95	0.92	0.98	0.10
corp	0.30	0.2	0.3	0.1	0.37
ice	0.9	0.0	0.0	0.0	0.0
prod	0.2	0.0	0.0	0.0	0.16

Table 7: This Table shows sample query counts, as a normalized fraction of all received NXDomain queries, broken out per region.

In calculating the relative fraction of queries sent by each region (and to each applied-for gTLD) we can see differences in query affinities. Our general observation is that for most applied-for gTLDs, across regions, the results are somewhat consistent (regions tend to mirror each others' affinities), and deviations (like in Table 6) help to quantify when a region has an abnormal affinity for an applied-for gTLD. We quantify a regional affinity as any regional interest (i_r^{gTLD}) that is more than two standard deviations ($2 \times \sigma^{gTLD}$) away from the average interest across all regions I^{gTLD}. That is, we define an applied-for gTLD's overall interest I^{gTLD} as the average of all regional interests for that gTLD:

$$I^{gTLD} = \frac{\sum_{i=1}^{|R|} i_i^{gTLD}}{R}$$

Where R is the set of all regions. Others may choose to define regional affinity with a different constant factor (other than 2), and we just present this choice as a reasonable starting point.

The results of this approach lend a continuous metric of which regions may have abnormally high preferences for applied-for gTLDs (in general), and which gTLDs have heavy regional affinities. For example, Japan displays high affinity scores for tokyo ($12.06 \times \sigma^{.tokyo}$), osaka ($9.42 \times \sigma^{.osaka}$), and kyoto ($7.88 \times \sigma^{.kyoto}$); whereas the Netherlands has a high affinity for amsterdam ($8.75 \times \sigma^{.amsterdam}$). Similarly major brands appear to have higher affinities for their brand-applied-for gTLDs in their target markets: US: comcast ($3.55 \times \sigma^{.comcast}$), Brazil: uol ($4.94 \times \sigma^{.uol}$), CN: baidu ($11.06 \times \sigma^{.baidu}$), DE: lanxess ($4.98 \times \sigma^{.lanxess}$), or in France, sfr has an affinity score of $13.6 \times \sigma^{.sfr}$.

Some of the more pronounced affinities include: exchange, which has an affinity score of $13.90 \times \sigma^{.exchange}$ in Estonia. This seems noteworthy due to the possibility that a future collision with the exchange gTLD could affect email deliverability if query hits correspond to Microsoft Exchange deployments. Also interesting is the disproportionate affinity scores of $13.77 \times \sigma^{.love}$ and $12.95 \times \sigma^{.accountant}$, both from the U.S. Virgin Islands. Also, search, which draws a high affinity of $14.21 \times \sigma^{.search}$ score from South Korea. In Australia both win and iinet receive high levels of affinity, at $14.33 \times \sigma^{.win}$ and $12.82 \times \sigma^{.iinet}$, respectively. Several regions have high affinity scores for school (Australia, New Zealand) while in Germany, India, Belgium they exhibit higher localized affinities for alternatives or translations (schule, training, and college). Though interpreting the relative significance of these scores is somewhat qualitative, the metric helps isolate leading indicators for us to examine weak signals. That is, if query rate were the lone metric, one might have excluded accountant, even though it displays some high affinity traffic from the U.S. Virgin Islands because it is in the bottom half of traffic counts. Additionally, one might exclude tjx because it is ranked 908 in overall traffic ranking, but it has an affinity score of $10.41 \times \sigma^{.tjx}$ in Haiti. To illustrate some of the trends, Table 8 lists a snapshot of several of the regions in each continent (barring Antarctica) that show some of the highest regional affinities.

Periodicity: In addition to our candidate measurement of regional affinities, we also evaluate the possibility that applied-for new gTLD strings are already in use by automated systems, whose traffic may display measurable periodicity. That is, we speculate that some systems (such as, monitoring systems or embedded devices) may periodically beacon out DNS queries. So, we measured the inter-query time gap for each query to each applied-for gTLD from each AS, and then calculated the variance for each time series. Our intuition is that when queries are emitted at roughly the same rate over time, the variance in this value should be low. Under high concurrency, one might expect low variance scores (as query rates would approach the inverse of the TTL period: $\lambda = \frac{1}{TTL}$). Under lower concurrency, our hypothesis is that similarly low variance scores might suggest some degree of weak signal (perhaps automation). What we find from measurements is that some of applied-for gTLDs strings with the greatest query volume (such as corp) do indeed have very low variance scores from large ASes (such as Verizon, Rackspace, Deutsche Telekom, etc.). Our interpretation of this metric is that it offers an additional piece of evidence that some ASes may have a reliance on applied-for gTLD strings.

In addition, however, a more intricate set of interactions illustrates how complicated some of the effects of

Region / gTLD	σ^{gTLD}
Europe (EU)	
.page	12.65
.hot	10.83
.office	8.78
.epson	8.17
Macedonia (MK)	
.dance	11.40
.room	9.45
.promo	8.52
.arc	8.50
United States (US)	
.host	3.62
.comcast	3.55
.ice	3.50
Haiti (HT)	
.thd	11.70
.ril	11.45
.how	11.17
.church	10.78
Myanmar (MM)	
.vip	12.97
.kia	12.82
.university	12.32
Japan (JP)	
.bbt	13.15
.bet	12.66
.email	10.82
Angola (AO)	
.software	12.00
.security	8.62
.shop	8.36
.bcg	8.11
Nigeria (NG)	
.store	12.45
.pharmacy	11.49
.bible	10.07
.pictures	9.90
.mobile	9.84
Venezuela (VE)	
.ford	13.62
.barcelona	13.22
.gree	8.83
.movistar	8.76
Paraguay (PY)	
.click	10.83
.free	8.73
.frontier	6.95
.navy	6.59
Australia (AU)	
.win	14.33
.iinet	12.82

Table 8: This Table describes the regional affinities calculated from our methodology.

DNS resolution can be. A particularly poignant case emerges for a specific string under the `box` applied-for gTLD: `fritz box`. This domain name appears to be used by a specific brand of Small Office / Home Office (SOHO) servers called FRITZ!Box [33], which implement the Session Initiation Protocol (SIP) [56], and are popular in Europe. As a Voice over IP (VoIP) protocol, a bad interaction between SIP and DNS resolution could not only lead to telephony outages for subscribers but could *also* (in at least one particular case, discussed below) prompt crashes in VirtualBox [59] (a popular brand of virtual machine). Below the `fritz box` domain, a diverse set of third and fourth level labels can be observed to range from strings with no obvious meaning (such as `kmswiyfxcj`), to DNS service discovery (`_dns-sd _udp`), to more common strings (like `twitter`); but with measurably low variance in inter-query periodicity. Examination of these labels should raise concern when intersecting them with the general roles of home SIP servers.

One implication of the combination of DNS resolution failure and SIP calls from home users is that failure modes could disrupt subscribers' abilities to make emergency phone calls. This sort of failure could be brought about because SIP calls require DNS service discoveries for control messages to initiate calls. However, when DNS resolution is affected, even more general failures could be caused behind SOHO routers. In a ticket listed in the the bug tracking system for Oracle's VirtualBox [58], an interaction between several versions of FRITZ!Boxes, VirtualBox, and DNS led to kernel panics in virtual machines. The bug ticket includes the following capture of the default DNS configuration (`resolv conf`) for FRITZ!Boxes:

```
#
# This file is automatically generated
#
domain fritz box
nameserver 192 168 20 1
```

From this default configuration we can see that FRITZ!Boxes are configured to use `box` as an iTLD. The ticket [58] goes on to identify that a workaround for the kernel panic is to use Google's public DNS resolution (8.8.8.8), and this correlates well with our measurements. We observe that Google's ASN has measurably low variance in its periodicity for `box`. Further investigation into FRITZ!Box reveals that its configuration pages are all internally serviced from `fritz box`, and their configuration advice [22] notes that loading this page can both be problematic, and can cause browsers to incrementally send DNS queries as the user enters each part of a DNS domain name. Figure 10 illustrates the most popular SLDs under the `box` string.

Another interesting example is `sfr`, which we identified a regional affinity for with France (above). This

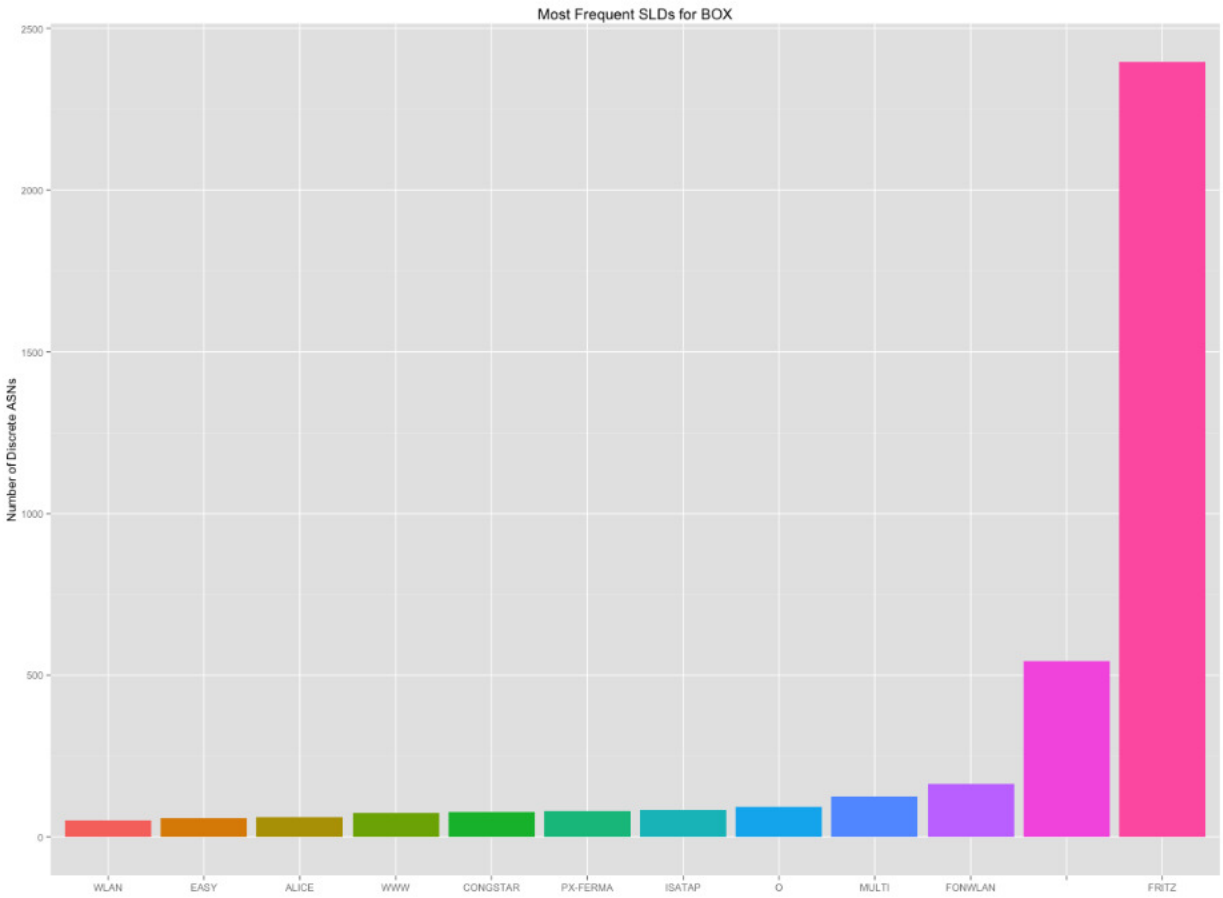

Figure 10: This Figure shows popular SLDs within `box`. The label `fritz` is the most frequent, but the threat vector for `isatap` is also visible.

string *also* has a measurably low variance score from an AS registered to the Societe Francaise du Radiotelephone S.A. These *two* pieces of evidence reinforce each other and may suggest that this string is already actively relied upon. Other periodic trends include `ice` from Microsoft, or `maif` from British Telecommunications. Each of these applied-for gTLD strings have varying traffic patterns across their constituent ASes, but this metric helps identify which ASes may be more systematically dependent.

4.3 Impact

We, next, use measurements of specific network protocols (WPAD, ISATAP, and DNS-SD queries) to estimate the *impact*, or degree to which Internet users may be vulnerable to subversion by the new gTLDs.

WPAD: Measurements of the WPAD label showed that 1,002 of the applied-for gTLD strings received requests of the form `wpad` ⟨`*`⟩ ⟨`gTLD`⟩. Based on HTML links observed in our IPS system, we have evidence that existing Web pages are at least one source of query traffic that drives user agents (such as Web browsers, some of which implement WPAD) to new gTLDs. Figure 11 shows the most requested WPAD gTLDs by distinct ASNs.

As previously stated, `corp` has been noted in several publications to be associated with private use. Further inspecting the WPAD NXDomain traffic for `corp` showed numerous major corporations present within the queries. Figure 12 shows the most popular SLDs for `corp` with a WPAD label also present in the FQDN. This Figure suggests that numerous corporations may be using some form of internal TLD namespace under the iTLD `corp`. Note the incidence of SLDs like "`AirBus`," which could indicate a dependency on NXDomains by a large aerospace manufacturer. Or, consider "`AD`," which has been used by some as an Active Directory control point [7]. Furthermore, our measurements show evidence that (due to the nature of WPAD's resolution look up behavior) if a registry operator were to begin responding to queries for `wpad corp`, many businesses would be directly vulnerable to risks such as those identified earlier. These risks are similar to those exposed by the actual incidents discussed in Section 3.1.

Perhaps, more alarming are the implications that could be drawn from the most popular SLDs under `cisco`, seen in Figure 13. Here we can see that not only is `wpad` the second most popular SLD, but `isatap` is number one. Moreover, DNS-SD-like labels (`_udp` and `_tcp`) round out the top seven, and other *current* gTLDs (such as `gov`, `net`, and `info`) are in the list. Further, there are strings like `user-pc`, `server`, and `owner-pc` (just to name a few) that may suggest DNS queries for internally-scoped names are leaking out to the global DNS root. This could, perhaps, enable a vendor to learn about internal network structures of the clients they sell to, if (in the future) `cisco` were to be operated by such an entity.

ISATAP: Measurements of the ISATAP label showed that 951 of the unique applied-for gTLD strings received requests of the form `isatap` ⟨`*`⟩ ⟨`gTLD`⟩. Figure 14 shows the most requested ISATAP gTLDs by distinct ASNs. This data shows that applied-for gTLD strings are being widely used in private networks.

DNS-SD: Measurements of the DNS-SD like FQDNs showed that 1,036 of the applied-for gTLD strings received requests of the form `_UDP` ⟨`*`⟩ ⟨`gTLD`⟩ or `_TCP` ⟨`*`⟩ ⟨`gTLD`⟩. Figure 15 shows the most requested DNS-SD applied-for gTLDs by distinct ASNs.

5 Security Analysis: How Pervasive is the Risk?

The goal of our security analysis is not to precisely quantify the degree to which any given string, region, AS, user, etc. is at risk for compromise or attack. Rather, we use our measurements as evidence of potential risks, and extrapolate the relative weights of the risk posed by each applied for gTLD, based solely on our measurements. We submit that our security analysis serves simply as quantitative evidence that certain risks *do* exist (without representing their conclusive acuteness). To this end, our *Risk Matrix* (Table 9) illustrates our measured risk vectors, and it is sorted according to the amount of evidence that we measured.

In this table, our three risk vectors for setting up proxied, tunneled, or other services are described in the columns for WPAD, ISATAP, and DNS-SD. The values in this table represent the number of unique ASes that were observed emitting queries for various applied-for new gTLD strings that might be exploited either intentionally by an adversary, or inadvertently. In addition, we list the relative fraction of all ASes that were seen querying for this applied-for new gTLD string versus the number of ASes that queried this string *for* these services, in order to represent the spread of the risk across constituent queriers.

In addition to this, we measure how many visible X.509 certificates exist in public crawls for these strings. We note that while we feel this measurement is valuable, SAC057 [43] illustrates that anyone can acquire a legitimate X.509 "Internal Name Certificate" from authentic CAs at any time. Therefore, concern is warranted, even in the absence of any observed internal named certificates for new gTLD strings. Nonetheless, we include measured evidence as a potential increased level of risk

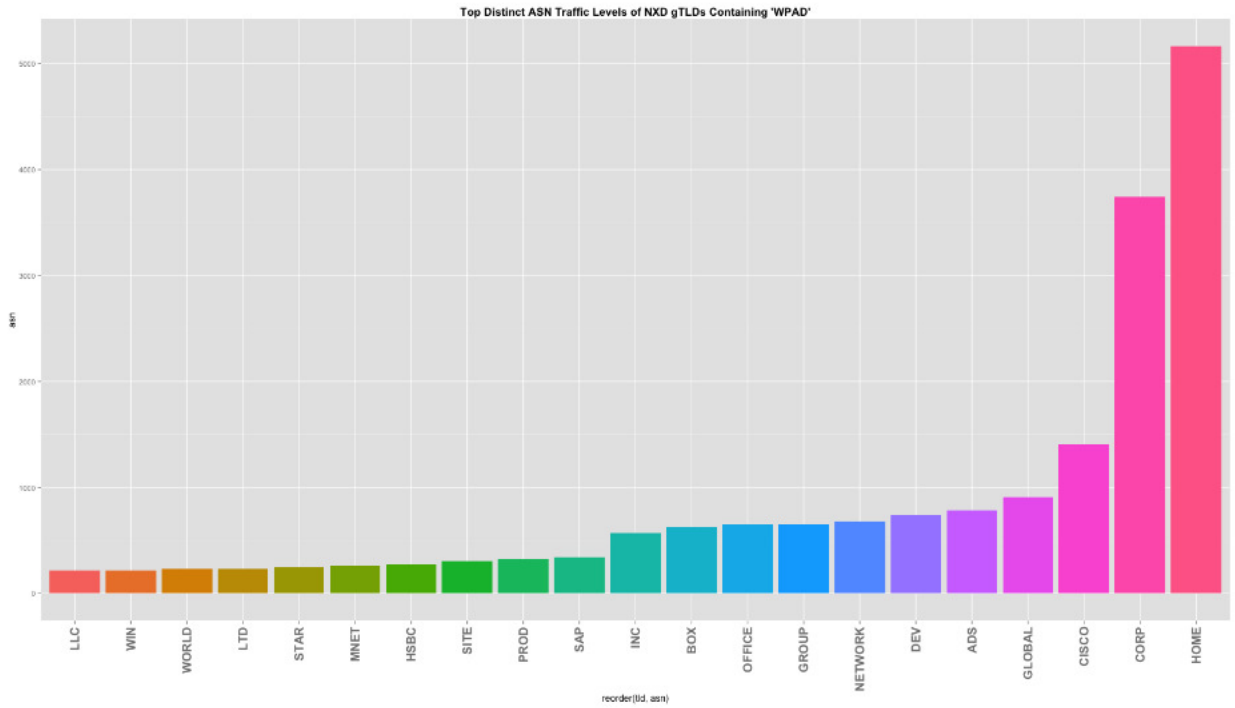

Figure 11: This Figure shows the top gTLD strings seen to have WPAD queries issued for them. Of note is that `home`, `corp`, and `cisco` round out the top three SLDs, respectively.

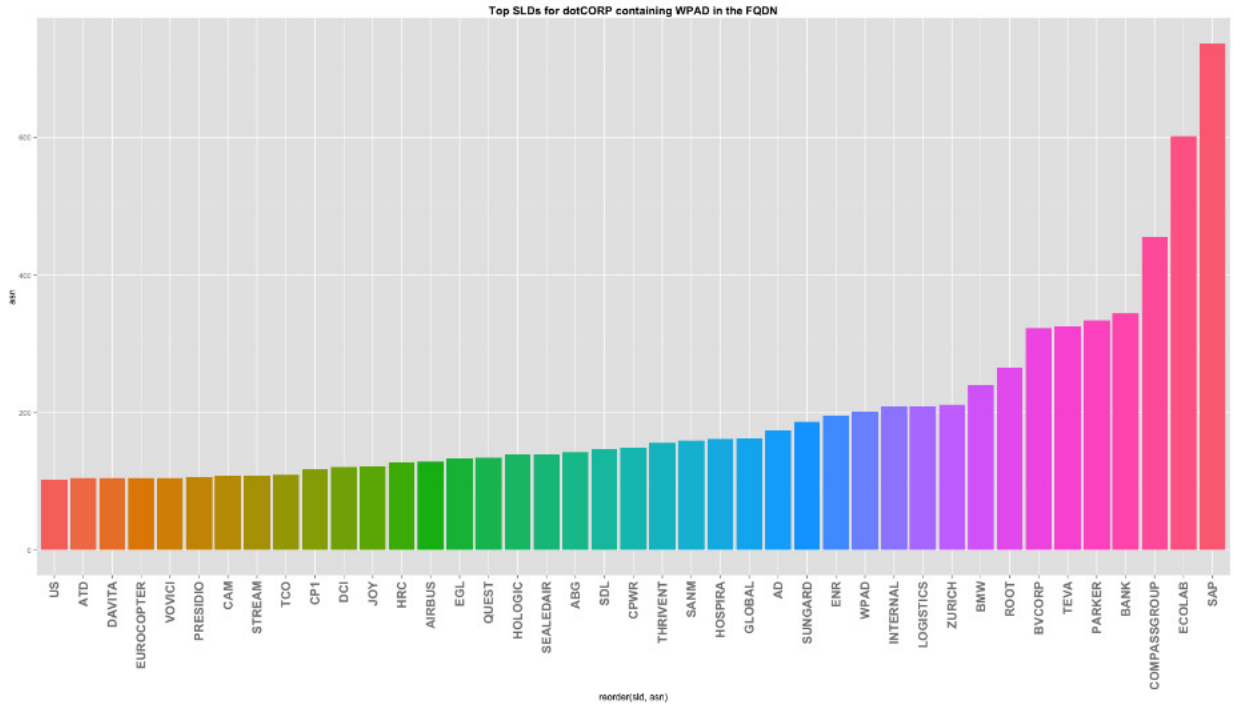

Figure 12: This Figure shows popular SLDs within `corp`.

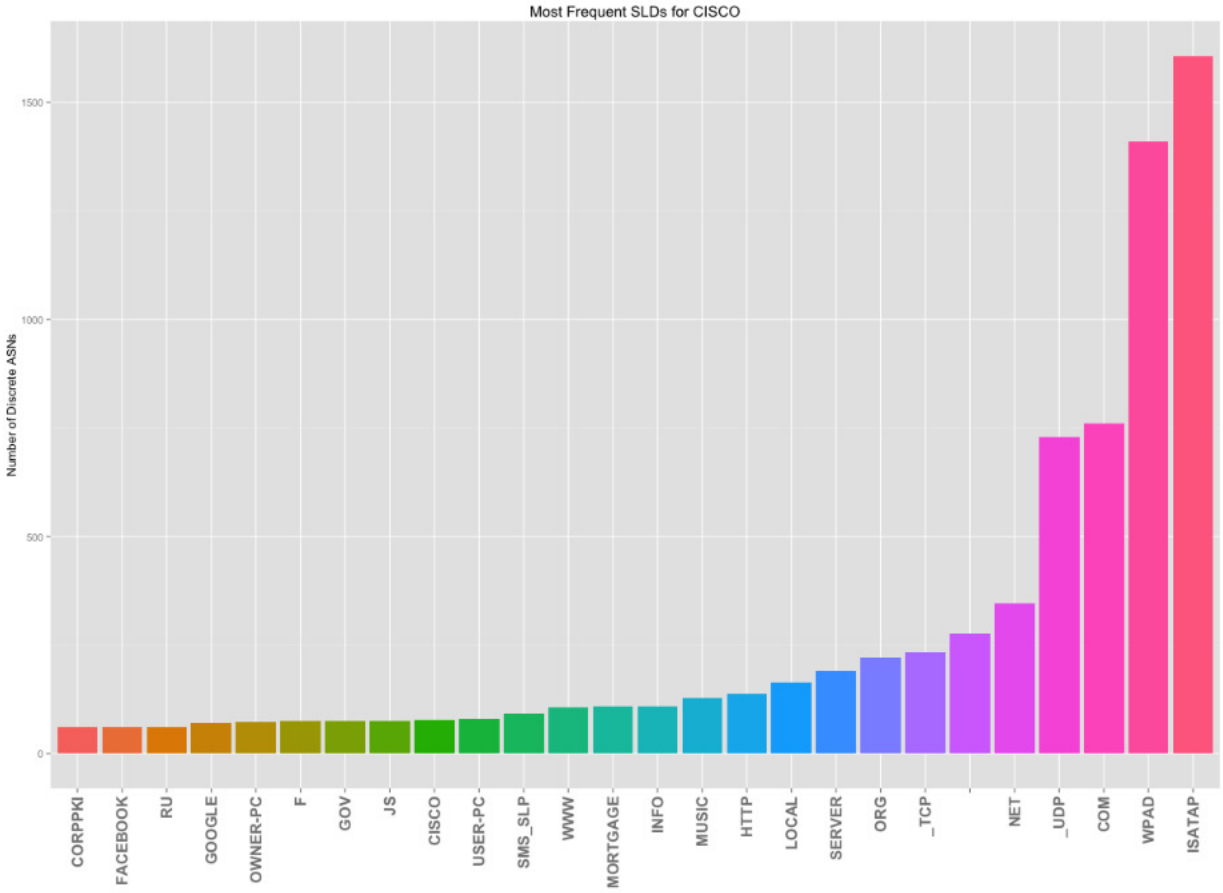

Figure 13: This Figure shows popular SLDs within `cisco`. The labels `isatap`, `wpad`, and `_udp` and `_tcp` could directly indicate threat vectors, but are certainly not the only SLDs that could be problematic.

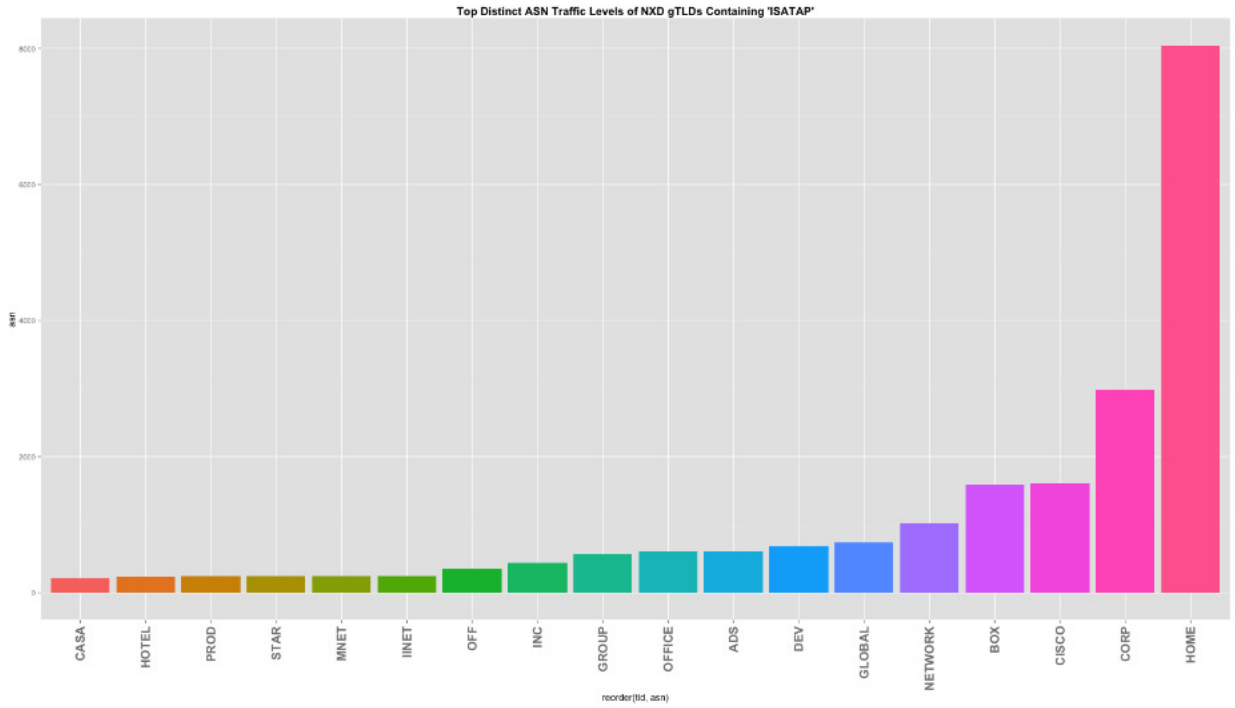

Figure 14: This Figure shows the most popular applied-for gTLDs for which ISATAP queries were seen.

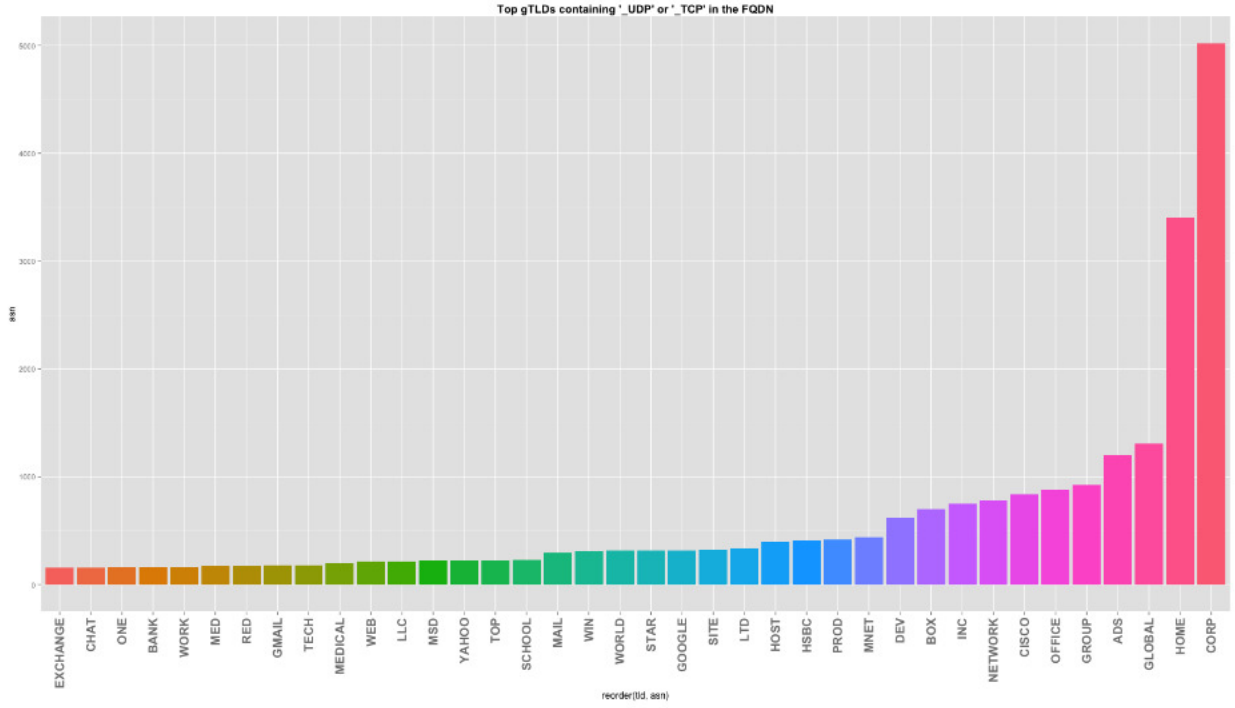

Figure 15: This Figure shows the most requested DNS-SD gTLDs by distinct ASNs.

gTLD	WPAD ASNs	ISATAP ASNs	DNS-SD ASNs	X.509 Certs	HTML Refs	# Risk Vectors	ASN Spread	Regional Affinities	Interisle Risk
MEDICAL	83	59	202	1	12	6	0.71	JP: 7.84 PR: 10.50	Uncalculated
CORP	3744	2984	5020	378	130	6	0.65	AP: 4.34 GT: 2.33 HN: 4.24 HR: 3.47 HU: 3.42 LV: 2.01 NI: 2.89	High
BOX	631	1588	702	53	36	6	0.65	MQ: 12.27 NA: 7.08	Uncalculated
HOTEL	112	233	128	1	39	6	0.61	RW: 3.15 UZ: 13.42	Uncalculated
NETWORK	679	1026	778	39	28	6	0.61	DK: 2.98 FI: 13.80	Uncalculated
GROUP	653	565	925	22	27	6	0.60	RW: 3.89 TG: 12.56	Uncalculated
GLOBAL	912	742	1305	21	18	6	0.60	AT: 3.15 FI: 2.04 GT: 5.63 KR: 6.06 MN: 2.79 SE: 7.73 SK: 2.36	Uncalculated
ADS	782	614	1199	79	43	6	0.57	FR: 6.05 RE: 11.71	Uncalculated
HOUSE	169	175	150	3	49	6	0.56	BN: 6.71 PR: 8.55 PY: 3.93 UY: 2.24	Uncalculated
OFFICE	648	610	884	659	33	6	0.55	EU: 8.78 GP: 3.20 HU: 3.34 MK: 3.30 PR: 3.62 UA: 2.39 UZ: 4.23	Uncalculated
OLYMPUS	131	94	127	3	2	6	0.53	AT: 2.19 CU: 9.68 SK: 3.74	Uncalculated
SCHOOL	156	192	232	2	39	6	0.53	AE: 3.83 AU: 3.06 MV: 9.96 NZ: 3.05 TW: 5.34	Uncalculated
GMBH	26	21	62	7	2	6	0.52	AT: 3.27 DE: 8.40 HR: 2.90	Uncalculated
DENTAL	42	34	37	2	5	6	0.51	AT: 9.95	Uncalculated
LLC	214	174	213	2	11	6	0.50	MN: 14.21	Uncalculated
SOLUTIONS	31	29	41	2	36	6	0.48	LV: 10.79	Uncalculated
CLINIC	35	33	43	1	4	6	0.47	CY: 11.90	Uncalculated
MOSCOW	18	34	21	2	3	6	0.46	AE: 6.01 CY: 4.93	Uncalculated
HSBC	274	68	409	9	7	6	0.46	GT: 5.59 HN: 12.30 SV: 2.28	Uncalculated
SECURITY	38	27	41	1	14	6	0.17	AO: 8.62 LT: 4.64 MM: 5.20	Uncalculated
SECURE	47	64	59	1	65	6	0.09	LV: 2.74 SK: 11.44	Uncalculated

Table 9: This is a snapshot of the overall Risk Matrix, calculated by measuring all of risk vectors

in our matrix as it may (if nothing else) indicate that attention has already been paid to a string and X.509 certificates already exist that could enable an adversary or facilitate an attack.

Our Risk Matrix also includes our measurement of the incidence of applied-for gTLD strings seen in HTML pages. As we discussed in Section 4.3, we consider the presence of these links as one possible risk vector. While many could be misconfigurations, some could also be links that are only intended to resolve within a corporation or development environment. Regardless of *how* they came to be in public pages, their presence can drive user traffic to strings that are (as yet) not delegated. The change in delegation status of the relative HTML links will impact the experience and underlying systems behavior for users, and we consider that a risk.

Finally, we considered regional affinities as an independent measure of risk (as we described earlier). This measurement is also described in our Risk Matrix.

While our analysis of this matrix is not a quantified metric of danger, we have sorted the overall results based on the number of our risk factors that each applied-for new gTLD string triggered, and the relative spread observed. Table 9 enumerates just those new gTLD strings that appeared to have the most measurable evidence of potential risks, under one candidate sorting scheme. This subset of risks is sorted based on how many risk vectors were observed for each new gTLD string (WPAD, ISATAP, DNS-SD, X.509 Internal Named Certs, etc.), and then how widely spread across the querying ASes those risks were observed (on a per-string basis). Our matrix contains other strings (which are much lower down in the list) that illustrate how intuitively more nuanced and less popular strings exhibit lower probabilities of collisions in the namespace than sexier and more common strings, such as `secure`.

6 Discussion

Our goal with this study is to illustrate evidence of potential issues that may arise with the delegation of the applied-for new gTLDs (and new TLDs in general). Beyond this, it is our belief that constructive recommendations can be issued and followed to help mitigate problems that may lie ahead, and we enumerate a candidate list here. At a high level, we simply recommend that ICANN implement those recommendations that were outlined in the "Signposts in Cyberspace" report [48], 'Scaling the Root" study [19], SAC045 [40], SAC046 [41], SAC057 [43], and SAC059 [42]. These are all work products which we have, in various capacities, participated directly in over the past decade. At a high level, these recommendations are:

1. Build a root server system instrumentation capability to accurately assess the systemic impact of applied-for strings, as well as to detect stresses on the root server system itself.
2. Develop technical and policy frameworks for braking and rollback of delegations, perhaps to include ephemeral delegations but only after recommendations above are effectuated.
3. As per above, assess potential user impact and notify potentially impacted parties of impending delegation, and provide mitigation recommendations to Internet users and operators that could be impacted.
4. Establish a communications plan to notify potentially impacted parties, vendors, and establish operations of, and build, a call center and supporting framework that outlines whom to contact if issues arise.
5. Evaluate liabilities and risks associated if issues arise, and what protections are in place for involved parties.

In addition, if they haven't done so already, we believe the ICANN Governmental Advisory Committee (GAC) and other such ICANN stakeholders may want to consider engaging with their respective agencies to explicitly address the question of if (and which) strings may be in use in their critical infrastructure and key resources (CIKR) [28, 14]. This would enable them to forewarn impacted parties, as well as potentially mitigate impacts prior to delegation of a given string. This is particularly vital given their context related to DNS ecosystem, and their proximity to the new gTLD program over the past several years. This is also important for the obvious reason that simply removing a delegated string (un-delegating) could possibly be an entirely inadequate remedy for damages that could result; for instance caching and various systemic effects could result in prolonging any impacts in addition to other residual effects.

7 Conclusion

In this study, we conduct one of the largest investigations of DNS root zone traffic to date, with DNS queries from up to 11 of the 13 root instances, dating back to 2006. In addition, we propose a novel methodology to gauge the risk posed by applied-for new gTLD strings, and quantify it using measurements of DNS, the World Wide Web, X.509 certificates, regional preferences, and inter-query timing analysis.

What we found was that quantifying the risk that applied-for new gTLDs pose to Internet users goes beyond simply evaluating query rates for, as yet, undelegated new gTLD strings. Indeed, we found several instances where automatic proxy protocols, X.509 internal

names certificates, and regional traffic biases could leave large populations of Internet users vulnerable to DoS and MitM attacks, immediately upon the delegation of new gTLDs.

Our measurements and quantification of risks exist as just candidate approaches. While we feel there is quantifiable evidence of risk, there is clearly room for alternate methodologies and this effort will certainly benefit from community input and more comprehensive analysis. However, we believe that this study constitutes the first attempt to conduct an interdisciplinary (and consumer impact) analysis of the new gTLDs in the global DNS.

One of the tangible benefits of this study has been quantitative analysis that has qualified some of the implications of unresolved recommendations. In this work, we have presented evidence that suggests that these unresolved recommendations have potentially damaging implications to general Internet consumers, corporations, and *public interest*. Additionally, we believe that the new gTLD program could pose very real risks to both the set of entities that have been charged with effectuating new gTLD delegations, and the set of those responsible for giving due consideration to (and implementation of) recommendations provided by ICANN's advisory committees and expert contributors, *if* those recommendations remain unresolved.

While in a 2005 National Research Council [48] study the number of recommended delegations was on the order of tens per year, we are not advocating any particular number. We are, however, advocating that instrumentation be in place and recommendations be enacted to support the safe introduction of new gTLDs.

We believe that further study and express focus on implementation of recommendations already provided is critical in progressing the new gTLD program in a safe and secure manner for all stakeholders. We believe that this work has demonstrated evidence that risks exist, to both the existing Internet user base, as well as to new gTLD applicants and services consumers. We believe recognition of this evidence and explicit consideration, planning, and appropriate resourcing for further study and resolution of outstanding recommendations is the most prudent and expeditious manner with which to move forward.

References

[1] DNS Namespace Planning. Support 254680, Microsoft. http //support microsoft com/kb/254680.

[2] Internet Corporation for Assigned Names and Numbers (ICANN). http //www icann org/.

[3] Naming conventions in Active Directory for computers, domains, sites, and OUs. Support 909264, Microsoft. http //support microsoft com/kb/909264.

[4] Public Suffix List. In *Mozilla Wiki*. https //wiki mozilla org/Gecko Effective_TLD_List.

[5] Novell Open Enterprise Server. Novell dns/dhcp services administration guide, Novel, October 2006. https //www novell com/documentation/oes/pdfdoc/dhcp_enu/dhcp_enu pdf.

[6] Bug 252342 - fix cookie domain checks to not allow .co.uk . In *Bugzilla@Mozilla* , April 2008. https //bugzilla mozilla org/show_bug cgi?id=252342.

[7] MAD dns naming ? In *Novell Forums*, July 2008. http //forums novell com/novell/novell-product-discussion-forums/netware/nw-other/dns-dhcp/336488-mad-dns-naming html?pagenumber=%3E.

[8] Diagnosing a 6038 Error in Identity Manager. Support, Novel, January 2010. https //www novell com/communities/node/9465/diagnosing-6038-error-identity-manager%3E.

[9] Google Chrome handles new TLDs badly. In *Domain Incite Blog*, May 2012. http //domainincite com/8978-google-chrome-handles-new-tlds-badly.

[10] New Generic Top Level Domains Applicant Guidebook. In *ICANN*, June 2012. http //newgtlds icann org/EN/APPLICANTS/AGB.

[11] Why does iTunes 10.7 try to contact the domain bogusapple.com? In *Apple SUpport Communities*, October 2012. https //discussions apple com/thread/4380270?tstart=0.

[12] Apple, Google and Microsoft still don't understand new TLDs. In *Domain Incite Blog*, January 2013. http://domainincite.com/11673-apple-google-and-microsoft-still-dont-understand-new-tlds.

[13] BEIJING - New gTLD Security Stability & Resiliency (SSR) Update. In *ICANN 46*, April 2013. http //beijing46 icann org/bitcache/c1f445deb97b93056b6d528bc82ad405b2a26212?vid=50079&disposition=attachment&op=download.

[14] E9-1-1 best practices, final report. In *CSRIC III: WORKING GROUP - 8*, March 2013. http //transition fcc gov/bureaus/pshs/advisory/csric3/CSRICIII_6-6-12_WG8-Final-Report_Pt2 pdf.

[15] How certificate revocation (doesn't) work in practice. Blog post, Netcraft, May 2013. http://news.netcraft.com/archives/2013/05/13/how-certificate-revocation-doesnt-work-in-practice.html.

[16] New gTLD Security and Stability Considerations. Technical Report 1130007 version 2.1, 2013. http //www verisigninc com/assets/gtld-ssr-v2 1-final pdf.

[17] Part 2 of 5; Internet Infrastructure: Stability at the Core, Innovation at the Edge. In *Between the Dots Blog of Verisign*, May 2013. http //blogs verisigninc com/blog/entry/internet_infrastructure_stability_at_the.

[18] Proposed Delegation of Invalid Names from SAC 045 and RFC 6762. Correspondence, Paypal, March 2013. http //www icann org/en/news/correspondence/hill-smith-to-chehade-crocker-15mar13-en.

[19] Jaap Akkerhuis, Lyman Chapin, Patrik Fltstrm, Glenn Kowack, Lars-Johan Liman, and Bill Manning. Scaling the Root: Report on the Impact on the DNS Root System of Increasing the Size and Volatility of the Root Zone. Technical report, September 2009. http //www icann org/en/groups/rssac/root-scaling-study-report-31aug09-en pdf.

[20] R. Arends, R. Austein, M. Larson, D. Massey, and S. Rose. DNS Security Introduction and Requirement. RFC 4033, March 2005.

[21] artemis group. .secure a safer Internet. https //artemis net/ncc-group.

[22] AVM Knowledge Base. FRITZ!Box user interface cannot be opened. http //service avm de/support/en/SKB/FRITZ-Box-7390-int/649 FRITZ-Box-user-interface-cannot-be-opened.

[23] Nevil Brownlee, KC Claffy, and Evi Nemeth. Dns measurements at a root server. In *Global Telecommunications Conference, 2001. GLOBECOM 01. IEEE*, volume 3, pages 1672–1676. IEEE, 2001.

[24] CAIDA and DNS-OARC. A Day in the Life of the Internet (DITL) . http //www caida org/projects/ditl/.

[25] S. Cheshire and M. Krochmal. DNS-Based Service Discovery. RFC 6763, February 2013.

[26] S. Cheshire and M. Krochmal. Multicast DNS. RFC 6762, February 2013.

[27] S. Cheshire and M. Krochmal. Special-Use Domain Names. RFC 6761, February 2013.

[28] The President's National Security Telecommunications Advisory Committee. Nstac report to the president on communications resiliency. April 2011. http //www ncs gov/nstac/reports/NSTAC%20Report%20to%20the%20President%20on%20Communications%20Resiliency%20(2011-04-19)(Final)(pdf) pdf.

[29] T. Dierks and E. Rescorla. The transport layer security (tls) protocol version 1.2. RFC 5246, 2008.

[30] D. Eastlake and A. Panitz. Reserved Top Level DNS Names. RFC 2606, June 1999.

[31] Peter Eckersley and Jesse Burns. An observatory for the ssliverse. In *Defcon 18*, 2010.

[32] International Organization for Standardization. *ISO 3166-1:2006 Codes for the representation of names of countries and their subdivisions - Part 1: Country codes*. ISO, Geneva, 2 edition, 2006.

[33] FRITZ!Box. AVM - Everything's Networked With FRITZ! http //www fritzbox eu/en/index php.

[34] Paul Gauthier, Josh Cohen, Martin Dunsmuir, and Charles Perkins. Web Proxy Auto-Discovery Protocol draft-ietf-wrec-wpad. Internet draft, IETF, December 1999. http //tools ietf org/html/draft-ietf-wrec-wpad-01.

[35] E. Gavron. A Security Problem and Proposed Correction With Widely Deployed DNS Software. RFC 1535, October 1993.

[36] Martin Georgiev, Subodh Iyengar, Suman Jana, Rishita Anubhai, Dan Boneh, and Vitaly Shmatikov. The most dangerous code in the world: validating ssl certificates in non-browser software. In *ACM Conference on Computer and Communications Security*, pages 38–49, 2012.

[37] Saikat Guha and Paul Francis. Identity Trail: Covert Surveillance Using DNS. In *Proceedings of the 7th International Symposium on Privacy Enhancing Technologies (PETS)*, Ottawa, Canada, Jun 2007.

[38] Vernita D. Harris. Addition of New gTLDs to the Root Zone. http //www icann org/en/news/correspondence/harris-to-kane-02aug13-en pdf.

[39] Kelly Jackson Higgins. New .secure Internet Domain On Tap. In *Dark Reading*, May 2012. http //www darkreading com/management/new-secure-internet-domain-on-tap/240000187.

[40] ICANN Security and Stability Advisory Committee (SSAC). Invalid Top Level Domain Queries at the Root Level of the Domain Name System. SSAC Advisory 045, November 2010. `http //www icann org/en/groups/ssac/documents/sac-045-en pdf`.

[41] ICANN Security and Stability Advisory Committee (SSAC). Report of the Security and Stability Committee on Root Scaling. SSAC Report 046, December 2010. `http //www icann org/en/groups/ssac/documents/sac-046-en pdf`.

[42] ICANN Security and Stability Advisory Committee (SSAC). Reponse to the ICANN Board Regarding Interdisciplinary Studies. SSAC Advisory 059, April 2013. `http //www icann org/en/groups/ssac/documents/sac-059-en pdf`.

[43] ICANN Security and Stability Advisory Committee (SSAC). SSAC Advisory on Internal Name Certificates. SSAC Advisory 057, March 2013. `http //www icann org/en/groups/ssac/documents/sac-057-en pdf`.

[44] J. Moss. ICANN's April Update on SSAC 046 & 057. ICANN Government Advisory Committee Advice, April 2013. `http //www icann org/en/news/correspondence/moss-to-falstrom-30apr13-en`.

[45] Patrick S. Kane. Joint Test Summary Report, RZM. 2.0. `http //www icann org/en/news/correspondence/kane-to-harris-30may13-en pdf`.

[46] O. Kolkman, A. Sullivan, and W. Kumari. A Procedure for Cautious Delegation of a DNS Name draft-kolkman-cautious-delegation. Internet draft, IETF, June 2013. `http //tools ietf org/html/draft-kolkman-cautious-delegation-01`.

[47] P. Mockapetris and K. J. Dunlap. Development of the domain name system. In *SIGCOMM '88*, 1988.

[48] National Research Council (Etats-Unis) . Committee on Internet navigation, the domain name system. Technical alternatives, policy implications, National Research Council (Etats-Unis) . Computer science, telecommunications board. Division on engineering, and physical sciences. *Signposts in cyberspace : the domain name system and internet navigation*. Washington, D.C. National Academies Press, 2005.

[49] Eric Osterweil, Dan Massey, and Lixia Zhang. Managing trusted keys in internet-scale systems. In *The First Workshop on Trust and Security in the Future Internet (FIST '09)*, 2009.

[50] Eric Osterweil, Danny McPherson, and Lixia Zhang. Operational implications of the dns control plane. *IEEE Reliability Society Newsletter*, May 2011.

[51] Eric Osterweil, Michael Ryan, Dan Massey, and Lixia Zhang. Quantifying the operational status of the dnssec deployment. In *IMC '08*, 2008.

[52] Eric Osterweil and Lixia Zhang. Interadministrative challenges in managing dnskeys. *IEEE Security and Privacy*, 7(5):44–51, 2009.

[53] P. Kane. Verisign Comments for New gTLD Board Committee Consideration of GAC Safeguard Advice. ICANN Government Advisory Committee Advice, April 2013. `http //forum icann org/lists/comments-gac-safeguard-advice-23apr13/msg00103 html`.

[54] Keith Parkansk. How To Set Up Linux DNS Services. `http //www aboutdebian com/dns htm`.

[55] Venugopalan Ramasubramanian and Emin Gün Sirer. Perils of transitive trust in the domain name system. In *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*, IMC '05, pages 35–35, Berkeley, CA, USA, 2005. USENIX Association.

[56] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, and E. Schooler. SIP: Session Initiation Protocol. RFC 3261, June 2002.

[57] F. Templin, T. Gleeson, and D. Thaler. Intra-Site Automatic Tunnel Addressing Protocol (ISATAP). RFC 5214, March 2008.

[58] Oracle VM VirtualBox. Ticket #11649 (new defect) NAT-related crash of ubuntu guest on OSX host. `https //www virtualbox org/ticket/11649?cversion=0&cnum_hist=14`.

[59] Jon Watson. Virtualbox: Bits and bytes masquerading as machines. *Linux Journal*, 9941, February 2008. `http //www linuxjournal com/article/9941`.

[60] D. Wessels and M. Fomenkov. Wow, That's a lot of packets. In *Passive and Active Network Measurement Workshop (PAM)*, San Diego, CA, Apr 2003. PAM.

