Donuts Name Collision Comment

Donuts appreciates the opportunity to respond to comments submitted on the name collision issues.

<u>Executive Summary</u>

So-called "name collision" has been presented to the community as a last-gasp argument for delay in the eight-year new gTLD process. This is regrettable, as name collision is not a threat to the Internet's infrastructure in the way it has been negatively portrayed.

This comment will demonstrate the opposite—that:

- "name collision" is an existing issue in .COM, without damage to the domain name system (DNS);
- the data shows that collision isn't the danger it may have been portrayed to be, and the Interisle report is deficient in its findings;
- "collision" existed prior to the introduction of other new gTLDs, again without damage to the security and stability of the DNS or the Internet;;
- mitigation (if needed) can be successfully handled on a case-by-case basis;
- there is no need to delay the program further or measure the root system more than is already planned; and
- the key to understanding name collision is to examine second-level domains, which the Interisle study did not do.

<u>The Scope of Any Perceived Problem</u>

Interisle—by its own admission—was not granted sufficient time to completely study the root system data available to it for name collision. Had the study been more thorough, it would have responsively demonstrated that the magnitude of collision in .COM is far higher than the rate of collision in every new round applied-for gTLD.

In order to make a fair comparison of the relative risk regarding collision, it's critical to point out that Verisign, as manager of the .COM registry, experiences collision at a rate of at least 2,000 names per day for the studied period in 2013, and at least 16,000 names per day for the study period in 2012 (see below for additional data).

Opponents of the new gTLD program, or those who have commercial interest in opposing namespace expansion, apparently express that collision and non-existing domain (NXD) traffic is acceptable in the current root—especially .COM—but, cynically, that such traffic in other gTLDs (such as .CORP and .HOME) actually threatens lives.[1] Such statements are not only needlessly alarmist, they attempt to create fear and uncertainty where none is warranted.

---

[1] http://domainincite.com/13221-verisign-says-people-might-die-if-new-gtlds-are-delegated

Notification of Applicants

Prior to the opening of the application window, the Security and Stability Advisory Committee (SSAC) issued SSAC 045[2] and recommended the applicant would be alerted "during the string evaluation process about the pre-existence of invalid TLD queries to the applicant's string." SSAC also recommended "ICANN should coordinate with the community to identify a threshold of traffic observed at the root as the basis for such notification."

ICANN has set the bar and notified the applicants as SSAC recommended, even if after the conclusion if Initial Evaluation. The SSAC has not recommended that the applied-for TLDs over the bar be delayed or restricted from root entry. Accordingly, advice was followed, and the SSAC has not advised delay.

History of Collision

As DNS creator Paul Mockapetris cites[3], the central issue is there are DNS facilities, and simultaneously there are programs that use the DNS. Some software or networks don't reliably collaborate with the DNS—and has been stated often, such is a business risk assumed by those producers. The DNS is not responsible for reverse-engineering to fit broken software. Any collision problem doesn't reliably lay with the existence of a new gTLD.

In fact, Mockapetris, in his authoritative experience, shares the fact that 200-some ccTLDs were added to the root without issue. The same position was reliably put forth by the New TLD Applicant Group (NTAG) in its 5 August 3013 comment on name collision, stating:

> "A Verisign analysis using data from January 2006, prior to the launch of several active TLDs, found that .xxx received more queries before delegation than any other new TLD. Despite having more queries than all of the TLDs currently under consideration in the 'Uncategorized Risk' category, .xxx was delegated in 2011. This TLD launched without incident, and no public complaints or technical issues have been identified since.
>
> In addition, most of the other TLDs listed in Table 1 of the Verisign report, including .asia, .kp, .ax, .um and .cw, also demonstrated much higher numbers of NXDOMAIN responses than all 279 of the 'uncategorized' strings, and again all were delegated with no noticeable impact. In fact, the least 'dangerous' current gTLD on the chart, .sx, had 331 queries per million in 2006. This is a higher density of NXDOMAIN queries than all but five proposed new TLDs. Again, .sx was launched successfully in 2012 with **none** of the problems predicted by these reports."

As Mockapetris states, worries about collision have existed since the introduction of the DNS. But, as he further says, and as the Board surely realizes, there is an unprecedented level of caution in the introduction of this round of TLDs compared to previous rounds, or to the introduction of ccTLDs. The introduction of previous TLDs with no impact bears out the fact that this issue is overstated.

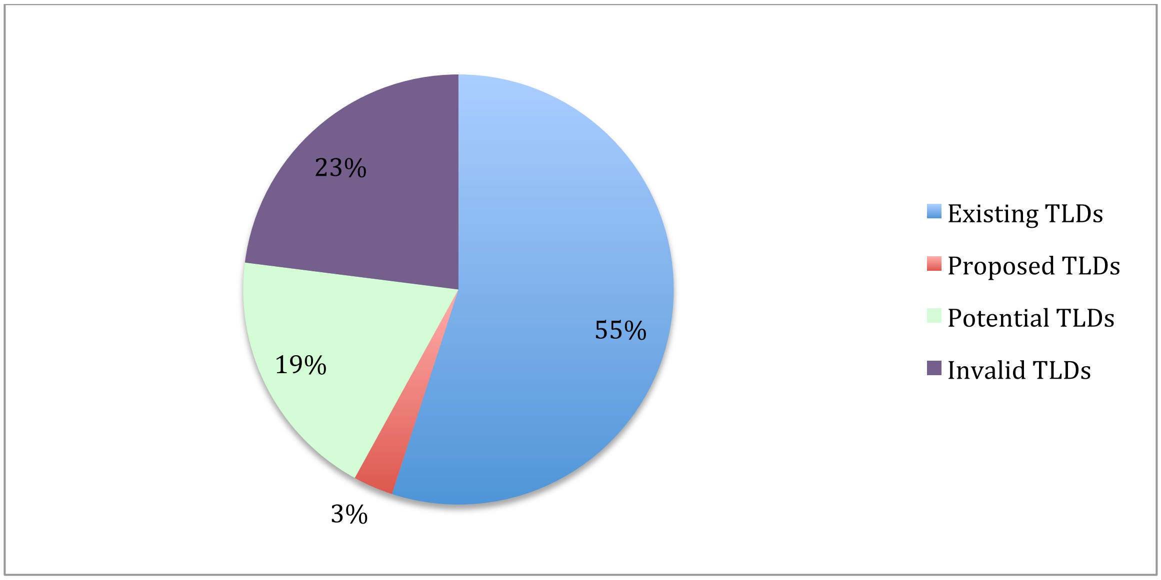
Deficiencies in the Interisle Study

This is the chart from the Interisle report that shows the queries to the root (both NXD and non-NXD for the study period (approximately one day) in 2013:

---

[2] http://www.icann.org/en/groups/ssac/documents

[3] http://www.icann.org/en/news/press/kits/video-mockapetris-15aug13-en.htm



However, according to our in-depth (with an appropriate amount of time to review the data) study of the same information, we found the following:

Item	Without accounting for TTL	Estimate with accounting for TTL
Existing TLDs	55%	78%
Proposed TLDs	3%	2%
Potential TLDs	19%	10%
Invalid TLDs	23%	12%
	100%	100%

Source: Applicant OARC data study (for 2013)

It appears that Interisle calculated its percentages based on the top 100 TLDs for which the root saw *any* queries (NXD or Non-NXD). Our study looked at *all* TLDs for which the root saw any queries.

If one accounts for the "time to live" (TTL, or the length of time that an ISP will retain the previous DNS answer), the NXD for applied-for names would be a smaller percent of the overall number of queries than shown in the Interisle report.

The Interisle study did not consider TTL as a factor when calculating ratios for the above pie chart. The TTL for non-existent SLD.TLD names is in fact *half* that for delegated TLDs. Existing TLDs are delegated, and obviously the others aren't. Because of the half figure for non-delegated TLDs, the system "asks" the root twice as often as the delegated existing TLDs, making an apples-to-apples comparison in a pie chart misleading. The DNS traffic for sparsely queried names at recursive level is nearly directly reflected at the root level. And this is especially true for non-existent TLDs compared to existing TLDs. Comparing those sparsely trafficked names to heavily trafficked names at the recursive level, those heavily trafficked names have much less of their traffic reflected at the root level. These facts not only makes NxD traffic for non-existent TLDs difficult to compare to traffic at legacy TLDs but also difficult to compare to other non-existent TLDs.

Had we had access to a large sample size of recursive DNS data, we would have been able to get a more accurate approximation, but what we do know is that the NxD traffic to applied-for TLDs is a smaller percent than what is in shown the Interisle report.

Data: Apples to Apples Comparison

It's critical that studies about potential collision focus on metrics that are appropriately comparable.

Measurements in .CORP, .HOME, or other so-called "dangerous" gTLDs can't reliably be measured against existing gTLDs because of a factor that has yet to be mentioned—that in .COM and other incumbent gTLDs, registrations constantly remove some of the NXD traffic. This in fact is a factor that could and should have been examined in the Interisle study.

It is very likely that .COM has NXD traffic at an order of magnitude far higher than .CORP; however, names in .COM are allowed to collide. It's unfortunately difficult to know with certainty, as we did not have reliable access to the .COM nameserver or to a significant sample of recursive data.

Examine the following data regarding existing gTLDs and proposed gTLDs, including the cited .HOME example:

	Percent	Total For category
Existing TLDs (not incl .com)	37.1%	
.com non-NxD	17.7%	55.0%
.com NxD	0.2%	
Proposed TLDs (not incl .home)	0.8%	3.0%
Largest proposed TLDs (.home)	2.2%	
Potential TLDs	19.0%	
Invalid TLDs	23.0%	
	100.0%	

Source: Applicant OARC data study (for 2013)

Donuts calculated .COM NXD traffic by examining each .COM SLD name that received traffic at the root (this is a significant number of SLD .COM names—more than 200 million unique .COM SLDs were tracked) for the period in question (we used the same period and the same OARC data as did the Interisle study) and noted whether or not that SLD.COM was in the .COM zone at the time. For any that were not, the .COM name servers would reply with NXD—we therefore totaled such queries and assigned them as NXD, since we do not have access to the .COM name-server logs. We are assured of the accuracy of this data.

None of the root-level NXD traffic attributed to .COM is for registered names, such as high-traffic sites Yahoo.COM, MSN.COM or Google.COM, as may be seen at the recursive level. ***All*** NXD traffic that we attributed to .COM comes from unregistered names.

As a delegated TLD, .COM has a TTL twice as long. Therefore for that entire TTL period, any NXD query from a recursive resolver (not the SLD.COM name, but .COM itself) would be cached; therefore the root would not see subsequent NXD queries during the cache period. Due to this, we estimate that the actual NXD queries that .COM is seeing is approximately 10x the level indicated above.

In 2013, .HOME queries accounted for nearly 75% of the total queries for all proposed TLDs. Our preliminary analysis shows that 92% of these queries (92% of the 75%) appear to be from Google's

Chrome browser querying for a random sequence of 10-character SLDs. These strings are random and Google Chrome is relying on the fact that it should return an NXD DNS response. To ensure at least one NXD response, Google Chrome will submit three of these random strings at a time, which significantly increases the query count. Many routers have ".HOME" as a search suffix, so a significant majority of these random 10-character string DNS queries end up at the root under the .HOME TLD. Since these strings are designed to not only be random but also not collide with existing registrations, the likelihood of collision and subsequent risk is low to none.

NXD traffic in 2013 is higher than in 2012 for applied-for TLDs. However, there is *less total traffic* in 2013 compared to 2012. Since the data above is from 2013, and this was after the TLD application window closed and the applied-for TLDs were revealed, its possible some of this additional NXD traffic is due to competitive gaming of the NXD traffic. It may be possible to measure this in the historical data if given enough time (by for example, examining source IP and other patterns in the query data). There are many NXD queries (1.7 million) for SLDs that appear to be random computer-generated SLDs (not including the random Chrome browser 10-letter anomaly, which is much larger) in the 2013 data compared to the 2012 data.

Size of the Existing "Problem"

	During the one day study (approx.) period in 2012	During the one day study (approx.) period in 2013
Number of .COM queries that came to the root for names that are not in the com zone (this is the number of NXD queries for .com coming to the root)	888,569,816	83,565,199
Number of unregistered SLD names that generated the queries above	136,905,037	6,296,463
Number of those names above that produced the NXD queries that were registered within the next month	548,188	58,755
Number of NXD queries those subsequent registered names above represent	28,671,101	1,629,429
Percent of the NXD that was subsequently registered (AND THEREFORE COLLIDED)	21%	26%

Source: Applicant OARC data study (for 2013)

The most significant statistic in the above table is that to reach the "danger" threshold outlined by Interisle, a proposed TLD needs to have received 50,000 queries during the study period. ***.COM received more than 80 million.***

A significant number of domain names in .COM are registered each month that have NXD traffic. In fact, 548,000 names in a month (on average, about 16,000 name collisions per day) is more names than most registrars have under management after 10 years of operation. These names are not being "tasted," since that was banned before 2012. Its very likely that some registrants are specifically targeting .COM names that have NXD traffic, and Verisign is allowing these name to be registered and therefore colliding with "pre-registered users" of these .COM names (those Internet users who use the name and cause the NXD traffic). These NXD names therefore are being specifically targeted for registration. We note that Verisign sells .COM NXD data for this targeting registration purpose.

If these second-level names were to be disallowed, as they were in the proposed TLDs (TLDs were not in the root during the study period), then .COM NXD traffic would *grow* at approximately 1.6 million queries per month (in 2013). Therefore, the NxD query traffic for .COM appears to be low not only due to the TTL, but also due to the fact that these SLDs names are being registered, reducing the NXD load in subsequent months. In other words, the NXD for .COM for the study period would have been even higher if during the previous months Verisign had not allowed NXD .com names to already collide.

.COM queries

Further to the point above, according to the following table, the number of total queries for NXDs decreased between 2012 and 2013, though .COM queries decreased at a faster rate.

	2012	**2013**	
Total queries	44.8	37.9	billion
.com queries	16.0	7.9	billion
.com as a percent of the total	36%	20.8%	

Source: Applicant OARC data study (for 2013)

By allowing these so-called collisions, Verisign collected $4 million in registration fees over the month following the single-day study period, which in fact reduced further collision.

It appears further that during the study period in 2012 that .COM name servers were unusually loaded, as compared to the much lower load seen in 2013. In 2012 .COM had 16 billion queries during the study period, compared to 8 billion (half as much) in 2013.

It is also observed that in 2013, during the study period, the J root received 7% percent of the traffic, but in 2012 the J root had 53%.

Distribution of NXD Traffic for Top-Queried SLDs

There's a smarter way to examine SLD NXD traffic across TLDs. The following table shows each heavily trafficked SLD in order, and the NXD counts for each across ALL applied for TLDs for the approximately one-day period in 2013 using the OARC data. It also shows the TLDs with the highest subdomain traffic.

For example, 99.98% of all NXD traffic that the "Den" SLD produces goes to "Den.ICE". If the .ICE registry blocks "Den" (or otherwise mitigates it) then any harm that "Den" SLD traffic *may* create is mitigated. As another example, the "google-10-characters" (as if they were one single SLD) saw the most NXD traffic compared to all other SLDs, and 97% of that traffic goes to .HOME, and 0.7% goes to .CORP. If those processes are blocked in just those two proposed TLDs (or otherwise mitigated, by Google for instance) then that similarly goes away. Further, 99.9% of the second-level name "sap" NXD traffic goes to ".CORP", while 92% of the SLD "org" traffic (a total of 13.2 million NXD hits) goes to .HOME. Therefore, based on 2013 data alone, blocking those second-level names in these two TLDs will very nearly eliminate 46% of all NXD traffic to **all** applied-for TLDs.

NxD Count for the SLD across ALL applied-for TLDs in the 2013 study period (approx 1 day)	SLD	The TLD where this SLD has the largest NxD traffic	The 2nd largest TLD for this SLD…	3rd TLD…	4th…	5th	6th	7th	8th	9th	10th	11th	12th	13th	14th	15th
		home	corp	cisco	iinet	business	inc	site	box	orange	office	ads	group	global	casa	wow
616,610,970	GOOGLE-10	97.20598%	0.7%	0.2%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
		home	site	cisco	iinet	kpmg	business	smart	samsung	ericsson	mail	off	casa	office	host	unicorn
97,441,158	com	91.85863%	3.1%	0.7%	0.5%	0.4%	0.3%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
		home	corp	site	hsbc	global	ads	inc	group	business	iinet	cisco	smart	star	sew	telefonic
72,107,496	2-CHAR	64.85888%	9.8%	3.9%	3.4%	2.2%	1.5%	1.4%	1.0%	0.7%	0.6%	0.5%	0.5%	0.5%	0.5%	0.4%
		home	casa	box	homes											
20,947,759	hitronhub	99.99768%	0.0%	0.0%	0.0%											
		ice	home	cisco	nab	audi	wow	tech	mail	stream	gal	mobile	site	bosch	gle	men
19,656,489	den	99.97824%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	site	smart	cisco	world	mail	business	iinet	blanco	off	search	samsung	clubmed	office	ericsson
19,407,868	net	79.28173%	12.7%	1.5%	0.7%	0.6%	0.5%	0.4%	0.4%	0.2%	0.2%	0.2%	0.2%	0.2%	0.1%	0.1%
		home	corp	cisco	tech	office	network	global	google	iinet	mnet	bom	casa	samsung	zip	site
19,294,864	INVALID-SLD	64.27320%	3.9%	3.4%	1.0%	0.9%	0.9%	0.8%	0.8%	0.7%	0.6%	0.5%	0.5%	0.5%	0.5%	0.4%
		corp	group	world	red	home	csc	sap	prod	office	dev	cisco	svr	help	box	security
18,341,704	sap	99.96337%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	cisco	iinet	business	office	samsung	casa	acer	dell	comcast	gold	unicorn	wow	network	work
17,839,407	ENDS-PC	93.94764%	2.1%	1.3%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
		home	site	cisco	iinet	mail	business	photos	casa	samsung	office	app	network	dev	orange	wow
13,287,215	org	92.29337%	2.4%	0.6%	0.6%	0.5%	0.4%	0.3%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
		corp	prod	home	bank	center	business	barclays	online	energy	loans	map	cisco	ing	one	dev
11,572,213	bank	99.93451%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		med	home													
10,309,104	ngha	99.99994%	0.0%													
		corp	home													
6,724,845	appriver	99.99996%	0.0%													
		network	ads	home	juniper	page	office	comcast	corp	dev	bar	here	site	iinet	cisco	now
6,669,197	home	91.02283%	4.0%	3.2%	0.4%	0.3%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

		box	home	house	bom	fox	olympus	cisco	srt	foundatio	site	off	book	casa	aol	media
6,553,276	fritz	99.98346%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	site	cisco	iinet	business	orange	casa	samsung	office	wow	ericsson	tiffany	gold	work	off
6,291,326	info	94.19806%	0.8%	0.7%	0.7%	0.5%	0.3%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
		home	iinet	now	zip	cisco	team	live	app	business	cab	yahoo	srt	doctor	mail	wow
6,232,786	3	99.85584%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	cisco	office	casa	site	corp	network	house	philips	family	work	acer	samsung	google	abc
4,373,272	wpad	60.01424%	7.1%	1.4%	0.8%	0.8%	0.8%	0.7%	0.5%	0.5%	0.5%	0.5%	0.4%	0.4%	0.4%	0.4%
		win	home	iinet												
4,314,744	detnsw	99.99875%	0.0%	0.0%												
		sap	ads	star	prod	home	exchange	network	terra	web	telefonic	mit	corp	office	life	dev
4,312,075	corp	39.22221%	38.2%	13.5%	2.2%	1.9%	0.6%	0.6%	0.6%	0.5%	0.4%	0.4%	0.3%	0.3%	0.3%	0.2%
		home	cisco	office	casa	iinet	corp	network	hotel	gold	business	samsung	philips	house	work	orange
4,226,477	isatap	71.69562%	5.4%	0.9%	0.7%	0.7%	0.5%	0.5%	0.4%	0.4%	0.4%	0.4%	0.3%	0.3%	0.3%	0.3%
		youtube	google	you	goo	goog	yahoo	home	baidu	gmail	hot	tech	amazon	college	hotmail	med
3,360,787	www	15.87955%	15.1%	11.2%	5.3%	3.3%	2.6%	2.1%	1.8%	1.7%	1.6%	1.5%	1.4%	1.4%	1.3%	1.0%
		corp	home	ecom	global	airtel	cal	bom								
2,998,139	ecolab	99.99857%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%								
		corp														
2,825,696	bvcorp	100.00000%														
		hsbc	prod	home	cal											
2,756,516	adroot	88.24219%	11.7%	0.0%	0.0%											
		corp	prod	cisco	home	wiki	web	bet	cam							
2,683,666	sungard	99.04299%	1.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%							
		corp	global	bet	ads	group	zone	home	team	top	clinic	roma	dev	how	zip	prod
2,586,550	root	57.89179%	36.3%	2.4%	1.3%	0.7%	0.6%	0.2%	0.2%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
		corp	home	family	dev	mail	vip	new	city							
2,532,228	teva	99.94321%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%							
		corp	dev	home	ibm	build	prod	call	zone	iinet	camp	bom				
2,455,899	zurich	76.41857%	23.5%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%				
		corp	home	supply	shop	iinet	work	orange	cam							
2,420,139	airbus	99.98529%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%							

		host	home	mail												
2,297,250	theplanet	99.99443%	0.0%	0.0%												
		home	ads	host	dev	cisco	smart	iinet	site	mail	office	network	corp	academy	show	active
2,201,453	local	53.55031%	29.7%	2.8%	1.9%	1.7%	1.1%	0.7%	0.6%	0.6%	0.5%	0.5%	0.5%	0.4%	0.4%	0.3%
		corp	home	healthcai	web	car	and	cisco	iinet	blog	zip	office	business	movie	house	you
2,162,699	parker	99.88426%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		corp	home	group	bmw	acer	dev	world	business	moto	iinet	car	vegas	club	cisco	auto
1,840,722	bmw	99.79687%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		corp	network	hsbc	home	global	house	gold	next	prof	red	inc	consultin	anz	law	medical
1,662,918	global	56.80262%	38.0%	3.8%	0.4%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
		corp	home													
1,599,425	compassgroup	99.99994%	0.0%													
		home	movistar	anz	office	box	corp	swiss	cisco	red	roma	vip	network	casa	iinet	now
1,508,834	router	55.65556%	41.8%	1.6%	0.1%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		corp	auto	home	prod	games	wiki	game	sky	site	xn--80as	downloa	business	style	orange	arab
1,373,477	gta	99.75005%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	iinet	site	cisco	global	tiffany	casa	wow	business	work	ericsson	network	gold	office	host
1,357,528	biz	93.91556%	0.8%	0.7%	0.7%	0.6%	0.6%	0.5%	0.4%	0.3%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%
		ltd														
1,276,736	pcare	100.00000%														
		corp	home	volkswag	app	pid	world	online	med	jmp	group					
1,250,056	atd	99.98240%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%					
		corp	zone	home	network	iinet	terra	host	cloud	mail	office	roma	law	mit	support	work
1,081,347	internal	91.76472%	2.9%	2.4%	1.2%	0.3%	0.3%	0.2%	0.2%	0.1%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%
		dev	cisco	home	house	xn--55qx										
1,055,272	kft	99.99706%	0.0%	0.0%	0.0%	0.0%										
		home	ice													
1,032,112	corega	99.99990%	0.0%													
		prod	home	chrome	taipei	cisco	red	today	clinic	business	work	llp	top	tata	off	sport
972,283	ad1	99.56000%	0.2%	0.1%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

		corp	office	home	dev	inc	mtn	google	etisalat	olympus	mail					
970,597	hsi	99.71152%	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%					
		global	home	moscow	dog	fox	business	box	pub	mail	college	you	mobile	iinet	black	xn--55qx
963,995	grey	99.81815%	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		mail	services	data	security	manager	web	design	home	matrix	ads	oracle	iinet	map	madrid	link
923,303	system	73.57910%	11.4%	2.8%	2.5%	2.3%	2.1%	1.5%	1.5%	1.3%	0.3%	0.1%	0.1%	0.0%	0.0%	0.0%
		corp	home	global	mail	ram	cisco	foo	bet	inc	page	services	mobile	gmbh	docs	dev
922,991	logistics	99.92947%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
		home	orange													
916,472	flybox	70.96562%	29.0%													
		corp	home	film												
879,392	sanm	99.99898%	0.0%	0.0%												
		corp	web	home												
866,037	hospira	99.99954%	0.0%	0.0%												
		home	sfr	cisco	business	site	network	map	iinet	itau	icbc	orange	house	wow	sbi	best
842,602	arpa	88.65265%	3.8%	1.5%	1.1%	0.5%	0.4%	0.4%	0.3%	0.3%	0.3%	0.2%	0.2%	0.2%	0.2%	0.1%
		global	star	network	home	monash	cbs	abc	training	rmit	wtc	open	bcn	life	loans	loan
834,738	student	96.15856%	1.6%	0.5%	0.2%	0.2%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%

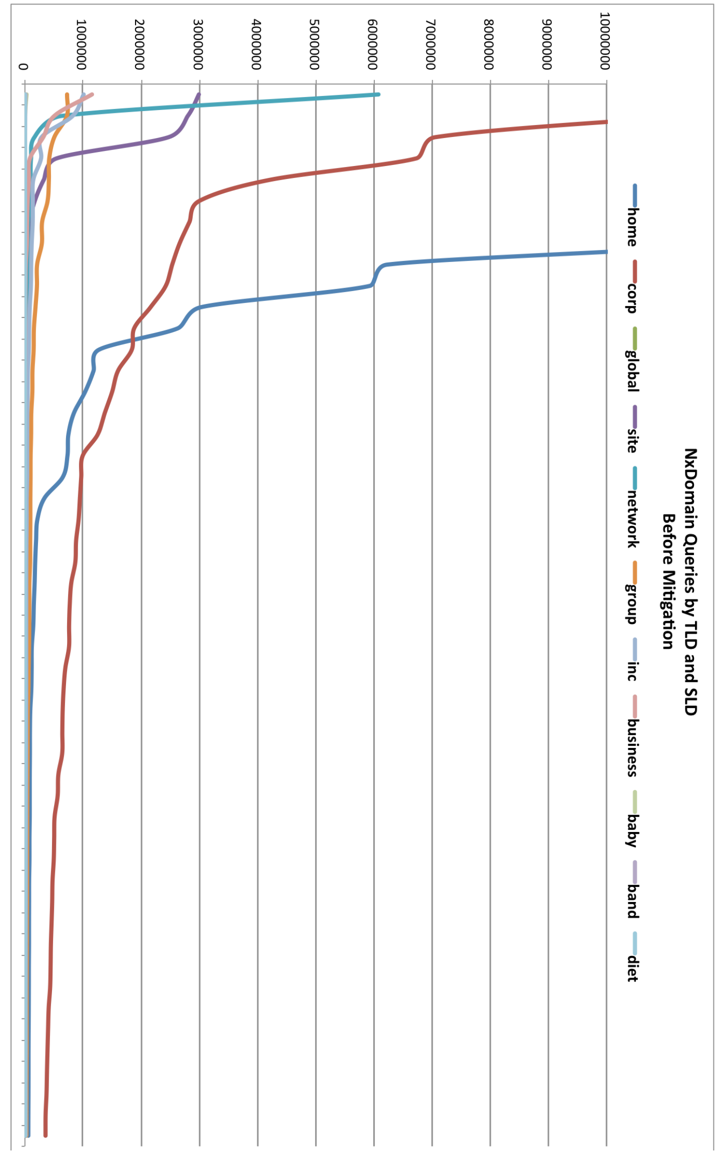
NXD Queries By TLD and SLD Before Mitigation

The following graph shows a selection of 11 TLDs for which Donuts has applied. The TLDs represent the top eight—by rank of NXD queries in 2013—and the bottom three (but still in the Interisle "20%") by the same Interisle ranking.

The NxD for each TLD is plotted in order of highest trafficked to least trafficked SLD, for the top 50 SLDs in each TLD. For example, "sap" is the highest trafficked SLD in the ".CORP" TLD, followed by "Bank.CORP". We also grouped SLDs, in effect, to comprise one SLD. These grouped SLDs are in all caps below the graph. For example, all two-letter SLDs into "2-CHAR", and the 10-random-character Google/Chrome SLDs into "GOOGLE-10". We've done so because either the entire group is already blocked from registration, such as 2-CHAR, or easily blocked such as GOOGLE-10. As you can see, the NXD traffic for the top TLDs are heavily weighted toward the very few head SLDs. Blocking these SLDs mitigates any potential issue until other mitigation methodologies are employed (such as source IP mitigation).

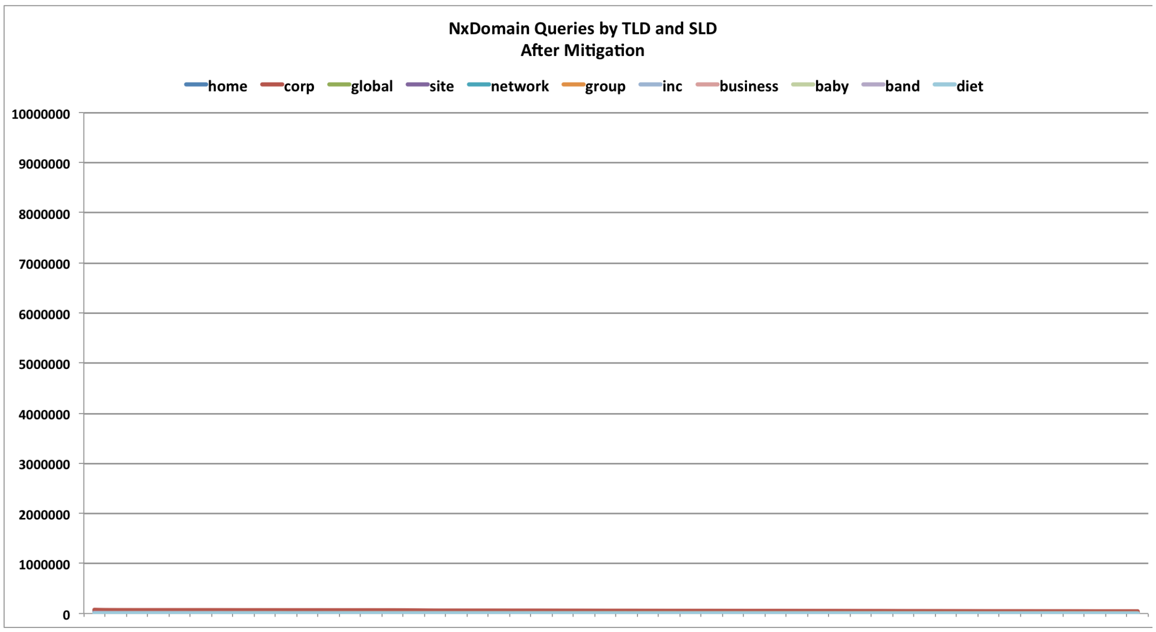
We comprised the following groups of blocked SLDs:

1) Invalid SLDs, such as TLDs that contain an underscore like "srv". These are already blocked.
2) ASCII Two-character, such at "uk". These are also already blocked by ICANN regulations.
3) All TLDs already reserved or have special rules/control by ICANN, for example, "france", or "example"
4) All SLDs that are prohibited as TLDs, for example "local", and "arpa"
5) All TLDs that are major TLDs, such as ".com" and ".org"
6) A list of about 32 SLDs that are technology terms, such as "router", and "lan"
7) A list of major brands that are heavily trafficked across all applied-for TLDs, such as "sap", "google", and "msn"
8) All SLDs that are a number less than or equal to 255, such as "10"
9) A small handful of SLD strings removed on a TLD-by-TLD basis.



TLD	.home	.corp	.global	.site	.network	.group	.inc	.business	.baby	.band	.diet
SLD 1	GOOGLE-10	sap	2-CHAR	com	home	2-CHAR	2-CHAR	GOOGLE-10	INVALID-SLD	master	GOOGLE-10
SLD 2	com	bank	grey	2-CHAR	global	winshbos	GOOGLE-10	2-CHAR	GOOGLE-10	engeband	com
SLD 3	2-CHAR	2-CHAR	root	net	INVALID-SLD	essilor	psi	com	wpad	GOOGLE-10	wpad
SLD 4	hitronhub	appriver	student	GOOGLE-10	GOOGLE-10	pernod-ricard	mediatek	net	www	www	m
SLD 5	ENDS-PC	GOOGLE-10	draeger	org	mtsintra	winsltsb	services	INVALID-SLD	search	harrycallagha	2-CHAR
SLD 6	net	ecolab	hkmg	rsag	2-CHAR	sorun	leviton	org	isatap	4dmusicsk	isatap
SLD 7	INVALID-SLD	bvcorp	kcg	INVALID-SLD	private	GOOGLE-10	quala	ENDS-PC	com	INVALID-SLD	asyapc12
SLD 8	org	sungard	ttg	site	mtsfe	tuv	cbr	info	2-CHAR	ipiranga	www
SLD 9	3	teva	magna	google	gma	mapei	smarttech	isatap	baby	cubatao	INVALID-SLD
SLD 10	info	airbus	nexansdomai	info	mercury	giordano	gome	arpa	isearch	corridordigita	org
SLD 11	isatap	parker	GOOGLE-10	edu	baleno	tkse	soteha	oracle	server	jamieruddick	ENDS-PC
SLD 12	wpad	zurich	rockwool	wpad	com	galenica	mti	wpad	pounds	stafa	diet-p4
SLD 13	biz	bmw	frd	zenworks	wpad	corporate	anapharm	local	ENDS-PC	wpad	forum
SLD 14	local	compassgrou	blackbaud	rchk	moralenet	fnac	innovatia	biz	casa002	ryantaubert	kimkins
SLD 15	corega	root	olympus	linux	corp	rcs	INVALID-SLD	unifi	a	top	net
SLD 16	router	gta	nbty	psbe	eastern-tele	luxottica	ahc	serwer	zion	willshare	video1
SLD 17	arpa	atd	bax	zet	foster	ray	apptis	www	honey	charlieparrad	otc
SLD 18	retracker	internal	skylines	gs08proxy	welzorg	ecobank	tecta	sbsmen	mbit	icegrillsjp	comalerts
SLD 19	flybox	hsi	lgdisplay	local	chemeketa	enersys	crossview	edu	comwww	paranagua	health
SLD 20	edu	global	INVALID-SLD	dcsr	atempo	s-net	upwindsolutic	biuro	saandy	i	vivinter
SLD 21	home	logistics	ssnc	mis	mcc	amorg	q1labs	irac5030-eg	happybaby	bootybakery	health-2145
SLD 22	gov	sanm	laerdal	itellium	holdings	cicc	health	netopsie	f	2-CHAR	stb
SLD 23	server	hospira	enelint	localhost	isatap	ktn	heartland	server	proxyserver	bang-sbs	facebook
SLD 24	none	sdl	applied	biz	ipex	bulgari	gilson	dhcppc0	hp552857	projectr	psy-ab
SLD 25	http	cpwr	cst	gov	superfos	ima	phf	http	forum	isatap	louisvuittonb
SLD 26	mobi	ini	spoc	pecem	tms-internal	arcelik	siatech	sbs2008	user	andrewmfilm	info
SLD 27	html	INVALID-SLD	ads	schule	cadg	telekomaustri	zotos-adnet	tpr	susi	wearecanvas	proof
SLD 28	setup	quest	nobianet	http	core	bonfiglioli	zogenix	cfgllc	you	julianaricherd	forumup
SLD 29	lan	rackspace	crucell	tct	jagex	grontmij	comgmkt	lan	net	fytch	best
SLD 30	unifi	longtai	gov	hamel	cdr	schnellecke	mlt	sekretariat	shopping	monsterrally	verizon
SLD 31	bthomehub	sealedair	xtralis	isatap	internal	kier	fosbel	user-kompute	save	cultclassicrec	tips
SLD 32	music	abg	kordsa	kaukokiitodor	net	rbi	alloy	dhcppc1	vicchou	ashellinthepit	special
SLD 33	c	main	allseas	servidor	aravo	daisy	sidlee	komputer	home	homestuck	life
SLD 34	mortgage	stream	unitedlex	www	guest	gpm	electrocraft	ania	eeyore	tehearthcamp	judith
SLD 35	corp	delta	nktcables	dialyserbogot	composite	seetec	enterasys	admin-kompu	selina	lainerec	http
SLD 36	google	hrc	rud	customer	sd33	banqueatlanti	standard-insu	rse	edu	joverdose	energu
SLD 37	livebox	joy	easynet	umantis	unic	promatic	fcbc	business	love	com	com173
SLD 38	www	davita	mclagan	domgb	bfg	leman	rideau	laptop	gabriela	news	x
SLD 39	null	dci	sapmnt	tnt	org	aoig	komag	qandam	city	robfel	obat
SLD 40	facebook	sys	atea-et	model10	amcomp	attends	idealease	biuro1	e69447m1f17	thehumbletyr	johannes-ottc
SLD 41	name	iusacell	metito	br0	qatar	allyes	auntmarthas	jesco	babyo	drfrankenrhyr	heirloom-hp
SLD 42	zonhub	quanta	ssg	bin	local	viscofan	yorktowne	html	sex	fredyaddaden	hehe
SLD 43	f	enr	horibaabx	drcee	aacilocal	microgen	fitipower	dhcppc2	otc	ENDS-PC	g
SLD 44	m	astrium	recagroup	sensenet	nacg	ferretti	com	et0021b7a275	ymail	sddx	freyhans
SLD 45	registry	hosting	nex	sanappl1	artix	jss	acumends	agnieszka	saves	thatkidfromru	fatloss4idiots
SLD 46	corppki	thrivent	vax	cron	calibre	tereos	spgroup	lx735	luan	mugaritz	dukan
SLD 47	workgroup	onefirm	novenco	dh1la	ocp	simatelex	foseco	dhcppc4	sipinternal	grandmaster	dongs
SLD 48	int	tmrk	ipte	rcp	lwlchp	planit	bjs	google	sip	rad1	com208
SLD 49	asia	tsol	jebsen	aero	hello	kinepolis	ftt	dell	mediaserveri	attention	alf

Note not ALL of the above SLDs need to be blocked in each and every TLD to provide the “after” NxD mitigation show below:



TLD	.home	.corp	.global	.site	.network	.group	.inc	.business	.baby	.band	.diet
SLD 1	aninvalidpoi	drtn	freynet	pecem	artix	monteco	ltd	paulina	search	engeband	asyapc12
SLD 2	intra	lasercard	hites	schule	calibre	bit	axcient	npiec9a7464	baby	harrycallagh	diet-p4
SLD 3	https	csodsandbo:	shared1	http	ocp	jetta	svi	andrzej	isearch	4dmusicsk	kimkins
SLD 4	pro	rel	tzmosa	tct	lwlchp	nova	ipsoftcom	renata	server	ipiranga	video1
SLD 5	tracker	broker	torm	hamel	hello	marmedsa	allos	eris-tor-gru-	pounds	cubatao	otc
SLD 6	br0	beauty	savona	kaukokiitod	shop	mprise	beckwithele	npi0026551!	casa002	corridordigit	comalerts
SLD 7	user	rosinv	nworks	servidor	binder	pacorini	flc	biuro3	a	jamieruddicl	health
SLD 8	nohost	amlaw	american-ap	dialyserbogo	gbi-ad	greenflex	glaciermedia	a	zion	stafa	vivinter
SLD 9	iphone	vfa	comtek	customer	united	sphere-root	casious	ricoh	honey	ryantaubert	health-2145
SLD 10	e	srggi	osf	umantis	stbedes	gigatms	base	desktop	mbit	top	stb
SLD 11	desktop	grupoq	sbw	domgb	ceomelb	ihk	atr	labor09	comwww	willshare	psy-ab
SLD 12	windows	domain	jsapdc	tnt	solaris	womcorp	systar	grzegorz-b4	saandy	charlieparra	louisvuittonl
SLD 13	dhcppc0	uafc	stage	model10	webspirit	icp	tempress	tt-03	happybaby	icegrillsjp	proof
SLD 14	admin	dorseylaw	proferoeuro	br0	lovebrand	maau	rsw-mtl	et0021b76a	f	paranagua	forumup
SLD 15	invalid	orklanet	michaelhill	bin	summit	lgb	nsi	allegro	proxyserver	i	best
SLD 16	mail	hymall	asesal	drcee	smts	isk	cbsa	zaplecze	hp552857	bootybakery	tips
SLD 17	images	tatenergo	outatime	sensenet	bristol	nap	presidium	sklep1	user	bang-sbs	special
SLD 18	computer	eisenmann	fgol	sanappl1	s	choueiri	gic	rafal	susi	projectr	judith
SLD 19	gmail	cvstarrco	stat2k	cron	stlorenz	fhp	medtox	et0021b706	shopping	andrewmfilr	http
SLD 20	pc2	dipcmi	pallomaro	dh1la	elas	finalyse	ampagency	beta-kompu	save	wearecanva	energu
SLD 21	server1	csodmgmt	gsnhost	rcp	yakult	rinet	culligan	wojtek	vicchou	julianaricher	com173
SLD 22	lifbcibllhkdh	airwatch	ggn	eth0	3dents	eckelmann	sovran	videos	home	fytch	obat
SLD 23	i	cfs2	hcpexchange	bogon	nhanhoa	bma	guajome	rejestracja	eeyore	monsterrally	johannes-ot
SLD 24	laptop	dealerdotco	belmay	enel	johnson	duzey	coral	emil-6be0be	selina	cultclassicre	heirloom-hp
SLD 25	error	neos	isotron	rogesa	biosound	linney	xrt	maria	love	ashellinthep	hehe
SLD 26	videos	247wpp	gtstar	sanappl3	yoa	bradmanlak	seaflex	lukasz	gabriela	homestuck	g
SLD 27	activationbb	imax	bpmgrp	gollmann	austwide	jts	mtl	leszek	city	tehearthcan	freyhans
SLD 28	n	tecsys	nexandomai	farsa	poli	maus	jla-usa	darek	e69447m1f1	lainerec	fatloss4idiot
SLD 29	b	cegos	enders	eth1	andu	magr	amobee	npi85e159	babyo	joverdose	dukan
SLD 30	servidor	meus	nexus-eleve	web	szrcs	pacifica	anabolic	jola	sex	news	dongs
SLD 31	xxx	raccorp	cnsme	ftp	pinnacleschc	crw-plaza	sbliusa	drugi	otc	robfel	com208
SLD 32	icongo	fpprod	hopheat	your	gilford	engtek	preload	dhcppc5	ymail	thehumblety	alf
SLD 33	s	buyabs	testgroup	claisse	insyndia	emtec	training	oem-kompu	saves	drfrankenrh	untuk
SLD 34	namics	bps	iittala	public	luminar	tat	zygatech	pc1	luan	fredyaddadc	reduced
SLD 35	ieeb	stibo	dstgs	kelowna-1	scexpress	kras	certifiedpow	kierownik	mediaserver	sddx	poor
SLD 36	dhcppc1	3pillar	chemical	building	bestthinking	minster	tec	kasa1	club	thatkidfrom	pnqheuyxxa
SLD 37	a	canternet	integralsys	adlsrv	bearcountry	being	piidea	stanowisko3	old	mugaritz	paleo
SLD 38	ufi	nmhgext	pronorm	rxddc	prefmost	lts	infovue	krzysiek	kitty	grandmaster	ofkqmlgo
SLD 39	hotmail	ncic	sede	lincoln	motion	dback1	ucfaa	place	greg	rad1	noakes
SLD 40	mybookwor	grupomag	mail	printer-srv	nwls	britannic	rhp	npief6245	cute	attention	nbaonline
SLD 41	wdtvlive	price	ulusoy	intracentre	student	partners	chanluu	its	fe-frontend	lemmino	mail
SLD 42	ipad	alior	nexus-maxir	bideco	sau-58	hollard	saunatec	et0021b7aa	zero	blog	jon
SLD 43	com0	ddefue001	naclpbiofl5	sp8	opegieka	nsg	cine-chroma	ania-kompu	premature	elninomusic	i
SLD 44	printer	associa	rhjint	lkvr	uks	icl	altmangroup	baza	ratanzs	email	hh1
SLD 45	d	freetranslati	erowacn	profab	lex	uniteam	surespan	firma	music	nobleoak	helmut
SLD 46	user-hp	hcmdomain	lila-logistik	home	bvnelas	sellner	spindance	ola	bindi	kickthepj	denismen
SLD 47	diskstation	cardone	ipcore	gmail	mtcwork	ruf	cwe	mirek	bb4	jayt	day
SLD 48	u	amscan	cgigroup	tracker	barnabas	gknplc	lloyd	grazyna	baby-kompu	setia	com217
SLD 49	php	wrgd	ucic	tsg	jgh	caribbean	barberind	ewa	vista	brickstudio	com215

As you can see the NxD traffic would be mitigated if the above nine types of SLDs were blocked across all TLDs. Note that the new "high-count" SLDs are now very esoteric strings with very low NXD traffic.

The .CORP and .HOME Examples

The applied-for .CORP and .HOME gTLDs have been singled out as particularly (potentially) troublesome. Unfortunately, the Interisle report did not examine *second-level names* for either of these gTLDs, which is more likely the source of any issue.

Consider the usage of these names by brand companies. For example, someone searching for Airbus.CORP may have collided already with another TLD, as another TLD could have been in the search suffix list. The existence, therefore, of .CORP *benefits* Airbus by preventing the what may already be leakage of traffic. End-user search via .CORP demonstrates demand for the name's usage even before its delegation to the root.

New gTLD traffic is similar to NXD traffic in .COM, .UK, and other existing TLDs. In the case of .CORP, some networks append this name in the event of an unsuccessful resolution. For example, if Google.COM weren't to resolve, even temporarily, a network could try Google.COM.CORP.

However, the reverse is not true. If a user typed in Gougle.COM, intending to reach Google, the community does not forbid such names from being registered or receiving NXD traffic. If .CORP were in the search suffix list, then Google.COM.CORP must resolve for any harm to occur Even then, as we said in our prior comment, if the Com.CORP registrant were prohibited from obtaining a certificate, a secure connection to a destination that was unintended would not occur.

In the instance of the term "corp," the Corp.COM registration exists. As its registrant admits, it generates significant error traffic, likely due to ".com" (or "Corp.COM") being in the search suffix list for a large number of users. If Corp.COM were to be deleted, Verisign would allow it to be registered again. This significant error traffic (that would otherwise be NXD traffic) to Corp.COM is not causing harm.

Subdomains, Source Traffic, and Brands

Further to the issue of second-level registrations, the careful blocking of but a few are a viable and effective mitigation for potential collision issues.

Blocking certain second-level names is effectively the same as not inserting the "problem" gTLD into the root, as it relates to collision. The registry simply replies to a query with the NXD as it would now.

In the instance of .HOME, setting aside the "Google-10", many of the second-level names that would cause collision are already disallowed by ICANN, including two-letter names, names with underscores, and certain terms.

Corollary to this point, it's critical to examine the source of NXD traffic to ensure the collision issue is not inappropriately overstated. For example, as Google itself will document, the Chrome browser defaults to a random 10-letter domain name search in the event of NXD entry—this accounts for over a third of all NXD traffic. If this issue is fixed, a significant percentage of such traffic will cease.

Brands cannot be discounted in this equation. ICANN rules already provide significant difficulty for non-brand holders to secure a branded second-level name that potentially could collide, thanks to the significant trademark abuse mitigation tools in place in new gTLDs that don't exist in incumbent gTLDs and ccTLDs. Certainly, it benefits the brand owner to have their names available to them, and to receive the traffic, rather than have them banned from registration in these gTLD outright.

Recommendations

Donuts recommends the following to address name collision:

- On a TLD-by-TLD basis, disallow a small list of second-level names that receive the disproportionately highest share of collision traffic, if not already blocked by ICANN or otherwise mitigated.
- Request that browsers stop generating random lookups (e.g., random 10-letter names).
- On a TLD-by-TLD basis maintain the ban on two-letter names for a period, or until the NxD traffic is mitigated at the second-level for these names
- Maintain the ban on names with underscores and other punctuation.
- Do not attempt to "instrument" the root by delaying new gTLDs to monitor collision—the data on collision clearly is available.
- Collect sources of problem traffic and address the issue directly with ISPs.
- Don't set up MX records for certain second-level names (e.g., some SLDs in .MAIL) until that is mitigated by other means.
- Disallow a few Internet terms (e.g., POP3, HTTP, etc.) to mitigate such NxD traffic in some TLDs until that NxD is mitigated by other means.
- Disallow existing TLDs as second-level names (e.g., COM, NET, BIZ, INFO) in certain TLDs for a period.
- Name collision mitigation must be addressed directly with applicants and not decided in a vacuum by the Board. Each TLD will have individualized needs for any perceived mitigation.

Conclusion

Our conclusion from the data is that no applied-for TLDs need mitigation, with the possible exception of a very few.

Had Interisle enjoyed the luxury of additional time, its data perhaps would have been more complete and illustrative.

Name collision, however, exists and will continue to exist in legacy TLDs. It does not cause a problem, though the volume is far greater in these existing TLDs than what could be expected in new gTLDs. There is no evidence that points to a need to limit or slow new gTLD development, and this is confirmed by extensive Initial Evaluation activity for security and stability. The Internet can handle the introduction of new TLDs, even TLDs with existing NXD traffic, just as it has done competently in the past.

As Paul Mockapetris stated, technology is not supposed to limit choice. Donuts respectfully requests the Board to honor this tenet of the new gTLD program.