Rapid Growth in Top Level Domains in the Domain Name System

SEM5720 - Assignment 1

Alexander D Brown (adb9)

Contents

1 Introduction

In recent years, the number of Top-Level Domain (TLD) in the Domain Name System (DNS) has been increasing. This is, in part, due to the introduction of Internet Corporation for Assigned Names and Numbers (ICANN), the aim of which was to promote competition in the registration of domain names.

Before the creation of ICANN in 1998 there were a total of eight generic Top-Level Domain (gTLD); TLDs which are not specific to a country or the infrastructure of DNS. These eight were intended to have specific uses (`com` for commercial entities, `gov` for government organisations, etc.) but since then a proportion of these have become truly generic.

In 2000 and 2004 ICANN successfully applied for and instated fifteen new gTLDs and have since gone on to create a program which reviews all applications for new gTLDs[1].

2 Expansion of TLDs

At time of writing there are a total of 364 TLDs, of which 295 are country-code Top-Level Domains (ccTLDs), 42 gTLDs (3 of which are restricted use), 15 sponsored TLDs, 11 used for testing and 1 (`.arpa`) dedicated to the infrastructure for DNS[2].

There are two factors which have spurred the growth of TLDs:

1. Introduction of new gTLDs by ICANN
2. Support for non-Latin characters in the DNS, allowing for Internationalized Top-Level Domains (ITLDs)

Until October 2009[3] TLDs could only consist of US-ASCII (or "Latin") characters. This changed with the approval of the "IDN ccTLD Fast Track Process"[4], which allows countries to apply for Internationalized Domain Name (IDN) ccTLDs which represents their specific country (or territory) name in non-Latin script enabling users to access domain names in their own language.

In more recent years, 2011 to be specific, IDNs have also been approved for use in gTLDs, allowing for a greater degree of freedom in international users.

In hand with this there has been a fair increase in the number of gTLDs before 2010 there were only 22 gTLDs available. On 12 January 2012, applications for new gTLDs opened and since then a total of 20 new gTLDs have been added, including several IDNs.

Even so, this is a drop in the ocean compared to the 1,930 applications ICANN received and it is anticipated that the number of gTLDs will increase further in coming years.

3 Non-technical Support of the Expansion of TLDs

There is much support in favour of the expansion of TLDs, much of which is non-technical, especially in marketing and branding. The ability to have self-descriptive domain names is obviously appealing to businesses, especially those with well known trademarks.

Berneke[5] gives ten reasons as to why gTLDs are important to businesses, explaining the notable factors such as the ability to have self-descriptive domains which clearly identify the area of a business.

There is the counter argument that TLDs are becoming obsolete[6], tools like search engines and increasingly "intelligent" web browsers are reducing the need for domain names. Though this is obviously true, there is still a need for a human-readable name to access websites; the machine-readable IP addresses are not memorable, or short, enough to easily distribute. Domain names also play a significant part in Search Engine Optimisation (SEO), although according to Google experts, the new gTLDs will not have any kind of preference over existing TLDs[7].

Another important factor Berneke details the ability to target a international demographic through the new non-Latin IDNs. For those demographics which use a non-Latin script, this is a vital element to reach the whole of the target market.

With the addition of new TLDs comes with a greater range of choice of domain names. Before this there were only 22 gTLDs and many of these are restricted to specific institutions and organisations, leaving only a handful of domains remaining for non-ccTLDs domains.

This has led to the overcrowding of the more popular gTLDs; `.com`, `.org` and `.net` which has, in turn, ensured that desirable domains are either sparsely available or expensive to purchase. The arrival of these new gTLDs is expected to alleviate at least some of this overcrowding, but the author expects that the popular domains will still suffer from this as entrepreneurial people snap up these domains at a cheap price and sell them when the market prices are higher.

A lot of major search engines are now placing increasing importance of geolocation[8] and although a lot of this can be expressed through ccTLDs there are some proposed gTLDs which will encapsulate more specific geolocation through city information with TLDs such as `.london` or `.nyc`. Of course the restrictions on this would have to be fairly tight and it may be almost impossible to police. Search engines are therefore likely to take the TLD information with a pinch of salt, but it may open up the market for products relating to country information, Cologne for example.

In hand with this, the ability to register a TLD for a professional career, a doctor for example, seems like a ideal way of stopping false advertising. These could prove to be easier to police, but again there is no guarantee that a person owning a domain with the TLD `.doctor` is indeed a doctor.

For businesses that can help perform segment focusing, isolating the target demographics to those TLDs which appeal to them, be it art or football.

Even with this, there is already a lot of misinformation and ignorance as to how the internet really works in the general population that these gTLDs may not be effective as hypothesised.

4 Technical Issues in the Expansion of TLDs

There have not been many technical issues with the root zone of DNS until fairly recently; two factors had helped in the stabalisation of this[9]:

1. There were a fairly small number of TLDs, allowing the size of the root to be around 80,000 bytes.

2. TLDs were absorbed very slowly into the root zone and these changes were relatively small.

Now, with the introduction of new gTLDs and ITLDs, there are going to be many more root zone entries and this increase is happening quite quickly.

Not only are there going to be more TLDs required, additional factors like the new IPv6 protocol (discussed in subsection 4.2) and the need for security in DNS through Domain Name System Security Extensions (DNSSEC)[10] (discussed in subsection 4.1). All these factors combined means that the size of a DNS query or response may not fit inside a 512 byte User Datagram Protocol (UDP) packet DNS was initially intended to handle.

There are several options to cope with this; one approach would be to send a UDP packet greater than 512 bytes, another would be to use Transmission Control Protocol (TCP) to handle the sending of packets which are larger than 512 bytes. DNS itself is able to handle both of these options, but neither are particularly desirable options.

Sending large UDP packets over a network may run into the problem of requiring fragmentation if size of the path Maximum Transmission Unit (MTU) is less than the size of the packet needing to be sent.

The other option, using TCP, has a large amount of overhead involved, both in terms of packet size in adding the information required to perform a connection in TCP and also in terms of network traffic for the three-way handshake and for acknowledging received packets.

Both these methods bring about problems when encountering firewalls. An incorrectly configured firewall might be set to not accept DNS packets with a size greater than 512 bytes via UDP (the National Institute of Standards and Technology estimate that DNSSEC will increase the DNS response size to be *at least* 2048 bytes), so a response from a DNSSEC signed response would be dropped by the firewall.

Another exchange would happen, with the resolve implicitly stating a buffer size of 512 bytes, but as the response cannot be reduced to this size, the root server would fragment the response into this size and set both the fragmented and truncated flags, to encourage resolver to retry with TCP.

The worst case outcome from this is that the misconfigured firewall would also drop a TCP DNS packet and query would remain unresolved signed DNS responses and will have caused a lot of traffic on the network.

4.1 DNS Security

In recent years, the need for security in the DNS has become apparent, simple attacks like DNS Spoofing, where the response from a DNS query is not for the correct server, but instead for a different server, typically the attackers. One method of doing this is DNS Cache Poisoning[11], where the principal of saving lookup time by caching results leads to the wrong result being cached for as long as the Time To Live (TTL) of the cache. This even affected servers from the root zones when a name server provides both an authoritative and recursive name service, where an attack on the recursive side would lead to bad data given to computers wanting an authoritative answer; the net result of which is that one could insert or modify domain data inside a TLD.

The formal solution to this is to introduce security to the DNS through public key cryptography to sign the responses given by the recursive DNS lookup, allowing the response to be verified by the client. DNSSEC is the mechanism used to perform this process.

4.2 DNS Support for IPv6

The DNS cannot be easily extended to support IPv6 addresses, since it is assumed by clients that a 32-bit IPv4 address will be returned and not a 128-bit IPv6 address.

A new resource record was implemented in DNS to accommodate IPv6 addresses, the AAAA record, by RFC3596[12].

The main issue brought about by these changes in DNS is the increase of size in the query/response messages due to the four times increase in number of bit required to represent the IP address.

4.3 IDNs and Issues with Unicode

There is also the question of the safety of using a non-Latin encoding (e.g. Unicode) where certain characters may look either very similar or even indistinguishable to one another[13]. This can lead to the spoofing of known domain names which attackers could use for malicious purposes.

Of course, the registration of IDNs at the root zone is restricted to prevent this[14], but that has not meant that the implementation of IDNs has not caused issues with other areas involving non-Latin encoding, The Homograph Attack[15], for example.

References

[1] (2013) New Generic Top-Level Domains: About the Program. ICANN. Accessed 18/12/2013.

[2] (2013) Root Zone Database. IANA. Accessed 17/12/2013. [Online]. Available: http://www.iana.org/domains/root/db

[3] (2013) Internationalized Domain Names. ICANN. Accessed 19/12/2013.

[4] *Final Implementation Plan for IDN ccTLD Fast Track Process*, ICANN, 5 Nov. 2012.

[5] L. Berneke. (2013, Feb.) 10 Reasons Why the New Internet Extensions (new gTLDs) Are Important.

[6] B. Leiba, "The Good and the Bad of Top-Level Domains," *Internet Computing, IEEE*, vol. 13, pp. 66–69, 9 Jan. 2009.

[7] M. Cutts. (2012, 14 Mar.) Accessed 21/12/2013. [Online]. Available: https://plus.google.com/+MattCutts/posts/4VaWg4TMM5F

[8] Linkdex, "The Impact of Ranking Variability on Natural Search," 26 Jul. 2012.

[9] B. Manning, "Infrastructure challenges to DNS scaling," *The Internet Protocol Journal*, vol. 14, no. 4, pp. 9–14, Dec. 2011.

[10] D. Eastlake, "Domain Name System Security Extensions," Internet Requests for Comments, RFC Editor, RFC 2535, Mar. 1999. [Online]. Available: http://www.rfc-editor.org/rfc/rfc2535.txt

[11] K. Davies. (2008, Oct.) DNS Cache Poisoning Vulnerabilities: Explanation and Remedies. [Online]. Available: http://www.iana.org/about/presentations/davies-viareggio-entropyvuln-081002.pdf

[12] S. Thomson, C. Huitema, V. Ksinant, and M. Souissi, "DNS Extensions to Support IPv6," Internet Requests for Comments, RFC Editor, RFC 3596, Oct. 2008. [Online]. Available: http://www.rfc-editor.org/rfc/rfc3596.txt

[13] J. Kaufman. (2013, 14 Jun.) Is Unicode Safe? Accessed 17/12/2013. [Online]. Available: http://www.jefftk.com/p/is-unicode-safe

[14] (2011, 2 Sep.) Guidelines for the implementation ofinternationalized domain names. ICANN.

[15] E. Gabrilovich and A. Gontmakher, "The Homograph Attack," *Communications of the ACM*, vol. 45, no. 2, p. 128, Feb. 2002.