May 2019
Geoff Huston

Report: ICANN DNS Symposium

By any metric, the queries and responses that take place in the DNS are highly informative of the Internet and its use. But perhaps the level of interdependencies in this space is richer than we might think. When the IETF considered a proposal to explicitly withhold certain top level domains from delegation in the DNS the ensuing discussion highlighted the distinction between the domain name system as a structured space of names, and the domain name system as a resolution space where certain names are instantiated to be resolvable using the DNS protocol. It is always useful to remember that other name resolution protocols exist, and they may use other parts of the domain name space. Having said that, the recent ICANN DNS Symposium was almost exclusively devoted to the name space associated with the DNS resolution protocol, and this protocol.

The DNS protocol represents an inconsistent mix of information leakage and obscurity. When a name resolution query is passed through a forwarder or a recursive resolver the identity of the original source of the query is not preserved. Name resolution is a hop-by-hop process that hides the end user's identity in address terms. At the same time the full query name is used throughout the resolution process, which exposes the end user's DNS traffic in all kinds of unanticipated places. Oddly enough we've seen recent changes to the protocol specification that attempt to reverse the effect of both of these measures!

The anonymity of the end user in DNS queries was compromised with the adoption of the Client Subnet extension. The ostensible rationale was to improve the accuracy of DNS-based client steering, allowing an authoritative name server to respond with the content address that would optimize the user experience. But when one looks at the number of Client Subnet enabled authoritative servers on a country-by-country basis the countries which feature at the top of this list include the United States, Turkey, Iran, China, Taiwan and the United Kingdom. Some 10% of users use recursive resolvers that will add effectively gratuitous client information to the query. It seems that use of the client subnet extension has gone far beyond the original objectives of using the DNS to perform content steerage, as David Dacon pointed out in his keynote presentation to the symposium.

At the same time, we've seen moves to seal up the gratuitous information leaks in the DNS. The use of full query names when performing name server discovery is a major problem in the DNS, and the operators of the root zone servers tend to see a wealth of information relating to terminal names as a result, as do the operators of the top-level domains. The adoption of query name minimization by recursive resolvers effectively plugs that leak point, and the resolver only exposes the precise extent of information that it needs to expose in order to complete the various steps in the iterative name server discovery process.

The EU NIS Directive

The introduction of the GDPR regulations in the EU and the adoption of similar measures in other national environments has gone a long way to illustrate that the Internet's actors are not beyond conventional regulatory purview. There is a relatively EU directive, concerning the operation of "essential services" and the imposition of various requirements on the operators of such essential services, with hefty fines for non-compliance with the measures and also for serious outages of the essential service, as Jim Reid pointed out. The usual suspects of transport, banking, health care, financial markets and similar

services are all part of this measure, but there is the inclusion of digital infrastructure in this directive, which appears to sweep in top level domain registries and DNS service providers. What makes a DNS service "essential" is an interesting question. How to measure such criticality when much of the information is provided in local caches is also an interesting question.

Working out a set of objective metrics to define an "essential" part of the DNS infrastructure seems like a rather odd requirement, but to implement this NIS directive we may see work in this area. In any case the bottom line is very clear. The name space is part of a set of essential public services, and it demands far more than a "best available effort" response by DNS service providers.

Measuring "DNS Magnitude"

If parts of the DNS are considered to be an essential service then we may want to have some kind of metric that measures the use or impact of a domain name, as compared to other domain names. This leads to efforts to measure what has been termed "DNS Magnitude".

The DNS name resolution infrastructure is basically a collection of caches. The whole approach is to ensure that as often as possible your DNS queries are directly answered from a nearby cache. The queries that seen at the authoritative servers are essentially cache misses. This confounds various attempts to measure the use of any domain name. If the name uses an extended cache time (TTL) then the number of cache misses will drop. If the use pattern of a name is highly bursty again the cache will be very effective and the authoritative server will see a small cache miss rate. So how can one use the query data seen at an authoritative name server to measure some aspect of the popularity of a domain name if the effective query rate is so dependent on the name's TTL settings?

The work presented by Alex Mayrhofer of nic.at starts with the assumption that the number of queries is of less value than the number of discrete hosts. He cites the extreme example that 100,000 queries from the same host address is a lesser indicator of domain impact than a single query from each of 100,000 hosts. The basic idea is that if the shared name server sees a certain number of hosts making queries, then the relative magnitude of any particular domain name is the ratio of the number of hosts performing a query for this name as compared to the size of the total host set.

The work uses a log scale to capture details of the "long tail" that exist in such metrics, so the refined metric is the log of the host seen querying for a domain compared to the log to the size of the overall host set. The metric appears to be reasonably independent of TTL settings, but it does assume a wide distribution of DNS recursive resolvers, which appears to be an increasing dubious assumption as the large open DNS resolvers gather more momentum. One can only guess what QNAME minimisation will have on this work, as the query rate would be unaltered by the full domain name is occluded from the upper level DNS servers.

Dark Deeds in the DNS

It is no secret to either the people who undertake dark deeds on the Internet or to those trying to catch them that the DNS is one of the few systems that is universally visible. So, it's no surprise that domain names are used to control botnets. Much time and effort has been spent studying DNS and the ways in which the DNS has been coopted to control malware. Stewart Garrick of Shadowserver presented on the Avalanche investigation, a multi-year law enforcement effort that spanned a number of countries. Some 2.5M domain names were blocked or seized during the investigative process.

There are various forms of blacklists that are intended to help service providers in denying oxygen to digital miscreants. One of these, SURBL, was described at the symposium. It uses a DNS-based reputation database where a client can append the common surbl.org suffice to the name and query the DNS for an A record. If the query returns an address within the loopback address prefix then this dns name has been listed as blocked by the operators of this service.

As Paul Vixie explained, SURBL is a specific instance of a more general approach to Response Policy Zones in the DNS that have existed for many years as a standard for DNS firewall policies. The firewall operates via a DNS zone and firewall rules are published, subscribed to, and shared by normal DNS zone transfer protocol operations. A recursive resolver can be configured to subscribe to a response policy, and resolution operations for firewalled names result in a NXDOMAIN response being generated by the recursive resolver. Implementations of this approach exist for Bind, Knot, Unbound and PowerDNS. More information on this approach can be found at https://dnsrpz.info.

Domain Attacks

Much has been said in recent times about the weakest link in the secured name environment, namely the link between the registrar and the name holder. If this relationship can be breached and unauthorised instructions can be passed to the registrar, which in turn are passed to the registry and make their way into the zone file, then the resources that lie behind the name can be readily compromised by trusting applications. One service operator, PCH, was compromised in this manner, and Bill Woodcock shared some details of the attack process. The subversion of the name matched a local holiday shutdown window. An earlier attack had exposed a collection of EPP (Extensible Provisioning Protocol) credentials. The rogue instructions to change the target's name servers were passed into the system via a compromised EPP credential. With control of the domain it was then possible to obtain a domain validated name certificate immediately, using a CA that did not perform DNSSEC validation, even though the domain was DNSSEC-signed. This then allowed a remote mail access server (IMAP) to be compromised and IMAP account credentials to be exposed, together with mailboxes, and all other material sitting in various mail stores. Because the DS records were not altered in this particular attack, other information that required a validation check on the domain name was not exfiltrated. If the attack had also changed the DS records it may have exposed more assets.

The attack was a well-rehearsed and rapidly executed set of steps, so other defense mechanisms, such as certificate logs ("certificate transparency") offer little in the way of substantive defense here. In this particular case the use of DANE to perform certificate pinning would've been of material assistance, particularly if the TLSA record in DANE referenced the zone's KSK public key, but this particular case was a NS delegation change without a DS record change. Had the attacker also changed the DS record then DANE would not have been helpful. A similar comment can be made about CAA records and other forms of in-band pinning

More generally, if the registrar/customer relationship is vulnerable, then many other aspects of name security are also vulnerable. If the attacker can alter both the delegation records and the zone signing key data in the parent zone, then there is very little for applications to fall back on to detect the attack and correctly identify the new information as bogus. It seems that in today's name environment that registrar/customer relationship is not well protected in many cases, and minimum practices of two factor authentication would be a necessary and practical minimum. The other aspect of such attacks is speed of execution. Deliberately slowing down the process of change of records in the parent zone through registry lock practices does offer some tangible benefit.

As usual there is no magic cure-all defense here, and careful selection of name registrars, coupled with constant monitoring is an essential minimum these days.

DNS over HTTPS

Any DNS meeting would not be complete without extended mention of DNS over HTTPs and the Symposium was no exception. However, I have covered this topic in some detail in recent posts, so I'll skip making any further comment here!

Meeting Materials

The full agenda and presentation materials for the 2019 symposium can currently be found at https://archive.icann.org/ids

Disclaimer

The above views do not necessarily represent the views or positions of the Asia Pacific Network Information Centre.

Author

Geoff Huston B.Sc., M.Sc., is the Chief Scientist at APNIC, the Regional Internet Registry serving the Asia Pacific region.

www.potaroo.net