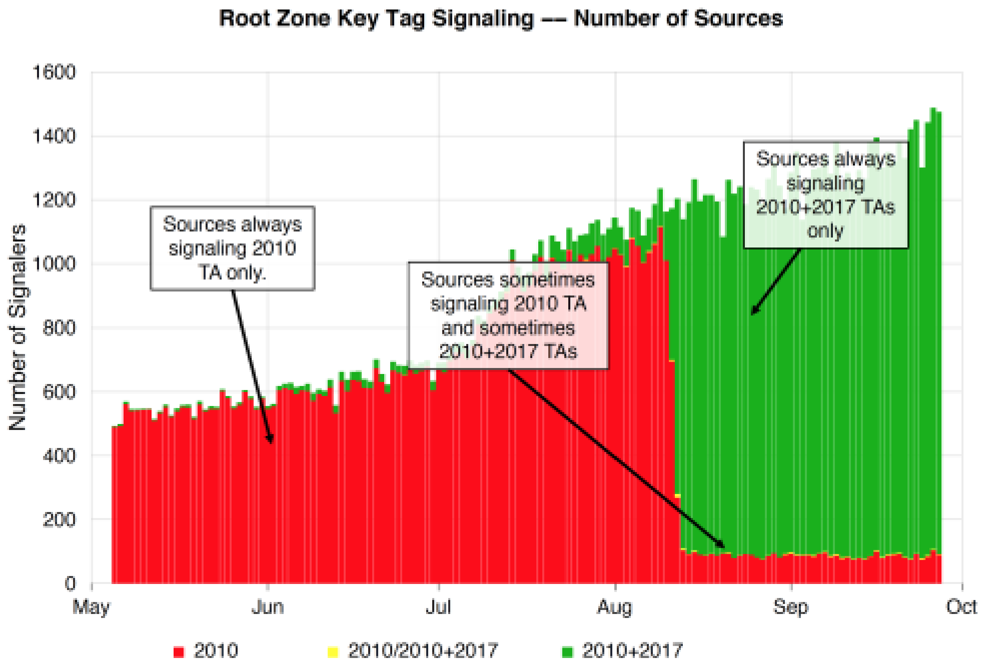
The ISP Column
A column on things Internet

Geoff Huston
October 2017

DNS OARC 27 Meeting Report

The DNS OARC meetings are an instance of a meeting that concentrates on the single topic of the DNS, and in this case it delves as deep as anyone is prepared to go! It's two days where too much DNS is barely enough!

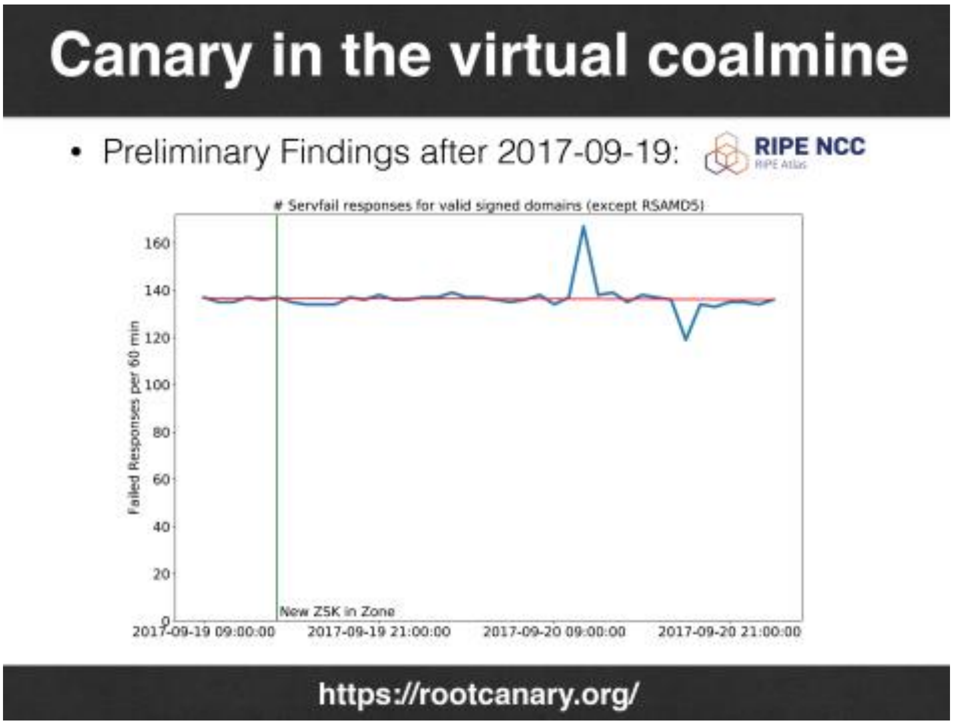
The hot topic of the meeting was the news that the proposed roll of the Key-Signing-Key of the root zone of the DNS, originally scheduled for October 11, was to be postponed. Part of the reason behind that postponement was presented to the meeting by Duane Wessels in his presentation on analysis of the data collected by the recently implemented Trust Anchor Signal mechanism, described in RFC8145. The data set, while small, shows a clear indication of DNSSEC-validating resolvers adding the new key into their local trusted key stash following the RFC5011-defined hold down period, but worryingly also shows a second set of resolvers that have not followed the automatic key roll procedures and still trust only the old key. I have written about this at length in a related article, and I won't repeat it here. (http://www.potaroo.net/ispcol/2017-10/notksk.html)



Duane Wessels - A Look at RFC 8145 Trust Anchor Signaling for the 2017 KSK Rollover

Scott Rose of NIST reported on the profile of DNSSEC used in the .gov domain space. Some 84% of US Federal .gov domains publish DNSSEC-signed zones, while the number drops to 20% if you include the entirety of the .gov delegated domain set. The is still a significant number of zones using RSA/SHA-1, but the larger set is now using RSA/SHA-256. One domain had shifted to use ECDSA P-256, perhaps due to its relatively recent inclusion in the allowed set of crypto protocols by the .gov registrar. There is a significant level of use of NSEC3 in DNSSEC-signed zones, although best practice of changing both the NSEC Salt values and varying the number of iterations in line with KSK changes is not followed very well. Perhaps a description of what constitutes the best operational practice for ZSK management, coupled with Salt and iteration management for NSEC-3 signed zones would be useful.

Moritz Muller reported on the status of the so-called 'root canary' project, a measurement exercise that was intended to closely monitor the progress of DNSSEC-validating during the planned roll of the KSK. They use data gathered active probes using Atlas probes and scripted tests using the Luminati network. What I found interesting is the degree of sample bias, where some 42% of Atlas probes are using DNSSEC-validating resolvers, yet a far smaller proportion of resolvers (7%) are seen with the Luminati tests. Small sample sets (and large sets too for that matter) always have the accompanying concern of sample bias, and these figures appear to give rise to that concern in this case. I was hoping for a greater level of insight into DNSSEC deployment here, but I suspect that measurement scale could be an issue here.



Canary in the virtual coalmine

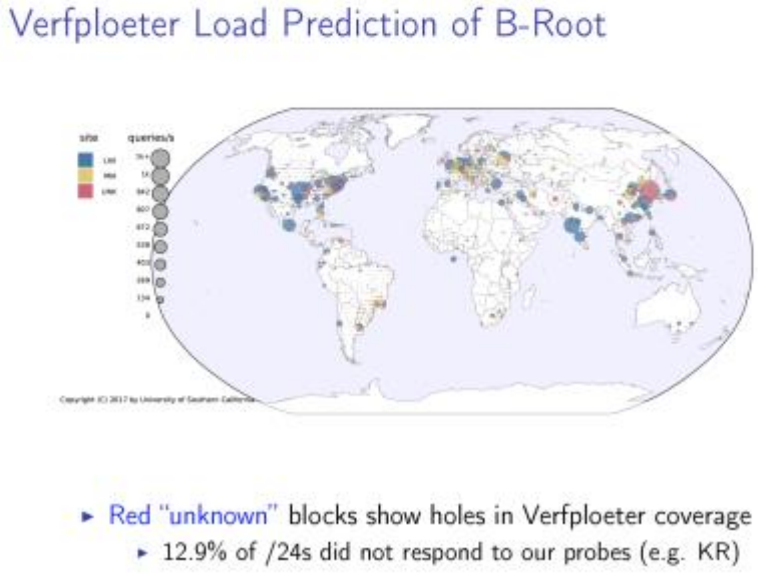
- Preliminary Findings after 2017-09-19:

https://rootcanary.org/

Moritz Muller – The Root Canary Project

Samir Jafferali of Linkedin reported on the use of "Dual DNS" to improve resiliency. This was not a talk on the use of two authoritative nameservers for a zone! These days a large content provide may choose outsource their DNS to a specialised DNS service provider that will perform a variety of services, include using the DNS to 'steer' users to particular content publication points in order to improve overall service levels for delivery of the content. This talk was about the issues in using two such providers for the same content. One thought arising from this presentation is the internet draft to serve 'stale' cache data in the case where the authoritative servers are uncontactable (draft-tale-dnsop-serve-stale). This draft raises the question whether its better for recursive resolvers to serve what is acknowledged to be stale DNS data from its cache, or serve no data at all for these uncontactable domains. The draft suggests a separate timer for this zombie state, and suggests a value of 7 days. It's yet another instance of the compromises that exist in the DNS. If you want to effect changes quickly in the DNS the challenge is to flush out the stale information and replace it with the updated information. This is particularly the case when trying to change the nameservers for a zone. At the same time you also want resiliency, particularly in the face of attack, so having recursive resolvers continue to serve data even with the authoritative name servers are unavailable can be very useful!

Most folk would agree that Anycast is "tricky" and Wes Hardaker's presentation of Verfploeter as a tool to look inside anycast catchments certainly bears this out. It's one thing to send a packet to an any cast address, but what if you sent a packet that had a source of an anycast address? If you use a packet that invites a response (ping is a good example here) and you send to a wide variety of destinations (such as one destination address for every active /24 in the deployed IPv4 Internet) then you can map a pretty complete picture of the catchment regions of an any cast constellation. Of course, this does mean that you actually have to deploy anycast listeners at the locations you want to test. So to some extent this is not a 'what if' exercise, but is the next step where you are evaluating prototype anycast deployments and wanting to understand how the anycast deployment partitions the network. The difference between the Atlas probes and this more comprehensive mapping are certainly of interest, and the bias of Atlas to measure Western Europe and parts of North America is clearly evident. Anycast is a solution to a number of problems - be they DDOS resiliency, load shedding, and replication to name three. Verfploeter will not answer all of these questions, but it does produce pretty maps that are operationally meaningful!



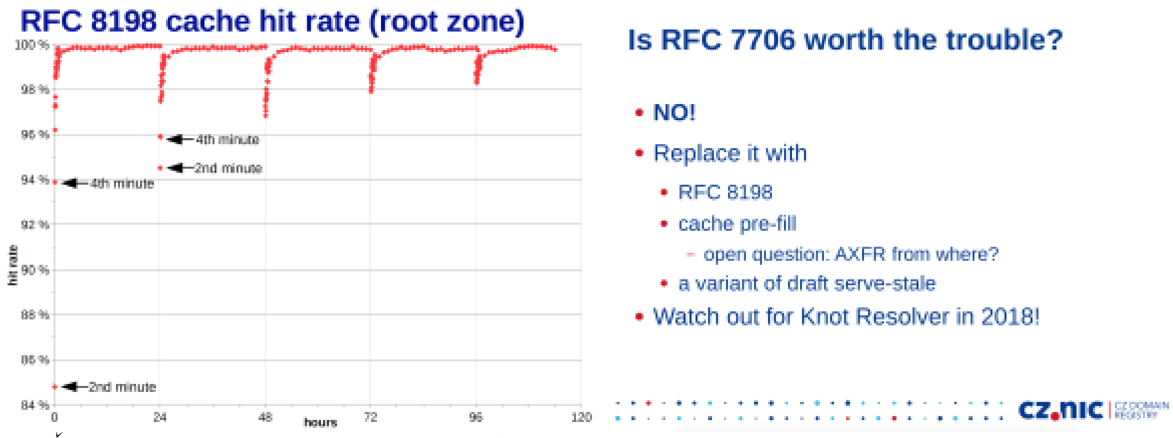
Verfploeter Load Prediction of B-Root

- Red "unknown" blocks show holes in Verfploeter coverage
  - 12.9% of /24s did not respond to our probes (e.g. KR)

Wes Hardaker - Verfploeter: Broad and Load-Aware Anycast Mapping

Steven Lagerholm reported on the progress within the T-Mobile network to switch off IPv4 completely and offer end users an IPv6-only service. There are two dominant mobile platforms in today's world: the Android platform which supports a IPv4-mapped-into-IPv6 approach of 464XLAT and the Apple iOS platform which supports an IPv6-only platform and leaves it to the operator to perform a protocol mapping from IPv6 to IPv4 at the network boundary (the larger framework of the DNS and session protocol translation is commonly referred to as "DNS64"). The presentation highlighted the inevitable compromise that DNS64 entails. Some sites are simply not visible in this environment. The larger issue is that we appear to faced with a transition which will not be completely flawless. We can expect collateral damage here and at this point consumers are now faced with the choice of choosing to use a provider that still provides IPv4 support in some form or fashion, or using a provider that is not capable of providing universal connectivity. (It is unclear if the former provider that still provides some NATted IPv4 is actually capable of supporting universal connectivity, but perhaps that's another story for another time.) Some of the issues facing a DNS64 provider is that other folk still assume that their dual stack service will be used by dual stack clients. This means that if the IPv6 part of their service is incorrectly configured in any way then the IPv4 service will mask this out. However, DNS64 will latch onto an advertised IPv6 connection and will not fall back to IPv4. This mismatch of expectations and behaviour is behind many of the connectivity issues being faced by DNS64 operators.

The root zone name service continues to be a topic of conversation in this meeting. The default behaviours of sending queries to the root service, both for good reasons and bad then to make the root servers the nexus of a potential tsunami of traffic! One perennial topic of conversation is how to jump out. How can we 'splay' the traffic that is currently being directed to the root servers and get that traffic to be absorbed by recursive resolvers. One approach is to turn the recursive resolvers into unauthoritative secondaries for the root zone, and respond to queries in the same way a root server would respond. RFC7706 describes how. Another approach is to incrementally learn the contents of the root zone and improve the effectiveness of the local cache. RFC8198 on aggressive NSEC caching describes how to do this. Petr Špaček of cz.nic has looked at both approaches and analysed their relative effectiveness. There is no doubt that NSEC caching is really efficient, in that it will learn the zone very quickly, keep it in the cache for 24 hours and serve both known and unknown TLD names without further reference to the root zone servers. This is considered to be more effective than local root service, and a combination of nsec caching and some form of serve-stale to continue service across any form of root unreachability event appear to be the most effective approach to improving a recursive resolver's performance and resilience according to Petr's work.



Is RFC 7706 worth the trouble?

- NO!
- Replace it with
  - RFC 8198
  - cache pre-fill
    - open question: AXFR from where?
  - a variant of draft serve-stale
- Watch out for Knot Resolver in 2018!

Petr Špaček - Running DNS root on localhost and its implications

A few meetings ago DNS attacks and defence occupied everyone's attention, and we heard many presentations on Rate Limiting approaches. It appears that some 17% of zone servers exhibit some form of rate limiting behaviour as they attempt to respond to burst queries, according to an active probing measurement study undertaken by Casey Deccio.

I have written a number of times about the broader domain name space and how to allow the various users of names, include names in the delegated root zone used by the DNS, names used with other name resolution protocols that still look like domain names, and other forms of locally scoped names to coexist without causing confusion, or unintended information leaks by users. Warren Kumari contributed his idea of a common local use and locally scope domain name, proposing to reserve a top level label of ".internal" for this purpose. It's easy to criticise many of the proposals that have been aired in this space, and while this latest proposal is by no means a panacea for the domain space, it does have some merit, and is worthy of further consideration.

There was a brief, but meaningful moment when Dan Mahoney turned off the DNSSEC Lookaside Validation trusted key cache. Prior to the signing of the root of the DNS this was a useful mechanism to publish the apex keys of signed zones in the DNS, but with the root having been signed for over six years its time for these hacks to be put to bed! And it was duly consigned to history!

If the DNS is your thing, then these OARC meetings are a very rewarding two days of total DNS immersion!

All the presentations from this meeting can be found in the OARC 27 meeting web site at: https://indico.dns-oarc.net/event/27/contributions

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

Disclaimer

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