December 2022
Geoff Huston

The DNS at the IGF

I don't normally make the effort to attend the Internet Governance Forum gatherings these days. It seems to me that this forum continues to struggle for relevance. In my view it has never been able to realize an effective engagement with the set of actors who make up the supply side of the Internet environment and it has never pretended to take on the role of advocacy for consumers and end users. At the same time, it hasn't been able to engage in the broader geopolitical issues that lurk just below the surface. If you are looking for answers to the question of how we can resolve the issue of today's digital behemoths and their overarching control of much of the digital ecosystem, then the IGF really is not all that helpful for your quest.

I wrote about this back in 2018 when musing on the topic: "Has Internet Governance become Irrelevant" and, if anything, the slightly hopeful tone at the end of that article has been proved to be unwarranted over the ensuring four years.

In any case, I was invited to participate in a session at IGF 2022 that was devoted to the workings of the DNS. I'd like to share my contribution to this session with my thoughts on where the DNS is headed.

The session brief reads: "The DNS is receiving increased attention from policy-makers and standards setting bodies for its central role in the functioning of the Internet. From the DNS4EU proposal which seeks to create an EU-based recursive DNS service, to local and regional conversations about the potential impacts of DNS encryption, domain names infrastructure and governance have become new sources of contention. But what does the data say on these issues? And perhaps as importantly, what data is missing to develop evidence-based policies around the DNS that protect users' trust on the Internet?"

The DNS lies in a relatively obscure part of the Internet. Unlike browsers and the World Wide Web, or the social network applications, the DNS is not exactly prominent, or even visible to users. The operation of the DNS name resolution protocol operates in a manner that confounds even the end client. It is extremely challenging to trace where and why DNS queries are propagated through the DNS infrastructure and where DNS answers come from and why. The simple question of who gets to see your online activity, in the guise of you and your DNS queries, is often very challenging to answer, Yet, even though its inner workings are obscure to the point of impenetrability, as a protocol its task is simple: the DNS resolution system takes names and translates them into network addresses. All this might seem innocuous enough, but there are a few aspects about this function which have been used and abused by many over the years, and this lies at the heart of today's issues with the DNS.

This particular protocol can trace a history back to the 1970's. The initial specification of the protocol was published in the RFC series some 35 years ago in 1987 as RFC 1034 and RFC 1035, based on earlier work on the specification of data objects used to query name servers that was initially published in 1978 with Internet Experimental Note 61. The DNS followed the pattern used by many other network protocols of the time, in that it was *open*. That is to say its payload, who is asking and what name are they asking about, was not encrypted. It was also *trustful*, in that it did not bother to authenticate whom it was talking to, and a client simply believed in whatever answers were elicited from its query. In defence of what today would be considered an obvious shortcoming, at that time we weren't constructing the final version of a future global communications infrastructure. This was just a small-scale experiment in packet networking. The DNS protocol as it emerged 35 years ago was in retrospect overly trustful to be point of being naively gullible, and any determined adversary that intrude upon the DNS query traffic could observe and tamper. But this was a research project. Why would they ever want to do this in any case?

When the Internet started to assume a more central role in the public communications realm, then the DNS came along with it, and quickly became a point of vulnerability. If I could see your DNS queries and tamper with DNS answers then I could misdirect you, or I could claim that sites and servers did not exist when in fact they did. I could poison your cache with gratuitous information in DNS responses that you were prepared to believe. In all this, you would be none the wiser because, as we have already noted, the inner workings of the DNS are totally opaque to its users.

However, tampering with the DNS is not just a tool for bad actors and bad actions. Many national regimes have used their regulatory and judicial powers to compel Internet Service Providers to actively censor the DNS by intercepting queries for certain DNS names and synthesising a DNS response that claims that the name does not exist or misdirects the end user to a different service point. This is very widespread today. But perhaps more disturbing, at least for some members of the technical community (RFC 7258) that form the core of the IETF, was the Snowden revelations of 2013 which showed that the Internet was being used by a number of national agencies, including some US agencies, to perform mass surveillance. Everything that happens online starts with a call to the DNS. Everything. If I was able to observe your DNS query stream, then there are no secrets left for you. I really do know everything you are doing online and with whom!

The technologists' response to the Snowden papers has been to erect a new set of protections around the DNS. DNS messages are encrypted, sources of DNS information are authenticated, DNS queries are trimmed of all extraneous information, and DNS content is verifiable. Tampered DNS responses can be recognised as such and discarded. These days we are looking at perhaps the most complete measures with two-layer obfuscation, such that no single party can correlate who is asking and the name that they are asking about. It's not that such information is well hidden - it's that it does not exist in any such form any more once it leaves the application on the user's device. What exactly is "DNS Data" begin referred to in the session brief in this obfuscated world? Where might we find it? The answer is that there is none!

The result is that DNS is going dark. Very dark.

It's unclear what this means in the long run. Do bad actions and actors go undetected? Do we lose our visibility into network management? What is a "secure" network and how do we secure it using traditional techniques of network perimeter traffic inspection when all the network traffic is opaque? If we can't see inside the DNS anymore, then how can we tell if (or when) the DNS has been captured by one or two digital behemoths? How can public policy makers, market regulators and market actors assess the competitive "health" of the DNS as an open and efficient market for providers and consumers where the market itself heads into deliberately dimmed obscurity.

There is much to think about here about whether the reaction to the original perceived abuse is causing its own set of issues that are commensurate with the original trigger issues that started us down this path.

Already, DNS query data is incredibly hard to find. It's easy to talk about the provisioning part of the DNS, but extraordinarily hard to find out how the DNS is being used. I know this only too well as a

researcher in this space. The privacy implications are just too great to make this data available, and obfuscating it makes it largely useless! Our efforts have had some limited success in exposing query patterns and behaviours but it's a window that is shutting down bit by bit day by day.

Where we are heading is an outcome that there will be nothing left to see in the DNS - no data, nothing! And in my view no policy or regulation can materially alter this trajectory. What we are talking about here are the actions and behaviour of applications. Trying to exercise some regulatory impost on the way that the DNS protocol behaves is about the same in my mind as attempting to regulate the fine-grained behaviour of Microsoft Word or the Chrome browser.

In many ways it has been a convenient coincidence of motives for both the large operators in today's Internet (Google, Apple, etc) and their perception of user preference that there is an apparent new found regard for user privacy. With the ascendency of the application level as the dominant factor in the Internet ecosystem there is a strong aversion by applications to allow the network or the platform to gain any insight at all into the behaviour of the application, or the content of the application's transactions. The QUIC protocol is a good example of loading the entire function of transport and content drivers into the application and hiding absolutely everything from the platform and the network.

The DNS is heading in the same direction, where with tools such as resolverless DNS over HTTPS and DNSSEC we can remove end user DNS queries entirely and have the server pre-provision DNS information via server push. If you had thought of the DNS as a common piece of network-level infrastructure, then that view is being superseded by the view of the DNS as an application artefact.

The implications of this combination of increased opacity in the DNS and a shift of the DNS from common infrastructure to application artefact inevitably head into consideration of areas of splintering and fragmentation, as applications customise their view of the space of names for their individual purposes. There is the prospect, admittedly a distant one at the moment, of declining residual value in a common general purpose name space. As all this operates behind a veil of encrypted and obscured DNS traffic, it is going to be highly challenging to try and prevent such market forces of destructive entropy from forcing an inevitable outcome here on the Internet as a whole.

To try and provide a response to the question as to what data is missing to develop evidence-based policies around the DNS that protect users' trust on the Internet, then for me the answer is not exactly encouraging. We really have no generally available data to use for this purpose today, and the pressures for ever-increasing diligence in the handling of such collected data and the shift to more effective encryption and obfuscation in DNS queries provide more than ample disincentives to collect and disseminate such data to policy makers in any case in the future.

Disclaimer

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

Author

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

www.potaroo.net