Independent Submission
Internet Draft

D. Lazanski
Last Press Label
M. McFadden
Internet policy advisors, ltd

Intended status: Informational
Expires: April 24, 2023

Oct 24, 2022

Protocol and Engineering Effects of Consolidation
draft-lazanski-consolidation-05.txt

Status of this Memo

Abstract

This document contributes to the continuing discussion on Internet consolidation. Over the last several years there have been many types of discussions around consolidation at a technical level, a economic or market level and also at an engineering level. This document aims to discuss recent areas of Internet consolidation and provide some suggestions for advancing the discussion.

Table of Contents

1. Introduction

The Internet itself is a decentralised network of networks. Resilience, security and best effort delivery of data and information on all layers of the Internet works best in a decentralised manner. But over the last several years there have been discussions on how the Internet is becoming "centralised" and "consolidated".

Internet consolidation is "the process of increasing control over internet infrastructure and services by a small set of organizations." [2] Let us consider two general categories of concentration: "player" and "layer". By player concentration, we mean the aggregating of a market to a small number of providers for a particular service. Layer concentration means the combining of functions within a given layer. An example of "player" concentration would be a relatively small number of email service providers who offer billions of users email service. Or the number of web search providers or even web browser offerings. [3]

As defined in draft-nottingham-avoiding-Internet-centralization-05 "centralization" as the ability of a single entity or a small group of them to exclusively observe, capture, control, or extract rent from the operation or use of an Internet function. Furthermore, "centralisation" as noted in the Internet of three Protocols is that one or two or three single protocols are being used for everything rather than one protocol for one operation as is a guiding principle of protocol design until now.

The Internet is being centralised and, thus, consolidated on all layers of the Internet and it is essential to recognise the technical, political and economic reasons for this happening. The rest of this draft will focus on different aspects of the issue of consolidation.

2. Background to Consolidation Issues and the Role of Standards

The Internet is being consolidated at all layers, from the application layer to the network layer. Today's traffic over the Internet is primarily derived from search and content companies. The current set of Internet protocol standards, often originating from

work at one of these companies, has facilitated the push to a more consolidated Internet.

In years past, standards have been viewed as a means to prevent barriers to entry. During the 1980s, AT&T was required to abide by standards as part of the consent decree that resolved antitrust litigation, leading to the ability of anyone to connect a telephone to its network. By 1994 standards were recognized as a means to prevent technical barriers to trade (TBT) during the Uruguay Round of the World Trade Organization.

Since 2019 a number of organisations and individuals published on the subject of consolidation. Most notably the Internet Society focused on consolidation as a key topic for their 2019 Global Internet Report [4] Additionally, the Journal of Cyber Policy published a special edition on Consolidation. Topics in this special issue included market concentration and security, DNS consolidation, supply chains, interoperability and Internet architecture. [5]

Discussion of consolidation primarily focuses on Internet services and data. However, there is more to consolidation than just applications and services. The application layer is directly user facing and, as a result, is what users experience. But the underlying infrastructure and protocols are also going through consolidation as they develop. The transport layer protocol development is focused on the end-to-end encryption model which, when implemented, forces data to the end points. Thus, the data is managed at the endpoints only and not managed throughout its entire transit. This results in a limited number of available options for the use of the data.

The QUIC protocol is an example of the consolidation between layers of the Internet. Designed and deployed as a transport layer protocol, it effectively replaces TCP at the network layer while also adding improved security. The result is the merging or consolidation of layers. QUIC should improve efficiency and delivery of applications, but also forces all data to be managed at the endpoint, which in this case is a browser or application, making it more difficult or impossible to manage traffic at the network layer.

2.1. Process

Another key point to make is that the process of standards development impacts the outcome and adoption of the standard. This is key for consolidation. Though Internet protocol development

should be multistakeholder, standards development is subject to vested interests, personal approaches and commercial realities. [6] Developing protocols, and standards more generally, takes time, much discussion and a bottom-up approach. However, commercial organisations have different goals in the process of trying to standardize protocols. Larger organisations have more resources dedicated to protocol and standards development. Larger organisations with staff specifically dedicated to standards tend to have the ability to push for their proposals and their protocols. There is no coincidence that these companies are the ones that have facilitated consolidation on a commercial level and are facilitating consolidation on a protocol level.

3. Overarching Issues Related to Consolidation

In A Taxonomy of Internet Consolidation, a new draft by Mark McFadden, a set of typologies for Internet consolidation is clearly and concisely outlined. Four main areas, namely economic, traffic and infrastructure, architectural and service and application consolidation create a taxonomy that can be used to articulate the different issues and challenge of consolidation.

Consolidation has led to the use of a few, large online platforms which is facilitated by choice and market consolidation. But consolidation also has led to the development of a protocols manage traffic flow and capture data. The over-reliance on one or two entities for delivery of that use a small handful of protocols has led to technical dependencies on these protocols.

"Did the IETF create a better internet when it approved DoH? There's a lot of disagreement about that, but what has upset many is that DoH was a surprise - the IETF standardised it without consulting some who it was likely to affect," it says in RFC 8890 [7] However, there was little multistakeholder consultation and discussion prior to the adoption of DoH. This was more of a rapid development and deployment process, without the market driving the use cases and uptake. By forcing the concentration of the data at the endpoint, the data is consolidated into the service provider at that endpoint.

3.1. Economic

According to the Internet Society's 2019 report, consolidation is broadly defined as, "economic activities that either support the Internet or are fundamentally dependent on the Internet's existence."[8]

One focus of consolidation is around the concentration of power - consumer, technical and financial - into a handful of large Internet companies. The first point of engagement with any of these companies, including Facebook and Google, is through consumer applications. The ability to easily understand consolidation at an application layer, because of the widespread and common use of Facebook and Google, has caused the focus of consolidation and anti-competitive issues from policy makers and politicians to be at the application layer.

However, consolidation also has upsides. Consolidation allows for economies of scale, investment in infrastructure and the ability for small and medium enterprises to buy and use services, like cloud storage, content distribution networks and security technology, without having to build them from the ground up every time. However, the lack of market diversity means a lack of competition which, in turn means a lack of innovation and a lack of consumer choice.

New entrants and competition in a market creates options for consumers that potentially pulls them away from popular websites and applications. When a market is not competitive or viable, regulation and anti-trust measures can intervene to remedy a consolidated market which is tending towards or has achieved monopoly status. Legal and regulatory intervention, however, tends to create its own set of issues as seen through several decades of EU intervention in big tech starting with Microsoft in 2004. Unintended consequences with regulatory or legal intervention may skew the market even further.

3.2. Security

Consolidation of protocol development has facilitated the secure, end to end encryption of information going over networks in recent years. New technologies such as DNS-over-HTTPS (DoH), Oblivious DNS over HTTPs (ODOH)and DNS-over-TLS (DoT) standardised through the IETF allow for confidential look up of DNS queries. However, it has required updates onto many DNS servers and operating systems. The implementation of these protocols enable circumvention of DNS filtering which ISPs offer for protection from malicious websites and software on the network.

This is a form of market consolidation based on development choices by several large companies. These development choices are often technically opaque without transparency of what happens when updates

take place, resulting in more difficulty when trying to troubleshoot security issues.

The development of these protocols, while providing increased privacy and addressing issues concerning government surveillance, have forced other unintended consequences which is promoting consolidation.

Consequences of the security of the global Internet are evident. On June 8, 2021, a global outage of Fastly, a content delivery network (CDN), was caused by a software update which included an undiscovered bug. [10] While this was resolved within a working day, one of the main causes of the outage was a consequence of the limited number of CDNs running services in the cloud. Other CDNs, which resolved traffic via Fastly for redundancy, were also taken down as a result of the Fastly outage. This dependency is caused by consolidation and a concentration of infrastructure. A highly consolidated CDN network facilitates a less secure environment because of the weakening of resilience [11]

On 22 June 2022 Cloudflare suffered an outage that lasted just over an hour and impacted 19 data centers. Though the outage was due to a misconfiguration that was quickly resolved, the impact of the outage renewed calls for a critical look at decentralising the Internet. A handful of cloud and infrastructure providers are responsible for global connections. This outage was a reminder of the need to think about resilience and security in global Internet connectivity. [12]

4. Implications of Consolidation on Internet Architecture

4.1. Changing Internet architecture

The phenomenon of consolidation may be in the eyes of the beholder. A government may see market failure or a need for regulation. [13] A civil society advocate may see it from the point of view of privacy or free speech. For the purposes of this draft we view it from the perspective of the underlying architecture of the public Internet

Consolidation in the Internet's architecture is not a new development. The approach of providing intermediaries to deliver service or content rather than the more traditional end-to-end approach has been in place for more than a decade. However, it is possible to argue that the architecture of the Internet has changed dramatically in the last decade.

The architecture of the Internet is always changing. New services, applications and content mean that the market creates new ways to deliver them. Consolidation clearly has economic, social and policy issues, but it is important to understand how consolidation affects the underlying architecture of the Internet. The impact of intermediaries on architecture is often not obvious

The use of intermediaries in the Internet's architecture may include the use of third parties to provide services, applications or content. In the early days of the Web, this was evident when rendering a web page that included content from multiple sources. In today's Internet the intermediaries are not so obvious. Authentication servers, content distribution networks, certificate authorities, malicious content protection and DNS resolution services are all examples of tools provided to the Internet by intermediaries - often without the knowledge or approval of both endpoints.

Having intermediaries embedded in the architecture is a different effect from having them embedded in the service infrastructure. The domination by a few companies of the content and application layer is largely an economic effect of scale. On the other hand, there is a prevalent belief that the Internet puts intelligence at the edge. While that may have been true in the past, it is hard to argue that this is a feature of the contemporary Internet.

There is a suggestion that the network simply provides for the transport of data. There are almost no network connections like that in today's Internet. A consumer's view of the Internet is limited by unseen intermediaries of many types. A consumer on the Internet seldom makes choices about those intermediaries: they are simply part of the fabric that makes up the Internet.

Almost all important parts of the architecture have been affected by consolidation: DNS resolution, access service, transit provision, content distribution and authorization. Consolidation in these areas has a direct effect on engineering and protocol design.

4.2. End to end principle redux

The end-to-end principle is the idea that reliability and trustworthiness reside at the end nodes of networks rather than in the network itself. In other words, the idea was that the network

itself was dumb and intelligence was at the edge or end. However, Internet architecture is evolving in such a way that this principle is changing.

Networks and the devices on the networks act as access consolidators. While, in the past, the network was a simple transporter of bits, today's networks see intermediaries consolidating both access and the delivery of information (e.g. streaming media). For example, 5G will allow for different services, systems and use cases at a very specific level. Network slicing in 5G will concentrate services like video on demand into concentrated - and consolidation - areas on a network. [14] In other words, as specific types of services are relegated to a segregated part of a network, the availability and access of that service is limited to accessing a specific network. Depending on the type of device or maturity of the network infrastructure available at the point of the attempted access, options for access might be limited. If a network slice on 5G is where a specific service is located, for example, but it is only possible to use a 3G mobile network, then the service is unavailable. Thus, the service is only available on a consolidated part of the mobile network.

Another change is how the layers of the Internet, as discussed in the QUIC example, are consolidating. Differentiation among layers is fading fast with the development of applications which require network access and control.

Rapidly, the end-to-end principle is becoming the edge-to-edge principle. The layers of the internet are morphing into several consolidated layers and it is becoming difficult to differentiate between the end or edge, and also nearly impossible to ensure the reliability of the internet because of it. But the important part of this is the network is not dumb. Data processing, storage and highly evolved services (including custom data and metadata processing at the edge) means that the 'dumb' network is no longer dumb.

If the number of organizations that provide those "network services" that we rely upon is small, our dependence is higher. In extreme cases of engineering, we put ourselves at risk of engineering a single point of failure. But also if organisations can't and won't enter the market, the market is left with very few options and choices. In other words, if a handful of organisations enable end to end encryption and those same organisations also offers services at the edge, then only a handful of organisations provide the entire value chain. This is consolidation.

The trend toward highly specific and concentrated processing, as well as the drive for highly customised applications and services will drive the Internet away from an end-to-end principle. This will create not a network of networks, but a mesh. If the mesh is dependent on a small number of very large providers through consolidation, we will have engineered a single source of failure into the Internet.

5. Implications of Consolidation on Protocol Design

5.1. Does Protocol Design Really Affect Consolidation

As noted in "Internet of Three Protocols" draft, "One of the guiding principles of designing a protocol in the original Internet community was the protocol is not complete when everything possible has been added, but rather when everything possible has been removed." This is so that security, scalability, resilience and observability can be ensured. However, the recent trend has been towards having a few protocols, but having those protocols do all things.

So, in effect, the protocols themselves are becoming consolidated. The point of protocol design is not to develop all things on one protocol, but to have a protocol that improves the sustainability of the Internet.

5.2. Case Studies in Consolidation and Protocol Design

5.2.1. DNS over HTTPS (DOH)

The development of encrypted DNS, specifically DNS-over-HTTPS (DoH), has been driven by a desire to show full end-to-end encryption of network connections. The protocol was completed and the DoH working group wound up in March 2020 despite the absence of both resolver discovery and selection mechanisms. This may be addressed in the future.[15] Client software is developing with interim discovery solutions which almost always favour the large, cloud-based resolver operators. This is leading to a situation where users are being presented with a very small number of pre-configured resolver options irrespective of their location - in some client software as few as three or four options may be presented. [16] Currently, there are many thousands of DNS servers operating without DoH.

It is likely that most of the DNS traffic will be consolidated onto a handful of global operators, if multiple options for discovery

mechanisms are not developed. The impact that such a loss of diversity of providers may have on the long-term resilience of DNS should not be underestimated. [17] Nor should the attractiveness of these potential network chokepoints to attack be overlooked either to access consolidated data or launch an attack from. One danger is that if DNS traffic is concentrated onto a small handful of global operators and is 'automatically-on' the result would be default adoption by the vast majority of the Internet's clients. The suggestion that there were mechanisms for users to opt-out would not matter in the face of statistics that regularly show that users almost never change default settings. Currently, the deployment approach for DoH is opt-in . For CDNs, DoH default-on would disrupt and render CDN geolocation designed to manage traffic flows more efficient closer to the desired delivery location. Thus, protocol design decisions that are enshrined in default settings will become the norm. In this case, default on, which facilitates consolidation, will become standard.

By routing the DNS over HTTPS, it becomes much easier to track user activity through the use of cookies. Therefore a protocol that was developed to enhance user privacy and security could actually undermine both: privacy through the use of cookies and security by consolidating DNS traffic onto far fewer resolver operators that are far more attractive targets for malicious actors of various types.

5.2.2. Encrypted Server Name Indication (eSNI)

Options to encrypt the Server Name Indication (SNI) have been explored in the TLS working group but to date it has not been possible to develop a solution without shortcomings. This flaw in the encrypted SNI (eSNI) options under evaluation required a rethink in the approach being taken. The solution now proposed, Encrypted Client Hello (ECH, previously called ECHO) assumes that private origins will co-locate with or hide behind a provider (CDN, application server etc.) which can protect SNIs for all of the domains that it hosts.[18] Whilst there is logic in this approach, the consequence is that the would-be standard encourages further consolidation of data to aid privacy. What it does not appear to consider is the attractiveness of this larger data pool to an attacker, compared with more dispersed solutions.

eSNI can be implemented by a "fronting" service which protects a hidden service behind it. Because the client will not verify the identity of this fronting service, server spoofing attacks are possible. Indeed, the fronting service could be pressured by attackers. The fronting service then becomes a rich source of

information about client connections and an attractive attack surface for adversaries.

5.2.3. Privacy Pass

The Privacy Pass protocol provides a set of cross-domain authorization tokens that protect the client's anonymity in message exchanges with a server. This allows clients to communicate an attestation of a previously authenticated server action, without having to reauthenticate manually. The tokens retain anonymity in the sense that the act of revealing them cannot be linked back to the session where they were initially issued.

For Privacy Pass to succeed clients must be able to acquire tokens that they can later redeem with greater privacy and anonymity. This document does not discuss the goals of privacy or anonymity. Instead, it identifies a problem related to the upper bound in number of servers that affects the Privacy Pass ecosystem.

"Server centralization" is the strict limit or upper bound in the number of servers available from which a client can acquire a token for later redemption. The current architecture of privacy pass strictly limits the number of participants (so-called Attesters or Issuers). The current architecture suggests a non-protocol approach to addressing the centralization problem (through a multi-stakeholder governance model) and also suggests a different approach where a quorum of parties acted in a way where clients would have more opportunities to switch between attestation participants.

However, neither of these approaches is required by the Privacy Pass architecture document and the centralization problem created by the specification of the protocol is left to implementations to solve.

6. Potential Technical Risks

There are a number of potential risks to the security, stability and performance of the Internet and many of them are well articulated in DoH Implementation Risks [19], but some notable ones are:

1. Significant operational shift of the global Internet from a highly distributed to a centralised system. This would impact both security and resilience.

2. Decreased stability due to the fact that a centralised system will have higher fragility, fewer points of failure and greater impact on the system when it does fail.

3. Increased security issues caused by the reduction of number of recursive DNS operators. [20] Lack of distributed and recursive DNS creates a lack of redundancy for when security attacks hit parts of the Internet.

4. Loss of security threat visibility due to degraded ability to use DNS blocklists and overall network management for malware, phishing, spam, DDoS and etc if DNS management is consolidated into a few operators.

5. Reduced diversity in the Internet ecosystem. Diversity creates greater redundancy, resilience and agility to respond to attacks, outages and network issues.

7. Security Considerations

While this document does not describe a specific protocol, it does discuss the evolving architecture of the Internet. Changes to the Internet's architecture have direct and indirect implications for the Internet's threat model.

Specifically, the changes to the end-to-end model (see section 4.2 above) have inserted new interfaces which must be reflected in security considerations for new protocols.

8. IANA Considerations

This memo contains no instructions or requests for IANA.

9. Conclusions

This document seeks to further continue the discussion on consolidation. As argued above, Internet consolidation is happening at different places and different layers of the Internet and ongoing discussions, particularly in DINRG group.

10. References

10.1. Normative References

[1] Considerations on Internet Consolidation and the Internet Architecture [draft-arkko-iab-internet-consolidation-02].

[2] IBID

[3] Google has over at least 80% worldwide market share. https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

[4] https://www.internetsociety.org/tag/consolidation

[5] Centralised Architecture in Internet Infrastructure [draft-arkko-arch-infrastructure-centralisation-00]page 5.

[6] Dominique Lazanski, Governance in international technical standards-making: a tripartite model, Journal of Cyber Policy, 4:3, 362-379, 2019. https://www.tandfonline.com/doi/full/10.1080/23738871.2019.1696851

[7] RFC 8890, The Internet is for End Users. Nottingham, Mark. August 2020. https://www.rfc-editor.org/info/rfc8890

[8] Consolidation In the Internet Economy, Internet Society, 2019. https://future.internetsociety.org/2019/consolidation-in-the-internet-economy

[9] https://seekingalpha.com/article/4544613-cloudflare-on-the-right-trajectory

[10] Fastly Blog, June 8, 2021. https://www.fastly.com/blog/summary-of-june-8-outage

[11] The Deeper Root Cause of the Fastly and Akamai Outages, CircleID, June 28, 2021 https://www.circleid.com/posts/20210628-the-deeper-root-cause-of-the-fastly-and-akamai-outages/

[12] https://www.site24x7.com/blog/6-lessons-from-cloudflares-june-2022-outage

[13] See Google, antitrust and how to best regulate big tech, The Economist, 7 October 2020 https://www.economist.com/business/2020/10/07/google-antitrust-and-how-best-to-regulate-big-tech

[14] What is Network Slicing? https://5g.co.uk/guides/what-is-network-slicing/

[15] DNS over HTTPS (doh) https://datatracker.ietf.org/group/doh/about/

[16] At the time of writing, the Firefox browser presents a list of three pre-configured resolver options to North American users: Cloudflare, NextDNS and Comcast.

[17] Cloudflare DNS goes down taking a large piece of the Internet with it, 17 July 2020. https://techcrunch.com/2020/07/17/cloudflare-dns-goes-down-taking-a-large-piece-of-the-internet-with-it/

[18] TLS Encrypted Client Hello draft-ietf-tls-esni-07 https://tools.ietf.org/html/draft-ietf-tls-esni-07

[19] https://datatracker.ietf.org/doc/draft-livingood-doh-implementation-risks-issues/

[20] https://hbswk.hbs.edu/item/evidence-of-decreasing-internet-entropy-the-lack-of-redundancy-in-dns-resolutionbymajor-websites-and-services

11. Acknowledgments

Many thanks to all who discussed this with us, especially Jason Livingood and Eliot Lear.

This document was prepared using 2-Word-v2.0.template.dot.

Authors' Addresses

Dominique Lazanski
Last Press Label
London, UK

Email: dml@lastpresslabel.com

Mark McFadden
Internet policy advisors ltd
Chepstow, Wales, UK

Email: mark@internetpolicyadvisors.com