Internet Engineering Task Force
Internet-Draft
Intended status: Informational
Expires: April 26, 2019

J. Arkko
Ericsson
B. Trammell
ETH Zurich
M. Nottingham

C. Huitema
Private Octopus Inc.
M. Thomson
Mozilla
J. Tantsura
Nuage Networks
October 23, 2018

Considerations on Internet Consolidation and the Internet Architecture
draft-arkko-iab-internet-consolidation-00

Abstract

Many of us have held a vision of the Internet as the ultimate distributed platform that allows communication, the provision of services, and competition from any corner of the world. But as the Internet has matured, it seems to also feed the creation of large, centralised entities in many areas. This phenomenon could be looked at from many different angles, but this memo considers the topic from the perspective of how available technology and Internet architecture drives different market directions.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 26, 2019.

Table of Contents

1. Introduction 2
2. Consolidation 3
2.1. Economics 4
2.2. Data- and Capital-intensive Services 4
2.3. Permissionless Innovation 5
2.4. Fundamentals of Communication 5
2.5. Technology Factors 5
3. Action 6
4. Contributors 8
5. Acknowledgements 8
6. Informative References 8
Authors' Addresses 8

1. Introduction

Many of us have held a vision of the Internet as the ultimate distributed platform that allows communication, the provision of services, and competition from any corner of the world. But as the Internet has matured, it seems to also feed the creation of large, centralised entities in many areas.

Is Internet traffic consolidating, i.e., moving towards a larger fraction of traffic involving a small set of large content providers or social networks? It certainly appears so, though more quantitative research on this topic would be welcome.

This phenomenon could be looked at from many different angles, but this memo considers the topic from the perspective of how available technology and Internet architecture drives different market directions. How are technology choices and fundamentals of communication affecting some of these trends?

Our engineering remit is to focus on technology, but of course we also want to understand the implications and externalities of the technical arrangements we design. Technology affects economics and vice versa. The Internet technology community continues to make decisions that have ramifications on Internet systems, just as we are subject to forces that affect them.

As technologists, one question we have is whether there are changes in technology that would help reduce technically-driven large-player advantages.

This memo reviews areas where consolidation may be occurring in the Internet, and discusses the potential reasons for this. Section 2 discusses consolidation and the reasons behind the creation of larger entities, and Section 3 looks at some actions that might alleviate the situation.

Note: If you are interested on this or other architecture-related topics, please subscribe to the IAB architecture-discuss mailing list as one forum for discussion.

2. Consolidation

Consolidation is driven by economic factors relating to scale and ability to reach a large market over the Internet. In general, an efficient market such as the Internet tends to enable winners to take large market shares.

The most visible aspects of this involve well-recognised Internet services, but it is important to recognise that the Internet is a complex ecosystem. There are many underlying services whose diversity, or lack thereof, are as important as that of, say, consumer-visible social networks. For instance, the diversity of cloud services, operating systems, browser engines is as important as that as of application stores or the browsers themselves.

Of course, the Internet allows plenty of choice both in these and other areas. Too many or too few choices create different kinds of problems.

It would be useful to break these general factors and observations down a bit further. In particular, it is useful to distinguish market or economic factors from technical factors.

2.1. Economics

Scaling benefits are natural for many types of businesses. And many Internet-based businesses can potentially serve a very large customer base, as the cost of replicating and delivering their service to new customers or areas is small.

However, typically the network effect has an even more pronounced impact. Each additional user adds to the value of the network for all users in a network. In some applications, such as the open web, this value grows for everyone, as the web is a globally connected, interoperable service for anyone with a browser can use.

There is an important distinction between different applications of the network effect, however. Consider email as another example; anyone with an account at any email server can use it globally. However, here we have seen much more consolidation into few large email providers, both due to innovative, high-quality services but also because running email services by small entities is becoming difficult; among other things due to spam prevention practices that tend to recognise well only the largest entities.

In some other applications, such as social media, the services have a more closed nature. The value of being a customer of one social media service depends highly on how many other customers that particular service has. Hence, the larger the service, the more valuable it is. And the bigger the value difference to the customers, the less practical choice they have in selecting a service.

In some cases, these developments also allow asymmetric relationships to form, with the customers having less ability to affect the service than they would perhaps wish.

2.2. Data- and Capital-intensive Services

The scaling advantages are only getting larger with the advent of AI- and machine learning -based technologies.

The more users a service has, the more data is available for training machine learning models, and the better the service becomes, bringing again more users. This feedback loop and the general capital-intensive nature of the technology (data and processing at scale) makes it likely that the largest companies are ahead in the use of these technologies.

2.3. Permissionless Innovation

The email vs. social media example also highlights the interesting roles of interoperability and the "permissionless innovation" principle -- the idea that a network can be simple but still powerful enough that essentially any application could be built on top of it without needing any special support from anyone else. Permissionless innovation has brought us all the innovative applications that we enjoy today, on top of a highly interoperable underlying network, along with advances in video coding and other techniques used by applications.

Paradoxically, if the underlying network is sufficiently powerful, the applications on top can evolve without similar pressures for interoperability, leading to the closed but highly valuable services discussed above. We call this the Permissionless Completeness Problem.

2.4. Fundamentals of Communication

There are also fundamental issues. For instance, speed of light; low-latency services can fundamentally only be provided through

globally distributed data centers. These are often provided built by large organisations, although collaborative and data center or cloud computing service approaches also exist.

A similar issue has arisen in recent years around large-scale denial-of-service attacks, and how various entities can deal with them. While the largest attacks affect all players (see, for instance, the Dyn attacks in October 2016), it is also true that large cloud- and content delivery providers can better deal with such attacks due to their scale. This is one reason that attracts many network services to such providers.

2.5. Technology Factors

One of the key questions is whether we are seeing developments that are driven by economic factors or whether fundamental reasons or lack available technology drives particular models. For instance, centralised solutions might desirable due to business incentives, or they might be necessary because there is no distributed, collaborative solution.

For instance, some technical issues have historically not been easy to solve, such as e-mail spam, which has lead to reliance on non-technical solutions. Today, it is becoming increasingly difficult to run your own mail services, essentially forcing many organisations and individuals to employ larger providers. The issues relate

directly to size of entities; no one can afford to disconnect from the largest providers. But as a small entity, there is little leverage to convince peer entities or various supporting white/blacklist entities to deal with you properly.

Many Internet services are based on gathering data about users, and using that data for, for instance, targeted advertisements. More data from more users makes it possible to run a service more accurately or with better results; here again scale brings advantages.

Another trend is that more and more content is becoming available locally, from a content delivery or provider function directly on your own ISP's network. We predict that eventually most content will be delivered this way, reducing the role that global IP connections

across the Internet play. By some metrics this has already happened; what practical - positive or negative - impacts might this have on the Internet technology?

There are also security tradeoffs. Large entities are generally better equipped to move to more recent and more secure technology. For instance, the Domain Name System (DNS) shows signs of ageing but due to the legacy of deployed systems, has changed very slowly. Newer technology developed at the IETF enables DNS queries to be performed confidentially, but its deployment is happening mostly in browsers that use global DNS resolver services, such as Cloudflare's 1.1.1.1 or Google's 8.8.8.8. This results in faster evolution and better security for end users.

However, if one steps back and considers the overall security effects of these developments, the resulting effects can be different. While the security of the actual protocol exchanges improves with the introduction of this new technology, at the same time this implies a move from using a worldwide distributed set of DNS resolvers into, again, more centralised global resolvers. While these resolvers are very well maintained (and a great service), they are potentially high-value targets for pervasive monitoring and Denial-of-Service (DoS) attacks. In 2016, for example, DoS attacks were launched against Dyn, one of the largest DNS providers, leading to some outages.

3. Action

Are there assumptions about the Internet architecture that no longer hold in a world where larger, more centralised entities provide big parts of the Internet service? If the world changes, the Internet and its technology/architecture may have to match those changes.

It appears that level the playing field for new entrants or small players brings potential benefits. Are there technical solutions that are missing today?

Of course, it may well be that technology improvements are hard to come by. Nevertheless, recognising the risks of consolidation in both current and proposed future technologies is the first step in proactively avoiding those risks where possible.

Assuming that one does not wish for regulation, technologies that support distributed architectures, open source implementations of currently centralised network functions, or help increase user's control can be beneficial. Federation, for example, would help enable distributed services in situations where smaller entities would like to collaborate.

Similarly, in an asymmetric power balance between users and services, tools that enable the user to control what information is provided to a particular service can be very helpful. Some such tools exist, for instance, in the privacy and tracking-prevention modes of popular browsers but why are these modes not the default, and could we develop them further?

It is also surprising that in the age of software-defined everything, we can program almost anything else except the globally provided, packaged services. Opening up interfaces would allow the building of additional, innovative services, and better match with users' needs.

Silver bullets are rare, of course. Internet service markets sometimes fragment rather than cooperate through federation. And the asymmetric power balances are easiest changed with data that is in your control, but it is much harder to change when someone else holds it. Nevertheless, the exploration of solutions to ensure the Internet is kept open for new innovations and in the control of users is very important.

What IETF topics that should be pursued to address some of the issues around consolidation?

What measurements relating to the developments centralization or consolidation should be pursued?

What research -- such as distributed Internet architectures -- should be driven forward?

4. Contributors

Much of the text in this memo is from a blog article written by Jari Arkko, Mark Nottingham, Christian Huitema, Martin Thomson, and Brian Trammell for the Internet Architecture Board (IAB), and from a blog article written by Jari Arkko and Brian Trammell APNIC and RIPE.

5. Acknowledgements

The authors would like to thank IAB members, Geoff Huston, Gonzalo Camarillo, Mirjam Kuehne, Robert Mitchell, Olaf Kolkman, and many others for interesting discussions in this problem space.

6. Informative References

[RFC1958] Carpenter, B., Ed., "Architectural Principles of the Internet", RFC 1958, DOI 10.17487/RFC1958, June 1996, <https://www.rfc-editor.org/info/rfc1958>.

Authors' Addresses

Jari Arkko
Ericsson
Kauniainen 02700
Finland

Email: jari.arkko@piuha.net

Brian Trammell
ETH Zurich

Email: ietf@trammell.ch

Mark Nottingham

Email: mnot@mnot.net

Christian Huitema
Private Octopus Inc.

Email: huitema@huitema.net

Martin Thomson
Mozilla

Email: martin.thomson@gmail.com

Jeff Tantsura
Nuage Networks

Email: jefftant.ietf@gmail.com