Draft Report of DINRG Workshop on Centralization in the Internet, June 3, 2021

Christian Huitema
Private Octopus
USA

Geoff Huston
APNIC
Australia

Dirk Kutscher
HS Emden/Leer
Germany

Lixia Zhang
UCLA
USA

ABSTRACT

This is a report on the IRTF DINRG Workshop on Centralization in the Internet, held on June 3rd 2021, focusing on painting a broad brush landscape of the Internet centralization problem space: its starting point, its driving force, together with an articulation on what can and should be done.

1 INTRODUCTION

There is a general consensus among the networking community that the Internet consolidation and centralization trend has progressed rapidly over recent years, as measured by the structural changes to the data delivery infrastructure, the control power over system platforms, application development and deployment, and even in the standard developments. This trend has brought impactful technical, societal, and economical consequences. When the Internet first rolled out as a decentralized system 40 years back, few people, if any, could have foreseen how it looks today. How has the Internet evolved from there to here? What have been the driving forces for the system's consolidation? From a retrospective view, was there anything that might have been done differently to influence the course the Internet has taken? And most importantly, what can and should be done now to mitigate the consequence of centralization? Although there are significant interests in these topics, there has not been much structured discussion on how to address these important problems.

The IRTF Research Group on Decentralized Internet Infrastructure (DINRG)[1] organized a workshop "Centralization in the Internet" on June 3, 2021 with the objective of starting an organized open discussion on the above questions [1]. Although counter measures to the centralization problem seem needed urgently, this workshop aimed to take a step back: before jumping to exploring potential approaches to steer the Internet away from centralization, we need to discuss *how* the Internet has evolved and changed, and *what* have been the driving forces and enablers. We believe that a sound and

---

[1]https://datatracker.ietf.org/rg/dinrg/about/

evidence-based understanding is the key towards devising effective remedy and action plans. In particular, we would like to deepen our understanding on the relationship between the architectural properties and economic developments.

This workshop consisted of two panels, each panel started with an opening presentation, followed by panel discussions, then open-floor discussions. There was also an all-hand discussion at the end. Three hours of the workshop presentations and discussions show that this Internet centralization problem space is highly complex and filled with intrinsic interplays between technical and economic factors.

This initial report aims to draw a broad brush picture of the problem space. We hope that this big picture view will help the research group, as well as the broader IETF community, reach a clear and shared high level understanding of the problem, and from there to identify what actions are needed, which of them require technical solutions, and which of them are regulatory issues but require technical community to provide inputs to regulatory sectors in order to develop effective regulation policies.

This report roughly follows the structures of the panels. Section 2 summarizes Panel 1's opening presentation by Geoff Huston, which shows convincingly that, from the industrialization history to today's Internet, economies and opportunities of scale drive industry players towards consolidation of resources and control. Section 3 summarizes Panel 2's opening presentation by Christian Huitema, which elaborated on a number of specific factors that played a role in the Internet's evolution from its initial decentralized rollout to where we are today. Because the discussions at Panel-1, Panel-2, and the all-hand discussion at the end of workshop covered broad and overlapping topics, we summarize all the discussions in Section 4. Section 5 presents a high level of workshop summary which enumerates identified major issues in the Internet centralization problem space.

Appendix A contains the list of workshop submissions, grouped by their characteristics. It is impossible to incorporate all the issues discussed at the workshop into a single report without making it too long or diluting the focus. We use this first report to summarize the major ingredients/factors in a high-level picture of the Internet centralization, and record

in Appendix B the issues not included in the main body of the report, which are all highly relevant and important and may be topics for future investigations.

This report aims to provide a faithful summary of the workshop and to reflect diverse views of the participants. We also inject observations and commentaries at various places in the discussion section, with a hope to offer additional food-for-thought and to suggest potential topics for future studies.

2 PANEL 1 PRESENTATION: IS BIG REALLY BAD?

Geoff Huston offered the opening presentation for Panel 1 on "Is 'Big' Necessarily 'Bad' "? (see [2] for presentation slides). He drew lessons from history, using US economy as an example to articulate the root cause of centralization. About one and half centuries ago, US went through a phase of industry centralization which was enabled by a liberal labor market, large capital funding (from Europe), and the opening of the railroads for transportation which transformed many smaller regional markets into a single very large national market. The impact of the resulting industrial superpowers went well beyond national boundaries over the ensuing decades. A number of large scale US enterprises dominated the emerging global market, and US domestic economy was a major beneficiary of this position through much of the twentieth century. Their rapid expansion also overwhelmed government regulatory measures of the time, allowing these emerging enterprises to script their own preferred regulatory ruleset. When the industrial incumbents get to define the terms of trade, the inevitable outcome is the suppression of competition, and the transformation of dominant positions by such actors into entrenched monopolies.

The well-known 1890 Sherman Anti-Trust Act authorized US federal government to prohibit companies from colluding or merging to form an effective monopoly, in order to help workers and smaller enterprises by encouraging competition. The law was applied in 1910 to Standard Oil, American Tobacco, and General Electric to curb their concentrations of power. However, breaking up these monopolies led to unintended consequences, including national economic panic and a depression in the following year, and the creation of vigorous political lobbying by the large industrial enterprises to bring their interests to the political process and to bind the interests of politicians to their interests.

Huston went on to explain that today's Internet industries simply followed the footstep of the past giants, with massive concentration in their chosen area of activity and massive lobbying. One big difference is that the past giants exploited the labor force to accumulate profits, while today's Internet giants moneytise the accumulated large pools of personal data via advertising markets. With today's Internet practice, individual users simply cannot profit from their own personal data even if they would like to; the data is intrinsically valuable in terms of its aggregate volume in conjunction with its individual specificity, which implies that only large enterprises can amass such a critical mass of profile data. This has become a positive feedback where the accumulated value of personal data can underwrite the investment in services that are offered to users at heavily discounted prices, or more commonly as free service, which in turn attract more users and enable further data harvesting. These popular and free services lead to further service concentration. The wealth of these enterprises lies not in the value of physical goods, nor in the value of the services they provide, but in the volume and accuracy of the user preference data that they have amassed. A term "surveillance capitalism" has been coined to describe this new enterprise model.

To date, the Internet giants have benefitted from cavalier attitude towards data security and privacy concerns. The absence of regulatory imposts regarding data security and due attention to user privacy concerns resulted in the absence of and effort to safeguard user interests on the part of these enterprises. The EU's GDPR effort represents a welcome change in this regard, however the financial fines used in this measure may be too low to effectively curb the violation behavior from the very largest of the enterprises. The potential downside of such regulatory measures is the unintended consequence that they act as a disincentive to smaller new entrants in the market and further entrench the position of the largest incumbents.

Huston also alerted us an even bigger danger, Google's dominance in search market, at this time it occupies 92% of world wide search market.[2] Search has now replaced all forms of reference libraries. Search engine has become the ultimate decider of public argument, the ultimate tool for researchers, the ultimate source of all information. Whenever one has a question, no matter what kind, "*Google answers all questions. Google tells one what to think*", Dominant search engines are both incredibly empowering, allowing the world's knowledge at one's fingertips, and incredibly threatening as they can shape the society's views in all subject areas, from today's fashions to world politics.

Observing the creation processes of industrial giants from history to present, one may conclude that the appearance of dominant industrial players is a natural outcome out of the compounded gains from economies of scale. In addition, big players can grow way faster than public's understanding of their consequences, let alone the regulation development. when a market player grows big enough, it becomes possible for it to make its own rules.

---

[2]https://gs.statcounter.com/search-engine-market-share

The above answered the question of whether being big is bad. However, Huston also reminded us that being big is not *all* bad. Big players can bring national economic gains. They also possess big capital to invest into future technology advances, thus prohibiting companies of big sizes may destroy potential opportunities. In short, being big can have positive effects, and breaking big companies may lead to economic pains.

On the other hand, enterprises seek to maximize their own profits, and shareholders' interests do not match public interests in general. Therefore, the public sector needs to regulate big players to prevent their excessive aggregation of power from interfering with competition and consumers interests. However, it remains a big challenge to (1) determine *what* and *how* to regulate to allow the society to both minimize the economical impact of the regulations and maintain the gain from economy of scale, and (2) to gain a clear understanding of *what results* one may expect from the defined regulations.

3 PANEL 2 PRESENTATION: HOW DID WE GET HERE?

Christian Huitema presented opening talk in Panel 2 on *How Does Centralization Start?* (See [3] for presentation slides). He pointed out that the Internet centralization trend started many years back with companies investing efforts into unsolved problems. As the Internet growth exploded during 90's, information search became an immediate need. Early attempts include Archie, Altavista, and *Yahoo!* to meet users' needs, and a relative latecomer Google quickly developed new search technologies and advanced past the competition. There were also plenty of earlier diverse efforts on identities, contacts, and presence, they were overtaken by centralized services such as AOL, MySpace, and eventually Facebook won the race.

While search and social networking apps started with centralized servers, email service, as the first popular Internet app, started with predominant distributed deployment. Although some Internet portal service providers, such as AOL and *Yahoo!*, provided email service to residential customers, majority of institutions and companies deployed their own email servers. However, as time went, the emerging IT giant Google was able to invest and provide high quality email service, in particular leveraging its scale to support effective spam mitigation, and eventually consolidated a large portion of email service by Gmail in many regions. Web hosting is another similar case, with many institutions and companies running their own web servers at the beginning. Again as time went on, increasing network security threats, such as viruses and DDoS attacks, together with shortage of IT manpower, resulted in web services being consolidated by dominant players such as Cloudflare and GoDaddy. Big web

hosting companies can leverage competitive advantage from economies of scale, and can invest in additional capacity to mitigate DDoS through building castles with strong walls (well protected data centers). Enterprises today feel very exposed if they set up their own online servers as a critical part of their service provision, as they could be brought to a halt by a DDOS attacks, not to mention possible compromises of the server software and user data.

Once those early application providers started their business, multiple factors have driven them to move their offerings up the protocol stack and centralize further. First, it is costly and time consuming to deal with multiple OSes, multiple versions to coordinate development efforts. The problem becomes much easier if one can i) develop applications running on centralized servers, and ii) control the client platforms, which enables one to simply ship the app code whenever a new feature is added.

Second, the emergence of big data and machine learning has facilitated centralization. Some service providers (SPs) have started with free services to attract users. To serve users well requires that one knows the users well. Gaining more users lead to bigger data collections, which enables development of better services, hence more customers, creating a positive feedback circle: more users $\implies$ more data for ML $\implies$ better services $\implies$ higher revenues and more users. One example that shows the importance of big data is BING's initial lack of success, despite Microsoft's big investments. Possessing user data that a newcomer did not have, enabled Google to tailor search results for individual users in a way that BING could not match.

On the flip side, many of these large scale free services are financed by advertisement, which in turn requires that the advertisement platform amass a collection of user profiles gain the ability for targeted advertisement to maximise the effectiveness. The proliferation of such free services leads to greater reliance on a surveillance economy to sustain this business model. The more these SPs know about you, the better services they can provide, but also the more control they have over you through selective information provision, blurring the line between service and influence. Recent years have seen plenty evidences of such influence as documented in the book "The Age of Surveillance Capitalism" [4].

Huitema observed that the playground of centralized services have changed over time. For example, Microsoft Windows used to dominate desktop market, but it is no longer the case. One might explain this by a few relevant factors, for example the company missed the server market. It also published APIs to promote interoperability with other systems, say enabling people to develop their own version of Powerpoint, a decision that seems (inadvertently) to have

weakened Windows' dominance[3]. However, one can also see the consolidation of overall OS market by moving to open source Linux, except the OSes for phones which settled with Android and iOS [4].

Huitema concluded his presentation by pointing out that decentralized competitors can face uphill battles against centralized providers. Although surveillance capitalisms exploit user privacy to fund free services and attain monopoly to maximize profits, the solution space is largely a political one. Furthermore, decentralized competition requires standards, and standard development is costly in both efforts and time. In contrast, it is far easier, simpler, and faster for monopoly service providers to develop new apps and add new features.

4 WORKSHOP DISCUSSIONS

The workshop discussed many different aspects of centralization: What are the different relevant aspects of centralization (*Categorizing Centralization*)? What factors causes or contributes to centralization (*Factors of Centralization*)? What are future perspectives (*Looking Ahead*)?

We added some editors' comments in italics, marked with the word "*Comment:*".

4.1 Categorizing Centralization

Several workshop attendees brought up the notion that centralization can be sorted into multiple categories.

- Operational centralization, as we have described.
- development consolidation, as measured by the people and organizations that are developing network protocols. Today there seems to be a small number of organization, concentrated in a few coutntries that are developing Internet protocols.
- Centralization and consolidation at different protocol layers. For example, consolidation of transport protocols (e.g., QUIC), consolidation of DNS services, and of course consolidation of applications such as e-mail service.

On the other hand, some participants pointed out that, the passing years have brought changes in the value chain. Networking started from dominance of carriage, then moved to dominance of platforms, and today's dominance is application services. As technologies advance over time, lower layer services became commodity services, and the locus of value and money are shifted up the protocol stack, where one can exploit centralization with minimized cost. Today, application themselves take over the control of everything, creating their own ecosystem; centralized control of DNS

operations and QUIC development could be viewed as part of that ecosystem. Therefore, the observed consolidation symptoms in different categories may all share the same root cause.

4.2 Factors of Centralization

4.2.1 Centralization Is Driven by Economy of Scale.

The workshop participants largely agreed that economy of scale has played a center role in the Internet consolidation. Without regulatory restrictions, markets naturally consolidate when economies of scale come into play. More specific driving factors include:

- The economy of scale enables one to generate the same service outcome with far lower production costs, and consume fewer resources for each instance of the service transaction.
- A large user pool produces big data which helps improve service customisation for each user, letting bigger companies gain an edge over smaller competitors.
- Centralized application developments reduce the number of platforms, hence substantially reduce the cost in development and maintenance, and circumvent interoperability issues.
  Consolidated development and operational efforts also help mitigate technical expertise shortages.
- Most of all, monopoly players can dictate to the market the terms of the service and the service price bought to the market, which causes longer term stagnation of the market and increased inefficiency within the market, which acts as a drag on further innovation.

4.2.2 Applications Got Centralized; Security Threats Further Intensified Centralization.

Both panel speakers pointed out that the playground of networking has changed in fundamental ways, driven by multiple factors. Economy of scale is the first and foremost one, as we discussed already. However, we should not overlook security, or lack of it, which is also a big factor that contributed to application and service centralizations.

The TCP/IP protocol stack did not come with security built-in, and naive IP devices are trivially compromiseable, creating fertile ground for malign exploitation. Universal IP connectivity has been massively abused by DDoS attackers, and all applications today demand secure connections. Therefore, big players have built fortress, the cloud services, with strong walls to deal with DDoS, and demand all users connect to clouds via TLS connections. Today's networking picture seems to be remotely mimicking the one from the 70's, where all user terminals connected to main frame computers via dialup lines.

---

[3]This example shows that publishing APIs could be a worthwhile issue to consider in future decentralization efforts. The same suggest is made in [5].
[4]Although Android's clear linkages to Linux and its open source model also suport a more general case that the OS platform market has largely consolidated upon a small collection of open source instances.

By and large, today's network applications run in clouds; none runs over direct user-to-user or device-to-device communications, with Apple's Airdrop as a noticeable exception.

> *Comment: Airdrop runs over secure connections between iDevices, because Apple installs user identities and corresponding crypto credentials on those devices. In today's Internet, on the other hand, only servers have identities and crypto credentials, but not users and user devices in general. Users only have application provider assigned identifiers, such as a Gmail address or a Facebook ID, and can only be authenticated by those application providers. Consequently, the only means of secure user-to-user communication is to go through cloud-based authentication services.*

4.2.3 Regarding "Allowing Anyone to Send to Anyone Else".
The workshop discussions touched on the issue about whether IP's original model of "enabling anyone to send packets to any other IP address" is one of the major causes of today's security threats. Indeed, IP enabled any host to send packets to any other host, and similarly TCP/UDP let any process connect to any other processes. In its early days, this universal reachability enabled the Internet to grow rapidly and spread to the entire world. Today, unfortunately, this same feature enabled DDoS attacks at global scale, and email spams and phishing attacks fill the majority of today's total email traffic.[5]

The workshop participants pointed out that source IP addresses can easily be spoofed, and there is no effective wide deployment to shut down spammers. Not being able to clearly identifying who is sending data is a problem that we have not worked out effective fixes.

> *Comment: The value of universal reachability has not changed over time. What has changed is TCP/IP's operational environments, from being a playground for research community to political and economical battlefields of worldwide scale. It is not the universal reachability model, but the lacking of security in network communications itself, that is problematic and mandates cloud-based solutions, per-service fortresses and hence centralization.*

4.2.4 Network Security Is Failing.
The above discussion is a reflection of the ineffectiveness of the existing security solutions, which is also one of the major problems raised by the workshop submissions. Vanilla TCP/IP devices have no security, which resulted in massive device breaches that led to large-scale DDoS attacks. Consequently the scale of attacks overwhelms the protection capabilities of individual organizations, driving everyone

---

[5]https://dataprot.net/statistics/spam-statistics/

to outsource their services. Large IT companies, like Amazon, Microsoft, Google etc., can afford the significant capital and labor costs associated with building fortresses to protect their assets, as well as taking in all their customers. There seems a shared pessimism that the possibility of reversing this trend is not in any foreseeable future.

Why has security failed? The participants offered several reasons.

- Premature service releases ("rush to ship") often pay less attention to secure measures than they should, as security features are not attractive product differentiators that are visible to customers immediately. The cost of missing a deadline is immediate, but the cost of lacking security measures may only show up at much later time, and it is often the case that somebody else bears that cost.
- As a technical community, we do care about getting security done right. But getting security right is really hard, and interoperable security is even harder to support decentralized systems and applications. In contrast, it is much easier to develop closed, and hopefully less vulnerable, systems and applications.
- The computer community at a whole is yet to be able to get to next level of software quality that can withstand attacks from well-resourced adversaries, such as the case of nation state sponsored attackers.

> *Comment: 40 years of Internet development has trained a skilled network technical community who know how to forward packets even in very large scale. The same network community is yet to gain expertise in developing security systems.*

4.3 Looking Ahead
4.3.1 Are We Doomed?
As both panel speakers pointed out,

- personalized services require data, which leads to today's practice of services providers collecting all personal data; and
- personal data has value, but only at scale. Individuals cannot monetise their own profile data, yet when this data is amassed in conjunction with the profile data from millions or even billions of other users, the data collection becomes extremely valuable.

These observations project a pessimistic view of the future, suggesting an improbability of having users control their own data. If one believes that data is the new oil for digital economy, then letting application monopolies control data seems an insurmountable barrier towards decentralization.

> *Comment: The key question is "who controls my data". Today it is the application providers, e.g., Google. If one is to imagine an alternative solution for tomorrow, it could be the user oneself – this could be done by*

> *storing the data either in user devices and/or in cloud storage in encrypted forms, which is accessed with users permissions to personalize services.*

4.3.2 Can Network Protocols Prevent Centralization?

A conclusion that one may derive from the discussion on the factors of centralization is that network protocol design *alone* may not stop the market's movement towards centralization.

The Internet started as a decentralized network, where IP's distributed routing enabled a network of decentralized *connectivity*, and individual organizations set up their own servers for the few applications at the time (e.g. DNS, ftp, email), i.e. creating a decentralized world. Many people assumed that the system would stay in that decentralized way.

Looking back, we note that protocols simply facilitate the movement of packets from one place to another, without restricting where this "other place" may be – as packet carriers, protocols do not dictate where packets go, it is application deployments who make that decision. We further note that, in the early days of Internet, many organizations ran application servers to provide *services for their users*; they were not doing it as *revenue generating* business. Once network application services started being revene-generating businesses, economy of scale drives them towards consolidating all, because bigger sizes make it more affordable for service providers to invest into better services, and makes outsourcing services attractive to organizations both in affordability and in quality.

> *Comment: In the absence of effective regulations, it is the economy, not network protocols, that dictates the future direction of a system's evolution. It seems that the networking community did not recognize this fact early on, and the lack of this recognition led to complacent with the consolidating changes at their early stages.*

5 SUMMARY

This report aims to provide a high level summary of the workshop contributions and discussions at and after the workshop. All the workshop contributions are summarized in Appendix A, and the complete workshop recording is available at [6]. We note that Internet consolidation has been observed across all aspects of the Internet, ranging from network access to infrastructure and to application deployment. With limited time, however, the presentations and discussions at this workshop mostly focused on the aspect of data and application centralization.

We use this summary to capture the most important observations from the workshop. Reflecting on those observations, we articulate the lessons to be learn and new insights to be derived. We hope that these lessons and insights can help aid the community's efforts in mitigating centralization.

These lessons and insights are from the authors' perspective, and presented here to seek feedback from the community through open discussions. Our first goal is to reach a clear and shared understanding on the problem space, before diving into the exploration of specific technical solutions.

5.1 Observations

Today's centralization is the result of unregulated economy of scale. This is the high level answer to the question of how we got here, which is not affected by the technical specifics. Similar to the industrial revolution history as described by Huston, the Internet revolution happened so quickly, the technical community seemed blinded on exactly where the train was heading to, and the regulatory sector fell behind. Consequently, the market has largely been left to run on its own to maximize large corporations profits, which runs into conflicts with end users' interests in privacy and sovereignty.

We note that economy of scale motivates corporations to grow big, and that being big in size itself is not the problem. Today's centralization concern is not about the size of those cyber giants, but rather, the control power and the influence they possess over users, and over the society as a whole.

Importance of Security. We observed that security threats are one of the key drivers for unwanted centralization of infrastructures, in particular DDoS attacks whose power is far beyond individuals protection power. Currently, only centralized systems seem to have capabilities to absorb them.

Another security factor that drives centralization is the limitation of the existing Internet security framework. It is a web-server-focused security framework with the web PKI that seems to have led to a server-biased communication style. Users in general do not have their own identities, but by necessity every user is assigned a unique ID by their app providers. As a consequence, even local neighbors have to go to centralized platforms and their authentication services to communicate.

An additional problem is that enabling secure communications (servers with valid certificate chains, implementing security building blocks correctly etc.) is perceived (most likely correctly so) as complex and essentially un-manageable without expert knowledge and service-provider-level scale.

Decentralization by blockchain? The blockchain movement, offen referred to as an enabler for *Permissionless Innovations*, has proposed to move Internet out of centralization by replacing larger parts of the Web and additional infrastructure for a new system based on Blockchain technologies. In addition to a number of economical and technical issues that have been brought up by many people (see [7], [8] as two recent ones), this movement aims to use a technical solution to mitigate an economical problem. As this report argued

that economical forces have driven the Internet itself from a decentralized start to a consolidated state today, a lesson that the Blockchain community may want to pay attention to.

5.2 Reflections

This workshop on centralization in the Internet succeeded in deepening our understanding of the problem space. We recognize that technology alone is not responsible for the observed centralization, it is the uncontrolled economic force to blame. Early Internet applications were largely operated in a decentralized manner when they were provided as user services, and before adversaries recognized the value of this new cyberspace. Their move to revenue-generating business triggered the market force to drive further development towards consolidation, and security threats further intensified the move. Given centralization is fundamentally an economic problem, it cannot be mitigated away solely by technology solutions.

The society thrives on the balance between economy, regulation, and technology. Today, we see an imbalance which is tilted to economy, with the regulations facing challenges of figuring out exactly what should be regulated and how, and the technologies falling behind security threats. To effectively mitigate centralization, we need to hit the nail on the head. To that end, we view that effective regulation and legislation is a deciding factor in curtailing unconstrained market, and that the technical community holds the responsibility to inform the regulatory sectors of what/how to regulate, and work with them in concert to effectively carry out new regulations by providing new technical solutions that can curtail DDoS threats instead of merely absorbing them (which requires the power of centralized systems), and that can enable direct user-to-user communications without reliance on centralized services.

The topic of Internet centralization has caused attention in the IETF community over the last few years, and inspired several ongoing efforts. As an example, draft-nottingham-avoiding-internet-centralization [9] enumerated the negative impacts of control centralization and suggested new technical solutions to mitigate them. We applaud all investigation efforts into new technical solutions, and note that a key challenge to all new solutions is their wide adoption by the market, which do not happen *automatically* if they focus on brining benefits to users and society at large, but do not bring tangible benefits to the existing control parties. We hope that this report contributes to an open discussion to move the community to a shared understanding on the problem space, from there to derive effective next steps to progress towards decentralization.

Economic forces tend to perpetuate the never-stop trend towards capital concentration and infrastructure centralization, and the lessons we have learned from the Internet development – both technical and economical – should help us gain new vigilance to watch out future (re-)centralization, and to start mitigation efforts at its early stage by providing input to regulators and by adjusting technical solutions to meet the new challenges.

As our departing words: The Internet centralization problem will not solve itself. The networking community needs to take actions, now.

REFERENCES

[1] DIN Research Group. *DINRG Workshop on Centralization in the Internet: Workshop Material.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/session/dinrg, June 2021. This web page includes Workshop Agenda, Workshop presentation slies, and all workshop submissions.

[2] Geoff Huston. *Is "Big" Necessarily "Bad"?* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-centrality-in-the-internet-geoff-huston-00, June 2021.

[3] Christian Huitema. *How Does Centralization Start?* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-christian-huitema-how-does-centralization-start-00, June 2021.

[4] Shoshana Zuboff. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* Online: https://www.amazon.com/Age-Surveillance-Capitalism-Future-Frontier/dp/1610395697, January 2019.

[5] Jari Arkko. *The influence of internet architecture on centralised versus distributed internet services.* Online: https://www.tandfonline.com/doi/pdf/10.1080/23738871.2020.1740753?needAccess=true, January 2020. Published in JOURNAL OF CYBER POLICY.

[6] DINRG. *DINRG Workshop on Centralization in the Internet: video recording.* Online: https://youtu.be/1kbSbVjb1ZU, June 2021.

[7] David Rosenthal. *Can We Mitigate Cryptocurrencies' Externalities?* Online: https://blog.dshr.org/2022/02/ee380-talk.html, February 2022. Stanford Technical Seminar.

[8] Cory Doctorow. *Comprehensive synthesis of the technological, ecological and political critique of blockchainism.* Online: https://pluralistic.net/2022/02/14/externalities/#dshr, February 2022.

[9] Mark Nottingham. Centralization, decentralization, and internet standards. Internet-Draft draft-nottingham-avoiding-internet-centralization-05, IETF Secretariat, July 2022. https://www.ietf.org/archive/id/draft-nottingham-avoiding-internet-centralization-05.txt.

[10] Trinh Viet Doan. *Measuring Web Centralization.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-measuring-web-centralization-trinh-viet-doan-00, June 2021.

[11] Henning Schulzrinne. *Centralization is about Control, not Protocols.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slidinterim-2021-dinrg-01-sessa-centralization-is-about-control-not-protocols-henning-schulzrinne-00, June 2021.

[12] Dominique Lazanski. *Protocol and Engineering Effects of Consolidation.* Online: https://datatracker.ietf.org/doc/draft-lazanski-consolidation/, June 2021.

[13] Geoff Huston. *Centrality in the Internet.* Online: https://datatracker. ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-centrality-in-the-internet-geoff-huston-00, June 2021.

[14] Christian Huitema. *We are probably doomed to Internet consolidation, but we should still try do something.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-we-are-probably-doomed-christian-huitema-00, June 2021.

[15] Lixia Zhang. *How We Got from There to Here: Searching for the Root.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-searching-for-the-root-cause-lixia-zhang-00, June 2021. An updated version can be found at https://lixiazhangucla.blogspot.com/2021/06/internet-centralization-how-we-got-from.html.

[16] Jari Arkko. *Mitigation Options against Centralization in DNS Resolvers* . Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-mitigation-options-against-centralization-in-dns-resolvers-jari-arkko-00, June 2021.

[17] Trinh Viet Doan. *Decentralization != Equality?* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-decentralization-equality-trinh-viet-doan-00, June 2021.

[18] Thomas Hardjono. *Decentralized Data Infrastructures for the New Digital Economy.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-decentralized-data-infrastructures-for-the-new-digital-economy-thomas-hardjono-00, June 2021.

[19] Phillip Hallam-Baker. *Autonomy is the real goal.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-autonomy-is-the-real-goal-phillip-hallam-baker-00, June 2021.

[20] Charles Eckels. *Focus on Standards Adoption.* Online: https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/materials/slides-interim-2021-dinrg-01-sessa-focus-on-standards-adoption-charles-eckels-00, June 2021.

A   WORKSHOP SUBMISSIONS

We invited contributions to the workshop that were distributed to participants as input before the workshop. The contribution covered multiple aspects (technical and economic) as well as historic developments.

We have categorized the contributions into these categories: *Rather Problem-Oriented*, *Both Problem- and Solution-oriented*, and *Rather Solution-oriented* below and have provided brief summaries[6] and links to the respective documents in the IETF datatracker.[7]

A.1   Rather Problem-oriented

Trinh Viet Doan: *Measuring Web Centralization [10].*

Abstract: For this paper, the author studied and measured Content Delivery Infrastructure (CDI) centralization and discussed metrics for Internet/Web centralization. The paper suggests CDI penetration as one possible metric to investigate Web centralization.

Henning Schulzrinne: *Centralization is about Control, not Protocols [11].*

Abstract: This paper is discussing *Operational Centralization.* Many common "consumer" applications, i.e., applications widely used by non-technical users, are now provided by a very small number of companies, even if that set of companies differ across geographic regions, or rely on a very small number of implementations even if the applications are largely standards-based. The author argues that a common set of reasons other than protocol design explain this, and thus protocol design may not be the most effective way to counter this development – unless the protocol design anticipates the problems and includes them in the design. The author emphasize that what matters is decentralization of control, not protocol design.

Notes: Dealing with bad actors requires special, expensive expertise, promoting centralization. Decentralized systems need to demonstrate that they can offer users protection against unwanted, harmful or illegal content, at reasonable cost. Decentralized systems increase the total cost of trust. Legal compliance also favors centralization.

Any attempts to reduce centralization have to take non-technical and operational causes into account, to avoid pseudo-solutions. The technical community has tried technical solutions to achieve decentralization in the past; it would be helpful to understand why they have (largely) failed.

Mandated interoperability, through antitrust enforcement, can lower the barrier to entry and reduce provider lock-in, requiring protocols for data portability and interoperability.

Dominique Lazanski and Marc McFadden: *Protocol and Engineering Effects of Consolidation [12].*

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 amd provide some suggestions for advancing the discussion.

Notes:
- technical, economic, engineering areas of Internet consolidation
- summarizing previous discussions, including IAB DEDR workshop
- "layer" vs. "player" consolidation
- econonmic concentration happens everywhere
- can lead to consolidation
- also important to look at consolidation below application layer, e.g., QUIC

---

[6]The summaries are mostly the paper abstracts/introductions.

[7]https://datatracker.ietf.org/meeting/interim-2021-dinrg-01/session/dinrg

- consolidation not always bad: CDN, Cloudflare
- skeptical about regulatory/legal intervention
- DoH/DoT promoting consolidation
- discussion about Internet architecture: intermediaries as a form of consolidation, but consolidation happening in other parts, too: DNS, CDN
- e2e principle may be on a downturn, instead: "edge-to-edge"
- discussion about implications of consolidation on Protocol Design
- questioning IETF folklore about multistakeholder, open participation etc.
- DoH, ECH and Privacy Pass

Geoff Huston: *Centrality in the Internet [13].*

Abstract: In classical public economics, one of the roles of government is to detect, and presumably rectify, situations where the conventional operation of a market has failed. Of course, a related concern is not just the failure of a market, but the situation where the market collapses and simply ceases to exist. Perhaps markets are more than enablers simple transactions between a buyer and a seller. Karl Marx was one of the first to think about the market economy as a global entity, and its role as an arbiter of resource allocation in society. When we take this view, and start looking for potential failure points, one of the signs is that of "choke points" where real investment levels fall, and the fall is masked by a patently obvious masquerade of non-truths taking the place of data and facts. Any study of an economy involves understanding the nature of these choke points. Telecommunication services are not an isolated case but can be seen as just another instance of a choke point in the larger economy. Failure to keep it functioning efficiently and effectively can have implications across many other areas of economic activity.

Notes:
- It's just Economics
- Innovation and Transformation
- Have to understand content economy
- It's big necessarily bad?
- Historical perspective
- Inefficient big business vs. economies of scale vs. too big to fail
- (in-)effectiveness of regulation
- Consolidation
- Examples: DoH, QUIC

A.2 Both Problem- and Solution-oriented

Christian Huitema: *We are probably doomed to Internet consolidation, but we should still try do something [14].*

Abstract: The Internet consolidation and centralization trend has progressed steadily since the 1990's, bringing impactful societal and economic consequences. Periodically, technologists wonder whether this is due to some mistake of their part, maybe a failure to standardize key technologies. Not much ever results from these concerns, because centralization is pushed by powerful economic trends.

Notes:
- centralization caused by powerful economic trends
- monopololies as profit generators, incentives for massive capital injections
- works especially for under-served niches – where standards are (still) lacking
- difficult to induce change with technological progress when market is already dominated
- advertisement business especially prone to concentration – difficult for newcomers to enter when success depends on vast knowledge of users. Massive data bases are major competitive advantage.
- political force may be required, and there are first signs it may be applied
- In that case, it would be good to have new, better technologies and standards

Lixia Zhang: *How We Got from There to Here: Searching for the Root Cause[15].*

Abstract: Recent years have witnessed a growing number of efforts and activities in developing solutions to "decentralize the Internet" – clearly the Internet today is no longer what it is used to be. However, in order to find effective means to move forward, the first step is to understand how we got here. Unprecedented Internet growth took most, if not all, people by surprise and unprepared. Only in retrospect, one gets to see a bit clearly what has happened. It seems to me today's centralization resulted from an imbalance among the three factors: economy, network architecture, and regulation.

Notes:
- three factors: economy, network architecture, and regulation; they went out of balance, why
- Internet is about its applications. Today's Internet no longer enable distributed applications as it did in the past
- Users' identity locked into cloud-based platform – users cannot communicate securely directly – p2p is dead
- cloud appears as only way to communicate securely and defend against attacks (e-mail, DNS, web)
- suggesting new framework for Internet security, enabling direct, secure user-to-user communication at the edge

Jari Arkko: *Mitigation Options against Centralization in DNS Resolvers [16].*

Abstract: Centralization and consolidation of various Internet services are major trends. While these trends have

some benefits - for instance in deployment of new technology - they also have serious drawbacks in terms of resilience, privacy, and other aspects. This contribution focuses on the question of centralization related to DNS resolver services.

Notes:

- discussing DNS centralization through DoT/DoH
- suggesting better solutions for DNS privacy

Trinh Viet Doan, Heikki Hämmäinen, Jörg Ott, and Yiannis Psaras: *Decentralization ≠ Equality? [17].*

Abstract: In the last decade, the Internet has been perceived to become increasingly centralized, as most of the exchanged information is stored, processed, and served by so called "hypergiants". Their large-scale, global infrastructures are able to provide benefits in terms of availability, performance, and security thanks to their resource richness, which provides scale benefits and further drives centralization. However, these increasing centralization trends have lead to concerns regarding user privacy and data sovereignty. As such, the networking community has expressed interest in (empirically) studying the motivation, extent, and implications of such centralization trends in the Internet ecosystem from technological, societal, economical, and legal points of view.

Notes:

- discussing motivation and potential consequences of decentralization
- inequality in P2P/blockchain systems with respect to resource availability and benefits
- centralized systems may provider more equal services, also to less resourceful users
- suggesting discussion of motivations and implications of Internet (de)centralization with associated trade-offs

A.3 Rather Solution-oriented

Thomas Hardjono: *Decentralized Data Infrastructures for the New Digital Economy [18].*

Abstract: The TCP/IP Internet has provided the *communications infrastructure* for billions of people worldwide, with numerous social and economic benefits to humanity. However, as society becomes increasingly data-driven, we face a number of challenges regarding data ownership, centralization and privacy. A new paradigm for trusted *decentralized data infrastructures* is needed that balances the creation of value for data owners with the protection of user privacy

Notes:

- society becomes more data-driven
- many challenges around data ownership, centralization, privacy
- suggesting decentralized data infrastructures ("Data cooperatives")

Phillip Hallam-Baker: *Autonomy is the real goal [19].*

Abstract: The Mathematical Mesh 'Mesh' is a cryptographic security framework that has the goal of making each individual user of the Internet their own ultimate trust authority. Originally planned as a means of achieving a fully decentralized, zero-trust environment as the workshop CFP appears to envisage, it was realized that a less rigid, less ideological approach would better serve the real goal of autonomy.

Notes:

- trust cannot be eliminated, only managed
- ability to minimize the extent to which other parties are trusted and for what risks that is important
- paradox of the Web is that the attempt to decentralize the publication of ideas and knowledge has instead led to an unprecedented concentration of power over information
- the Web gives anyone with access the ability to publish but only a very small number have the ability to be heard
- DNS has long been recognized as the chief centralization point of the Internet and Web
- Mathematical Mesh to give users back control of their identity

Charles Eckels: *Focus on Standards Adoption [20].*

Abstract: The Internet was built on standards developed and deployed in a decentralized yet cooperative fashion, giving rise to a Decentralized Internet that provided a means to connect to a large number of users. This inspired more deployment and collaboration on new standards. But as the Internet grew, so did the number and complexity of standards on which it was built. New standards lacked corresponding efforts to promote adoption and use. Proprietary overlays that were easy to use became very attractive, with large conferences, dedicated support staff, and user forums promoting their adoption and leading to a centralization of the Internet.

Notes:

- more standards (adoption) instead of proprietary overlays
- more international hackathon events

B OTHER ISSUES DISCUSSED AT THE WORKSHOP

Below is a list of important observations and issues discussed at the workshop, as well as suggestions on next step in mitigation actions, that did not get integrated into the main body of this report. We record them here for potential future studies.

B.1 Increased Protocol Complexity

Multiple workshop participants pointed out that network protocols have become increasingly complex over time, making it less feasible for individuals to grasp and contribute, or

to design new solutions. The complexity effectively raises the barrier to entry by a broader community.

Standardization of lower layers leads to higher implementation and deployment cost. Big companies get around standards by developing closed apps, monopolizing apps with much lower cost.

Workshop participants generally agree that protocols get more complex, and standard developments take longer. These problems need and should be addressed. However, one may ask the question of whether these are the root cause of the (centralization) problem. If not, then addressing them alone wound not help much with mitigating the centralization problem.

B.2   Is Technology Part of the Solution Space?

It seems that none of the driving factors to centralization is a technical problem, or at least none of them is largely a technical problem.

If that is the case, is technology part of the solution space? Or what roles can technology play in the solution space?

Do the people at the right pay-grade understand the problem?

Need explanation to policy makers what standards are needed, and what results they may bring – this is what IETF can and should help with.

B.3   Potential Topics for Next Steps

1. could look at every step of Internet development: what could have been done differently?
2. investigation into how/why security failed.
3. To address the problem that individuals can't monetise their own data today, invent new solutions of user controlled data sharing and advertisement usage for tomorrow.
4. kicking off work on measuring/ascertaining centralization in selected areas

DNS
DNS was once decentralized & delegated,
- study what has driven it towards centralization
- e.g., DYN
- Design factors in DNS?
- Zooko's Triangle[8]
- political influence
- take DNS deployment as an example
- try to find out about people's motiviation to outsource deployment

---

[8]https://en.wikipedia.org/wiki/Zooko%27s_triangle

Certificate Authorities: Let's Encrypt
- centralization of PKI
- how decentralized was PKI in the first place

Other Topics
- application of security infrastructure into BGP
- Mail
- CDN

B.4   Economics and antitrust

One should always look to see whether consolidation would benefit consumer welfare. Chicago school of economics had the notion that the old model was not sufficiently focused on consumer benefit.
- argument for concentration in telco business was always: consumers would benefit economically
- often hard to argue against
- Similar argument could be made for Internet services, e.g., Facebook. Users seem to like these services
- it used to be that the amount of capital required to create a national telecom services was beyond the capacity of private capital markets, therefore public investments
- now new world: national interest is quoted again and again, but not with money
- money is sitting in private investment vehicles
- those are heavily concentrated
- VC interest: requires billions of investments, not millions
- relatively lax and free moving capital markets with aggregating capital into free moving vehicles that produce outcomes
- all other player get massive barriers, can't aggregate capital to compete
- real change requires understanding regulation of international movements of capital
- state and national issues need to play a role in order to regulate to create different market outcomes
- until this is done, anything we do on a technical level or on a subject-matter regulatory level is never going to work
- part of this actually a much bigger question than just this industry
- Actual question is how do we create large-scale capital movement to further or fuel national communities rather than having a small number of international bodies run their own agenda
- don't see clean solution rather than going to the root cause

B.5 Countering centralization and possible outlook

- What counters centralization in todays world is actually legislation
- Governments playing national interest card
- Google taking away advertizing business from traditional press but cannot be banned
- traditional publishing business forcing legilators to impose taxes on Google, and subsidize incumbent business
- expecting carriage industry to do the same in the next 5 to 6 years
- "national interest" argument (critical infrastructure, businesses)
- expecting rise of ITU-T and international mechanisms to reinforce national regulatory structures
- as a counter to international centralization (Google is dominant across all markets)
- one large Internet is over: future of a few Google- or Facebook-like ecosystems or lots of little national and regional ones

- so counter to centralization is actually nationalization rather than decentralization
- may want to avoid fatalism
- there are some more achievable things you can do such as privacy laws, e.g.,GDPR
- regulatory mechanisms: data portability, restrictions on end-to-end new markets
- Sherman Act-like laws: making it illegal for companies like google to dominate adjacent markets (e.g., Google moving from Search to everything else)
- more local problem: as a community we have been complicit (techno-optimism)
- illusion that going to back to peer-to-peer protocols will make a difference
- we have to be a lot more skeptical against those arguments because they will translate into the policy domain (politician adopting techno-optimism)
- "new Google or new Facebook will come along"
- consolidation will not be as transient as it was in the past
- Google dominates all markets. May trigger national protections, make the day of global Internet over.