Workgroup: Network Working Group
Internet-Draft:
draft-nottingham-avoiding-internet-
centralization-14
Published: 14 September 2023
Intended Status: Informational
Expires: 17 March 2024
Authors: M. Nottingham

Centralization, Decentralization, and Internet Standards

Abstract

This document discusses aspects of centralization that relate to Internet standards efforts. It argues that while standards bodies have limited ability to prevent many forms of centralization, they can still make contributions that assist decentralization of the Internet.

About This Document

This note is to be removed before publishing as an RFC.

Status information for this document may be found at https://datatracker.ietf.org/doc/draft-nottingham-avoiding-internet-centralization/.

Source for this draft and an issue tracker can be found at https://github.com/mnot/avoiding-internet-centralization.

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 https://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 17 March 2024.

Table of Contents

1. Introduction

One of the Internet's defining features is its lack of any single point of technical, political, or economic control. Arguably, that property assisted the Internet's early adoption and broad reach: because permission is not required to connect to, deploy an application on, or use the Internet for a particular purpose, it can meet diverse needs and be deployed in many different environments.

Although maintaining that state of affairs remains a widely espoused goal, consistently preserving it across the range of services and

applications that people see as "the Internet" has proven elusive. Whereas early services like NNTP and e-mail had multiple, interoperable providers, many contemporary platforms for content and services are operated by single, commercial entities without any interoperable alternative -- to the point where some have become so well-known and important to people's experiences that they are commonly mistaken for the Internet itself. [Komaitis]

These difficulties call into question what role architectural design -- in particular, that overseen by open standards bodies such as the IETF -- can and should play in controlling centralization on the Internet.

This document argues that while decentralized technical standards may be necessary to avoid centralization of Internet functions, they are not sufficient to achieve that goal because centralization is often caused by non-technical factors outside the control of standards bodies. As a result, standards bodies should not fixate on preventing all forms of centralization; instead, they should take steps to ensure that the specifications they produce enable decentralized operation.

Although this document has been discussed widely in the IETF community (see Appendix A), it represents the views of the author, not community consensus. Its primary audience is the engineers who design and standardize Internet protocols. Designers of proprietary protocols and applications can benefit from considering these issues, especially if they intend their work to be considered for eventual standardization. Policymakers can use this document to help characterise abuses that involve centralized protocols and applications and evaluate proposed remedies for them.

Section 2 defines centralization, explains why it is often undesirable but sometimes beneficial, and surveys how it occurs on the Internet. Section 3 explores decentralization and highlights some relevant strategies, along with their limitations. Then, Section 4 makes recommendations about the role that Internet standards can play in controlling centralization. Section 5 concludes by identifying areas for future work.

2. Centralization

In this document, "centralization" is the state of affairs where a single entity or a small group of them can observe, capture, control, or extract rent from the operation or use of an Internet function exclusively.

Here, "entity" could be a person, group, or corporation. An organization might be subject to governance that mitigates

centralization risk (see Section 3.1.3), but that organisation is still a centralizing entity.

"Internet function" is used broadly in this document. Most directly, it might be an enabling protocol already defined by standards, such as IP [RFC791], BGP [RFC4271], TCP [RFC793], or HTTP [HTTP]. It might also be a proposal for a new enabling protocol, or an extension to an existing one.

Because people's experience of the Internet are not limited to standards-defined protocols and applications, this document also considers centralization in functions built on top of standards -- for example, social networking, file sharing, financial services, and news dissemination. Likewise, the networking equipment, hardware, operating systems, and software that act as enabling technologies for the Internet can also impact centralization. The supply of Internet connectivity to end users in a particular area or situation can exhibit centralization, as can the supply of transit between networks (so called "Tier 1" networks).

This definition does not capture all types of centralization. Notably, technical centralization (where, for example, a machine or network link is a single point of failure) is relatively well-understood by engineers, and can be mitigated, typically by distributing a function across multiple components. As we will see, such techniques might address that type of centralization while failing to prevent control of the function falling into few hands. A failure because of a cut cable, power outage, or failed server is well-understood by the technical community, but qualitatively different from the issues encountered when a core Internet function has a gatekeeper.

Likewise, political centralization (where, for example, a country is able to control how a function is supplied across the whole Internet) is equally concerning, but not considered in depth here.

Even when centralization is not currently present in a function, some conditions make it more likely that centralization will emerge in the future. This document uses "centralization risk" to characterise that possibility.

2.1. Centralization Can Be Harmful

Many engineers who participate in Internet standards efforts have an inclination to prevent and counteract centralization because they see the Internet's history and architecture as incompatible with it. As a "large, heterogeneous collection of interconnected systems" [BCP95] the Internet is often characterised as a "network of networks" whose operators relate as peers that agree to facilitate

communication, rather than experiencing coercion or requiring subservience to others' requirements. This focus on independence of action is prevalent in the Internet's design -- for example, in the concept of an "autonomous system".

Reluctance to countenance centralization is also rooted in the many potentially damaging effects that have been associated with it, including:

* *Power Imbalance*: When a third party has unavoidable access to communications, they gain informational and positional advantages that allow observation of behavior (the "panopticon effect") and shaping or even denial of behavior (the "chokepoint effect") [Judge] -- capabilities that those parties (or the states that have authority over them) can use for coercive ends [FarrellH] or even to disrupt society itself. Just as good governance of states requires separation of powers [Madison], so too does good governance of the Internet require that power not be consolidated in one place without appropriate checks and balances.

* *Limits on Innovation*: A party with the ability to control communication can preclude the possibility of "permissionless innovation" -- the ability to deploy new, unforeseen applications without requiring coordination with parties other than those you are communicating with.

* *Constraints on Competition*: The Internet and its users benefit from robust competition when applications and services are available from many providers -- especially when those users can build their own applications and services based upon interoperable standards. When a centralized service or platform must be used because no substitutes are suitable, it effectively becomes an essential facility, which facilitates abuse of power.

* *Reduced Availability*: Availability of the Internet (and applications and services built upon it) improves when there are many ways to obtain access. While service availability can benefit from the focused attention of a large centralized provider, that provider's failure can have a disproportionate impact on availability.

* *Monoculture*: The scale available to a centralized provider can magnify minor flaws in features to a degree that can have broad consequences. For example, a single codebase for routers elevates the impact of a bug or vulnerability; a single recommendation algorithm for content can have severe social impact. Diversity in functions’ implementation leads to a more robust outcome when viewed systemically, because "progress is the outcome of a trial-

and-error evolutionary process of many agents interacting freely." [Aligia]

* *Self-Reinforcement*: As widely noted (see, e.g., [Abrahamson]), a centralized provider's access to data allows it the opportunity to make improvements to its offerings, while denying such access to others.

The relationship between these harms and centralization is often complex; it is not always the case that centralization will lead to them, and when it does, there is not always a direct and simple tradeoff.

For example, consider the relationship between centralization and availability. A centrally operated system might be more available because of the resources available to a larger operator, but their size creates greater impact when a fault is encountered. Decentralized systems can be more resilient in the face of some forms of failure, but less so in other ways; for example, they may be less able to react to systemic issues, and might be exposed to a larger collection of security vulnerabilities in total. As such, it cannot be said that centralization reduces availability in all cases; nor does it improve it in all cases.

This tension can be seen in areas like the cloud and mobile Internet access. If a popular cloud hosting provider were to become unavailable (whether for technical or other reasons), many people's experience of the Internet might be disrupted (especially due to the multiple dependencies that a modern Web site often has; see [Kashaf]). Likewise, a large mobile Internet access provider might have an outage that affects hundreds of thousands of its users, or more -- just as previous issues at large telephone companies precipitated widespread outages. [PHONE]

In both cases, the services are not technically centralized; these operators have strong incentives to have multiple redundancies in place and use various techniques to mitigate the risk of any one component failing. However, they generally do rely upon a single codebase, a limited selection of hardware, a unified control plane, and a uniform administrative practice -- each of which might precipitate a widespread failure.

If there were only one provider for these services (like the telephone networks of old), they would easily be considered as centralized in a way that has significant impact upon availiability. However, many cloud providers offer similar services, and in most places there are multiple mobile operators available. That weakens the argument that there is a link between centralization and their availability, because the function's users can switch to other

providers, or use more than one provider simultaneously; see Section 4.4.

These circumstances suggest one area of inquiry when considering the relationship between centralization and availability of a given function: what barriers are there to switching to other providers (thereby making any disruptions temporary and manageable) or to using multiple providers simultaneously (to mask the failure of a single operator)?

Another example of the need for nuance can be seen when evaluating competitive constraints. While making provision of various Internet functions more competitive may be a motivation for many engineers, only courts (and sometimes, regulators) have the authority to define a relevant market and determine that a behavior is anti-competitive. In particular, market concentration does not always indicate competition issues, so what might be considered undesirable centralization by the technical community might not attract competition regulation.

2.2. Centralization Can Be Helpful

The potential harms of centralization listed above are widely appreciated. Less widely explored is the reliance on centralization by some protocols and applications to deliver their functionality.

Often, centralization is present due to technical necessity. For example, a single, globally coordinated “source of truth” is by nature centralized -- such as in the root zone of the Domain Name System (DNS), which allows human-friendly naming to be converted into network addresses in a globally consistent fashion.

Or, consider IP address allocation. Internet routing requires addresses to be allocated uniquely, but if a single government or company were to capture the addressing function, the entire Internet would be at risk of abuse by that entity. Similarly, the Web's trust model requires a Certificate Authority to serve as the root of trust for communication between browsers and servers, bringing centralization risk that needs to be considered in the design of that system.

Protocols that need to solve the "rendezvous problem" to coordinate communication between two parties who are not in direct contact also require centralization. For example, chat protocols need to coordinate communication between two parties that wish to talk; while the actual communication can be direct between them (so long as the protocol facilitates that), the endpoints' mutual discovery typically requires a third party at some point. From the perspective of those two users, the rendezvous function has centralization risk.

Even when not strictly necessary, centralization can create benefits for a function's users and for society.

For example, it has long been recognised that the efficiencies that come with economies of scale can lead to concentration. [Demsetz] Those efficiences can be passed on to users as higher quality products and lower costs, and might even enable provision of a function that was not viable at smaller scale.

Complex and risky functions like financial services (e.g., credit card processing) are often concentrated into a few specialized organizations, where they can receive the focused attention and expertise that they require.

Centralization can also provide an opportunity for beneficial controls to be imposed. [Schneider2] notes that "centralized structures can have virtues, such as enabling publics to focus their limited attention for oversight, or forming a power bloc capable of challenging less-accountable blocs that might emerge. Centralized structures that have earned widespread respect in recent centuries – including governments, corporations, and nonprofit organizations – have done so in no small part because of the intentional design that went into those structures."

This can be seen when a function requires governance to realize common goals and protect minority interests. For example, content moderation functions impose community values that many see as a benefit. Of course, they can also be viewed as a choke point where inappropriate controls are able to be imposed, if that governance mechanism has inadequate oversight, transparency, or accountability.

Ultimately, deciding when centralization is beneficial is a judgment call. Some protocols cannot function without a centralized function; others might be significantly enhanced for certain use cases if a function is centralized, or might merely be more efficient. In general, though, centralization is most concerning when it is not broadly held to be necessary or beneficial, when it has no checks, balances, or other mechanisms of accountability, when it selects "favorites" which are difficult (or impossible) to displace, and when it threatens the architectural features that make the Internet successful.

3. Decentralization

While the term "decentralization" has a long history of use in economics, politics, religion, and international development, [Baran] gave one of the first definitions relevant to computer networking, as a condition when "complete reliance upon a single point is not always required."

Such technical centralization (while not a trivial topic) is relatively well understood. Avoiding all forms of centralization -- including non-technical ones -- using only technical tools (like protocol design) is considerably more difficult. Several issues are encountered.

First and most critically, technical decentralization measures have at best limited effects on non-technical forms of centralization. Or, per [Schneider1], "decentralized technology alone does not guarantee decentralized outcomes." As explored below in Section 3.1, technical measures are better characterised as necessary but insufficient to achieve full decentralization of a function.

Second, decentralizing a function requires overcoming challenges that centralized ones do not face. A decentralized function can be more difficult to adapt to user needs (for example, introducing new features, or experimenting with user interface) because doing so often requires coordination between many different actors. [Marlinspike] Economies of scale are more available to centralized functions, as is data that can be used to refine a function's design. All of these factors make centralized solutions more attractive to service providers, and in some cases can make a decentralized solution uneconomic.

Third, identifying which aspects of a function to decentralize can be difficult, both because there are often many interactions between different types and sources of centralization, and because centralization sometimes only becomes clear after the function is deployed at scale. Efforts to decentralize often have the effect of merely shifting centralization to a different place -- for example, in its governance, implementation, deployment, or in ancillary functions.

For example, the Web was envisioned and widely held to be a decentralizing force in its early life. Its potential as an enabler of centralization only became apparent when large Web sites successfully leveraged network effects (and secured legal prohibitions against interoperability, thus increasing switching costs; see [Doctorow]) to achieve dominance of social networking, marketplaces, and similar functions.

Fourth, different parties might have good-faith differences on what "sufficiently decentralized" means based upon their beliefs, perceptions and goals. Just as centralization is a continuum, so is decentralization, and not everyone agrees what the "right" level or type is, how to weigh different forms of centralization against each other, or how to weigh potential centralization against other architectural goals (such as security or privacy).

These tensions can be seen, for example, in the DNS. While some aspects of the system are decentralized -- for example, the distribution of the lookup function to local servers that users have the option to override -- an essentially centralized aspect of the DNS is its operation as a name space: a single, global "source of truth" with inherent (if beneficial) centralization in its management. ICANN mitigates the associated risk through multi-stakeholder governance (see Section 3.1.3). While many believe that this arrangement is sufficient and might even have desirable qualities (such as the ability to impose community standards over the operation of the name space), others reject ICANN's oversight of the DNS as illegitimate, favoring decentralization based upon distributed consensus protocols rather than human governance. [Musiani]

Fifth, decentralization unavoidably involves adjustments to the power relationships between protocol participants, especially when it opens up the possibility of centralization elsewhere. As Schneider notes in [Schneider2], decentralization "appears to operate as a rhetorical strategy that directs attention toward some aspects of a proposed social order and away from others", so "we cannot accept technology as a substitute for taking social, cultural, and political considerations seriously." Or, more bluntly, "without governance mechanisms in place, nodes may collude, people may lie to each other, markets can be rigged, and there can be significant cost to people entering and exiting markets." [Bodo]

For example, while blockchain-based cryptocurrencies purport to address the centralization inherent in traditional currencies through technical means, many exhibit considerable concentration of power due to voting/mining power, distribution of funds, and diversity of codebase. [Makarov] Over-reliance on technical measures also brings an opportunity for latent, informal power structures that have their own risks -- including centralization. [Freeman]

Overall, decentralizing a function requires considerable work, is inherently political, and involves a large degree of uncertainty about the outcome. If one considers decentralization as a larger social goal (in the spirit of how the term is used in other, non-computing contexts), merely rearranging technical functions may lead to frustration. "A distributed network does not automatically yield an egalitarian, equitable or just social, economic, political landscape." [Bodo]

3.1. Decentralization Strategies

This section examines some common strategies that are employed to decentralize Internet functions, along with their limitations.

3.1.1. Federation

Protocol designers often attempt to address centralization through federation: designing a function in a way that uses independent instances who maintain connectivity and interoperability to provide a single, cohesive service. Federation promises to allow users to choose the instance they associate with and accommodates substitution of one instance for another, lowering switching costs.

However, federation alone is insufficient to prevent or mitigate centralization of a function, because non-technical factors can create pressure to use a central solution.

For example, the e-mail suite of protocols needs to route messages to a user even when that user changes network locations or becomes disconnected for a long period. To facilitate this, SMTP [RFC5321] defines a specific role for routing users' messages, the Message Transfer Agent (MTA). By allowing anyone to deploy an MTA and defining rules for interconnecting them, the protocol avoids a requirement for a single, central server in that role; users can (and often do) choose to delegate it to someone else, or can run their own MTA.

Running one's own MTA has become considerably more onerous over the years, due in part to the increasingly complex mechanisms introduced to fight unwanted commercial e-mail. These costs create an incentive to delegate one's MTA to a third party who has the appropriate expertise and resources, contributing to market concentration. [Holzbauer]

Additionally, the measures that MTAs use to identify unwanted commercial e-mail are often site-specific. Because large MTAs handle so many more addresses, there is a power imbalance with smaller ones; if a large MTA decides that e-mail from a small one is unwanted, there is significant impact on its ability to function, and little recourse.

XMPP [RFC6120] is a chat protocol that demonstrates another issue with federation: the voluntary nature of technical standards.

Like e-mail, XMPP is federated to facilitate rendezvous of users from different systems - if they allow it. While some XMPP deployments do support truly federated messaging (i.e., a person using service A can interoperably chat with someone using service B), many of the largest do not. Because federation is voluntary, some operators captured their users into a single service, deliberately denying them the benefits of global interoperability.

The examples above illustrate that, while federation can create the conditions necessary for a function to be decentralized, it does not guarantee that outcome.

3.1.2. Distributed Consensus

Increasingly, distributed consensus technologies (such as the blockchain) are touted as a solution to centralization. A complete survey of this rapidly changing area is beyond the scope of this document, but we can generalize about its properties.

These techniques attempt to avoid centralization by distributing the operation of a function to members of a sometimes large pool of protocol participants. Usually, the participants are unknown and untrusted, and a particular task's assignment to a node for handling cannot be predicted or controlled. They typically guarantee proper performance of a function using cryptographic techniques (often, an append-only transaction ledger).

Sybil attacks (where a party or coordinated parties cheaply create enough protocol participants to affect how consensus is judged) are a major concern for these protocols, because it would have the effect of concentrating power into the hands of the attacker. Therefore, they encourage diversity in the pool of participants using indirect techniques, such as proof-of-work (where each participant has to show a significant consumption of resources) or proof-of-stake (where each participant has some other incentive to execute correctly).

While these measures can be effective in decentralizing a function's operation, other aspects of its provision can still be centralized; for example, governance of its design, creation of shared implementations, and documentation of wire protocols. That need for coordination is an avenue for centralization even when the function's operation remains decentralized. For example, the Ethereum "merge" demonstrated that the blockchain could address environmental concerns, but only through coordinated community effort and governance -- coordination that was benign in most eyes, but nevertheless centralized. [ETHEREUM]

Furthermore, a protocol or an application composed of many functions can use distributed consensus for some, but still be centralized elsewhere -- either because those other functions cannot be decentralized (most commonly, rendezvous and global naming; see Section 2.2) or because the designer has chosen not to because of the associated costs and lost opportunities.

These potential shortcomings do not rule out the use of distributed consensus technologies in every instance, but they do merit caution

against uncritically relying upon these technologies to avoid or mitigate centralization. Too often, the use of distributed consensus is perceived as imbuing all parts of a project with "decentralization."

3.1.3. Operational Governance

Federation and distributed consensus can both create the conditions for the provision of a function by multiple providers, but cannot guarantee it. However, when providers require access to a resource or cooperation of others to provide that service, that choke point can itself be used to influence provider behaviour -- including in ways that can counteract centralization.

In these circumstances, some form of governance over that choke point is necessary to assure the desired outcome. Often, this is through the establishment of a multi-stakeholder body: an institution that includes representatives of the different kinds of parties that are affected by the system's operation ("stakeholders") in an attempt to make well-reasoned, legitimate, and authoritative decisions.

The most widely studied example of this technique is the governance of the DNS name space, which as a “single source of truth” exhibits centralization. That source of truth is overseen by the Internet Corporation for Assigned Names and Numbers (ICANN), a global multi-stakeholder body with representation from end users, governments, operators, and others.

Another example is the governance of the Web's trust model, implemented by Web browsers as relying parties that have strong incentives to protect user privacy and security, and Certificate Authorities (CAs) as trust anchors that have a strong incentive to be included in browser trust stores. To promote the operational and security requirements necessary to provide the desired properties, the CA/Browser Forum was established as an oversight body that involves both parties as stakeholders.

These examples are notable in that the governance mechanism is not specified in the protocol documents directly; rather, they are layered on operationally, but in a manner that takes advantage of protocol features that enable the imposition of governance.

Governance in this manner is suited to very limited functions like the examples above. Even then, setup and ongoing operation of a governance mechanism is not trivial, and their legitimacy may be difficult to establish and maintain (see, e.g., [Palladino]); by their nature, they are vulnerable to capture by the interests that are being governed.

4. What Can Internet Standards Do?

Given the limits of decentralization techniques like federation and distributed consensus, the voluntary nature of standards compliance, and the powerful forces that can drive centralization, it is reasonable to ask what standards efforts like those at the IETF can do to accommodate helpful centralization while avoiding the associated harms -- while acknowledging that the distinction itself is a judgment call, and inherently political.

The subsections below suggest a few concrete, meaningful steps that standards bodies can take.

4.1. Bolster Legitimacy

Where technical standards have only limited ability to control centralization of the Internet, legal standards (whether regulation, legislation, or case law) show more promise, and are actively being considered and implemented in various jurisdictions. However, regulating the Internet is risky without a firm grounding in the effects on the architecture, informed by a technical viewpoint.

That viewpoint can and should be provided by the Internet standards community. To effectively do so, its institutions must be seen as legitimate by the relevant parties -- for example, competition regulators. If the IETF is perceived as representing or being controlled by "big tech" concerns or only US-based companies, its ability to guide decisions that affect the Internet will be diminished considerably.

The IETF already has features that arguably provide considerable legitimacy; for example, open participation and representation by individuals rather than companies both enhance input legitimacy; a well-defined process with multiple layers of appeals and transparency contributes to throughput legitimacy, and a long history of successful Internet standards provides perhaps the strongest source of legitimacy for the IETF -- its output.

However, it is also widely recognized the considerable costs (not just financial) involved in successfully participating in the IETF have a tendency to favour larger companies over smaller concerns. Additionally, the specialised and highly technical nature of the work creates barriers to entry for non-technical stakeholders. These factors have the potential to reduce the legitimacy of the IETF's decisions, at least in some eyes.

Efforts to address these shortcomings are ongoing; see, for example, [RFC8890]. Overall, bolstering the legitimacy of the organization should be seen as a continuous effort.

When engaging in external efforts, the IETF community (especially, its leadership) should keep firmly in mind that its voice is most authoritative when focused on technical and architectural impact.

4.2. Focus Discussion of Centralization

Centralization and decentralization are increasingly being raised in technical standards discussions. Any claim needs to be critically evaluated: as discussed in Section 2, not all centralization is automatically harmful, and per Section 3, decentralization techniques do not automatically address all centralization harms -- and they may bring their own risks.

However, standards participants rarely have the expertise or information available to completely evaluate those claims, because the analysis involves not only technical factors, but also economic, social, commercial, and legal aspects. For example, economies of scale can cause concentration due to the associated efficiencies [Demsetz], and so determining whether that concentration is appropriate requires a detailed economic analysis that is not in scope for a technical standards body. Furthermore, claims of centralization may have other motivations; in particular, they can be proxies for power struggles between actors with competing interests, and a claim of centralization might be used to deny market participants and (perhaps more importantly) users the benefits of standardization.

Therefore, approaches like requiring a "Centralization Considerations" section in drafts, gatekeeping publication on a centralization review, or committing significant resources to searching for centralization in protocols are unlikely to improve the Internet.

Similarly, refusing to standardize a protocol because it does not actively prevent all forms of centralization ignores the very limited power that standards efforts have to do so. Almost all existing Internet protocols -- including IP, TCP, HTTP, and DNS -- fail to prevent centralized applications from using them. While the imprimatur of an Internet Standard is not without value, merely withholding it cannot prevent centralization.

Discussions should thus be very focused and limited, and any proposals for decentralization should be detailed, so their full effects can be evaluated. [Schneider1] implores that proposals to decentralize be "really, really clear about what particular features of a system a given design seeks to decentralize" and promotes borrowing remedies from more traditional governance systems, such as separation of powers and accountability.

When evaluating claims that a given proposal is centralized, the context of those statements should be examined for presuppositions, assumptions, and omissions. One framework for critical interrogations is offered by [Bacchi], which can be adapted for centralization-related discussions:

1. What is the nature of the centralization that is represented as being problematic?

2. What deep-seated presuppositions or assumptions (conceptual logics) underlie this representation of the "problem"?

3. How has this representation of the problem come about?

4. What is left unproblematic in this problem representation? Where are the silences? Can the "problem" be conceptualized differently?

5. What effects are produced by this representation of the “problem”?

6. How and where has this representation of the “problem” been produced, disseminated, and defended? How has it been and/or how can it be disrupted and replaced?

4.3. Target Proprietary Functions

Functions that are widely used but lacking in interoperability are ripe for standardisation efforts. Targeting prominent and proprietary functions (e.g., chat) is appropriate, but so are smaller efforts to improve interoperability and portability of specific features that are often used to lock users into a platform; for example, a format for lists of contacts in a social network.

A common objection to this approach is that adoption is voluntary; there are no "standards police" to mandate their use or enforce correct implementation. For example, specifications like [ACTIVITYSTREAMS]) were available for some time without being used in a federated manner by commercial social networking providers.

That objection ignores that while standards aren't mandatory, legal regulation is. Legal mandates for interoperability are increasingly proposed by policymakers as a remedy for competition issues (see, e.g., [DMA]). This appetite for regulation presents an opportunity for new specifications to decentralize these functions, backed by a legal mandate in combination with changing norms and the resulting market forces [Lessig].

That opportunity also presents a risk, if the resulting legal regulation is at odds with the Internet architecture. Successfully

creating standards that work in concert with legal regulation presents many potential pitfalls, and will require improved and new capabilities (especially liaison), and considerable effort. If the Internet community does not make that effort, it is likely that regulators will turn to other sources for interoperability specifications.

4.4. Enable Switching

The ability to switch between different function providers is a core mechanism to control centralization. If users are unable to switch they cannot exercise choice or fully realize the value of their efforts, because, for example, "learning to use a vendor's product takes time, and the skill may not be fully transferrable to a competitor's product if there is inadequate standardization." [FarrellJ]

Therefore, standards should have an explicit goal of facilitating users' switching between implementations and deployments of the functions they define or enable.

One necessary condition for switching is the availability of alternatives; breadth and diversity of implementation and deployment are required. For example, if there is only a single implementation of a protocol, applications that use it are vulnerable to the control it has over their operation. Even Open Source projects can be an issue in this regard if there are factors that make forking difficult (for example, the cost of maintaining that fork). Section 2.1 of [RFC5218] explores some factors in protocol design that encourage diversity of implementation.

The cost of substituting an alternative implementation or deployment by users is another important factor to consider. This includes minimizing the amount of time, resources, expertise, coordination, loss of functionality, and effort required to use a different provider or implementation -- with the implication that the standard needs to be functionally complete and specified precisely enough to allow substitution.

These goals of completeness and diversity are sometimes in tension. If a standard becomes extremely complex, it may discourage implementation diversity because the cost of a complete implementation is too high (consider: Web browsers). On the other hand, if the specification is too simple, it may not enable easy switching, especially if proprietary extensions are necessary to complete it (see Section 4.7).

One objection to protocols that enable easy switching is that they reduce the incentives for implementation by commercial vendors.

While a completely commoditized protocol might not allow implementations to differentiate themselves, they provide opportunities for specialization and improvement elsewhere in the value chain [Christensen]. Well-timed standards efforts leverage these forces to focus proprietary interests on top of open technology, rather than as a replacement for it.

4.5. Control Delegation of Power

The users of some functions might realize substantial benefits if they are provided by a third party in communication. Adding a new party to communication can improve:

* *Efficiency*: Many functions on the Internet are more efficient when performed at a higher scale. For example, a content delivery network can offer services at a fraction of the financial and environmental cost that someone serving content themselves would otherwise pay, because of the scale they operate at. Likewise, a two-sided market platform can introduce sizeable efficiencies over pair-wise buyer/seller trading [Spulber].

* *Simplicity*: Completely disintermediating communication can shift the burden of functions onto endpoints. This can cause increased cognitive load for users; for example, compare commercial social networking platforms with self-hosted efforts.

* *Specialization*: Having a function consolidated into a few hands can improve outcomes because of the resulting specialization. For example, services overseen by professional administrators are often seen to have a better security posture and improved availability.

* *Privacy*: For some functions, user privacy can be improved by consolidating their activity to prevent individual behaviors from being discriminated from each other.[Chaum] Introduction of a third party can also enforce functional boundaries -- for example, to reduce the need for users to trust potentially malicious endpoints, as seen in the so-called “oblivious” protocols (e.g., [RFC9230]) that allow end users to hide their identity from services, while still accessing them.

However, if that new party is able to make their participation "sticky" -- for example, by leveraging their position in the network to require use of an intermediary, by exploiting their access to data, or because it is difficult to switch to another provider of the function -- there is a risk of centralization.

Most often, third parties are added to functions as "intermediaries" or in designated "agent" roles. Designing such functions with

thoughtful constraints on these roles can prevent at least the most egregious abuses of such power.

When adding new parties to a function, two guidelines have proven useful: first, third parties should only be interposed into communication when at least one of the primary parties takes a positive action to do so. Second, third parties should have their ability to observe or control communication limited to what is necessary to perform their intended function.

For example, early deployments of HTTP allowed intermediaries to be interposed by the network without knowledge of the endpoints, and those intermediaries could see and change the full content of traffic by default -- even when they are only intended to perform basic functions such as caching. Because of the introduction of HTTPS and the CONNECT method (see Section 9.3.6 of [HTTP]), combined with efforts to encourage its adoption, those intermediaries are now required to be explicitly interposed by one endpoint, and they only have access to basic routing information.

See [I-D.thomson-tmi] for more guidance on protocol intermediation.

The term "intermediary" is also used (often in legal and regulatory contexts) more broadly than it has been in protocol design; for example, an auction Web site intermediates between buyers and sellers is considered an intermediary, even though it is not formally an intermediary in HTTP (see Section 3.7 of [HTTP]). Protocol designers can address the centralization associated with this kind of intermediation by standardising the function, rather than restricting the capabilities of the underlying protocols; see Section 4.3.

4.6. Enforce Boundaries

Most Internet protocols and applications depend on other, "lower-layer" functions and their implementations. The features, deployment, and operation of these dependencies can surface centralization into functions and applications built "on top" of them.

For example, application protocols require a network to function, and therefore a degree of power over communication is available to the network provider. They might block access to, slow down, or change the content of a specific service for financial, political, operational, or criminal reasons, creating a disincentive (or even removing the ability) to use a specific provider of a function. By selectively hindering the use of some services but not others, network interventions can be composed to create pressure to use those other services -- intentionally or not.

Techniques like encryption can discourage such centralization by enforcing such boundaries. When the number of parties who have access to the content of communication is limited, other parties who handle it but are not party to it can be prevented from interfering with and observing it. Although those parties might still prevent communication, encryption also makes it more difficult to discriminate a target from other users' traffic.

4.7. Consider Extensibility Carefully

The Internet's ability to evolve is critical, allowing it to meet new requirements and adapt to new conditions without requiring a “flag day” to upgrade implementations. Typically, functions accommodate evolution by defining extension interfaces, which allow optional features to be added or change over time in an interoperable fashion.

However, these interfaces can also be leveraged by a powerful entity if they can change the target for meaningful interoperability by adding proprietary extensions to a standard. This is especially true when the core standard does not itself provide sufficient utility on its own.

For example, the SOAP protocol's [SOAP] extreme flexibility and failure to provide significant standalone value allowed vendors to require use of their preferred extensions, favoring those who had more market power.

Therefore, standards efforts should focus on providing concrete utility to the majority of their users as published, rather than being a “framework” where interoperability is not immediately available. Internet functions should not make every aspect of their operation extensible; boundaries between modules should be designed in a way that allows evolution, while still offering meaningful functionality.

Beyond allowing evolution, well-considered interfaces can also aid decentralization efforts. The structural boundaries that emerge between the sub-modules of the function -- as well as those with adjacent functions -- provide touchpoints for interoperability and an opportunity for substitution of providers.

In particular, if the interfaces of a function are well-defined and stable, there is an opportunity to use different providers for that function. When those interfaces are open standards, change control resides with the Internet community instead of remaining in proprietary hands, further enhancing stability and enabling (but not ensuring) decentralization.

4.8. Reuse What Works

When centralization is purposefully allowed in an Internet function, protocol designers often attempt to mitigate the associated risks using technical measures such as federation (see Section 3.1.1) and operational governance structures (see Section 3.1.3).

Protocols that successfully do so are often reused to avoid the considerable cost and risk of re-implementing those mitigations. For example, if a protocol requires a coordinated, global naming function, incorporating the Domain Name System is usually preferable to establishing a new system.

5. Future Work

This document has argued that while standards bodies have little means of effectively controlling or preventing centralization on the Internet through protocol design, there are still concrete and useful steps they can take to improve the Internet.

Those steps might be elaborated upon and extended in future work; doubtless there is more that can be done. New decentralization techniques might be identified and examined; what we learn from relationships with other, more effective regulators in this space can be documented.

Some have suggested creating a how-to guide or checklist for dealing with centralization. Because centralization is so contextual and so varied in how it manifests, this might best be attempted within very limited areas; for example, for a particular type of function, or a function at a particular layer.

The nature of centralization also deserves further study; in particular, its causes. While there is much commentary on factors like network effects and switching costs, other aspects such as behavioural, cognitive, and social and economic factors have received comparatively little attention, although that is changing (e.g., [Fletcher]).

6. Security Considerations

This document does not have a direct security impact on Internet protocols. That said, failure to consider centralization might cause a myriad of security issues; see Section 2.1 for a preliminary discussion.

7. Informative References

[Abrahamson]

Abrahamson, Z., "Essential Data", Yale Law Journal, Vol. 124, No. 3, 2014, <https://www.yalelawjournal.org/comment/essential-data>.

[ACTIVITYSTREAMS] Prodromou, E., Ed. and J. Snell, Ed., "Activity Streams 2.0", W3C CR CR-activitystreams-core-20161215, W3C CR-activitystreams-core-20161215, 15 December 2016, <https://www.w3.org/TR/2016/CR-activitystreams-core-20161215/>.

[Aligia] Aligia, P. D. and V. Tarko, "Polycentricity: From Polanyi to Ostrom, and Beyond", Governance: An International Journal of Policy, Administration, and Institutions, Vol. 25, No. 2, p. 237, April 2012, <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0491.2011.01550.x>.

[Bacchi] Bacchi, C., "Introducing the 'What's the Problem Represented to be?' approach", Chapter 2, Engaging with Carol Bacchi, 2012, <https://library.oapen.org/bitstream/handle/20.500.12657/33181/560097.pdf?sequence=1#page=34>.

[Baran] Baran, P., "On Distributed Communications: Introduction to Distributed Communications Networks", 1964, <https://www.rand.org/pubs/research_memoranda/RM3420.html>.

[BCP95] Alvestrand, H., "A Mission Statement for the IETF", BCP 95, RFC 3935, October 2004.

[Bodo] Bodó, B., Brekke, J. K., and J.-H. Hoepman, "Decentralization: a multidisciplinary perspective", Internet Policy Review, Vol. 10, No. 2, June 2021, <https://doi.org/10.14763/2021.2.1563>.

[Chaum] Chaum, D. L., "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms", Communications of the ACM, Vol. 24, No. 2, February 1981, <https://dl.acm.org/doi/10.1145/358549.358563>.

[Christensen] Christensen, C., "The Law of Conservation of Attractive Profits", Harvard Business Review, "Breakthrough Ideas for 2004", February 2004.

[Demsetz] Demsetz, H., "Industry Structure, Market Rivalry, and Public Policy", Journal of Law and Economics, Vol. 16, No. 1, April 1973, <https://www.jstor.org/stable/724822>.

[DMA] The European Parliament and the Council of the European Union, "Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on

contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act)", OJ L 265/1, 12.10.2022, September 2022, <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R1925>.

[Doctorow] Doctorow, C., "Adversarial Interoperability", October 2019, <https://www.eff.org/deeplinks/2019/10/adversarial-interoperability>.

[ETHEREUM] Ethereum, "The Merge", February 2023, <https://ethereum.org/en/upgrades/merge/>.

[FarrellH] Farrell, H. and A. L. Newman, "Weaponized Interdependence: How Global Economic Networks Shape State Coercion", International Security, Vol. 44, No. 1, p. 42, 2019, <https://doi.org/10.1162/ISEC_a_00351>.

[FarrellJ] Farrell, J. and C. Shapiro, "Dynamic Competition with Switching Costs", UC Berkeley Department of Economics Working Paper 8865, January 1988, <http://dx.doi.org/10.2307/2555402>.

[Fletcher] Fletcher, A., "The Role of Behavioural Economics in Competition Policy", March 2023, <http://dx.doi.org/10.2139/ssrn.4389681>.

[Freeman] Freeman, J., "The Tyranny of Structurelessness", Berkeley Journal of Sociology, Vol. 17, 1972, <https://www.jstor.org/stable/41035187>.

[Holzbauer] Holzbauer, F., Ullrich, J., Lindorfer, M., and T. Fiebig, "Not that Simple: Email Delivery in the 21st Century", July 2022, <https://www.usenix.org/system/files/atc22-holzbauer.pdf>.

[HTTP] Fielding, R., Ed., Nottingham, M., Ed., and J. Reschke, Ed., "HTTP Semantics", STD 97, RFC 9110, DOI 10.17487/RFC9110, June 2022, <https://www.rfc-editor.org/rfc/rfc9110>.

[I-D.thomson-tmi] Thomson, M., "Principles for the Involvement of Intermediaries in Internet Protocols", Work in Progress, Internet-Draft, draft-thomson-tmi-04, 8 September 2022, <https://datatracker.ietf.org/doc/html/draft-thomson-tmi-04>.

[Judge] Judge, K., "Intermediary Influence", University of Chicago Law Review, Vol. 82, p. 573, 2014, <https://scholarship.law.columbia.edu/faculty_scholarship/1856>.

[Kashaf] Kashaf, A., Sekar, V., and Y. Agarwal, "Analyzing Third Party Service Dependencies in Modern Web Services: Have We Learned from the Mirai-Dyn Incident?", October 2020, <https://dl.acm.org/doi/pdf/10.1145/3419394.3423664>.

[Komaitis] Komaitis, K., "Regulators Seem to Think That Facebook Is the Internet", August 2021, <https://slate.com/technology/2021/08/facebook-internet-regulation.html>.

[Lessig] Lessig, L., "The New Chicago School", Journal of Legal Studies, Vol. 27, June 1998, <https://www.journals.uchicago.edu/doi/10.1086/468039>.

[Madison] Madison, J., "The Structure of the Government Must Furnish the Proper Checks and Balances Between the Different Departments", The Federalist Papers, No. 51, February 1788.

[Makarov] Makarov, I. and A. Schoar, "Blockchain Analysis of the Bitcoin Market", National Bureau of Economic Research, Working Paper 29396, October 2021, <https://www.nber.org/papers/w29396>.

[Marlinspike] Marlinspike, M., "Reflections: The ecosystem is moving", May 2016, <https://signal.org/blog/the-ecosystem-is-moving/>.

[Musiani] Musiani, F., "Alternative Technologies as Alternative Institutions: The Case of the Domain Name System", The Turn to Infrastructure in Internet Governance, 2016, <https://link.springer.com/chapter/10.1057/9781137483591_4>.

[Palladino] Palladino, N. and N. Santaniello, "Legitimacy, Power, and Inequalities in the Multistakeholder Internet Governance", 2020, <https://link.springer.com/book/10.1007/978-3-030-56131-4>.

[PHONE] "Computer Failure Paralyzes Region's Phone Service", June 1991, <https://www.washingtonpost.com/archive/politics/1991/06/27/computer-failure-paralyzes-regions-phone-service/0db94ac7-89f0-446e-ba33-c126c751b251/>.

[RFC4271] Rekhter, Y., Ed., Li, T., Ed., and S. Hares, Ed., "A Border Gateway Protocol 4 (BGP-4)", RFC 4271, DOI

10.17487/RFC4271, January 2006, <https://www.rfc-editor.org/rfc/rfc4271>.

[RFC5218] Thaler, D. and B. Aboba, "What Makes for a Successful Protocol?", RFC 5218, DOI 10.17487/RFC5218, July 2008, <https://www.rfc-editor.org/rfc/rfc5218>.

[RFC5321] Klensin, J., "Simple Mail Transfer Protocol", RFC 5321, DOI 10.17487/RFC5321, October 2008, <https://www.rfc-editor.org/rfc/rfc5321>.

[RFC6120] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 6120, DOI 10.17487/RFC6120, March 2011, <https://www.rfc-editor.org/rfc/rfc6120>.

[RFC791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <https://www.rfc-editor.org/rfc/rfc791>.

[RFC793] Postel, J., "Transmission Control Protocol", RFC 793, DOI 10.17487/RFC0793, September 1981, <https://www.rfc-editor.org/rfc/rfc793>.

[RFC8890] Nottingham, M., "The Internet is for End Users", RFC 8890, DOI 10.17487/RFC8890, August 2020, <https://www.rfc-editor.org/rfc/rfc8890>.

[RFC9230] Kinnear, E., McManus, P., Pauly, T., Verma, T., and C.A. Wood, "Oblivious DNS over HTTPS", RFC 9230, DOI 10.17487/RFC9230, June 2022, <https://www.rfc-editor.org/rfc/rfc9230>.

[Schneider1] Schneider, N., "What to do once you admit that decentralizing everything never seems to work", Hacker Noon, October 2022, <https://nathanschneider.info/articles/DecentralHacker.html>.

[Schneider2] Schneider, N., "Decentralization: an incomplete ambition", Journal of Cultural Economy, Vol. 12, No. 4, 2019, <https://osf.io/m7wyj/>.

[SOAP] Mitra, N., Ed. and Y. Lafon, Ed., "SOAP Version 1.2 Part 0: Primer (Second Edition)", W3C REC REC-soap12-part0-20070427, W3C REC-soap12-part0-20070427, 27 April 2007, <https://www.w3.org/TR/2007/REC-soap12-part0-20070427/>.

[Spulber] Spulber, D. F., "Solving The Circular Conundrum: Communication And Coordination In Internet Markets", Northwestern University Law Review, Vol. 104, No. 2,

2010, <https://wwws.law.northwestern.edu/research-faculty/clbe/workingpapers/documents/spulber_circularconundrum.pdf>.

Appendix A. Acknowledgements

This document was born out of early discussions with Brian Trammell during our shared time on the Internet Architecture Board.

Special thanks to Geoff Huston and Milton Mueller for their well-considered, thoughtful, and helpful reviews.

Thanks to Jari Arkko, Kristin Berdan, Richard Clayton, Cory Doctorow, Christian Huitema, Mallory Knodel, Eliot Lear, John Levine, Tommy Pauly, and Martin Thomson for their comments and suggestions. Likewise, the arch-discuss@ietf.org mailing list and Decentralized Internet Infrastructure Research Group provided valuable discussion and feedback.

No large language models were used in the production of this document.

Author's Address

Mark Nottingham
Prahran
Australia

Email: mnot@mnot.net
URI: https://www.mnot.net/