INSIDE THE INTERNET

Nick Merrill [1]* *& Tejas N. Narechania* [2]*

Conventional wisdom—particularly in the legal literatures—suggests that competition reigns the inside of the internet. This common understanding has shaped regulatory approaches to questions of network security and competition policy among service providers. But the original research presented here undermines that long-held assumption. Where the markets for internet traffic exchange (and related services) have long been thought to be characterized by robust competition among public network services providers, our findings suggest that these markets have consolidated, relying instead on primarily private infrastructure. These trends raise a host of concerns across matters of network reliability, online speech, and consumer choice, among others. Indeed, some recent high-profile internet outages reflect some of these concerns. And so we consider how the internet's regulatory *infrastructure might respond to these new facts regarding the internet's interior* network *infrastructure. Specifically, we call for regulation to enhance visibility into the internet's interior and to assure a regime of fair carriage for all the internet's users.*

TABLE OF CONTENTS

[1*] Research Fellow, Center for Long-Term Cybersecurity, University of California, Berkeley.

[2*] Robert and Nanci Corson Assistant Professor of Law, University of California, Berkeley, School of Law. For helpful comments and suggestions, we thank Mat Ford, James Grimmelmann, Chris Hoofnagle, Tian Kisch, Khushali Narechania, Delia Scoville, Scott Shenker, Erik Stallman, Rebecca Wexler, as well as audiences at the Digital Life Initiative at Cornell Tech and the University of California, Berkeley, School of Law. For outstanding research assistance, we thank Jennifer Sun. We also thank the editors of the Duke Law Journal Online for their careful edits and thoughtful suggestions.

Preliminary Draft — Duke Law Journal Online (forthcoming 2023)
email tnarecha@law.berkeley.edu *for latest draft*

INTRODUCTION

On June 8, 2021, the internet seemed to come to a standstill. Suddenly, amazon.com wouldn't respond. Neither would CNN. Pinterest, Reddit, Spotify, Twitch were all down. HBO was inaccessible. Even the official website of the United Kingdom's government—gov.uk—was offline. Sources online speculated that a coordinated cyberattack had caused this sudden series of outages.[3]

In truth, these websites failed simultaneously because a simple error at Fastly, the world's second-largest content delivery network (or CDN for short), unsettled the internet's software supply chain.[4] But the internet is meant to be resilient—to avoid these sorts of cascading, catastrophic failures. In its original architecture, the internet was designed to route around such a problem at any one network services provider.[5]

How, then, could a relatively simple error at one internet services company (in this case, a CDN) metastasize into such a significant issue? Addressing this question requires a look into the internet's topology, alongside the governance and market structures that attend to the internet's interior. Many consumers understand the basics of the internet's edges: We know, for example, that we need a computer (an Apple Macbook, perhaps) with an internet connection (say, Comcast's Xfinity) to access a website (such as Google). But most know far less about how a user's request for Google's services traverses the *middle* of the internet, from Comcast's network to Google's servers and back.

In this Article, we set out new details regarding the internet's interior workings, drawing in part on the original internet measurement research developed by one of us (namely, Merrill). Conventional wisdom—particularly in law and policy contexts—suggests that competition reigns the markets at the middle of the internet.[6] But the findings outlined here suggest that such

---

[3] Ryan Browne, *What is Fastly and why did it just take a bunch of major websites offline?*, CNBC (June 8, 2021), https://www.cnbc.com/2021/06/08/fastly-outage-internet-what-happened.html.

[4] Clare Duffy, *Two Obscure Service Providers Briefly Broke the Internet. It Could Happen Again.*, CNN (June 17, 2021), at https://www.cnn.com/2021/06/09/tech/fastly-cdn-internet-risk/index.html; see also Summary of June 8 Outage, FASTLY (June 8, 2021), https://www.fastly.com/blog/summary-of-june-8-outage.

[5] See infra Part __.

[6] See, e.g., Jonathan E. Nuechterlein & Philip J. Weiser, DIGITAL CROSSROADS (2d ed. 2013) 183–84 ("By most accounts, transit services are highly competitive today. One reason is that … conventional backbone providers now compete not only with one another, but also with alternative mechanisms [including] CDNs."); see also infra notes __ and accompanying text.

competition is now far less robust than typically assumed.[7] Moreover, other trends in the internet's interior point not only towards *consolidation*, but also towards the increasing *privatization* of the internet's constituent networks.[8] Viewed together, these new facts pose significant, but overlooked, internet access, security, and reliability challenges (evinced, for example, by the June 8 outage).

In view of these findings, we contend that regulators must revisit the governance regimes for what has sometimes been known, perhaps too simplistically, as "the market for internet traffic exchange."[9] Specifically, we advocate for new transparency and regulatory regimes to help address the concerns arising out of the privatization and consolidation in these markets. Centralized infrastructures often require centralized risk management, particularly in network contexts.[10] But the opacity and secrecy that attends to the internet's increasingly privatized interior confounds attempts to perform risk analyses and guarantee connectivity in the face of a natural disaster or human error, and so we advocate for expanded disclosure mandates as one part of a more comprehensive federal risk management regime.[11] Moreover, this consolidation at the internet's interior renews debates (familiar to the network neutrality context) regarding consumer choice and speech, and so we make further, if tentative, recommendations for regulating these intermediary markets.[12]

This short Article proceeds in four Parts. In the first, we describe a conventional, if dated, understanding of the internet's core. This conventional wisdom regards the markets for internet transit (and related services) as characterized by robust competition among network services companies offering, essentially, public carriage of internet content.[13] Moreover, we describe how this view has shaped the regulatory environment thus far. In the second Part, we challenge this conventional wisdom, drawing on the original internet measurement research developed by one of us (Merrill) at the University of California, Berkeley, School of Information's Daylight Lab (which Merrill directs). In particular, this research suggests that the market for network services

---

[7] See infra Part ___.

[8] See infra Part ___.

[9] See, e.g., *Restoring Internet Freedom*, Declaratory Ruling, Report and Order, and Order, 33 FCC Rcd. 311, 99 ¶ 164 (2018), https://www.fcc.gov/document/fcc-releases-restoring-internet-freedom-order [hereinafter RIFO].

[10] Kevin Stine, Stephen Quinn, Gregory Witte, Robert Gardner, *Integrating Cybersecurity and Enterprise Risk Management (ERM)*, NAT'L INST. OF STANDARDS & TECH. COMPUT. SEC. RES. CTR. (Oct. 2020), https://doi.org/10.6028/NIST.IR.8286 (emphasizing the value of centralizing risk management in the cybersecurity context).

[11] See infra Part ___.

[12] See infra Part ___.

[13] See, e.g., C. Edwin Baker, Merging Phone and Cable, 17 Hastings Comm. & Ent L.J. 97, 131 (1994); Tim Wu, Why Have A Telecommunications Law? Anti-Discrimination Norms in Communications, 5 J. Telecomm. & High Tech. L. 15, 40 (2006).

at the inside of the internet has shifted over time to a more consolidated set of providers. These providers, moreover, have turned decisively towards relying on private infrastructure. In the third Part, we describe the security and competition concerns (among others) that attend to this new network and market structure. And so, finally, we consider how our regulatory infrastructure ought to respond to these changes in the internet's infrastructure.

COMPETITION INSIDE THE INTERNET?

We begin with some brief historical context regarding the modern internet.[14] In early conceptions, the internet was envisioned as a point-to-point network: Content, hosted by users in their homes and offices, was globally accessible via decentralized networks—essentially, local internet service providers (or ISPs) such as America Online (AOL) or Comcast—that were interconnected via intermediary networks (such as WorldCom).[15] Content requests would traverse a local ISP, one or several intermediate networks (ascending a stack of tiered providers, from Tier-3, to Tier-2, to Tier-1 providers, and then back down again), until finally reaching the content host's ISP and the host itself.[16] Internet access and a spare computer were all that it took to visit—and, critically, create—a website.[17]

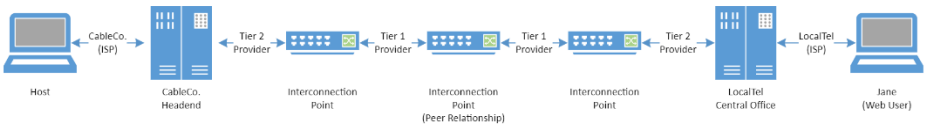

Figure 1*. A graphical representation of the point-to-point vision of the internet. Adapted from Tejas N. Narechania & Erik Stallman,* Internet Federalism*, 34* HARV. J.L. & TECH. *547 (2021).*

---

[14] See generally, Paul Dourish, *Protocols, Packets, and Proximity: The Materiality of Internet Routing*, *in* Signal Traffic: Critical Studies of Media Infrastructures (Lisa Parks & Nicole Starosielski eds., 2015).

[15] See Tung-Hui Hu, *Truckstops on the Information Superhighway: Ant Farm, SRI, and the Cloud*, J. NEW MEDIA CAUCUS, at http://median.newmediacaucus.org/art-infrastructures-hardware/truckstops-on-the-information-superhighway-ant-farm-sri-and-the-cloud/.

[16] Jonathan E. Nuechterlein & Philip J. Weiser, DIGITAL CROSSROADS (2d ed. 2013); see also *Protecting and Promoting Open Internet*, Report and Order on Remand, Declaratory Ruling, and Order, 30 FCC Rcd 5601 (7) ¶¶ 196 (2015), https://docs.fcc.gov/public/attachments/FCC-15-24A1.pdf [hereinafter 2015 OIO].

[17] Internet service from your ISP entitles you to an IP address, a publicly accessible numerical address that uniquely identifies your computer on the global Internet. By associating that IP address with a human-readable name (like a ".com") through the Domain Name System (DNS)–think a phonebook–others can look up your website without having to memorize your IP address (or update their records when your IP address changes).

So-called Tier-1 providers, operating in a competitive market for bandwidth, form the core of this model of the internet. Collectively, Tier-1 providers can reach any location on the internet without having to purchase carriage from another (lower Tier) provider. Some Tier-1 networks sell bandwidth—i.e., capacity—on their networks to smaller Tier-2 and Tier-3 providers (and many of these smaller providers have also built their own networks). And though one might worry about the power of Tier-1 providers to collusively fix prices in downstream bandwidth markets (including those encompassing sales to Tier-2 and Tier-3 providers), several surveys of the relationships among Tier-1 providers suggests that they operate in a competitive and efficient market for bandwidth, one which keeps providers honest and prices low for commodity bandwidth.[18]

The competitive market for bandwidth at the internet's core has led policymakers to believe that the internet's core is "efficient."[19] These networks are agnostic as to the content carried, and capacity can be bought by any internet user at a market-clearing price. In short, this competition (it is said) resolves concerns about price and potential discrimination.

This belief in an efficient core, however, hinges on the assumption that Tier-1 providers (and the market in which they operate) matter as much today as they did in the 1990s, when the internet was newer. But the model of a point-to-point internet, traversing providers agnostic about the content they carry and serve, eventually proved insufficient for today's modern and more scalable consumer internet in at least two ways.

First, this architectural model meant that requests for geographically distant content suffered from high latency. Such latency originally meant that websites would load comparatively slowly (at least by today's standards), causing browsers to "time out" before the content could load, or causing users to lose interest and abandon the request.[20] But as so-called "Web 2.0" applications transformed the internet from a set of static documents into a more dynamic experience—e.g., email inboxes that refresh automatically, or news feeds that are continuously updated—latency hindered significantly the development and utility of such

---

[18] See Dennis Weller & Bill Woodcock. Internet Traffic Exchange: Market developments and policy challenges, OECD Digital Economy Papers, No. 207, OECD Publishing.

[19] AT&T Services Inc., Comments In the Matter of Restoring Internet Freedom (July 17, 2017) at 47 ("All of these commercial relationships have always been unregulated, and the interconnection marketplace has always functioned efficiently…"); see also *Restoring Internet Freedom*, Declaratory Ruling, Report and Order, and Order, 33 FCC Rcd. 311, 99 ¶164 (2018), https://www.fcc.gov/document/fcc-releases-restoring-internet-freedom-order.

[20] See, e.g., Steve Lohr, For Impatient Web Users, an Eye Blink Is Too Long to Wait, N.Y. TIMES (Feb. 29, 2012), https://www.nytimes.com/2012/03/01/technology/impatient-web-users-flee-slow-loading-sites.html.

applications.[21] For example, geographic diversity across ISPs' users, along with the inherent, physical limits of internet facilities, made it difficult or impractical to deliver streaming video or audio reliably for all users through such a widely-distributed network.

Second, though this architectural model prized dispersed and distributed internet content, an increase in cyberattacks proved the truism that security is a collective good,[22] highlighting the value in shared defense.[23] For example, distributed denial of service attacks—or DDoS attacks, which bombard a service with a large amount of traffic thereby making it unavailable to legitimate users—rose in prominence throughout the late 1990s and early 2000s.[24] Such attacks, which are comparatively easy to deploy, became widespread and disruptive. In 2000, for example, sixteen-year-old Michael Calce (known online as *Mafiaboy*) brought down CNN, Yahoo, Amazon, Dell, eBay, and FIFA with a DDoS attack.[25] Such attacks can be difficult for individual websites to defend against: DDoS attacks use numerous endpoints to launch attacks, usually by "hijacking" individual computers that typically source legitimate traffic (e.g., a request to visit CNN.com), but are infected with malware that bombards a target (e.g., CNN) with repeated internet requests.[26] Such a sudden influx of traffic can overwhelm that target, depleting its bandwidth and computational capacity, thereby making the site inaccessible to legitimate viewers. Moreover, because these malicious requests are camouflaged as legitimate ones, content hosts have difficulty

---

[21] See, e.g., Andrea Cardaci, Luca Caviglione, Alberto Gotta, and Nicola Tonellotto, *Performance Evaluation of SPDY over High Latency Satellite Channels*, in PERSONAL SATELLITE SERVICES, (Riadh Dhaou et al. eds. 2013). Moreover, some studies have found that latency is a determinant of consumer trust in, say, a retail website or a banking application, suggesting that the public's willingness to adopt these advances has depended on the nature and quality of network access throughout the internet's structure. See Gerard Ryan and Mireia Valverde, *Waiting for service on the internet: Defining the phenomenon and identifying*, 15 Internet Res. 220 (2005) (citing Sung-Joon Yoon, *The antecedents and consequences of trust in online-purchase decisions*, 16 J. Interactive Marketing 47 (2002)); cf. Infrastructure Investment and Jobs Act, Pub. L. No. 117–58, 135 Stat. 1199 (defining minimum basic requirement for broadband internet access service, or broadband carriage, as including "a latency that is sufficiently low to allow reasonably foreseeable, real-time, interactive application").

[22] Deirdre K. Mulligan & Fred B. Schneider, *Doctrine for Cybersecurity*, 140 J. AM. ACAD. ARTS & SCI. 70, 80 (2011) (stating "[o]ur doctrine of public cybersecurity is rooted in the thesis that cybersecurity is a public good")

[23] Cf. Mazaher Kianpour, Stewart James Kowalski, Harald Øverby, *Advancing the concept of cybersecurity as a public good*, Simulation Modelling Practice and Theory 116 (2022), https://doi.org/10.1016/j.simpat.2022.102493. .

[24] See, e.g., Ketki Arora et al., *Impact Analysis of Recent DDoS Attacks*, 3 Int'l J. Comp. Sci. & Eng'r. 877, 882 (2011) (noting a 1,000-fold increase in DDoS attacks from 2003 to 2011).

[25] Rick Davis, *The History and Future of DDoS Attacks*, CYBERSECURITY MAGAZINE (Jan. 15, 2021), https://cybersecurity-magazine.com/the-history-and-future-of-ddos-attacks/..

[26] Commonly, attackers construct "botnets" from compromised machines. However, DDoS attacks have also been launched through collective, volunteer action, as was the case in Anonymous's attacks against the Church of Scientology. See generally, GABRIELLA COLEMAN, HACKER, HOAXER, WHISTLEBLOWER, SPY: THE MANY FACES OF ANONYMOUS (2014).

separating bad requests from good ones, thereby making it difficult to end the attack without taking the site offline altogether.[27] Victims of DDoS attacks, however, can better address these incidents by sharing internet-traffic–related intelligence. Indeed, the most effective defenses against DDoS attacks rest upon large-scale observation of network traffic: A centralized observer (or federated network of observers) can share intelligence about troublesome sources of internet traffic (thereby helping to separate legitimate requests from malicious ones).[28]

One response to these concerns implicates content delivery networks (or CDNs). CDNs duplicate—i.e., cache[29]—these internet companies' content in localized servers across the internet to minimize latency, while also offering some collective defense against cyberattacks (among other things, too). Hence, where the internet was originally conceived of as a widely distributed network of computers, with individual users connected to each other (as described above), CDNs offer an alternate distribution model. A user would no longer need to traverse a series of networks to connect to a computer maintained by, say, the NBA to obtain the latest basketball statistics. Instead, the NBA can direct these statistics to a CDN in advance of users' requests for them—and users would access that content through that same CDN (with the NBA paying, of course, for computational capacity and network bandwidth). Moreover, that CDN could distribute the NBA's content across its geographically dispersed network of servers—moving, say, Golden State Warriors-related content closer to its facilities in California and Boston Celtics-related content to Massachusetts—thereby reducing latency and enabling more dynamic content (such as video highlights). In either case, users access nba.com—but how nba.com is hosted varies: in the earlier example, nba.com is housed at maintained in one location (say, NBA HQ in NYC); but in our new paradigm, nba.com is distributed and replicated across a network of servers maintained by some third-party.[30] Moreover, a CDN that served the NBA—and the NFL and the NHL and MLB—might have a wider view of internet traffic, and so would be better able to detect and mitigate an attack directed at any one these leagues, based on an understanding of the patterns across all of them.

---

[27] Understanding Denial-of-Service Attacks, Cybersecurity and Infrastructure Agency, https://www.cisa.gov/uscert/ncas/tips/ST04-015 (last updated Oct. 28, 2022).

[28] See CloudFlare, What Is DDoS Mitigation, https://www.cloudflare.com/en-ca/learning/ddos/ddos-mitigation/ ("Cloudflare's network runs Internet requests for millions of websites, creating an advantage in analyzing data from attack traffic around the globe.")

[29] 2015 OIO at ¶¶ 197–98.

[30] As you might imagine, configuring this geographic dispersal in real-time is automated by algorithms, increasingly ones powered by machine learning (commonly termed artificial intelligence, or AI). The challenges for policy of an internet whose core infrastructure is increasingly characterized by such algorithms is discussed at greater length *infra* __.

On one view, then, CDNs helped to make the competitive market for internet traffic exchange even more so. The Federal Communications Commission, for example, describes CDNs as one class of participant in the general market for internet traffic exchange: In 2018, the agency described the market as "emerging and competitive," and CDNs as one "innovative . . . alternative" to other modes of traffic exchange, including the traditional modes transit described above (i.e., to Tier-1 providers and back again), as well as other alternatives (such as direct interconnection, among others).[31]

The Commission's most recent statement builds on a long line of precedent that understands the market for internet traffic exchange as robustly competitive. In 2010, for example, when the Commission decided to issue network neutrality rules in view of consolidation-related concerns in the access network market (i.e., the market for retail broadband subscriptions, such as those users might purchase from Comcast or Verizon), it bluntly noted that it was treating the interconnection market—i.e., the market for traffic exchange inside the internet—as beyond the scope of those rules, implying that the distinct competitive conditions in these respective markets—competition in the market for internet traffic exchange, but greater consolidation in the markets for internet access—justified the differential treatment.[32] Likewise, in 2015, when the Commission reinforced those network neutrality rules, it took only cautious steps in the direction of superintending the market for internet traffic exchange, noting the apparent competition among a wide range of services and service providers in the market (including, e.g., several Tier-1 providers, several CDNs, as well as a ranger of transit service providers).[33] And, finally, the Commission's most recent decision to re-deregulate this market is based on a continued view that market discipline, through the "competitive pressures in the market for Internet traffic exchange," are more efficient than regulatory intervention.[34] Moreover, the Commission's largely consistent statements over the past decade regarding the state of this market seem to reflect a view shared by a wide range of scholars,[35]

---

[31] RIFO at ¶¶ 168–69.

[32] *Preserving the Open Internet Broadband Industry Practices*, Report and Order, 25 FCC Rcd 17905 (21), 39 ¶ 67 n.209, ¶ 113 n.345 (2010), https://docs.fcc.gov/public/attachments/FCC-10-201A1.pdf. It is true that 2010 OIO rule has been criticized on the grounds that network neutrality should rationally extend through, at least, interconnection at the edge. See Jonathan E. Nuechterlein & Philip J. Weiser, DIGITAL CROSSROADS (2d ed. 2013) [pincite]. . Such critiques are certainly fair, but they do not necessarily extend to interconnection (and related) agreements at the interior of the market. If, however, these conclusions are based on incorrect understandings regarding the technical and economic structures of the interior of the internet, then we should revisit them.

[33] 2015 OIO at ¶¶ 197–98.

[34] RIFO at ¶ 170.

[35] Jonathan E. Nuechterlein & Philip J. Weiser, DIGITAL CROSSROADS (2d ed. 2013) 183–84 ("By most accounts, transit services are highly competitive today. One reason is that … conventional backbone providers now compete not only with one another, but also with alternative mechanisms [including] CDNs."); Kevin Werbach, *Only Connect*, 22 BERKELEY TECH. L.J. 1233, 1253–1254 (2007).

policymakers,[36] and market participants.[37] In all, the market for traffic exchange has long been thought to be characterized by a number of different classes of services—transit, CDNs, etc.—as well as a number of providers within each class.

CONSOLIDATION INSIDE THE INTERNET

As noted, one view—a dominant view, it seems—is that competition reigns the market for internet traffic exchange. Some recent incidents, however, might give us reason to question that longstanding assumption. In 2021, for example, we saw at least two high-profile internet outages—each of which might be traced back to an error or glitch at one of these CDN providers.[38] Yet, if the internet's interior was robustly competitive, we might be surprised that an error at any one provider could cause such widespread headaches. Such competition should facilitate the redundancy that is inherent to the internet's design. In short, these incidents may suggest that today's internet more closely resembles a centralized network with few central, critical points-of-failure, rather than the decentralized map of alternative traffic paths that many imagine when considering the internet's structure. Fortunately, this speculation raises a testable question: How consolidated, really, is the market for internet traffic exchange? Our novel study, described below, helps to answer this question.

Methodology

We can begin to address this question with methods and tools used by the community of internet measurement scholars. To compute these results, one of us (Merrill, who directs the Daylight Lab at the University of California, Berkeley, School of Information) collected data about the use of particular CDN

---

[36] See e.g., Response to Written Questions Submitted by Hon. John Thune to Hon. Ajit Pai, Hearing on Oversight of the Federal Commc'ns Comm'n Before the S. Comm. on Com., Sci., and Transp., 114th Cong. 153 (2015) ("Indeed, the best evidence in the record suggests the free market for interconnection has been an unmitigated success, with transit rates falling 99 percent over the last decade."); Dissenting Statement of Comm'r Michael O'Rielly, 2015 OIO at 394 (remarking that the market for Internet traffic exchange is a "thriving, competitive market"); *Broadband Connectivity Competition Policy*, FTC Staff Report (June 2007), at 26 ("To date, market forces have encouraged interconnection among backbones and between backbones and last-mile ISPs."); RIFO at ¶ 169 ("We believe that market dynamics, not Title II regulation, allowed these diverse [alternative internet traffic exchange] arrangements to thrive").

[37] See, e.g., AT&T Services Inc., Comments In the Matter of Restoring Internet Freedom (July 17, 2017) at 47 ("All of these commercial relationships have always been unregulated, and the interconnection marketplace has always functioned efficiently…"); Comcast Corp., Reply Comments In the Matter of Restoring Internet Freedom (Aug. 30, 2017) at 37 (interconnection is a "well-functioning marketplace").

[38] See Jim Salter, *Today's massive Internet outage comes courtesy of Akamai Edge DNS*, Ars Technica (July 22, 2021), https://arstechnica.com/gadgets/2021/07/todays-massive-internet-outage-comes-courtesy-of-akamai-edge-dns/; Annie Palmer, *Dead Roombas, stranded packages and delayed exams: How the AWS outage wreaked havoc across the U.S.*, CNBC (Dec. 9, 2021), https://www.cnbc.com/2021/12/09/how-the-aws-outage-wreaked-havoc-across-the-us.html; see also Fastly outage, discussed *supra*.

providers across the world's top websites in conjunction with W3Techs, an organization that collects data, via surveys and technical means, about the use of various internet technologies.[39] Using a list of the world's top websites,[40] W3Techs used data traces to determine which CDNs, if any, those websites rely upon.[41] More specifically, such data traces inspect the internet responses to a user's request for a website. Such responses are typically composed of many packets; websites are not delivered as single files, but rather many as discrete components—i.e., packets—that are delivered from their sources and received and "assembled" by the requesting computer. Each of these packets contains a "header," which includes certain metadata about the website content enclosed within the packet. Among that metadata is the IP address that originated the packet. By cross-referencing these IP addresses to lists of known service providers, we can determine which packets were delivered by particular CDNs. From this data, the Daylight Lab can compute an overall picture of the market for CDN services over time. If Cloudflare, for example, is found to deliver 76% of the packets in the survey, we would estimate its market share at 76%. We have compiled historical data on the market for CDNs dating from January 2017 to December 2022. And the data presented here has been used by, for example, the Internet Society (a nonprofit organization founded by two of the internet's so-called "founding fathers," Vint Cerf and Bob Kahn[42]) to describe the state of the internet.[43]

Our "whole-web" figures use, as a baseline, the 15,000,000 most popular websites as measured by Chrome's User Experience Research dataset. Of those top 15,000,000 websites, analysis of packets delivered on behalf of those websites reveals that 23.6% use a CDN to deliver service. We treat these websites as the effective market for CDN services. Of that understanding of the market, 91% of

---

[39] We compiled this data in conjunction with the Internet Society, whose financial support funded aspects of this data collection. The code we used for all data collection and analysis is available at https://github.com/elsehow/taaraxtak

[40] The data reported in this paper was created in reference to the Chrome User Experience Report, which recent internet measurement research has found to be the most comprehensive and accurate list of top websites available. See A World Wide View of Browsing the World Wide Web, ACM Internet Measurement Conference (2022) https://dl.acm.org/doi/10.1145/3517745.3561418. Historical data generated before May 1, 2022 was created by reference to Alexa Internet rankings, a public resource, commonly used in other internet measurement research. See for example https://dl.acm.org/doi/10.1145/3278532.3278552 for use of Alexa rankings to understand CDN usage specifically. The service, founded in 1996 by Internet Archive steward Brewster Kahle, was acquired by Amazon in 1999. While Amazon will discontinue Alexa on May 1, 2022, the data we report in this paper was generated prior to its shutdown. See Pulling Rank: The Legacy of Alexa Internet, The Data Horde (Apr. 22, 2022), https://datahorde.org/pulling-rank-the-legacy-of-alexa-internet/.

[41]

[42] See, e.g., Adam Mann, *Father of the Internet, Vint Cerf, on Creating the Interplanetary Internet*, WIRED (July 5, 2013).

[43] Market Concentration, INTERNET SOCIETY: PULSE, at https://pulse.internetsociety.org/concentration

the packets delivered in response to requests are delivered by one of three CDNs: Cloudflare, Fastly, and Amazon.

We acknowledge some drawbacks of our methodological approach. For example, there is some uncertainty about how, exactly, to calculate the relevant market. As suggested, there is a long tail of less-popular websites—about three-quarters of all websites—that use no CDN at all. While these account for a large portion of all websites, they account for a vastly smaller portion of web traffic. And other internet traffic, such as for streaming services, proprietary CDNs, or CDN transit, is excluded from our baseline.[44]

But we do not think that this uncertainty about the precise baseline against which to measure a CDNs' share undermines our basic point. Our data captures the market for websites that require a CDN and do not have the capacity to uild and deploy one for themselves. As our analysis reveals, this is a sizeable proportion of the overall web (23.6% of the top 15,000,000 websites) that accounts for a substantial portion of web traffic. Moreover, any remaining uncertainty serves to reinforce our point, also elaborated below, that we require greater transparency into how, exactly, traffic flows across the internet for a variety of purposes.

Results and Analysis

The Rise of—and the Consolidation in—the CDN Market. In the last thirty years, CDNs have grown rapidly along dimensions of both size (i.e., volume of data served) and scale (i.e., number and variety of users). Before Akamai (a leading CDN) was founded in 1998, no website used a CDN. Today, *every* website in the top 1,000 websites uses a CDN, and over 99.9% percent of the top 10,000 websites do.

These results reflect the major change in the structure of the internet described above. To appreciate the magnitude of this change, compare Cogent Communications, (considered by many to be a Tier-1 provider) with Akamai (a leading CDN). In March 2022, coverage of Cogent's decision to terminate service to Russia (in the wake of the conflict in Ukraine) noted that Cogent carried roughly 25 percent of the world's internet traffic.[45] Akamai is responsible for roughly the same volume of traffic.[46] Yet the way these two entities handle

---

[44] See, e.g., infra note 48.

[45] Igor Bonifacic, *Internet backbone provider Cogent cuts off service to Russia*, Engadget (Mar. 5, 2022), https://www.engadget.com/cogent-communications-223135454.html.

[46] 2015 OIO at n. 491 (citing Akamai Comments at 4 ("At any given time Akamai delivers between 15-30% of all web traffic, resulting in over two trillion interactions delivered daily.")).

Keen readers may discern a discrepancy between the statistic reported here and our results presented later: Here, we note that Akamai is responsible for about 25% of the internet's traffic; whereas later we ascribe to Cloudflare (a competitor) over 80% the market (for a total exceeding 100%). What

traffic is substantially different. As described above, Cogent is agnostic about the bits that travel over its networks. To invoke a (perhaps tired and tortured) metaphor: Cogent offers the rough equivalent of basic postal delivery. Cogent's customers provide Cogent with packages (namely, packets) that have "to" and "from" addresses, and Cogent delivers those packages for a price. Akamai, by contrast, is in the business of logistics (including, but not limited to, postal delivery). Akamai offers to warehouse its customers' data, and in response to user orders (i.e., internet requests) Akamai generates packages, completes the "to" and "from" fields, and assumes responsibility for their delivery. For that last step—delivery—Akamai may (but need not—as we elaborate *infra*) purchase bandwidth from Cogent or other Tier-1 providers. Hence, Akamai and other CDNs now intermediate the relationship between the internet's core and its users. And given their broad popularity, they play this intermediary role for a vast proportion of internet activity. Even if, then, competition among Tier-1 providers may remain strong, the practical benefits conferred by CDNs means that most web properties must rely on the full array of logistics services offered by CDNs, and not the mere delivery services sold by Tier-1 providers. Indeed, the widespread prevalence of CDNs noted above may help to confirm their status as a practical necessity.

Hence, CDNs' customers are largely companies that use the internet to conduct business. Such companies encompass a wide range of products and services: some may be e-commerce websites; others might be professional services firms; still others might offer ad- or subscription-funded news, reporting, or commentary. To be sure, some large and technology-first businesses, such as Google and Netflix, operate their own proprietary CDNs.[47] But the majority of enterprises by number employ an outside CDN to facilitate their business.[48] And

---

gives? We can resolve the apparent dilemma by drawing a distinction between the *volume* of traffic served and the *share* of websites served. Akamai delivers lots of content—i.e., is responsible for a lot of traffic—for many large websites (such as eBay). By contrast, Cloudflare appears to deliver content for many websites, even if many of them are quite small. Stated simply, Akamai rules the top while Cloudflare rules the long tail.

We note one further caveat: The explanation we offer here is our best understanding of the data we have (including Akamai's self-reported traffic statistics). But, because of the limited visibility into global internet traffic (a problem of transparency we address more fully *infra*), we can only observe certain slices of observable data (namely, what the data that is ultimately delivered to us and other end-users) and must make inferences about the rest.

[47] *Cloud CDN*, GOOGLE.COM, https://cloud.google.com/cdn (last visited Feb. 5, 2023); *Open Connect*, NETFLIX.COM, https://openconnect.netflix.com/en/ (last visited Feb. 5, 2023).

[48] A comprehensive list of companies that run proprietary CDN infrastructure is difficult to come by. However, hypergiants like Google, Netflix, Meta, and Apple are known to deploy their own infrastructure. See Petros Gigis, et al., *Seven Years in the Life of Hypergiants' Off-Nets*, *in* ACM SIGCOMM 2021 CONFERENCE PROCEEDINGS, https://doi.org/10.1145/3452296.3472928 (2021). These providers also serve their own traffic. For an internet-based company looking to serve content they originate, they must either host it with a large company or pay one of a small handful of CDNs. The approximately 3.7M websites that use some CDN service, identified in our methods, fall into the latter

many smaller entities may not even understand that, as they build a website, they are shopping for a CDN: CDN services are often packaged into the widely-available (and widely-used) online web design services. If, for example, a small business employs Shopify to manage its webstore, that small business becomes reliant on whatever CDN(s) that Shopify has employed.[49]

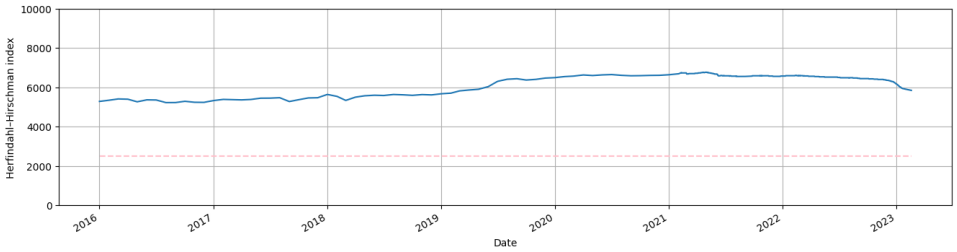

Figure 2*. The Herfindahl–Hirschman index (HHI) of the market for CDNs across the whole web, from 2016 to 2023. The dotted line shows the 2,500 threshold designating a highly concentrated market.*

Of these CDNs, Cloudflare is far and away the most dominant. Over seventy percent of websites using a CDN—over 99% of the top 10,000 websites, recall—rely on Cloudflare for such services.[50] After Cloudflare, Fastly serves about six percent of the market, and Amazon's Cloudfront serves just over five percent. In all, as *Figure 3* (and *Appendix Table 1*) suggest, only eleven providers control 99% of the CDN market. Economists and antitrust authorities sometimes measure market concentration using the Herfindahl–Hirschman Index (or HHI), which scales from 0 to 10,000. Values over 2,500 generally denote a "highly concentrated" market.[51] The market for CDN services currently weighs in at an HHI of 5,846 (*Figure 2*), and, indeed, has remained well above this benchmark for much of the regulatory history, described above, which treats this market as presumptively competitive.[52] But these market share statistics are strongly

---

category: we are able to identify them because they use a well-known commercial CDN to provide service.

[49] Some applications may run their own special-purpose CDNs available only to their customers. For example, Shopify runs a CDN specifically for user images. See https://cdn.shopify.com/. CDNs like this seem unlikely to meaningfully improve consumer choice; even if Shopify users can opt out of using Shopify's CDN for their images, they are still obligated to rely on Shopify's infrastructure for their application, including any CDNs on which Shopify relies.

[50] See supra note ___ for an explanation of an apparent—but resolvable—conflict in the data reported here.

[51] U.S. Department of Justice & FTC, Horizontal Merger Guidelines § 5.3 (2010).

[52] For reference, the markets for other cloud services—web hosting, for example—are far more competitive, with an HHI of only 153. HHI is computed as the sum of the squares of each firm′s market share. The resulting value ranges from 0 to 10,000. A high HHI indicates a few firms that dominate; a low

suggestive of a concentrated market. We acknowledge, as we must, that any complete analysis of market concentration would be more complicated, and would include more difficult questions of market definition and the substitutability of other options. But the practical benefits conferred by CDNs, together with the indicia of concentration (and pervasiveness) noted above, suggest that this market is far less competitive than the one usually ascribed to the Tier-1 providers at the core of the core of the internet. Stated otherwise, the single "market for internet traffic exchange" seems, instead, to be two markets: one characterized by the commodity bandwidth; and a second, "interpositioned" between this traditional core and its end-users, offering a wide range of "in network processing" services, including caching and security.[53]

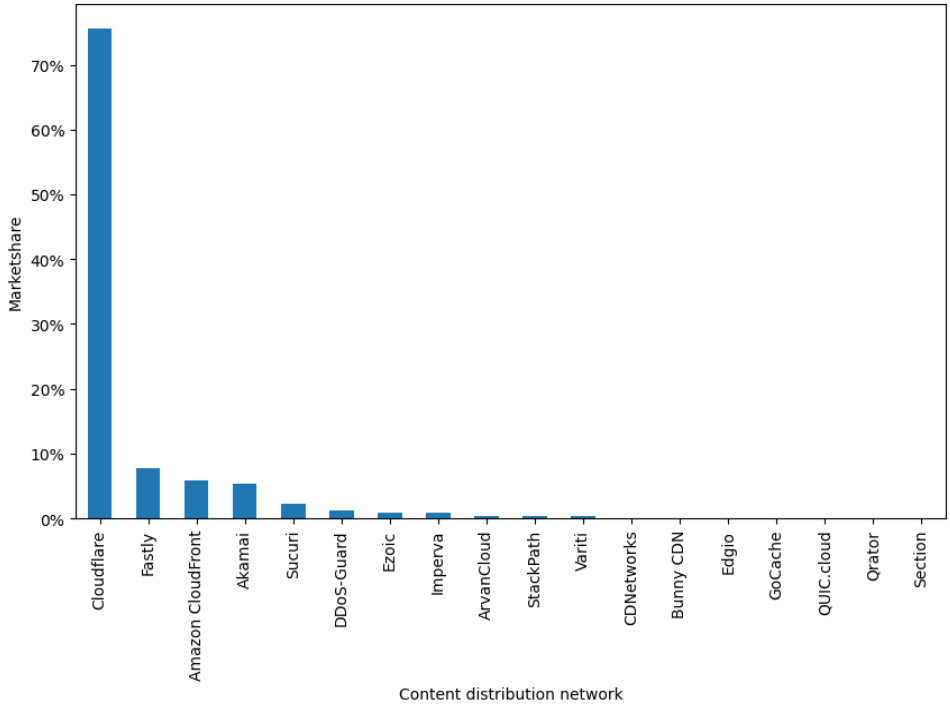

Figure 3*. The marketshare of CDNs across all websites in the sample that used a CDN to deliver service.*

Moreover, like some other internet infrastructure companies and data intensive services, there is a feedback effect to market consolidation, as CDNs benefit from both network effects and economies of scale. Consider Netflix,

HHI indicates many firms with small market shares. As a rule of thumb, the antitrust authorities have considered values above 2500 to indicate a low degree of competition.

[53] Scott Shenker, et al., *Creating an Extensible Internet Through Interposition* (manuscript on file with authors).

which, as noted, runs its own proprietary CDN. By observing geographic patterns in content consumption, Netflix can cache content files geographically close to the people most likely to watch it—it might, for example, cache episodes of *Sacred Games*, featuring some famous Bollywood actors, in India and in cities with relatively high populations of Indian-Americans—thereby yielding faster load times, and better experiences, for users. Some CDNs extend this process across a wide range of content, and, as a CDN grows, it gains an increasingly comprehensive view of global internet traffic—which, in turn, helps it to more efficiently cache content, and to better (i.e., more rapidly and accurately) respond to emerging attacks.[54] And while we have used relatively simple examples in our exposition here—NBA teams as suggestive of geography, or correlations between demographics and particular content—the truth is that much of the logic behind caching is automated, driven by machine learning algorithms that become more powerful as a CDN's scale and scope expand.[55] Hence, some CDNs even provide some services for free, offering the advantages of a more centralized architecture to smaller or newer companies, while continuing to grow their view of the internet traffic (thus fueling superior service to its paying and nonpaying customers alike).

Towards Private Networks. The results described above are suggestive of one further trend: Not only do these results highlight the rise of an increasingly concentrated CDN market, one that intermediates the relationship between the internet's traditional core and its users; they also highlight how the internet's tiers are "flattening."[56]

The flattening described here reflects a shift in the internet's topology. Recall that, under the first model of the internet we described, so-called Tier-1 and Tier-2 providers existed at the apex and middle (respectively) of an imagined hierarchy. Internet traffic was originally conceived of at starting at the bottom of this hierarchy with a local ISP, ascending a stack of tiered providers up to Tier 1 providers, and then back down again.

The internet measurement literature, combined with our original research above, casts further doubt on this model of the internet's core, and not only because individual business relationships with these bandwidth providers is increasingly intermediated by CDNs. Rather, recent findings show that internet users can access over 76% of the internet content they request without traversing

---

[54] Cf. Omer Yoachimik, Julien Desgats, Alex Forster, *Cloudflare Mitigates Record-Breaking 71 Million Request-Per-Second DDOS Attack*, THE CLOUDFLARE BLOG, at https://blog.cloudflare.com/cloudflare-mitigates-record-breaking-71-million-request-per-second-ddos-attack/.

[55] See Tejas N. Narechania, *Machine Learning as Natural Monopoly*, 107 IOWA L. REV. 1543, 1584–85 (2021) (explaining the "virtuous cycle" of machine learning-based applications, which become more accurate and effective the more data they collect and analyze).

[56] Todd Arnold, *Unpacking a flattened internet*, ASIA PACIFIC NETWORK INFORMATION CENTRE (APNIC) BLOG (Dec. 4, 2020), https://blog.apnic.net/2020/12/04/unpacking-a-flattened-internet/.

a Tier-1 or Tier-2 network *at all*.[57] Where internet requests used to hop from an ISP to a series of backbone providers and then to another ISP (and back again), most requests for internet content are now fulfilled simply by moving from an ISP to a CDN's proprietary network (and back). In this new, flatter internet, CDNs have brokered their own connections with users, using their own proprietary networks, thereby bypassing the Tier-1 providers that have traditionally made up the internet's core. Hence, though legal scholars and policymakers have long thought of the internet's interior as competitive—internet traffic exchange provisioned by diverse classes of infrastructure providers who compete with each other on terms of price, speed, and reliability—the story seems instead to focus increasingly on CDNs, including CDNs that use their own proprietary networks.

We readily acknowledge, as we must, that there is a long tail of websites and internet endpoints for which the old model of transit still rings true—about three-quarters of websites, as noted, use no CDN provider at all—and so we do not mean to suggest that these transit services are obsolete altogether. But for the majority of internet *traffic*—to emphasize, 76% of all traffic that end-users consume—these providers play practically no role.[58] And this is because the most popular websites, responsible for the vast majority of the internet's economic and social value, rely upon CDN services. For these applications and websites, internet traffic moves directly from a consumer's ISP to the CDN's private network and back again. We can hardly overstate the significance of this change for the internet's structure: A system once characterized by a robust array of competing network services providers operating on public networks has been replaced by a concentrated set of ISPs interconnected with a concentrated set of CDNs, relying on its own private network.

THE CONSEQUENCES OF THIS NEW CORE

The effective core of the internet has thus shifted away from public carriage over the networks of Tier-1 providers and towards the private networks of CDNs. This "new core" is critically different from the old core composed of a competitive market of Tier-1 providers. Tier-1 providers sell a fungible service—

[57] Todd Arnold, et al., *Cloud Provider Connectivity in the Flat Internet*, *IN* ACM INTERNET MEASUREMENT CONFERENCE PROCEEDINGS at 2, https://doi.org/10.1145/3419394.3423613 (2020). The methods Arnold et al used *traceroutes* to understand the connectivity between IP addresses. Using a corpus of traceroutes, Arnold et al. assembled a topology (essentially, a network diagram) to understand which IP addresses were reachable via which routes. Cross-referencing this topology against the IP addresses of known cloud providers, Arnold et al were able to quantify the heirarchy-free reachability of all cloud providers, determining that these providers can reach 76% of the internet without traversing the "heirarchy" of Layer-1 and Layer-2 networks.

[58] We say practically because we do not know—and cannot measure—how data arrives to the locations from which it is served. That is, to return to our NBA example, before a fan in California can watch a locally-served video highlight of a game played in, say, New York, the video recording must travel, at least once, from New York (where it was originally recorded) to California (where it is stored, i.e., cached).

the ability to deliver packets from one address to any other. The services CDNs provide are less fungible: they are the aggregate of a CDN's capacity to deliver traffic *and* its ability to (algorithmically) securely manage that traffic.[59] Indeed, as noted, CDNs rely on proprietary models, often powered by machine learning, to both cache and filter traffic. The result is an internet's "core" whose behavior is less predictable and scrutable to outside observers, including both users and regulators. When one requests a given website, what data will be delivered? From where? Are customers in low-income areas served differently from those in high-income areas? Are customers in low-income areas more likely to have their traffic blocked or throttled as "suspicious"? In the old model of the internet, the internet's "core" would have little say in such decisions, as competition among Tier-1 providers forced them to prioritize the efficient delivery of packets. Such answers would be found, instead, with the applications and websites themselves (each of which contained its own filtering, prioritization, and security logic). But today's privatized core now takes greater control over such matters, offering a service that is more like private carriage than anything we've seen inside the internet so far.[60]

We reiterate that this new model has helped to deliver better and more secure services to a wide range of the world, as it has enabled new applications such as streaming audio and video, even under some capacity-constrained conditions.

But it is also true that this new model is not costless. Rather, the consolidation of "core" internet services among private companies (using private networks) has produced two main externalities.

First, CDNs' consolidated and private infrastructure give rise to *central points of failure* that resist scrutiny and oversight. The private networks that are internal to CDN providers are opaque: They are black boxes to outsiders, confounding efforts at oversight and risk management.[61] When the internet was characterized by providers offering public carriage, we could more easily map, visualize, and assess the internet's infrastructure. But we now know far less now about how the internet's various interconnected networks fit together than before, largely because we have a very limited understanding of how CDNs route

---

[59] Indeed, just as we have described the CDNs as intermediating the traditional relationship with the "market for internet traffic exchange," scholars from the communications and computer networking community have described CDNs as offering the "interposition of in-network processing," i.e., the algorithmic management of traffic for caching and cybersecurity (among other) purposes. See Lloyd Brown et al., Creating an Extensible Internet Through Interposition.

[60] The way these models work (and sometimes fail) to filter traffic has raised questions among internet researchers about civil rights and equal access. See generally, Anne Jonas & Jenna Burrell, *Friction, snake oil, and weird countries: Cybersecurity systems could deepen global inequality through regional blocking*, 6 BIG DATA & SOCIETY 1 (2019), https://doi.org/10.1177/2053951719835238..

[61] See Petros Gigis, et al., *Seven Years in the Life of Hypergiants' Off-Nets*, *in* ACM SIGCOMM 2021 Conference Proceedings, https://doi.org/10.1145/3452296.3472928 (2021).

traffic internally (i.e., on their private networks for private carriage) and over to one another. Without some window into the CDNs' private networks, our public security and reliability efforts are frustrated.

Second, the increasing consolidation in this market gives rise to *centralized points of control.* These CDNs can (and sometimes do) filter the content available to internet users, sometimes in ways that are invisible to these users and often in ways that consumers may not avoid.[62] These decisions, moreover, are not captured by existing regulatory frameworks, thus giving rise to possibilities for abuse by CDNs and creating targets of opportunity for bad actors.[63]

Central Points of Failure

We begin with the risk, noted above, that consolidation yields central points of failure. Our opening example, in which an error at Fastly led to a cascade of problems for properties across the internet, helps to highlight this risk. So what, exactly, caused these widespread failures—even among websites that had no (apparent) commercial relationship with Fastly? Here, a closer look at the software supply chains that help to form the internet's content is instructive. Individual webpages (numbering in the *billions*) are built in real-time atop *thousands* of different software supply chains—assembling HTML and Javascript code alongside specialized fonts and images, among other resources, many of which rely upon some combination of *hundreds* of common web development tools. And many of these tools and resources are housed and delivered by one of the *eleven* (or so) CDN providers described above. Hence, disruptions at the base of this inverted pyramid—i.e., at one of the CDNs—can send ripple effects through the entire supply chain, yielding large-scale and difficult-to-predict patterns of failure.

---

[62] See Pengxiong Zhu, et al., *Characterizing Transnational Internet Performance and the Great Bottleneck of China*, *in* PROCEEDINGS OF THE ACM ON MEASUREMENT AND ANALYSIS OF COMPUTING SYSTEMS (Mar. 2020) https://dl.acm.org/doi/pdf/10.1145/3379479.

[63] Cf. Karen Kopel, *Operation Seizing Our Sites: How the Federal Government Is Taking Domain Names Without Prior Notice*, 28 Berkeley Tech. L. J. 859 (2013) (describing how Immigrations and Customs Enforcement (ICE) can use court orders to make online content inaccessible, if only in a relatively blunt way).

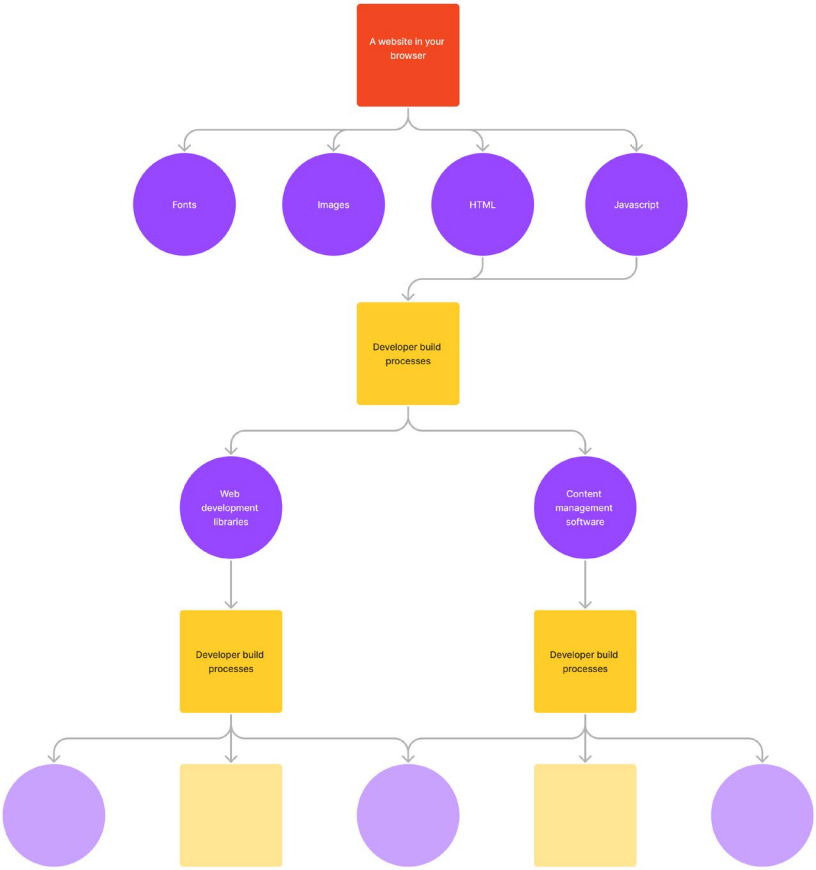

Figure 4. *A schematic diagram of the software supply chain. Websites that appear in your browser rely on a variety of static assets (like images and fonts) as well as software assets (like build tools). Those software assets themselves rely on development processes, which themselves rely on software assets, creating a recursive supply chain of "nested" dependencies that can be dozens if not hundreds of layers deep. Assets in purple represent reliances on content that is likely stored with one (or more) CDN provider(s). A failure to deliver any of these assets could cause all downstream products to become unavailable, behave unpredictably, or become impossible to update.*

Moreover, these failures can extend far beyond the scope of any clear, existing commercial relationships. Recall that Fastly has captured only about 5% of the CDN market; and yet an error there gave rise to effects felt far beyond such a footprint. How? Again, we can look to the internet's software supply chains for

an answer: Some websites will not load—they will break altogether—if even one component in the supply chain does not fall into place. For example, even if the Financial Times does not directly employ Fastly's services for its consumer-facing news sites, something in its software supply chain might. It might, for instance, require a font hosted on a CDN, and if that CDN suffers from some error, then users may be unable to load *any* Financial Times content. For want of a font, the whole site is lost.[64]

Consolidation among CDNs seems to also be entangled with consolidation in other aspects of the internet's core, such as the Domain Name System (or DNS). The DNS for example, maps human-readable "domain names" (e.g., nytimes.com) to machine-readable IP addresses. Although the DNS is itself a decentralized protocol, recent research shows that these DNS services are also increasingly centralized. Among the most centralized are Akamai, Amazon Web Services, and Cloudflare—three of the same organizations that dominate the market for CDN services.[65] How is it that this ostensibly different, and ostensibly decentralized component of the internet's core is increasingly centralized among the same providers? It is for good reason, as CDN providers include DDoS protection in their DNS service, offering cost savings and convenience, alongside cybersecurity protections at various layers of the technical stack. But this increasing consolidation reproduces the concerns for fragility (among others) described throughout this Part.

As noted, consolidation can cause problems at any one CDN to echo across the web's software supply chain in unpredictable ways. These problems are compounded by the private nature of the CDNs' networks. We, the public, know very little about how individual CDNs are physically connected to each other. Stated otherwise, we not only lack information about exactly what particular CDNs host (and for whom), we also lack information regarding the internal network connection within and across distinct CDNs. Indeed, even Amazon was affected by Fastly's outage, even though, as noted, Amazon has its own CDN. Hence, even large providers that maintain their own infrastructure have upstream dependencies that may rely on other providers. This interdependence illustrates the complexity of this network of reliance, and the fact that providers of all scales are systemically vulnerable to one another.

Since the internals of CDN networks are private (as are the mechanisms by which traffic is routed within them, as noted above), it is difficult to map the

---

[64] Such problems, moreover, can happen at any point in the supply chain. Or the error might happen upstream of the user: the developers who build the Financial Times website may rely on some tool, or set of tools, themselves stored on CDNs. If that goes down, no one will be able to produce the Financial Times–or any of the other websites that rely on that tool.

[65] See Samantha Bates, John Bowers, Shane Greenstein, Jordi Weinstock, Yunhan Xu, Jonathan Zittrain, *Evidence of Decreasing Internet Entropy: The Lack of Redundancy in DNS Resolution by Major Websites and Services*, 1 J. QUANTITATIVE DESCRIPTION: DIGITAL MEDIA 1, 28 (2021).

logical routes in the software supply chain. That makes it difficult for risk managers within organizations, such as cybersecurity professionals, to manage risk within their firms. It also makes it difficult for public sector risk managers, such as Cybersecurity and Infrastructure Security Agency (CISA) officials, to manage sector- or economy-wide risk. In short, this opacity imposes tremendous structural challenges on any attempt to map these physical and logical routes—and to plan for troubles on those routes.[66]

Consider, for example, Hurricane Sandy. In 2012, Sandy's landfall caused widespread damage to physical infrastructure, giving rise to outages among Tier-2 and Tier-3 providers in the New York metro region, and thereby making certain internet addresses unroutable (i.e., unavailable online) for extended periods of time.[67] Put simply, Sandy knocked New York off the internet. More unexpectedly, Sandy's landfall had cascading effects in far-flung places such as Brazil and Russia, which researchers measured by examining changes in the traffic patterns between other of the internet's constituent networks.[68] If such localized outages give rise to such far-flung and unpredictable effects, imagine what might happen if even larger providers are affected. And that is part of the point: We must only imagine what can happen, since our public visibility into these networks is so limited. Fastly's outage—accidental and short—offers a preview of the possible effects; but a more sophisticated cyberattack could inflict more widespread and longer-lasting troubles.

In all, the opacity of CDNs' private networks prevents the public (and its representatives) from establishing risk profiles and mitigation strategies regarding the internet's infrastructure. If Fastly had a partial outage, what would be inaccessible? If Cloudflare had a complete outage, how much damage would it do? Imagine an error at—or, worse, attack on—Cloudflare instead. Given the expansive nature of Cloudflare's scope, if its systems go offline (for whatever reason), the scale of the outage would be tremendous. But what *exactly* would be affected? Our limited view into the complex interdependencies among web properties means that we can barely anticipate, let alone prepare for, such a scenario. While these risks begin with technical, engineering, and security failures, the fact that we cannot anticipate and prepare for them is a policy failure. Our limited insight into these providers' inner workings (their customers, what they do for those customers, and how their physical datacenters connect to one another and to other firms') confounds our ability to establish specific and actionable disaster plans.

---

[66] See Petros Gigis, et al., *Seven Years in the Life of Hypergiants' Off-Nets*, *in* ACM SIGCOMM 2021 Conference Proceedings, https://doi.org/10.1145/3452296.3472928 (2021).

[67] Marguerite Reardon, *Hurricane Sandy disrupts wireless and Internet services*, CNET (Oct. 30, 2012), https://www.cnet.com/tech/mobile/hurricane-sandy-disrupts-wireless-and-internet-services/.

[68] Geoff Huston, *Superstorm Sandy and the Global Internet*, RIPE LABS (Dec. 3, 2012), https://labs.ripe.net/author/gih/superstorm-sandy-and-the-global-internet/.

Central Points of Control

The growing intermediation of the internet's core by CDNs has not only complicated our public efforts at, say, disaster planning, it has also, thanks to the increasing consolidation of the CDN market, the global internet is now too more subject to a limited number of *central points of control*.

Recall that one advantage of the shift towards CDNs is a benefit to cybersecurity: In addition to providing bandwidth and computational capacity at scale, CDNs also offer collective security to web properties by observing and responding to global traffic trends. But that collective security has come with a cost to individual access: CDNs can—and sometimes do—act as a gatekeeper to internet content. CDNs can filter incoming requests to any of their customers—CNN, NYTimes, gov.uk—in a fine-grained way, e.g., filtering content to particular users, originating from particular countries, and so on.

Indeed, security measures implemented by CDNs sometimes prevent populations in the Global South from accessing websites: Some automated systems treat traffic from countries like Ghana (or, more precisely, traffic that is *estimated* to originate from countries like Ghana) as presumptively suspicious.[69] The job of CDNs is to understand the risk-reward tradeoff for any traffic they observe. If it is the case *both* that traffic from Ghana is more likely to be malicious than traffic from the U.S., *and* that traffic from Ghana is less likely to be lucrative than traffic from the U.S., then the CDN will be more (perhaps much more) likely to block Ghanaian than U.S. traffic, and some notion of security (from the economic perspective of the firm that pays the CDN for service) has been achieved. The effects on Ghanaians are much more severe. An internet once imagined by policymakers as a life-raft of economic opportunity for developing nations instead acts as a moat that excludes them from an ostensibly global market for goods and services.

These security measures also have the effect of limiting the spread of new, potentially useful technologies. Tor, for example, is a privacy-protecting browser.[70] But because it is often used for (and has thus become associated with) illicit purposes, some CDNs have gotten in the habit of regularly challenging Tor-based requests for content, rendering the browser practically unusable.[71] To be

---

[69] Anne Jonas & Jenna Burrell, *Friction, snake oil, and weird countries: Cybersecurity systems could deepen global inequality through regional blocking*, 6 BIG DATA & SOCIETY 1 (2019), https://doi.org/10.1177/2053951719835238.

[70] *History*, TOR, https://www.torproject.org/about/history/ (last visited Feb. 5, 2023).

[71] In particular, Cloudflare detects traffic coming from a known Tor endpoint. Associating Tor traffic with fraud and denial of service attacks, it guards the page with a CAPTCHA, attempting to slow down potential attacks. Matthew Prince, *The Trouble with Tor*, CLOUDFLARE BLOG (Mar. 30, 2016), https://blog.cloudflare.com/the-trouble-with-tor/. When used correctly, Tor provides meaningful privacy guarantees. It helps people worldwide evade state censorship. But Tor is much less widely used than it could be, in part because Cloudflare makes it such a pain to browse the Internet with.

sure, there are equities on both sides: On one hand, users may use Tor to access medical websites in order to privately obtain information about a sensitive diagnosis, or to privately seek out abortion-related care in locales where doing so may lead to substantial liability[72]; on the other, Tor's frequent association with bad actors might justify a security concern.

Indeed, this conflict recalls one of the earliest network neutrality controversies, in which Comcast blocked access to BitTorrent, a service which can be used for legitimate purposes (including, in one notable example, Bible study), but was also frequently associated with copyright infringement. And so the question here—as it was there—regards the allocation of decisionmaking authority. Who should decide which applications and content are allowed: the internet's users, its infrastructural providers, or some other entity altogether? And how do we exercise discipline over the decisions made by infrastructure providers that face little meaningful competition? Indeed, even Cloudflare's chief executive, Matthew Prince, raised such questions after banning 8Chan (a online community strongly associated with neo-Nazi groups and others advocating violent white supremacy), noting that while he thought it appropriate for Cloudflare to terminate service for platforms that "directly inspire tragic events and are lawless by design," doing so thrusts the company into the "incredibly uncomfortable . . . role of content arbiter," and suggesting that public governance structures are better suited to resolving these disputes.[73]

But so far, no such governance structures have emerged—even as such conflicts persist. In late 2022, for example, Cloudflare was once again under pressure to stop providing CDN services to Kiwifarms, a close cousin of 8Chan that had "become notorious for waging online harassment campaigns against [LGBTQIA+] people, women, and others."[74] At first, Cloudflare explained that it would not discontinue service for Kiwifarms, elaborating its "view that cyberattacks not only should not be used for silencing vulnerable groups, but are not the appropriate mechanism for addressing problematic content online," and so it would continue to provide Kiwifarms with defenses from cyberattacks and other CDN services.[75] But only a few days later, Cloudflare reversed course in what it called an "extraordinary decision" made in view of "an unprecedented emergency" arising out of increasingly threatening content on Kiwifarms' website.[76] No matter whether one thinks Cloudflare got it right at first, or after its reconsideration, the essential point is that only Cloudflare controlled Kiwifarms'

---

[72] See, e.g., Texas Heartbeat Act, Tex. S.B.8, 87th Leg., Tex. Health & Safety Code, ch. 171, sec. 204 (2021).

[73] Matthew Prince, *Terminating Service for 8Chan*, CLOUDFLARE BLOG (Aug. 4, 2019), https://blog.cloudflare.com/terminating-service-for-8chan/.

[74] Casey Newton, *How Cloudflare Got Kiwi Farms Wrong*, THE VERGE (Sept. 6, 2022).

[75] Matthew Prince & Alissa Starzak, *Cloudflare's abuse policies & approach*, CLOUDFLARE BLOG (Aug. 4, 2022).

[76] Matthew Prince, *Blocking Kiwifarms*, CLOUDFLARE BLOG (Sept. 3, 2022).

online destiny (notwithstanding its own view that Cloudflare lacks "the political legitimacy to determine generally what is and is not online by restricting security or core Internet services"). So too for Tor: When Cloudflare blocks (or effectively blocks) Tor traffic, that's the end of Tor. So much of the web relies on Cloudflare, and so Tor users lack meaningful access to the internet's most popular destinations. That gives rise to self-fulfilling prophecy, as Cloudflare's decision gives Tor little practical use outside of the so-called "dark web"—and no one but Cloudflare had any meaningful input over the development of this competing browser technology.

*

In short, CDNs are a hidden vector of consolidated power. Likely few internet users even know that CDNs exist, let alone that a tremendous proportion of their traffic routes through them.[77] But their services have positioned them as key players inside the internet, where they function as sites of control and targets of opportunity. Just as network neutrality (among other policy efforts) is a response to a competition problem, the growing consolidation among CDNs suggests that similar responses may be appropriate. Our network neutrality debates have focused on the edges of the internet, where some ISPs enjoy monopoly status and so may block, throttle, or prioritize traffic with impunity; but we should look inside the internet, too, and consider who should decide what entities can access and traverse the inside of the internet, and how such decisions ought to be made.

WHAT CAN WE DO?

Challenges in scaling the internet's original decentralized design led to the emergence of a more centralized structure characterized by CDNs, whose power over the internet has since increased immensely. Indeed, the CDNs' control over this infrastructure not only includes the power to distribute content or help prevent cyberattacks, it also encompasses decisionmaking power over various aspects of the internet—who can use it, and on what terms. Specifically, the CDNs' increasingly private infrastructure confounds our ability to reason publicly and strategically about the internet's physical structure. We, as a public, know comparatively little about the interrelated dependencies inside the internet, and so face significant troubles in planning for and addressing outages and errors

[77] If the website you're communicating with uses the Transport Layer Security (TLS) protocol, configured correctly, the CDN won't be able to see the content of your requests. But it will know that you're making requests, and that metadata itself can be revealing enough. As General Michael Hayden, Obama's director of the CIA and NSA, once said, "We kill people based on metadata." Ryan Goodman, *Video Clip of Former Director of NSA and CIA: "We Kill People Based on Metadata,"* JUST SECURITY (May 12, 2014), https://www.justsecurity.org/10318/video-clip-director-nsa-cia-we-kill-people-based-metadata/.

causes by attacks and natural disasters. Similarly, the CDNs' consolidated control over so much internet content grants these infrastructural providers gatekeeper power over the path between users and content providers.

We do not, to be sure, mean to suggest that the answer is a return to the old model of the internet. CDNs help to solve many important problems (e.g., latency) by providing critical caching and collective defense services. But that CDNs offered an improvement over the prior status quo need not imply that we must accept their shortcomings, or that we should ignore any new problems that these solutions to the old problems introduce. And so, in the following sections, we consider policy responses that might improve meaningful governance—via either public regulation or market discipline—over these mission-critical providers.

Transparency and Security

We begin with the risks that attend to the increasingly private nature of CDNs' internal networks. As noted, public regulators (and the public more generally) lack clarity on both hard infrastructure (i.e., cable connectivity patterns within and across providers) as well as aggregate traffic flow patterns (i.e., who sends traffic to whom, and to what degree that traffic matters to any area of concern). This opacity is the main structural barrier to appreciating systemic risks to the internet's stability and resiliency. Currently, neither national regulators nor the community of internet measurement and cybersecurity scholars can assess strategic contingencies, due, in large part, to this lack of sufficiently detailed connectivity data. Stated simply, we don't know how private networks operate internally, how they connect to one another, or how they relate to broader internet.

Government risk management agencies, like, say, the United States Cybersecurity and Infrastructure Agency (CISA), could better address risks to the internet's resiliency with such information. With data, such agencies could prepare for certain cyberattacks, and could prepare disaster plans that, for example, prioritize certain web hosts, data centers, internet exchange points, in order to restore service most quickly to the widest population; or it could find ways to prioritize service to certain essential facilities?[78] FEMA, for example, has

[78] The amount of reverse engineering the Internet measurement community performs simply to observe proxies of this connectivity is, relative to the centrality of this infrastructure to global trade, commerce, communication, and emergency response, incredible. See, e.g., Zesen Zhang, et al., *Inferring Regional Access Network Topologies: Methods and Applications*, ACM Internet Measurement Conference (2021), https://www.caida.org/catalog/papers/2021_inferring_regional_access_network_topologies/inferring_regional_access_network_topologies.pdf; Petros Gigis, et al., *Seven Years in the Life of Hypergiants' Off-Nets*, *in* ACM SIGCOMM 2021 Conference Proceedings, https://doi.org/10.1145/3452296.3472928 (2021)

detailed maps—roads, topography, and so on—that help it plan for natural disasters. CISA needs maps, too.

One simple and straightforward step is thus to mandate public disclosure of data center connectivity and traffic flows. This would require that CDNs tell us where their infrastructure is, what it connects to, and, within appropriate bounds of user privacy, a general description of the sorts of content it serves.[79] Such information is critical to making national security assessments about our infrastructure's robustness under various failure models. It is also inspired by recent successes elsewhere in cybersecurity policy. Some evidence suggests that mandating breach notifications—that is, requiring companies to tell affected parties about data security incidents—have helped to mitigate such security failures.[80] Indeed, some bills working their way through Congress require companies to disclose details about certain cyberintrusions. Our approach differs, however, by requiring that critical providers (such as CDNs) disclose this information *ex ante*, in order to improve preparedness generally, rather than simply requiring that such providers report problems *ex post* in an effort to seek out and deter threats.

Hence, policymakers should require that providers such as Cloudflare report on their physical infrastructure,[81] including how that infrastructure is interconnected (e.g., private or rented long-haul fiber), how its own systems connect to other infrastructure, and what algorithms or logic govern routing among facilities. If, say, Fastly has a datacenter in Cheyenne and another in Asheville, which is more critical to protect or restore? Or if Cloudflare uses certain algorithms to route traffic, how might that logic help public officials decide which internet exchange point to restore service to first?

Indeed, managing systemic risk in an increasingly algorithmically governed internet will be a core matter of public concern for this new private core. Arriving at suitable answers will almost certainly require input from computer scientists, civil society, and, of course, industry. But disaster planning is—and ought to be—the role of a public agency, such as CISA. Unlike industry, which makes decisions

---

[79] Our formulation here gives rise to an immediate question: What, exactly, does it mean to remain "within appropriate bounds" of user privacy. We address this in more detail infra notes ___ and accompanying text. And we summarize the main point here, which is that CDNs have access to data that may be necessary for, say, certain disaster planning scenarios—and so CDNs should share that information with public officials, but only to the extent necessary to engage in such disaster planning. And where there are trade-offs to be made between personal privacy and public preparedness, we would much prefer that those choices be made by publicly accountable officials in an open and transparent process than by a private company guided primarily by private incentives that may or may not align with broader public goals.

[80] See Deirdre K. Mulligan & Fred B. Schneider, *Doctrine for Cybersecurity*, 140 PROTECTING THE INTERNET AS A PUBLIC COMMONS 4, 74 (2011); Aniket Kesari, *Do Data Breach Notification Laws Work?*, SSRN (Aug. 30, 2022), http://dx.doi.org/10.2139/ssrn.4164674.

[81] Consider this map of Cloudflare datacenters. *The Cloudflare global network*, CLOUDFLARE, https://www.cloudflare.com/en-ca/network/ (last visited Feb. 5, 2023).

in relative secrecy and which is governed by private incentives that need not align with public goals, federal agencies must conduct their work in the open, must be responsive to public input, and may be held to account for their decisions by elected leaders and the voting public. CISA, however, cannot effectively perform this role without better information about the internet's new core. Policymakers should make sure CISA gets it.

Fair Carriage and Gatekeeper Power

As noted, the consolidated nature of the new "market for internet traffic exchange" gives CDNs the power to control users' access to content. Moreover, CDNs' exercise of this power (intermittent as it may be) is not governed by any public standards. As Matthew Prince, Cloudflare's CEO noted, it may seem easy, in any one given (egregious) case, to "deplatform" an entire service from the internet, but it is much more "hard [to] defin[e] the policy [to] enforce transparently and consistently going forward."[82] And, as noted above, it is troublesome that such power over internet access—perhaps the most important modern utility—sits entirely with an entity that can render its decisions in the dark and that is guided by private, rather than public, incentives. Even if the decisions rendered by, say, Cloudflare strike us (so far) as good, correct, or public-minded, they remain so only for now—and we may rightfully wonder how long such private power will be vested in trusted actors and used for purposes with which we may broadly agree.[83]

As noted above, CDNs use algorithms, often informed by machine learning, to route and filter internet traffic. But these models may systematically disadvantage certain segments of the population for no reason other than their national origin, or that they value privacy more than other internet users. Algorithmic redlining is, of course, not new,[84] and many prior examples of problematic systems have been uncovered through audits and reviews. But, in this context, access to (or transparency over) these models has not yet been forthcoming—one further reason to consider the disclosure reforms we describe above.

Even Cloudflare, for example, has admitted some discomfort with its de facto power over internet speech, explaining that questions about content standards "are real societal issues that need politically legitimate solutions" (all

---

[82] Matthew Prince, *Terminating Service for 8Chan*, CLOUDFLARE BLOG (Aug. 4, 2019), https://blog.cloudflare.com/terminating-service-for-8chan/.

[83] Indeed, recent transactions and news reports suggest that some powerful individuals and entities have sought control of other private speech channels to amplify certain voices. See Zoe Schiffer & Casey Newton, *Elon Musk's reach on Twitter is dropping — he just fired a top engineer over it*, THE VERGE (Feb. 9, 2023).

[84] See concern about Amazon Prime's delivery radius: David Talbot, *Amazon Prime or Amazon Redline?*, MIT TECH. REVIEW (Apr. 25, 2016), https://www.technologyreview.com/2016/04/25/71105/amazon-prime-or-amazon-redline/.

while (understandably) relying on comparatively opaque algorithmic systems to secure and protect network systems).[85] And so we consider here some options that seek to wrest control over content from CDNs and instead subject it to forms of popular governance, including democratic control and market discipline.

We begin with the possibility of a fair carriage rule that prohibits CDNs from discriminating among internet users along dimensions such as national origin, or other protected characteristics, and generally requires that CDNs ensure access to lawful content by means of lawful applications, all while acknowledging that CDNs perform a beneficial security function and must be given some leeway to protect and secure the internet's constitutive networks.[86] In general terms, such a fair carriage rule prioritizes access to internet speech and content over a CDN's efforts to curate that speech.

We acknowledge that this formulation is somewhat open-ended, and leaves much to certain details. But, for our present purposes, that is sufficient. Our main point is not to fully elaborate the details of this fair carriage regime, but rather, to note that the carriage of internet traffic is presently subject to the whims of private industry, insulated from public oversight and democratic governance. We think, instead, that a public agency—the Federal Communications Commission, perhaps—should help to ensure that the internet's carriage practices and rules reflect public values rather than private incentives. Such a model helps ensure "a certain degree of democratic or quasi-democratic control over infrastructure that undergirds the modern world."[87]

Any effort to regulate the control CDNs exercise over internet content will echo in the debates over both network neutrality and content moderation. Advocates for network neutrality highlight the consolidation in local markets for internet access, contending that ISPs (such as Comcast) should not have the power to decide which streaming services, say, a user can access (all of them, or perhaps only the Comcast-owned Peacock).[88] Meanwhile, critics of legislated

---

[85] Matthew Prince, *Terminating Service for 8Chan*, CLOUDFLARE BLOG (Aug. 4, 2019), https://blog.cloudflare.com/terminating-service-for-8chan/.

[86] See James B. Speta, *Can Common Carrier Principles Control Internet Platform Dominance?*, 2022 Robert F. Boden Lecture, Marquette University School of Law, Northwestern Pub. L. Research Paper No. 22-29, at 4, https://ssrn.com/abstract=4228208 (2022). We note that James Speta's proposal for regulating infrastructural providers, such as Cloudflare, is aimed primarily at resolving questions of competition among user-facing platforms, such as Facebook and Twitter, on the theory that "applying [common carriage] rules to [these infrastructural] support layers could increase the diversity of platforms." By contrast, our proposal for regulating these providers is aimed at resolving questions of competition *among these providers themselves*.

[87] Daniel T. Deacon, *Institutional Considerations for the Regulation of Internet Service Providers*, 74 FED. COMM. L.J. 111 (2022); cf. USTA v. FCC, 855 F.3d 381 (D.C. Cir. 2017) (suggesting that the government may, upon showing that an intermediary exercises market power, regulate that intermediary's editorial discretion consistent with the First Amendment).

[88] See Timothy B. Lee, *Network neutrality, explained*, VOX.COM (May 21, 2015), https://www.vox.com/2015/2/26/18073512/network-neutrality.

standards for content moderation contend that the First Amendment guarantees platform providers the discretion to block offensive content, and that policymakers do not fully grasp the impossibility of the problem of moderating at scale.[89] In our view, the problems presented by CDNs are more closely related to those implicated in the network neutrality debate. Cloudflare's massive capacity, to be sure, presents some difficult problems of moderating at scale—and we do not mean to say that the CDNs have no First Amendment interest in the content that flows over its network.[90] But any of the CDNs' First Amendment concerns must be weighed against the speech interests of users—Ghanaian residents, for example; or those who wish to access sensitive content (information, say, on abortion access in certain states) discreetly.[91] Moreover, because only a few CDNs control these critical paths to the internet's most popular content—and do so in a way that consumers cannot readily avoid—new carriage rules for CDNs can both help to guarantee content access for the internet's users and ensure that CDNs do not leverage their gatekeeper power into adjacent markets.[92]

That is not to say that we should not do more to improve competition in the CDN market. We should. But the barriers to entry are high—CDNs require massive investments in data centers worldwide as well as in sophisticated network engineering and cybersecurity tools, and, as noted above, existing providers benefit from scale economies and network effects, leaving new entrants far behind. Even though entry into the CDN market is difficult, we may take steps to help improve competition among the market's existing players. Regulators might, for example, address switching costs among CDNs, such as (but certainly not limited to) egress fees, so that, for example, a provider of reproductive health content can more easily switch from a CDN that blocks private Tor connections to one that allows them.[93] We might even consider

---

[89] See, e.g., Brief of Respondents at 13, *Moody v. NetChoice, LLC*, No. 22-277 (Oct. 24, 2022).

[90] However, we equally do not concede that they do. Rather, as noted infra text accompanying note __, we simply assume that any such editorial interests must be considered against the speech interests of users. See, e.g, *Turner*, 512 U.S. 622 (1994) (Breyer, J., concurring) (noting the speech interests of both the intermediaries exercising editorial control and putative speakers and listeners)

[91] Indeed, the CDNs' apparently vast market power likely diminishes the strength of any of First Amendment challenge to new fair carriage rules. See, e.g., *Turner Broad. Sys., Inc. v. Fed. Commc'ns Comm'n* ("*Turner II*"), 520 U.S. 180, 189–90 (1997) (content-neutral regulation survives First Amendment challenge if it advances important governmental interests, such as eliminating restraints on fair competition, and doesn't burden substantially more speech than necessary); *Red Lion Broad. Co., Inc. v. Fed. Commc'ns Comm'n*, 395 U.S. 367, 390 (1969) (First Amendment's purpose in preserving free speech does not "countenance monopolization" of a market).

[92] See Tejas N. Narechania, *Network Nepotism and the Market for Content Delivery*, 67 Stan. L. Rev. Online 27 34–35 (2014).

[93] You may also be wondering: Can't providers just change CDN providers?" In practice, answers to this question depend on whom the CDNs' decisions affect. Providers like Cloudflare regularly block traffic originating in countries like Ghana, treating it as intrinsically suspicious. Anne Jonas & Jenna Burrell, *Friction, snake oil, and weird countries: Cybersecurity systems could deepen global inequality through regional blocking*, 6 BIG DATA & SOCIETY 1 (2019), https://doi.org/10.1177/2053951719835238.How many companies have stopped using Cloudflare in response? Very few, likely because few large tech companies

developing a public CDN option—a publicly-run service, definitionally subject to the First Amendment's prohibitions against speech discrimination—that can both discipline other CDNs' terms and rates through competition and give content providers another option.[94] Or policymakers might support the development of technical standards that enable a more competitive CDN market. Specifically, regulators might encourage extensible internet architectures that provision caching and security at the network layer, with the effect of structuring a more competitive, and responsive, market. But, in the meantime, regulators should not merely wait for competition to come to this market and should instead take action to ensure that internet carriage practices reflect public values.

CONCLUSION

Although many regard the inside of the internet as robustly competitive—a view that has shaped our regulatory approach to the "market for internet traffic exchange"—that view is flawed. Perhaps unsurprisingly, the internet's infrastructure has adapted to our more modern uses of the internet, and a concentrated set of CDNs now intermediate the relationship between the internet's users and its traditional "core." In many respects, this is good. CDNs offer advances in speed, reliability, and security.

But there are tradeoffs. While CDNs offer these advances, they come at the expense of transparency and gatekeeper control. We may want our internet infrastructure to deliver on several promises: access to (lawful) internet content that is ungated by intermediaries; privacy; and protection from cyberattacks. CDNs can implement these in different ways. CDNs might, for example, use automated processes that inspect internet content before deciding whether to carry it, thereby guaranteeing security and some modicum of privacy, but at the expense of the network neutrality norms that have long governed the internet's core. Or CDNs might require that internet users authenticate themselves before agreeing to carry traffic on equal terms, thereby ensuring network neutrality (at least for authenticated users) and security, but at the cost of privacy.[95]

---

have customers or engineers who notice. Likewise, how many companies have stopped using Cloudflare because they block Tor—itself a serious issue for people trying to circumvent Internet censorship globally? Again, given Cloudflare's persistent dominance, the answer seems to be "very few." Besides, switching reverse proxy providers isn't as easy as you might think. If you've ever used a 'standard' web library like Bootstrap or JQuery on your webpage, you probably used a version hosted on a reverse proxy—probably Cloudflare. Even if you stop using Cloudflare, the libraries you depend on might still. Switching all that stuff over can be a pain at best, and, at worst, could temporarily break your website. The incentives to stick with one's existing provider are high.

[94] Cf. Yotam Harchol, Dirk Bergemann, Nick Feamster, Eric Friedman, Arvind Krishnamurthy, Aurojit Panda, Sylvia Ratnasamy, Michael Schapira & Scott Shenker, *A Public Option for the Core*, 2020 PROC. ANN. CONF. ACM SPECIAL INT. GRP. ON DATA COMMC'N ON APPLICATIONS, TECHS., ARCHITECTURES, & PROTOCOLS FOR COMPUT. COMMC'N 377, 388

[95] See supra note __ (discussing the resolution of these trade-offs).

Resolving such trade-offs is likely to be a core internet governance question in coming years. Our specific response to these concerns is not to take us back to the old model of the internet, one less secure and less adapted to sorts of applications we are now accustomed to using. Instead, we imagine some policy reforms—new disclosure requirements and fair carriage rules—that directly address these emerging concerns.

Our primary focus for now, however, is on *who* decides on the rules that govern our internet infrastructure rather than on *what* those rules are in their details. At present, our core internet infrastructure is governed by private industry and guided by private incentives. We would much prefer that such rules and decisions come, in the spirit of the internet, by way of our systems of democratic governance and public participation.

APPENDIX

Provider	Market Share
Cloudflare	75.6
Fastly	7.7
Amazon CloudFront	5.9
Akamai	5.3
Sucuri	2.3
DDoS-Guard	1.3
Ezoic	0.9
Imperva	0.9
ArvanCloud	0.4
StackPath	0.4
Variti	0.3
CDNetworks	0.1
Bunny CDN	0.1
Edgio	0.1
GoCache	0.1
QUIC.cloud	0.1
Qrator	0.1
Section	0.1

Appendix Table 1*. Whole-web market shares of CDN providers as of February 17, 2023.*