Content may be King, but (Peering) Location matters: A Progress Report on the Evolution of Content Delivery in the Internet

Volker Stocker[1]
Georgios Smaragdakis[2]
William Lehr[3]
Steven Bauer[4]

***** September 29, 2016 *****

Paper presented at the 27th European Regional ITS Conference
Cambridge, UK, 7-9 September 2016

Abstract

Since the commercialization of the Internet, content and related applications, including video streaming, news, advertisements, and social interaction have moved online. It is broadly recognized that the rise of all of these different types of content (static and dynamic, and increasingly multimedia) has been one of the main forces behind the phenomenal growth of the Internet, and its emergence as essential infrastructure for how individuals across the globe gain access to the content sources they want. To accelerate the delivery of diverse content in the Internet and to provide commercial-grade performance for video delivery and the Web, content delivery networks (CDNs) were introduced. This paper describes the current CDN ecosystem and the forces that have driven its evolution. We outline the different CDN architectures and consider their relative strengths and weaknesses. Our analysis highlights the role of location, the growing complexity of the CDN ecosystem, and its relationship to and the implications for interconnection markets.

Keywords: CDN, Interconnection, Peering, QoS, QoE, Pricing, Internet Evolution

---

[1] University of Freiburg, Germany, *volker.stocker@vwl.uni-freiburg.de*
[2] MIT/TU Berlin, *gsmaragd@csail.mit.edu*; Georgios Smaragdakis was supported by the EU Marie Curie IOF "CDN-H" (PEOPLE-628441)
[3] MIT, *wlehr@mit.edu*. Dr. Lehr would like to acknowledge support from NSF Awards 1413973, 1547265 and the MIT Communications Futures Program (http://cfp.mit.edu).
[4] MIT, *bauer@mit.edu*

Table of Contents

1. Introduction

As Bill Gates famously remarked in 1996, content is king.[1] Since its emergence in the 1960s, the Internet has grown and evolved to be the dominant global platform for all electronic communications, and is widely recognized as essential infrastructure for society and the economy. The Internet ecosystem is rapidly evolving, complex and diverse. Providers of different content and applications include individuals, non-profits, and for-profit businesses. Increasingly, content of all types is shifting online – from text to audio to video; from advertising to entertainment media to educational materials; from static to dynamic; from passive to interactive; etc. Commercial providers of online content and applications have found that delays of even a few tens of milliseconds can cause significant adverse revenue effects. Only a couple of seconds of video buffering has a significant impact on users' abandonment rates. Faster and more reliable content delivery is considered a competitive advantage that can lead to higher revenues for e-commerce and streaming applications as well as enhance user engagement (cf. e.g., Kohavi et al., 2007; Dobrian et al., 2011; Krishnan and Sitaraman, 2013; and Singla et al., 2014). Accordingly, highly differentiated and heterogeneous distribution requirements of content and applications providers are in constant flux and pose fundamental challenges for content distribution, especially regarding security, delivery performance, i.e., Quality of Service (QoS), and cost efficiency.

Although the Internet's basic infrastructure has scaled remarkably well, its end-to-end, "best effort" design was premised on a communication paradigm based on passive traffic management that is not well-suited for meeting the distribution requirements of today's commercial content and application providers. While protocols that manage packet transmission on an end-to-end basis such as the Transmission Control Protocol (TCP) aim to achieve decentralized congestion management and fair resource allocations across competing flows, they display substantial deficiencies and undesirable limitations when it comes to distributing content. For example, TCP performance is known to degrade as the distance between two communicating hosts increases or if there is a packet loss along the path between source and destination (cf. Leighton, 2009). A number of applications may further take advantage of TCP's flow-rate fairness principles.[2] Additionally, as service provision is restricted to a single service class, QoS differentiations reflecting heterogeneous demand cannot be provided. Inefficient capacity allocations based on

[1] In an essay in 1996 entitled "Content is King," Bill Gates (1996) argued that "content is where I expect much of the real money will be made on the Internet, just as it was in broadcasting." He notes that "when it comes to an interactive network such as the Internet, the definition of 'content' becomes very wide" and he expects that "societies will see intense competition – and ample failure as well as success – in all categories of popular content – not just software and news, but also games, entertainment, sports programming, directories, classified advertising, and on-line communities devoted to major interests."

[2] For example, using so called swarming techniques, i.e., simultaneously operating hundreds of TCP connections, some applications are allocated disproportionate shares of traffic capacities thus causing "starvation" of other applications which do not use swarming. Discrimination potentials between heavy users and light users may arise (cf. e.g., Briscoe, 2007). While peer-to-peer file sharing applications (e.g., BitTorrent) are famous for using swarming techniques, modern web browsers typically operate several TCP connections simultaneously (e.g., in order to increase web site performance).

one-size-fits-all solutions may result.

While attempts have been made to expand the set of Internet protocols to enable better support for heterogeneous QoS requirements, it is extremely difficult to modify the basic Internet infrastructure since that requires widespread adoption and support for the new protocols. In the Internet, control is decentralized and dispersed amongst a myriad of different and often competing entities. The control of each entity is thus restricted to their corresponding networks. Although these entities may typically deploy enhanced capabilities for active traffic management within their networks, widespread inter-provider deployments are rare (cf. e.g., Stocker, 2015). Harmonizing traffic management practices to facilitate seamless interoperability across heterogeneous networks is problematic and coordinating the market-driven migration to a new Internet architecture turns out to be a daunting challenge.

In light of the difficulties inherent in upgrading the basic Internet, but cognizant of the demands of commercial content providers for better network support for delivering content online, new types of overlay value-added service providers, known as Content Delivery Networks (CDNs), have emerged to (partially) fill that gap. Corresponding overlay functionality is complementary to the Internet's basic packet transport infrastructure. As CDN innovations may be undertaken by individual service providers, CDNs are more agile and better able to swiftly adapt to changing market conditions. They can add new capabilities and services at a faster pace because CDN evolution does not require changes of the underlying basic Internet. Over time, significant evolution of CDNs has given rise to today's ecosystem which includes a complex array of CDN providers pursuing diverse business strategies and meeting diverse demand by content and application providers. CDNs play an important role in the Internet ecosystem and are reported to account for more than half of global Internet traffic. In addition to benefiting from the rapid growth in content traffic (principally, entertainment media associated with streaming video and audio, as well as downloads of video and audio files for offline consumption), the growth of CDNs has benefited from the widespread adoption of HTTP as the principal protocol for Web-based content. It is estimated that 60% of global Internet traffic is HTTP based, which includes much of the video streaming traffic (cf. e.g., Labovitz et al., 2010; Gerber and Doverspike, 2011; Cisco 2016; Sandvine, 2016; Poese et al., 2012; Richter et al., 2015; and Popa et al., 2010). The trends that have driven the growing importance of CDNs continue and are even re-enforced as more content-centric applications and digital media move online.

In the balance of this paper, we explore and describe the evolution and complexity of the CDN ecosystem. In Section 2, we discuss in greater detail how CDNs have evolved and present a taxonomy analyzing the relative strengths and weaknesses of alternative CDN business models and architectures. A key role of CDNs is to identify where the best location is to store content. CDNs serve a dual function of delivering content at lower resource costs while facilitating enhanced performance and control over how content is delivered to end-users thus ensuring a good Quality of Experience (QoE).[3] It turns out that choosing the best "location" has multiple interpretations that further help explicate the benefits of alternative CDN strategies. In Section 3,

[3] Quality of Experience is a holistic concept that describes the subjectively perceived quality of a user when consuming content or applications over the Internet.

we identify and discuss the many facets of location and derive implications for alternative strategies. In the following Section 4, we speculate about how the CDN landscape is likely to change. We conclude that the current complexity is likely to be sustained, although we expect to see growing efforts to integrate ISP and CDN functionality to take advantage of the mutual benefits to be realized from closer coordination. However, in many cases, we expect the coordination to be managed through contractual alliances rather than full vertical integration, in light of the strategic benefits to be had from retaining a degree of independence. Section 5 concludes and offers some thoughts on future directions for research.

2. On the Evolution of Content Delivery Networks

The basic transport model in the Internet relying on best effort packet delivery has worked remarkably well, allowing the Internet to scale far beyond the expectations of those who were involved in its original design. The communication paradigm embedded in the Internet's original design embraced the end-to-end principle. It focused on fostering universal connectivity and specified that most of the intelligent communications functionality (e.g., responsibility for solving congestion problems) should be located in the nodes or hosts at either end of the communication path and not be included within the network (cf. Saltzer et al., 1984). Although this may work well when end-users wish to communicate directly with each other, it is often not optimal when the goal is to access specific pieces of content rather than to transfer packets between particular end-nodes. For many types of content, multiple end-users may wish to access the same content, and users typically do not care from whence the content is delivered so long as the delivery provides a good QoE. Moreover, many types of content are relatively static, making it feasible to relocate and store copies at multiple locations, allowing users to access the content at the time and in the format or QoE of their choice.

Further, inherent to the basic end-to-end best effort paradigm is that it produces a single traffic class in which QoS levels result endogenously depending on available traffic capacities and the amount of data to be transmitted. In an increasingly heterogeneous ecosystem, this may lead to inefficient capacity allocations or impaired QoE. Thus, over time, the Internet suite of protocols have been expanded to include a range of capabilities that can be used to improve content delivery or offer other types of Quality of Service (QoS) differentiated services. These include support for multicast and protocols such as DiffServ and IntServ that support more than best-effort packet delivery.[4] As described above, although these capabilities exist, they are not uniformly available nor implemented in a standard way across ISPs.

CDNs emerged as overlay networks on the Internet that offered supplemental functionality to address the need for better options for content providers seeking to distribute their digital wares via the Internet.[5] CDNs served the dual function of lowering the resource costs required to

---

[4] Multicast allows a single packet that is to be sent to multiple destinations to be replicated only at branching points, thereby reducing the packet transport resources required (cf. Cisco, 2001). DiffServ and IntServ are two Internet protocol-based ways to provide differentiated QoS support for packet delivery (cf. Cisco, 2005).

[5] For a discussion of Internet overlays, see Clark et al. (2006).

distribute content, while facilitating enhanced performance and control over how content is delivered to end-users, thereby improving the QoE when end-users consumed the content.

The exponential growth in digital traffic and of the Internet into a global platform for all types of electronic communications has induced adaptations across the entire ecosystem. Beyond the growth and increasing importance of CDNs, ISPs that provide the basic transport infrastructure of the Internet have been upgrading their networks with enhanced capabilities using technologies like Software Defined Networking (SDN) and Network Function Virtualization (NFV) to enable them to offer much finer-grained, flexible/agile, and dynamic control over network resources.[6]

At the same time, and re-enforcing the incentives of ISPs to upgrade the capabilities of their networks, edge-based application and content providers expanded the range of their offerings to take advantage of more capable networks and a growing marketplace of more digital-savvy consumers, who are armed with a larger array of more capable devices (i.e., PCs, tablets, smartphones, smart-TVs, etc.). The range of online services and how they are used by consumers and businesses has greatly expanded, contributing to the rise of the Internet of Things, Big Data, Cloud Services, and 5G – to list just a few of the most popular buzzwords being used to characterize the evolution of the Internet ecosystem.

CDNs have evolved in response to and have contributed to driving these changes. Today's CDN landscape is both more complex and competitive, with a multiplicity of provider types vying for a share of the business. In the following sub-sections, we provide a taxonomy and discuss some of the strengths and weaknesses of alternative CDN business models and describe in greater detail how key Internet trends have impacted and been impacted by CDNs.

2.1. A Primer on CDN Service Provision

CDNs employ a scalable distributed architecture of servers that is overlaid on the Internet's basic packet transport infrastructure to enhance the speed, consistency, and reliability associated with delivering content to end-users, while at the same time, reducing the transport resources required to deliver content.

To provide this functionality, CDNs manage a set of cache servers that are strategically distributed across the Internet to replicate and store content in closer proximity to end-users. Positioning multiple copies of origin content at multiple locations across the Internet allows the CDN to better match incoming requests from end-users for content with the server that is best located in order to serve each end-user's request – where "best" often means the server closest to the end-user.[7] CDNs operate by employing sophisticated redirection mechanisms that intelligently match user requests for content with the most suitable server ("source"). In the case of CDNs with a network of dedicated servers, the redirection decision is typically done via the

---

[6] For discussions of NFV and SDN, see Telefonica (2014), ETSI (2012), or Metzler (2015).

[7] The "location" of the best server is usually the one that is the fewest number of network hops away from the requesting end-user; however, as we discuss further in Section 3, there are multiple ways in which the *best* location for serving content may be identified.

Domain Name System (DNS); whereas in peer-to-peer CDNs, a distributed naming scheme is utilized to redirect content routing requests.

CDN decisions of where to store content are heavily influenced by the nature and popularity of the content. A typical CDN maintains a hierarchy of distributed servers, with servers closest to the edge of the CDN storing the most popular content. When a user requests an object that is delivered by the CDN, the CDN decides on the appropriate edge server best able to respond to the request, redirecting the user's request to the selected server using the DNS system. If the edge server has a copy of the requested content, then it provides the content to the end-user; otherwise the request is passed up the server hierarchy to find the next "closest" server, which ultimately may require going to the origin server of the content producer (cf. Nygren et al., 2010).[8] Once the edge server retrieves the content object, it stores the object locally and delivers a copy to the end-user. In this way, content is replicated and distributed across the CDN's footprint of servers in response to changing demand. Additionally, content may be pre-cached in anticipation of requests by end-users or content providers.

Deciding what content to store in which servers and for how long to retain copies, and how to best manage requests for serving content is complicated and depends on the nature of the content, the preferences of the content provider, end-user demand for the content, what else is going on in the Internet, and the capabilities of the CDN provider. Some of the content may be static and need to be updated relatively infrequently, while other content may be dynamic; or, the origin servers may be located at different content providers (e.g., as might be the case for advertisements and background text); or, the content providers may have specialized security or stringent QoS requirements that necessitate specialized routing treatment by the CDN provider. As a concrete example, consider the challenge of serving content for a newspaper's or television news station's website, which includes both breaking news and long-lived stories, all of which may have very different user consumption patterns across the Internet.

Replicating the content in multiple locations facilitates server load balancing and protects against server failures or attacks (e.g., denial of service attacks launched against the origin server). CDNs continuously update where content is stored and where it is served from to dynamically adjust to changing traffic patterns in the Internet and fluctuations in the supply of and demand for the content served by the CDN.[9] By jointly managing multiple servers in real-time, the CDN can better balance server loads, enhancing the overall server capacity utilization efficiency (cf. e.g., Nygren et al., 2010; Maggs and Sitaraman, 2015). The distributed architecture of CDNs also facilitates capacity scaling, allowing the CDN to respond to flash crowds and denial of service attacks. Furthermore, distributing content across multiple servers enhances reliability by eliminating single points of failure.

---

[8] As will be explained in Section 3, the "location" of the next "closest" content server may be determined relative to different metrics in quality- or cost-space, depending on which delivery features are emphasized.

[9] Content providers may serve more or less dynamic content, and may pay for differing levels of service appropriate to the needs of their end-customers (or the budgets of the content providers). Commercial CDNs offer a range of services to match the desires of their content-provider customers to ensure a particular end-user QoE with fluctuations in the end-user demands for the content-provider's content.

One way CDNs reduce delivery costs and enhance performance is by reducing the distance data packets need to travel between CDN servers and content consumers. If ten users in a city want to access a particular piece of content for which the origin server may be across town or the globe, the ability to send a single copy to the city and cache it in a local server saves the need to transport ten copies that entire distance. Thereby, transport resources needed to distribute the content are reduced. Deploying servers close to the end-users helps overcome limitations of the Internet architecture including those imposed by transport protocols and Internet congestion, reducing end-to-end delays and enhancing the consistency of the end-users' experiences when accessing content (cf. Dilley et al., 2002).

CDNs further enhance their capabilities to serve content by employing proprietary protocols to distribute content across their network of cache servers; and can add functionality to handle content differentially on a per-request basis (e.g., to serve higher or lower-resolution content based on the identity of the requestor, the current status of the network, or the preferences of the content provider) (cf. e.g., Nygren et al. 2010; Winstein and Balakrishnan, 2013). By playing these roles, CDNs make up for the lack of functionality in the basic Internet.[10] The CDN assumes responsibility for dynamically optimizing how content is served, re-optimizing basic Internet protocols or substituting proprietary protocols. This is accomplished in a manner that is transparent to the end-users and the content providers, and thus saves them the need to download special software or otherwise modify how they interact with the Internet in order to address the limitations in the Internet's ability to provide QoS-differentiated services.[11]

As noted, CDNs typically operate as an overlay network that adds functionality and servers "on top" of the basic Internet. Typically, CDNs do not operate their own backbone transport infrastructure (although some do), nor do they typically lease dedicated connections to link content requestors and content providers (e.g., virtual private networks or private lines). Instead, CDNs rely on the public Internet for the delivery of content from their front-end edge servers to end-users, as well as for communications between the CDN's own servers. Typically, the decisions of how packets are routed in the data plane is left to the ISPs, and thus, is not directly managed by the CDN. In order to optimize packet delivery, CDNs continuously monitor the performance of alternative packet transport paths to maintain a dynamic map of the state of the Internet. These dynamic maps allow CDNs to determine where best to store and serve content.[12]

---

[10] As noted earlier, some of the functionality that CDNs offer, in principle, could be provided using standard Internet protocols, but lack of uniform support for these in the Internet, especially when traffic has to cross ISP management domains, renders it infeasible.

[11] For further discussion of the optimization techniques employed by CDNs, see Leighton (2009).

[12] Maintaining detailed maps giving real-time information on the state of the access and backbone networks that comprise the global Internet is complex and challenging. The underlying networks are run by independently managed and often competing ISPs that in general find it undesirable to share detailed network performance information. Thus, CDNs rely on their own measurement capabilities to keep abreast of changing network conditions.

Consequently, CDNs and access ISPs are in a symbiotic relationship. Both play a critical role in determining the end-users' QoE when consuming content. A number of researchers have characterized the access ISPs broadband platform as a two-sided market that matches broadband subscribers with content and applications.[13] Content providers need the ISPs to deliver their content to the broadband Internet subscribers and ISPs benefit from increased demand for broadband access services when subscribers can access desirable content. Both the ISPs and the CDNs have shared incentives in ensuring users have a good QoE, but may not agree on how best to achieve that or how to share the producer surplus generated. As CDNs, ISPs, and the markets for content evolve, the question often arises as to how best to apportion responsibility for a good QoE between CDNs and the ISPs, raising vertical and horizontal integration questions. We explore some of these issues further below.

2.2. Innovation and the Cost of CDN Service Provision

CDNs began to emerge in the mid to late 1990s. In the early days of CDN service provision, bandwidth costs (i.e., transit costs) were the major source of operating costs for CDN providers. Other cost categories such as the capital and operating costs of acquiring and maintaining servers (including energy to power and cool servers) and CDN personnel costs were less significant. Thus, in early CDN deployments the main concern was to minimize bandwidth costs. With the growth of the Internet, as a consequence of technical innovation, and as a result of changing market competition dynamics, transport bandwidth costs have fallen significantly. Consequently, managing bandwidth costs is relatively less important for CDN providers today, but remains a key component of the CDN value proposition.[14]

One of the factors that has contributed to the fall in transit costs is the growth and expansion of settlement-free (or, revenue-neutral) peering agreements in the Internet. This is part of a seismic shift in how the Internet is interconnected (cf. e.g., Faratin et al., 2008; Clark et al., 2011). Revenue-neutral peering paths may be used to bypass settlement-based transit agreements, thus eliminating transit costs that would otherwise have been incurred in delivering content from an origin server to an end-user. The growth of Internet Exchange Points (IXPs) (cf. Chatzis et al. 2013) and interconnection facilities such as Equinix have played an important part in this story as they provide platforms for direct interconnections between networks (cf. Labovitz et al., 2010). Further, as the Internet and the Web have grown, large private and public investments in fiber networks and transatlantic links have greatly expanded transport capacity in the global Internet, putting additional downward pressure on bandwidth pricing. As Norton (2014) reported, Internet transit bandwidth prices have fallen from $1,200 per Mbps in 1998 to $0.63 per Mbps in 2015 – representing an exponential decline in prices that continue to fall.

In spite of these general price decline trends, however, price differences between different regions continue to exist. While the declines in prices per Mbps have been significant around the

[13] For example, see Rochet and Tirole (2006), Evans and Schmalensee (2007), or Economides and Tåg (2012) for foundations of two-sided market theory and its applications to the Internet.

[14] Even though per-MB bandwidth costs have fallen, it is possible aggregate CDN bandwidth costs may have increased because the volume of CDN traffic has grown so significantly.

world, bandwidth prices in Asia and the Pacific may be up to three times higher than prices in Europe and North America (cf. Stronge, 2015; Telegeography, 2016). Moreover, bandwidth pricing is typically subject to significant volume discounting (i.e., large capacity transport paths are significantly cheaper on a per MB basis than are lower capacity paths). Thus, in spite of the reductions in bandwidth pricing, CDNs can still provide a valuable role in managing regional and volume-based bandwidth resources.

As a result of technical progress in accordance with Moore's Law, the cost of servers and storage have also fallen dramatically due to significant improvements in computing and communications technologies (cf. Armbrust et al., 2009). As a consequence, the cost structure of CDN service provisioning has changed in recent years. Cost components, such as payments for hosting services (to the ISPs, internet exchanges, or carrier hotels that provide the physical locations where servers are hosted) and the electricity used to power and cool the servers have grown to become significant components of CDN operating costs (cf. Qureshi et al., 2009; Greenberg et al., 2009).

At the same time that CDN costs have been changing, the customers for CDN services have become more demanding. In addition to continuing to need services for distributing static content, content providers and their customers have growing needs to serve a broader array of more dynamic and interactive content. To meet this demand and address the growing competition among CDN providers, some CDNs are expanding their offerings to include richer support for other types of applications, moving beyond traditional content delivery services. Indeed, in many cases, the rise of new types of media is straining legacy notions of what constitutes traditional media content distribution.

The delivery requirements of the diverse content and application providers served by CDNs and the end-users to whom content and applications must be served are becoming increasingly heterogeneous.[15] For example, a video streaming application requires high end-to-end bandwidth.[16] At the same time, video streaming services are less delay-sensitive and buffer technology can compensate for some amount of packet losses. In stark contrast, applications like web browsing, e-commerce, and interactive applications such as multiplayer online video games may be more sensitive to delays, but may not require as high data rates.

Additionally, different types of content may be associated with very different economics. For example, a security firm seeking to aggregate surveillance videos from properties they manage may have very different requirements for accessing the video than a commercial provider of entertainment content. Even if one focuses on commercial entertainment media, and further restricts attention to streaming video, such content can have very different distribution requirements. Lots of it may be static and cacheable. This includes movie or television programming libraries from Netflix, HBO, or Hulu. Other types such as live events (sports,

---

[15] For an overview over the requirements of different applications regarding delivery performance, see e.g., BEREC (2014) or Stocker (2015).

[16] For example, Netflix recommends 25Mbps for high definition 4K video streaming (cf. Netflix, 2016c).

concerts) and breaking news requires (near) real-time delivery.[17] The audiences for content can be quite heterogeneous also. For events like the Olympics, the audience is large and global, and near-real-time. Other entertainment content may have few if any end-users interested in viewing it, and such viewing as occurs may be widely distributed across geography and time.[18]

The variety of delivery requirements has substantial implications for the cost of CDN service provisioning. One-size-fits-all solutions are not cost-effective – either they fail to meet the demand of the most demanding services or over-provision capabilities for less demanding content. Some CDNs are exploring niche strategies, focusing on specific types of content, specific architectures, or specific markets to better balance requirements and costs. At the same time there are general-purpose CDNs that are seeking to offer a portfolio of services that are appealing to a broad class of content, content providers, and network environments. This increased complexity in the CDN ecosystem mirrors the growing complexity in Internet interconnection regimes. The two phenomena are interrelated since both reflect responses to the changing nature of Internet traffic and its continued exponential growth and the proliferation of user devices and contexts in which content must be delivered.

2.3. Evolution of the CDN Value Proposition

The increasing share of Internet traffic delivered via CDNs can largely be explained by the growth in rich multimedia traffic, including bandwidth-intensive entertainment video traffic; and the benefits that CDNs offer for better management of subscribers' QoE and for reducing the costs of content delivery for commercial entertainment content providers.[19] Demand for CDN services has increased as CDN capabilities have expanded.

As the number of consumers and businesses operating online expands and a growing share of their digital content is shifted online, the range of content to be delivered and the range of contexts (i.e., high and low resolution, big screen and small screen, static and dynamic, public and proprietary content to be delivered to/from mobile and fixed, locally and globally, etc.) in which content users may wish to access the content has expanded. Thus, the scope of new market opportunities to offer QoS-differentiated CDN services is widening. In addition to providing support for a growing variety of heterogeneous packet delivery requirements (e.g., with respect to optimized video resolution, reduced time-to-first-packet delivered or higher reliability and scalability), additional complementary services have become integral components in the service spectrum of many commercial CDN providers. These include support for enhanced security, additional cloud functionalities such as video-coding on the fly and other in-network computer

[17] Although, even live events are often cacheable for later viewing (or re-viewing) by subscribers who missed the live broadcast.

[18] The market for entertainment media tends to be long-tailed with the popularity of content falling rapidly as one moves from the most popular to less popular content. However, having the long-tail of choice (enabled by the Internet) can be valuable in aggregate, when a large number of choices that each attract limited attention are added together. See, for example, Brynjolfsson et al. (2003).

[19] See Lehr and Sicker (2016) for a discussion of implications of the growth of entertainment video on the Internet ecosystem and how it has changed traffic characteristics.

processing services, and support for market metrics and analytics. These complementary services have grown increasingly important as the marketplace for basic content delivery services has become more competitive and is increasingly available from multiple providers in the Internet ecosystem.[20]

The rise of CDNs has facilitated a symbiotic rise of new business models, including content providers offering Over-the-Top (OTT) entertainment services such as Spotify or Netflix. The emergence of new businesses that can make use of CDN services expands the market for CDNs, providing incentives for incumbent CDNs to expand their service offerings and for entry into the market by others.

The diversity of CDN options confronts content providers with a menu of options for how best to deliver content to their customers. At one extreme, there is the option of centralized hosting. A content provider can simply make content available at a single server and rely on the basic Internet to allow end-users interested in the content to access it. A slightly better solution is for content providers to take advantage of Internet-based application platforms like Facebook (social networking), BitTorrent (peer-to-peer), or Dropbox (file sharing) to post content easily to the Web and rely on the back-end support of those applications to ensure the content is accessible. While such options are readily available and require little in the way of added expense or effort, they provide only limited control over the QoE of end-users. For some content providers with limited budgets or audiences, the basic Internet services may be adequate. Others seeking to ensure a higher QoE for their customers, but unwilling or unable to pay for special services, may be able to rely on free but limited CDN services from providers such as CloudFlare (cf. CloudFlare, 2016).

Larger and predominantly commercial content providers may choose to avail themselves of the services of general-purpose CDN providers like Akamai or Limelight to develop customized service delivery plans tailored to their needs.[21] This middle-market of commercial content providers is the heart of the market for third-party CDN providers. These commercial content providers have sufficient budgets and interest in enhancing the QoE for their customers, while also managing their distribution costs. Commercial entertainment providers confront intense competition for the attention (advertising supported) and discretionary spending (e.g., for paid entertainment content) of their target audiences. Many engage in "windowing" strategies by which they seek to segment their target audience in order to facilitate price discrimination and

---

[20] For example, in 2015, Akamai generated more than half of its revenues from "Performance and Security Solutions". Cloud Security Solutions alone accounted for more than 11% of total revenues (cf. Akamai, 2016c, p. 33).

[21] Akamai, for example, has launched a program to provide their customers with means to better optimize their delivery. Using the Akamai Open platform, CDN services can be tailored to match the individual needs (cf. Akamai, 2016d).

thereby maximize the revenue potential to be captured from each segment.[22] CDNs offer services that can facilitate this goal.

The largest content providers such as Netflix and Google (YouTube), and increasingly companies like Facebook and Twitter, may be growing too large and have sufficiently specialized content-delivery needs that self-provisioning their own specialized CDN networks offers a better option. We discuss the incentives of the largest providers to vertically integrate into self-provisioning CDN further below.

In evaluating these options, content providers need to continually evaluate their make-vs-buy decisions. The growth of CDN options increases the options for outsourcing content delivery and reflects the increased complexity and changing nature of the Internet ecosystem. CDN services constantly evolve and allow their customers to flexibly scale and adapt to emerging challenges and to keep abreast of changing market, technical and policy trends such as changing regulations regarding network neutrality or privacy policies. The capability to adapt to these challenges helps ensure an enduring strategic value proposition for general-purpose CDN providers, especially those targeting the large middle-of-the-market. That is, those businesses wanting more than the basic Internet can provide, but not being so large as to be able to efficiently self-provision in-house CDN capabilities.

At the same time that content providers are evaluating their options for how best to deliver their content, the marketplace for incumbent CDN providers and entrants continues to shift. ISPs that had traditionally based their business models on providing bandwidth for basic packet transport services have seen their transit revenues erode in the face of falling transit pricing and restructured interconnection agreements. Across the ISP ecosystem, a number of ISPs are exploring their options for offering value-added services, and providing CDN services is one possibility. Among the largest operators of full-service networks such as Telefonica, AT&T, or Telecom Italia, investments in NFV and SDN technologies to "softwarize" their networks should enable them to lower costs and increase their ability to offer a full-array of fine-grained, dynamic and flexible Internet and cloud-based services.[23] Increasingly, these next generation ISPs are able

---

[22] Historically, the order of public release was used to implement price discrimination "windows" entertainment. Thus, books were first published in hard back at a high price, and then later as paperbacks at a lower price. Movies were first released to theaters, then to video rentals, and later for over-the-air broadcasting. With the evolution of media markets, traditional media windowing strategies have become much more complicated, but the general goal of bundling content with distribution channels to effect market segmentation to enable price discrimination remains important. For further discussion, see Lehr and Sicker (2016).

[23] The shift to software allows network operators to lower costs by enabling them to rely on commodity hardware; and allows network functionality to be de-localized (e.g., a single software switch can provide the call handling services that previously were provided by a hierarchy of local switches, and the logic and the actual switching do not need to be geographically collocated). Additionally, the shift to software increases the flexibility with which resources may be assigned, allowing the network infrastructure to be virtualized and then sliced into logical partitions that may support different levels of QoS and functionality. In its fullest realization of this development, the ISP can offer multiple tiers of cloud services ranging from wholesale access to core cloud resources (computing, storage, and transport) –

to offer substitutes for some functionality that CDN providers have historically provided as independent overlay networks. The evolution of these ISPs from legacy providers of telecommunication services to full-service providers of cloud services is blurring the boundary between where the Internet ends and overlays begin.

The mix of differing businesses for content and CDN providers on the one hand and ISPs on the other is resulting in a complex mix of vertical and horizontal business strategies and cross-linking organizational strategies. For example, ISPs with a limited geographic footprint cannot implement a CDN to serve global content providers unless they partner with other CDNs or ISPs. The matching between content provider needs and ISP capabilities can be met in numerous ways. Akamai, for example, addresses this challenge by maintaining an independent, global network of Akamai-owned servers that are hosted in ISP networks around the world. Akamai can mediate the relationships with the various ISPs on behalf of its customers, saving them the need to negotiate separate content delivery strategies for each of their markets. Akamai's ability to negotiate credibly with the ISPs is assisted, in part, because it is independent of the ISPs it negotiates with.[24] Other models for delivering content rely on creating federated CDN structures based on business agreements and partnerships that may span horizontally across multiple geographic markets or vertically across content providers or different ISPs. This includes a growing web of vertical agreements between ISPs and CDNs. For example, access ISPs that provide the broadband services that consumers use to access content are well-positioned to host the front-end servers close to the customer's locations; and those ISPs may be able to provide a range of services that can enhance a CDN's capabilities (e.g., packet prioritization, buffering).[25] In the next sub-section, we highlight examples of the different types of CDN architectures and models for collaboration and integration that are in use.

2.4. A Taxonomy of CDN Architectures

The CDN ecosystem has grown more complex. While commercial CDNs offer a portfolio of wholesale CDN services to third parties such as content providers, a number of large content providers have deployed in-house, application-specific CDNs. Some CDN providers have invested in deploying their own network servers, while others rely on peer-to-peer communications to distribute content.

In the following, we provide a survey of CDN strategies and the architectures being deployed to deliver CDN services in the market currently. These are summarized in Table 1. Key features

---

sometimes referred to as Infrastructure-as-a-Service; application or content hosting services – sometimes referred to as Platform-as-a-Service, or Software-as-a-Service, with the last providing the greatest level of management on the part of the ISP cloud provider.

[24] Imagine how Akamai's business position would change vis a vis other ISPs were it to be acquired by a large ISP with a near global footprint like Telefonica.

[25] Network Neutrality regulations may constrain the range of services ISPs may offer CDNs or content providers, but it is technically feasible for an ISP to offer QoS differentiated services, and depending whether these are offered on a discriminatory or common carrier basis may determine whether regulators allow such enhanced services to be offered.

that distinguish the different CDN strategies include the strategy for deploying servers (whether owned or leased, where situated, etc.) and the business models employed (what performance characteristics are emphasized, what key applications are the market focus, and the essence of the value proposition for the CDN clients). The first four examples identify multi-purpose CDNs, while the fifth is for specialized, single-purpose CDNs, and the last three are for new business models for delivering CDN services based on combining the resources of multiple market players.

Table 1: Taxonomy of CDNs

CDN Architecture	Examples of Providers	Deployment Strategy	Bandwidth	Latency	Business Model	Typical Applications
Datacenter-based	Limelight, CacheFly, CloudFlare	Servers at strategically connected facilities	High	Medium	Buy bulk resources	Video Streaming, static Web, software updates
Highly Distributed	Akamai	Servers at peering points and inside access networks	High	Very Low	General-purpose, provide global footprint, best quality	Various applications, including dynamic and interactive Web
Peer-to-peer	BitTorrent	Serverless, functionality at end-user equipment	Low	High	No investment in dedicated infrastructure	File sharing, bulk transfers
Hybrid	Akamai NetSession	Dedicated servers combined with functionality at end-user equipment	Low	High	Partial outsourcing of delivery to end-user equipment	Software updates, file sharing
Specialized	Netflix Open Connect, Google Global Cache, Amazon CloudFront	Specialized servers at peering points and inside access networks	High	Low	Reduce delivery costs for specialized service	Video delivery, specialized applications
Broker	Conviva, Cedexis	Relies on existing deployments of CDN functionality	Custom	Custom	Opportunistic cost management	Video and Web delivery
Licensed	Akamai AURA, Edgecast licensed CDN	Inside access networks	High	Very Low	Telco CDN, or ISP-CDN collaboration	All of above
Federated	Edgecast OpenCDN	Relies on existing deployments of CDN functionality	High	Low	Interconnection of CDNs to expand geographic footprint	All of above

Datacenter-based CDNs: Some CDN operators deploy large numbers of servers (i.e., so called server clusters or server farms) in a relatively small number of geographic locations. An advantage of datacenter-based CDNs is that hosting a large number of servers in a single location contributes to the realization of scale economies and management efficiencies that can significantly lower the costs of operating a CDN. Instead of incurring the fixed (per-site) costs associated with having to secure, power, and manage a large number of server sites, the CDN provider can focus on a smaller number of locations with a larger number of servers, and potentially higher capacity servers. The larger server installations can also benefit from volume discounting associated with bulk purchases of servers, power, and bandwidth capacity connections.

Many existing datacenters were established to provide access to general-purpose computing and storage resources rather than to provide CDN services. Key considerations in the location of datacenters include the cost of powering (and cooling) the servers; the real estate site costs; proximity to anchor tenants; and good connectivity to the Internet. Moreover, for reliability and competitive reasons, the Internet connectivity should support diverse routing to multiple ISPs.

For datacenters to be useful to support CDN services, the datacenters need to be strategically located to provide access to multiple upstream providers (e.g., content providers or transit ISPs) and access to peering locations. Datacenters located close to IXPs have the advantage of being able to peer with multiple networks at a single location. Thus, with a single installation of servers, the CDN can have access to a large number of end-users within a few network hops. However, having the servers located in only a few locations means that the distance packets may have to travel to content end-users may be quite far, requiring the packets to traverse multiple networks, affording the CDN limited control over the end-to-end QoE. In traveling a longer distance and across multiple providers, the packets consume additional transport resources and are more likely to be exposed to congestion that might be bypassed by CDNs with more distributed architectures. Consequently, these types of CDNs are best able to serve content with consumers who are localized near the datacenters, is delay-tolerant (and hence cacheble), or may be scheduled (e.g., to avoid peak congestion). This may include content for certain types of streaming video or software updates.

Limelight is one of the large CDNs that rely on this architecture. It maintains about 80 server clusters around the globe (cf. Limelight Networks, 2016). Smaller CDNs, such as CacheFly, Fastly, CloudFlare, MaxCDN use this architecture with up to tens of server clusters deployed around the globe.

Highly Distributed CDNs: An architecture that is significantly more expensive to build and maintain relies on highly distributed server deployments, spread across a large number of networks and at multiple peering locations. Some of these CDNs also operate dedicated backbone networks to tie their hierarchical server network together and to allow them to connect with high QoS to the origin servers of their content provider customers.

The performance benefits of this approach are substantial as a server can be closer, both in terms of network and physical distance, to a large fraction of the Internet population. Moreover, this provides the CDNs with more options for delivering fine-grained content delivery and application support services. The edge servers in this architecture can act as the front-end servers in the ISP networks where the connections of end-users are terminated. The hierarchy of CDN servers can coordinate using optimized protocols to deliver the content to the end user (cf. Leighton, 2009; Maggs and Sitaraman, 2015). In these highly distributed CDNs, employing dedicated servers, delivery performance is sufficient for the provision of a number of applications such as Web browsing, e-commerce, and interactive services. CDNs that follow this architecture typically rely on a common platform where multiple applications and services can be deployed on the CDN as modules (cf. Nygren et al., 2010). Large, general-purpose CDN providers such as Akamai use this architecture. For example, Akamai maintains about 220,000 servers deployed in more than 1,500 networks, and delivers more than 20% of global Web traffic (cf. Akamai, 2016a).

Because the fixed (and sunk) costs of establishing and maintaining a distributed CDN comprised of CDN provider-owned servers is quite substantial, the CDN operator needs a large customer base to render the model economically viable. However, the large size also makes it possible for such CDNs to take advantage of economies of scale and volume/bulk discounts associated with purchasing and provisioning many of the necessary inputs such as servers and back-office support. Additionally, the large CDNs are likely to have bargaining power when negotiating interconnection and hosting arrangements with ISPs.

Peer-to-Peer CDNs: The investment in server equipment and other resources comprises a significant share of the costs of owning and operating CDNs. In the 1990s and early 2000s peer-to-peer networks became popular. Peer-to-peer networks rely on the end-users to provide the servers used by the CDN. Members of the network seeking a particular file query the network to see if any of the other members have a copy, relying on crowdsourcing to share the files. These networks work best when a large number of users are all interested in the same collections of files. For popular files, hundreds of thousands to a million end-users may act as servers that participate in distributing the file (cf. e.g., Boorstin, 2015). Relying on end-user equipment to store and serve the content saves the CDN from having to invest in and manage the distributed content servers, but also means that the CDN has much less capability to manage the end-to-end experience (e.g., to control end-to-end latency). The CDNs that have been most successful with this architecture are either ones with limited commercial support (e.g., community-driven, non-profit), specialized content distribution needs, or are engaged in copyright infringement (where distributing ownership/control of the servers limits the CDNs vulnerability to copyright enforcement actions). However, some commercial products based on peer-to-peer CDNs exist. The best known commercial peer-to-peer service is offered by BitTorrent, Inc., the commercial counterpart of the popular free file sharing software. It is mainly used for bulk file distributions (e.g., medical data, scientific data, high definition videos) over the Internet.

Hybrid CDNs: Hybrid CDNs seek to economize on the expense of maintaining dedicated servers by outsourcing some of the replication and caching of content to servers maintained by end-users which serve as a CDN client and make the relevant content accessible to other CDN clients. The advantage of this architecture is that the availability of content increases as additional capacity and the number of replicated copies increases. Further, the CDN provider can save additional deployment and maintenance costs, while still retaining the control afforded by also maintaining dedicated servers and having some management control over the use of the end-user server functionality. However, because the CDN cannot reliably anticipate when end users may go offline, it is important that the CDN keep track of the state of each replica as well as to maintain a replica in an always-on server.

Additionally, accessing content stored on end-user servers may suffer from performance issues. Accessing content that is stored on an end-user's server requires traversing the access network twice, which can add to end-to-end delays (cf. Feldmann et al., 2010). Today, such architectures are typically used for large file downloads such as software updates. An example of a hybrid CDN architecture is the Akamai NetSession (cf. Zhao et al., 2013).

Specialized CDNs: In recent years, a number of large content providers that formerly relied on third-party CDN services for packet delivery have built their own application-specific CDN platforms. A major motivation for vertical integration is to take advantage of the scale, scope, and specialization economies that can be realized by bringing the CDN functionality in-house, assuming that the traffic volumes are sufficiently large. The specialization economies arise when the CDN architecture can be tailored to the specific requirements of the content-provider, avoiding having to provide support for unnecessary capabilities that a general-purpose CDN might need to offer, while potentially investing in even better special-purpose capabilities than the general purpose CDN might find to be economically viable. Furthermore, given the critical strategic importance of content delivery services for the content provider's business model, bringing those services in-house offers a range of strategic benefits. When the traffic a content provider needs to serve gets large enough, it may represent such a significant share of the traffic of any individual CDN to pose hold-up risks on both sides – neither the content-provider nor the CDN can tolerate sudden shifts in traffic patterns, creating joint-management risks that may be best managed through vertical integration.

One of the first content providers to deploy their own specialized CDN was Google. They deployed their own caches (Google Global Caches) in more than 1,000 networks (cf. Calder et al., 2013; Streibelt et al., 2013). Deploying servers deep within access ISP networks reduced the distance packets needed to travel to reach end-users and bypassed interconnections that may be subject to congestion. Additionally, Google's CDN infrastructure allowed Google to accelerate applications that run in the Google datacenters (cf. Flach et al., 2013). Although the cache deployment strategies may be similar to those used by the highly distributed third-party CDN providers, Google caches may be used exclusively for their own content. By some measures, Google, which includes YouTube, handles more than 25% of Web traffic.[26]

Recently, a number of other large content providers have elected to construct in-house CDNs. Among the largest ones are Netflix Open Connect (cf. Netflix 2016a) with servers in more than 233 locations worldwide, many of which are located inside access ISP networks (cf. e.g., Miller, 2016; Netflix, 2016a; Clancy, 2016). Sandvine reports that Netflix is responsible for about one third of the peak Internet traffic on fixed networks in the U.S. (cf. Sandvine, 2016). Another specialized CDN is Amazon's CloudFront (cf. Amazon Web Services, 2016) which has a presence in more than 50 locations around the globe.

Broker CDNs: Because many regions are served by a number of CDNs offering diverse pricing, brokers have emerged as intermediaries between content providers and CDN providers. These brokers help content providers mix-and-match services from multiple CDN providers to achieve cost-minimizing solutions that meet their global content delivery performance goals. CDN brokers maintain maps of available servers and run end-user experiments to assist content

[26] This estimate was provided by Craig Labovitz of Deepfield. In 2010, however, it was estimated that Google accounted for just 6% of Web traffic (cf. McMillan, 2013).

providers in the selection of CDNs that best fit the content providers' cost-performance requirements. CDN Brokers implement a meta-architecture for optimizing content delivery relying on existing cache deployments. One or more CDNs may be used simultaneously, with the assignment of user requests to specific CDNs made on a dynamic basis. Broker CDNs typically specialize on providing services for specific types of content or applications. For example, Conviva (cf. Krishnan and Sitaraman, 2013; Liu et al., 2012) is a CDN broker that specializes in delivering video streaming services, while Cedexis (cf. Cedexis, 2016) specializes in delivering Web browsing traffic.

Licensed CDNs: The development of fully proprietary CDN solutions is typically very expensive, requiring significant upfront investments in hardware and human capital. These costs increase both with the range of capabilities the CDN seeks to support, and the geographic scale of the CDN, many of which serve content-providers interested in distributing their content in multiple nations, if not globally.

This has posed a significant barrier to entry for ISPs that might otherwise have sought to bring the CDN functionality in-house as a service they could offer directly to content providers. Most ISPs lack infrastructure outside of their geographic service areas or inside the networks of competing ISPs with which to serve content to off-net subscribers. The desire for non-CDN market participants to participate in the provision of CDN services has given rise to a number of collaborative arrangements based on licensing agreements with existing CDN providers. For example, anticipating the interest that access ISPs have in enhancing the performance of content delivery to their subscribers while managing their transport costs, established CDN providers like Akamai (cf. Akamai, 2016b) and Edgecast (cf. BusinessWire, 2012; Verizon, 2016) introduced off-the-self wholesale CDN solutions. These solutions allow ISPs to deploy servers that can run software licensed from the CDN providers to enable those servers to act as CDN caches with functionality similar to that offered by the CDN provider's network. These ISP-based servers may be managed independently by the ISP seeking to provide CDN services on a standalone basis, or may be integrated with the CDN providers existing servers. In the latter model, the ISP servers act as front-end servers (delivering the content to end-users) while the CDN servers serve as the back-end servers coordinating delivery across the CDN's entire footprint.

A number of telco-based ISPs have elected to operate standalone CDNs based on bare licensing of the CDN software platform. More common, however, are ISP-CDN collaborative licensing agreements in which the CDN provider integrates the ISP servers into its (potentially global) network of servers, and directly manages the ISP's front-end servers. Akamai is involved in a number of such collaborative licensing agreements with ISPs around the globe (cf. e.g., Frank et al., 2013; Akamai, 2016b). Such collaborative licensing arrangements enable tighter integrated management of the CDN traffic, facilitating revenue sharing and making it easier to launch new applications. Additionally, such collaborations can alleviate information asymmetries. Improved performance can be achieved when the CDN and ISP jointly collaborate on identifying the best server from whence to provide content to specific user requests (cf. Frank et al., 2013).

CDN Federation: A CDN federation provides a framework for interconnecting and coordinating the exchange of traffic among independent CDNs with limited footprints, enabling members in the federation to offer services to content providers requiring wider geographic coverage than the participating CDNs could offer by themselves. CDN federations are based on horizontal collaborations between CDNs, and help reduce the costs of expanding server coverage. The IETF working group, CDNi, is focused on developing standards for federated CDNs (cf. Niven-Jenkins et al., 2012). The CDN provider Edgecast markets its OpenCDN as a commercial product that enables CDN federation.

An analysis of connection speeds provides useful insights into the benefits of alternative CDN deployment strategies.[27] First, it is worth noting that average latency[28] is expected to decrease as available connection speeds increase. Thus, the benefits of a highly distributed CDN architecture in terms of reducing latency is likely to be higher (lower) in access networks with lower (higher) connection speeds. Second, analyzing the origin of latency, i.e., whether it arises principally in the access network or in the backbone network, gives further insights on optimal server deployment strategies.[29] If latency in the access network dominates backbone latency, then those CDN architectures that more heavily rely on the access network, e.g., peer-to-peer and hybrid CDN architectures, are expected to perform worse than alternative CDN architectures. If, however, latency in the backbone network dominates, then CDN architectures that rely on server deployments inside the access network can be expected to deliver more substantial performance benefits.

3. The Multiple Facets of (Peering) Location

Earlier we noted the importance for CDN performance in correctly matching incoming user requests for content with the "closest" server from whence to serve the content. CDN providers offer different capabilities for identifying the "closest" server and this turns out to be another important factor to explain differences across CDNs. It also highlights how the market for CDN services is becoming more complex as the capabilities of CDNs expand.

[27] Studies on residential networks in the US (cf. FCC, 2015), the EU (cf. European Commission, 2015), and the UK (cf. Ofcom, 2016) illustrate significant differences between access technologies in both upstream and download speeds. Although available connection speeds have been increased during the last decade in the majority of examined countries, publicly released Internet connection speeds of users contacting the Akamai CDN (cf. Akamai, 2016e) emphasize the heterogeneity of available download speeds for Internet users located in different countries and regions.

[28] Here, latency is defined as the round-trip time of a data packet from the end-user's home to the closest measurement server and back.

[29] If a CDN edge server is located within the same metropolitan area as the end-user, latency in the access network typically exceeds backbone latency by tens of msec. In this case, latency in the backbone is negligible for QoE (cf. Pujol et al., 2014, figure 10). However, if the server is located in another metropolitan area, latency in the backbone may become significant and may even exceed access latencies. Depending on the distances data packets need to travel, backbone latencies typically vary from a few msec up to tens of msec, e.g., in the US between east coast and west coast, to more than 100 msec if servers are located in different continents (cf. e.g., AT&T, 2016; Pujol et al., 2014).

The selection of the location where a CDN exchanges traffic with other networks and the conditions at the corresponding peering or hosting locations affect both the performance and cost of content delivery. In this section we discuss how different CDN architectures exploit different notions of location and closeness.

3.1. Geographic Location

Location has a geographic or spatial dimension. The set of distributed servers belonging to the same CDN define its geographic coverage footprint. If an ISP and a CDN can exchange traffic at multiple locations, this can be beneficial in several ways. Packet delivery performance can be optimized, resilience can be enhanced, and optimizations of the traffic matrix may increase the efficiency of capacity utilization.

Optimizing the CDN footprint is based on the strategic decision on where to position CDN servers. For example, highly distributed CDN architectures enable the CDN to inject traffic into the terminating access network in close proximity to end-users.[30] When data is cached close to end-users, the data packets have to traverse fewer network hops, as well as fewer administrative hops (i.e., interconnection points between networks). Data communications between CDN edge servers and end-users are "regionalized" and the performance of packet delivery can thus be increased. However, optimizing the traffic matrix typically requires the CDN to identify optimal peering locations. Because CDN service provisioning typically relies on third-party access networks, the CDN may lack sufficient information regarding the condition of traffic on the networks of the access ISPs. To mitigate this deficiency, CDNs usually measure the performance of the available paths to end-users continuously. The CDNs use sophisticated algorithms to select the optimal paths. Direct interconnections between CDNs and access networks allow CDNs to bypass upstream providers (i.e., transit ISPs), thereby reducing the number of hops and packet delays. In general, shorter paths mitigate the effects of shortcomings of protocols like TCP, the performance of which diminishes disproportionally with the geographical distance between communicating hosts. Additionally, peering in multiple locations increases resilience and improves performance. Stroke points (i.e., congested links or link outages) can be bypassed and load balancing capabilities can be increased. Large CDNs make use of this and typically peer at a large number of geographical locations. For example, Google is peering in more than 100 cities (cf. Google, 2016) and Akamai is peering in more than 100 IXPs and private locations (cf. Hannigan, 2015).

It is worth noting, however, that geographic distance is not always a good proxy for the performance of content delivery. This can be illustrated by two examples. In CDN architectures such as peer-to-peer networks and hybrid CDNs, caches are located at the end-user. In this setting, two users can be very close to each other, but the content delivery performance can still

[30] The ability to control how a CDN hands traffic to the terminating ISP may prove useful in negotiating colocation and interconnection agreements (including peering) since it may be possible for the CDN to hand-off traffic in ways that do not impact the CDN customers' QoE, but may be more or less costly for the ISP to handle. Modern CDNs are capable of sustaining very high utilization along packet routes without any packet losses, and of shifting traffic among sources on short time scales.

be rather bad compared with the performance of alternative CDN architectures (e.g., datacenter-based or highly distributed architectures). Even if two clients are in close geographic proximity, packet delivery requires that each packet traverse parts of the access network twice, and since the access network is often where transmission bottlenecks may be encountered, traversing the access network twice can result in significant packet delays. In contrast, packets delivered by CDNs that deploy their servers at datacenters or within access networks have to traverse the access network only once. Therefore, simply focusing on the geographic distance might be misleading.

Another important observation is that network conditions may differ along paths in ways that may render a geographically longer path preferable (lower cost or lower congestion) in a given situation. For example, two CDN servers may be located in the same physical location (e.g., in the same carrier hotel), but one may be interconnected on the network side of the access ISPs network, while the other may be on the far side of a peering link into the access ISPs network. The latter server would be vulnerable to delays if the peering link is congested. According to recent studies, content delivery across a peering link can increase delays by up to 40 msec (cf. Luckie et al., 2014). When serving latency intolerant applications such as multiplayer gaming applications, such delays can result in significant performance impairments. While datacenter-based CDNs may suffer in these cases, agile CDN deployments that are highly distributed (e.g., licensed CDNs, specialized CDNs, federated CDNs, and CDN brokers) can utilize servers within networks and thus are typically better able to offer enhanced delivery performance.

3.2. Virtual Location

Internet routing protocols do not route packets based on the geographic (physical) location of the source or the destination, but based on their virtual location (i.e., the IP address of the destination). Referring to the example given in the previous section, this means that two servers that are in the exact same physical location but have different IP addresses that belong to different networks may be reached via very different paths. In contrast, two servers with IP addresses associated with the same provider may be hundreds of miles apart (e.g., one in New York and the other in Los Angeles) (cf. Freedman et al., 2005). A server's IP address determines how packets are routed to and from the server. If two network nodes in close geographic proximity having IP addresses that belong to different administrative domains (i.e., autonomous systems or ASes) exchange traffic, long paths that cross different geographic regions, or even national borders, might be involved. Delivery performance and QoE levels might be degraded substantially (cf. e.g., Spring et al., 2003; Gupta et al., 2014a).

The role of virtual location has been recognized and exploited to improve the performance of content delivery services (cf. e.g., Krishnan et al., 2009). In addition to traditional interconnection agreements, ISPs may provide CDNs with hosting services. Local IP addresses that belong to and are managed by the hosting ISP may be assigned to CDN servers. Some ISPs offer bundled services requiring that servers that are assigned local IP addresses must be collocated in the ISP's datacenters. Some CDN providers may opt for the hosting strategy to improve delivery performance.

When a significant share of the content is cacheable – as is the case with a lot of the long-form entertainment media,[31] there are multiple benefits that accrue from ISPs hosting the servers on their access networks and bypassing the need for the traffic to cross the peering or transit links that would be required if the CDN server were hosted off-net. In addition to improving the QoE associated with the delivery of the cacheable content, off-loading the traffic from the peering links is likely to produce benefits for other traffic and applications (including those that are less delay tolerant), while at the same time rendering the overall aggregate traffic flow across the peering link better able to achieve its requisite QoS without requiring differentiated services. Additionally, off-loading the peering traffic may allow the ISP to postpone making capacity upgrades not only at the interconnection points, but potentially also in the middle mile networks that would otherwise be required to haul the traffic to the edge-located CDN servers. Moreover, if the CDN provider and the ISP share information, then they can better plan how and when to update the servers, further facilitating management of costly interconnection capacity.[32]

3.3. Colocation Hubs

Today, so-called colocation hubs enable multiple ISPs and CDNs to interconnect and exchange traffic (cf. Lodhi et al., 2014). These are typically located at Internet Exchange Points (IXPs) (cf. Chatzis et al., 2013) or other interconnection facilities that enable private interconnections (cf. Giotsas et al., 2015). Although most of the traffic exchanged globally is via private interconnections, more than 30 Tbps of traffic are exchanged at IXPs, with most of that activity taking place in Europe.[33] There are more than 400 IXPs and more than 600 interconnection facilities around the globe and the numbers are increasing. In large IXPs such as DE-CIX in Frankfurt, LINX in London, and AMS-IX in Amsterdam, more than 600 networks can exchange traffic on a bilateral or multilateral basis. Moreover, hundreds of networks interconnect at large interconnection facilities like Equinix in New York. Being present in these large IXPs and peering facilities provides CDNs with various options for accessing a large number of networks and end-users while at the same time reducing both the geographic and network distance between source and destination (cf. e.g., Galperin, 2016). For example, Equinix maintains a number of interconnection facilities around the globe and states that peering in their facilities

[31] For example, popular videos or even entire video libraries can be easily replicated and stored on hosted CDN servers within the access network. Updates of these libraries can be timed strategically, i.e., off-peak periods can be used and distribution can take place several days in advance before the content is made available for end-users. Netflix, for example, follows such a strategy in order to pre-position video content on their cache servers (cf. Netflix, 2016b).

[32] For further discussion of the mutual benefits to be realized from CDNs with cacheable content hosting their servers on ISP networks, see Ager et al. (2010), Erman et al. (2009), Frank et al. (2013), Qureshi et al. (2009), and Chiu et al. (2015).

[33] Providing European-wide content delivery coverage typically requires using multiple ISPs since many lack multinational coverage, whereas in the U.S. there are multiple national ISPs that cover the entire U.S. market. Consequently, in the U.S. it is easier to negotiate bilateral peering arrangements with a few national ISPs to secure delivery to all of the broadband content users in the market; whereas in Europe, it would be necessary to negotiate a larger number of such interconnection agreements. This may partially explain why IXPs are, relatively speaking, more important in Europe.

makes if possible to access to 90% of the population in North America and Europe with 10 msec delays (cf. Equinix, 2016, p. 2).

The geographic location of colocation hubs is important since interconnection and colocation prices vary between regions and may even differ between hubs in the same city.[34] The growing number of hubs and IXPs in major cities (and increasingly, even in second tier locations) spurs competition, contributing to the downward pressure on bandwidth prices for transit and peering. For example, the recently proposed OpenIX initiative (cf. Chatzis et al., 2015) in major hubs in the U.S. is considered to have caused price declines in some peering facilities by up to 80% (cf. Temkin, 2016). Large hubs are located in major centers for commerce (e.g., New York, Frankfurt, London, and Hong Kong) and where major trans-oceanic fiber cables terminate (e.g., Amsterdam and Seattle) (cf. Durairajan et al., 2015; Giotsas et al., 2015). The density of networks and CDNs at the major hubs gives rise to network effects that result in scale economies. The scale economies and other benefits of interconnecting at hubs may be further enhanced through mergers and acquisitions.[35]

The rise of the hubs has contributed to the "flattening" of the traditionally hierarchical interconnection ecosystem. Beyond the fact that hubs provide facilities for interconnection and attract a large number of ISPs and CDNs, they also drive the adoption of more direct interconnection and hosting arrangements. Handing over traffic directly to terminating access networks reduces dependence on transit providers (i.e., by means of what is referred to as secondary peering or donut peering) and reduces interconnection complexity (cf. e.g., Reed et al., 2014). CDNs that collocate facilities at hubs and IXPs have more opportunities for direct peering and hosting, and can thus improve the performance of delivery services.

3.4. Innovation Hubs

There are also a growing number of innovative hubs offering support for a diverse array of interconnection arrangements. These include partial transit and paid peering, as well as public peering via route servers that enable hundreds of interconnects to be supported over a single port (cf. e.g., Faratin et al., 2008; Ager et al., 2012; Giotsas et al., 2013; Richter et al., 2014). The range of interconnection facilities may include cross-connects (i.e., typically based on dedicated circuit-switched connections between the interfaces of two peering networks), public peering (i.e., typically based on the use of the IXP's shared switching fabric), or, tethering (i.e., typically based on the use of a virtual local area network (VLAN) on the IXP's shared switching fabric) (cf. Giotsas et al., 2015). This diversity of mechanisms for physically interconnecting traffic can support complicated and nuanced traffic exchange arrangements (cf. Giotsas et al., 2014).

[34] For overviews over pricing diversity for colocation services and transit based on TeleGeography data, see, for example, Hjembo (2016) and Stronge (2015).

[35] For example, Equinix recently announced the acquisition of Telecity, one of the largest datacenter and peering facilities owners for $3.8 Billions (cf. Boyle, 2016). Last year, Digital Realty announced a $1.9 Billion acquisition of Telx (cf. Sverdlik, 2015).

Remote peering decouples the geographical location of the router from the geographical location of the interconnection.[36] A physical router presence is no longer necessary for networks to directly peer with other networks in many geographical locations. A CDN with servers in a single location may peer directly with multiple remotely-located networks via the services of a remote peering provider. Remote peering represents an increasingly attractive type of peering, especially for small networks, including small CDN providers. It provides a mechanism to expand a CDNs footprint and gain access to multiple markets at low cost. Limelight is one of the CDNs that takes advantage of this type of peering as well as other networks that have distributed servers such as LinkedIn. For example, more than 20% of the members in AMS-IX use remote peering (cf. Castro et al., 2014).

Owners of peering hubs support innovation by offering a number of free value-added services to their members. For example, the latest blackholing services can be used to mitigate attacks and ground unwanted ingress or egress traffic (cf. e.g., DE-CIX, 2016).[37] Several large IXPs, for example, DE-CIX, AMS-IX, and LINX already provide corresponding services. CDNs can take advantage of such innovations by peering at such locations. Moreover, new technologies such as SDN enable more advanced types of peering as well as sophisticated traffic engineering (e.g., load balancing, inbound traffic engineering and application-specific peering) (cf. Gupta et al., 2014b). Deploying innovative services at innovation hubs can make these capabilities available to all of the networks that interconnect at the hub. Table 2 provides an overview of some of the capabilities that are available at colocation hubs.

Table 2: Examples of Innovations available in Colocation Hubs

Innovation	Enables
Multilateral Peering	Scalable and complex peering arrangements
Blackholing	Ingress and egress attack traffic control and mitigation
Remote Peering	Reduce network equipment purchase and operating costs
Software Defined Networking (SDN)	Traffic engineering, application-specific routing, load balancing

[36] Remote peering allows interconnections for routers that are not physically located at the colocation hub/IXP but connected based on dedicated Ethernet or MPLS connections (cf. e.g., AMS-IX, 2016; Richter et al., 2014).

[37] A number of IXPs provide free blackholing services to mitigate the negative effects of DDoS attacks. Based on routing updates, blackholing allows that all adverse traffic towards the affected network can be dropped by the IXP (cf. e.g., DE-CIX, 2016; Dietzel et al., 2016).

4. Prospects for the future of CDNs and the Internet Ecosystem

The growth of the Internet has increased the demand for a wider-array of quality-differentiated content delivery services. It has also expanded the supply of wholesale platform services that have lowered the entry barriers for offering value-added services. This has encouraged new entry into the value-added provider markets and facilitated self-provisioning by a growing number of content providers and users. Although the overall market has grown, prices in legacy markets for basic transit or plain vanilla content delivery services have fallen.

Across the value chain, businesses are continuously confronting the make-vs-buy decision. The largest incumbent CDNs, best exemplified by Akamai, have sought to expand their services, coverage, and market share in a bid to realize scale and scope economies that will allow them to compete with a broad portfolio of service offerings, while also keeping costs low. At the same time, the largest ISPs have been upgrading their networks through softwarization to expand the portfolio of services ISPs can offer, while also helping to lower network costs. This includes firms like AT&T, Comcast, Telefonica, and Telecom Italia. Some of these are also integrating into content directly (e.g., Comcast acquired NBC; Verizon acquired Edgecast and Yahoo!). Smaller players are pursuing niche strategies or are forming federations in an attempt to approximate the capabilities and cost economics of the larger incumbents. Meanwhile a growing number of content-providers are taking advantage of CDN platform and other wholesale services to vertically integrate into self-provisioning their CDN services.

In this landscape there are likely to be multiple potential sweet spots. At the center of the market, we expect there to be a place for several, but probably a small number, of global scale, full-service CDN providers like Akamai that own and manage a global network of servers and retain the backbone networking capabilities to allow them fine-grained control. There are significant global scale and scope economies associated with managing a network like Akamai's and Akamai benefits significantly from incumbency and its first-mover advantages associated with its success in acquiring such a large installed base of customers. Replicating the global reach of Akamai's dedicated server network would be extremely challenging for another third-party CDN, yet there are a range of commercial content providers – including in the hugely important entertainment media – that want global-scale content delivery coverage. Such coverage cannot be matched by even the largest ISPs. Akamai is able to spread the fixed and sunk costs of developing additional services over its large customer-base. Additionally, its control over the routing of traffic likely provides significant leverage for Akamai in negotiating favorable colocation or interconnection terms with ISPs. Finally, because Akamai is not integrated vertically either into content or into providing last-mile Internet access services, it avoids potential channel conflicts when negotiating service agreements with competing content providers or ISPs. Although ISPs in non-overlapping geographic markets do not compete directly, many ISPs compete with other ISPs in the geographic markets they serve. Content providers typically want to reach all of the end-users in a market, not just the ones served by one ISP or the other.[38]

[38] If sufficiently valuable or if allowed by policymakers, content providers may be interested in exclusive ISP distribution contracts; but even in that case, to negotiate the best terms, they may wish to have the

Although it is hard to challenge Akamai in its dominance of general-purpose CDN service provisioning, there is still substantial room for niche strategies. First, smaller or new entrant CDNs may take advantage of the lower entry costs enabled by CDN platform providers (including Akamai) and various Infrastructure-as-a-Service or Platform-as-a-Service offerings from cloud service providers such as Amazon, Google, or the larger ISPs. Although they may not have as extensive capabilities as Akamai or have the scale of Akamai, it is feasible for them to operate on much smaller margins because the availability of scalable wholesale platform options renders most of their costs variable. By appealing to a niche market by application, geographic market, or by customer type (type of traffic, type of customer), niche providers may be able to exploit a competitive advantage.

Second, other entrants in the value-added CDN services market may be able to leverage another asset (e.g., existing datacenters, a favorable location to content providers or Internet exchange facilities) to enter CDN services at low cost, because most of the costs are recoverable from the firms' main line of business. The temptation to do so and the architecture that may be employed, in taking advantage of sunk or shared costs, may be sub-optimal relative to greenfield CDN deployments. However, identifying such instances may be difficult.

Third, large content providers may find sufficient benefits from reducing costs and in increasing control over how content is delivered to their end-users to make it desirable to vertically integrate into self-provisioning CDN services. As already noted, the largest content providers like Netflix may have been compelled to self-provision because the sheer volume of their needs exceeded the capabilities of all but the largest CDN providers. At the same time, their specialized needs (e.g., in the case of Netflix, serving content that was essentially 100% static and cacheable) may make it feasible for customized self-provisioning to realize substantial cost savings relative to those achievable by a general-purpose CDN provider. The viability of self-provisioning and the expansion of this model to a growing range of ever-smaller content providers is a direct consequence of the same growth in wholesale CDN and cloud platform services that has lowered entry barriers as noted above. Over time, as some of these content providers expand their business models and the range of services they offer, it is conceivable that they may expand the capabilities of their specialized CDNs, and may eventually start to offer CDN services to other content providers. Should they do so, they would confront potential channel conflict issues unless the other content providers were offering complementary content.

Fourth, access ISPs that already play a critical role in determining QoE and are feeling pressure as revenues for legacy transport services erode may seek to vertically integrate into value-added services. On the one-hand, the softwarization of ISP networks increases their capabilities to offer value-added services. Additionally, their proximity to end-users gives them a natural advantage in hosting and managing edge-located content caches (cf. Rayburn, 2014). Furthermore, to the extent they may control or manage consumer set-top boxes or provide other services to consumers, they may have additional advantages in determining how content might be cached or

---

ISPs compete for exclusivity. A CDN that is tied to a single ISP lacks the ability to cover the entire market or to play one ISP off against the other.

served to end-users (cf. Laoutaris et al. 2008; Valancius et al., 2009). In many ways, it may seem that as the basic functionality of the underlying infrastructure expands, the need for overlay CDN services would decrease. On the other hand, most ISPs do not have sufficient geographic coverage to match the needs of many content providers, and even in the markets where they do have coverage, they serve only a part of the market of end-users targeted by content providers. Nygren et al. (2010) studied the share of CDN traffic served by different ISPs and found that the largest ISP served only 5% of the traffic, while the majority of ISPs accounted for less than 1% of the CDN traffic.

Like the smaller or niche CDNs, ISPs that are seeking to vertically integrate into CDN services can enhance their offerings by interconnecting at IXPs or colocation hubs, or by joining federations to expand their geographic or market reach (horizontal integration). Such a strategy may make the most sense in markets like Europe (where there might be the need to establish Europe-wide content delivery coverage but where most of the ISPs are regional or national).

Alternatively, large CDNs may seek to partner with ISPs under licensing agreements of the sort discussed earlier. This may allow ISPs and CDNs, while retaining independence by avoiding exclusivity deals, to share capabilities in ways that are mutually beneficial. This may make sense in the U.S. where there are several competing national-scale ISPs (e.g., Verizon, Charter, Comcast, AT&T) – none of which seems especially well-positioned to supplant Akamai as the leading general-purpose CDN, but each of which may have strong incentives to expand into value-added services to leverage network upgrade investments.

As the sheer volume of cacheable content grows, rendering edge-based caching more attractive, and as the volume of interactive, dynamic and otherwise complex content also grows, the benefits of information sharing between ISPs (e.g., about real-time conditions in last-mile networks) and CDNs (e.g., about content delivery requirements and customer interests) increases. Closer integration and better information sharing between ISPs and CDNs can allow both to make more nuanced decisions about how to assign network resources and route traffic, resulting in win-win performance gains and network resource costs savings. For example, once deployed, NFV capabilities will allow ISPs to pop-up edge-servers on demand, allowing CDNs to better dynamically scale their server needs. Moreover, ISP-CDN collaborations may enable transparent caching, obviating the need for CDNs to deploy expensive middle-boxes. Today, there are tens of active ISP-CDN collaborations, e.g., between Akamai and large ISPs such as AT&T, Korea Telecom, Turkish Telecom, Deutsche Telekom, Telefonica, and Orange. The ISP and the CDN can exchange information about the state of the network and the servers available for serving requests using a variety of mechanisms. They can exchange information using the DNS, Border Gateway Protocol (BGP) or via a proprietary protocol (cf. Poese et al., 2010). ISP-CDN collaboration expands the number of available servers, and hence the granularity by which multiple content copies can be cached close to users and expands the scope for traffic engineering (which enhances the efficiency for cache management and resiliency) (cf. Frank et al., 2013).

As we continue to evolve towards next generation clouds and potentially 5G networks, the need for and benefits of collaboration are likely to increase. The vision of 5G calls for networks offering 1 msec latency capabilities which will be impossible to realize without much closer coupling of access and CDN networks. Moreover, as wireless access options become ever more important (for mobile content access), the need to supplement or replace basic Internet protocols with ones enabling cross-layer optimizations (a growing need in wireless) may further accentuate the benefits of ISP-CDN collaborations.

Summing up, we do not expect the landscape for CDNs to become less complex or to consolidate in the near future. We expect to see a continuing need for a few (one to three) global general-purpose CDN providers like Akamai continuing to expand the range of services offered by CDNs; while at the same time competing with a stream of new entrants and specialized CDNs, many of which will likely be transitory. Finally, although we do not think ISPs can fully replace CDNs, we believe that the benefits of closer collaboration of CDNs and ISPs will lead to increased vertical integration of ISPs into CDN services and partnerships and licensing agreements between CDNs and ISPs.

5. Conclusions and Directions for Future Research

In this paper, we reviewed how CDNs have evolved in conjunction with the broader trends driving changes in the Internet ecosystem. The demand from commercial content providers for more efficient (i.e., lower resource cost) and better performance network support (i.e., more consistent, reliable, and higher quality of end-user experience) for exponential growth in Internet traffic was the principal driver for the rise of CDNs. Over time, as the capabilities of other components of the ecosystem improved (i.e., better edge devices, more digital-savvy consumers, faster access and core networks), the market for value-added CDN services expanded. Over time, a complex landscape of CDN architectures and business models emerged to address these changes in the marketplace.

The diversity of CDNs is due to multiple causes, but is likely to persist. On the one hand, content providers differ widely in their needs for (and willingness or ability to pay for) content delivery services. Some content providers do not need or cannot pay for additional support, and rely on basic Internet functionality to distribute their content. At the other extreme, some content providers are either so large or have needs sufficiently specialized that self-provisioning an in-house CDN offers the lowest cost and best strategic option. In the middle are a wide range of third-party CDN providers ranging from incumbents to new entrants, regional to global providers, general-purpose or niche service providers, pursuing a range of business models. The business models typically reflect their efforts to leverage their perceived competitive advantages which in the case of incumbents includes their first-mover advantages in establishing large customer bases and global-reach CDNs. Newer entrants often seek to leverage existing assets such as datacenters or local ISP infrastructure, or to take advantage of the growing array of wholesale platform options for CDN or cloud services. Because entry costs are low, we expect competition at the lower end of the CDN market to remain intense, even if not very profitable for those so engaged. The market for full-service, general-purpose CDNs is likely to continue to be dominated by a small number (one to three) global providers like Akamai.

In light of the significant growth in content that is mostly entertainment and mostly cacheable (e.g., movie libraries that are primarily static content), there are significant compelling benefits to be realized by both CDN providers and ISPs from hosting the CDN servers inside the ISP networks. In addition to enhancing the end-user QoE (by avoiding the potential for congestion at interconnection links), it also enables the ISP to conserve interconnection capacity, which in addition to reducing investment costs also benefits the QoE for other applications. Moreover, sharing information between the CDN and ISP can allow them to jointly better manage traffic for cache updates to avoid congesting the interconnection links during peak periods.

Perhaps the greatest opportunities and challenges will arise as ISPs increasingly evolve toward cloud service providers, expanding beyond providing just basic packet transport services into providing computing and storage services, as well as other more advanced services such as security, network management, and other value-added services. At the same time, CDNs are increasingly expanding their capabilities to support more dynamic, interactive, and diverse types of content. The boundary between basic Internet functionality and value-added overlay functionality is increasingly being blurred. There are strategic and regulatory reasons, however, for keeping the ISPs and CDNs separate.

From a strategic perspective, CDNs risk channel conflict in their ability to negotiate last-mile delivery services with competing ISPs if they are too closely associated with particular ISPs. Other ISPs may prefer dealing with an independent CDN provider rather than an affiliate of a direct competitor. This may be less of an issue for ISPs in non-overlapping markets. From a regulatory perspective, the blurring of CDN and ISP business boundaries is likely to complicate efforts to regulate the provision of broadband Internet access services. For example, so-called Network Neutrality rules are focused on ensuring non-discriminatory traffic treatment for different types of content; but providing such discriminatory service is at the heart of what CDNs try to do and is a capability that many content providers can be expected to want in order to differentiate their offerings favorably from the offerings of competing content providers.

Additionally, as mobile content access becomes more important, that is likely to call for new types of CDNs. Providing support across diverse wireless access networks, operating in different portions of the radio frequency spectrum, will call for better support for cross-layer protocol optimizations.[39] Moreover, since mobile devices are often more likely to be power constrained and/or have smaller screens, there may be more call for down-coding the resolution of content or undertaking other energy saving strategies when delivering content to mobile devices.

As the Internet continues to evolve toward the Internet of Things and 5G, tighter integration of fixed and mobile, wired and wireless, transport, computing, and storage network resources is likely to be called for. In this environment, with Big Data analytic capabilities embedded in networks, there is likely to be expanded support for real-time and automated traffic management. Figuring out where critical functionality and network intelligence should reside and who should control it (end users, content providers, ISPs, or CDNs?) will pose an interesting challenge. We anticipate that these questions will not yield to simple, unitary solutions but will continue to support a diverse array of business models and CDN solutions.

[39] For example, TCP often proves sub-optimal when used over wireless links.

6. References

Ager, B., Schneider, F., Kim, J. and A. Feldmann (2010). 'Revisiting Cacheability in Times of User Generated Content', *INFOCOM IEEE Conference on Computer Communications Workshops*, pp. 1-6.

Ager, B., Chatzis, N., Feldmann, A., Sarrar, N., Uhlig, S. and W. Willinger (2012). 'Anatomy of a Large European IXP', *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication (SIGCOMM '12)*, pp. 163-174.

Akamai (2016a). 'Akamai Facts and Figures', available at: https://www.akamai.com/us/en/about/facts-figures.jsp [accessed: 20 Aug 2016].

Akamai (2016b). 'AURA Licensed CDN', available at: https://www.akamai.com/us/en/solutions/products/network-operator/licensed-cdn-solutions.jsp [acessed: 23 Aug 2016].

Akamai (2016c). Annual Report 2015, February, available at: http://www.ir.akamai.com/phoenix.zhtml?c=75943&p=irol-reportsannual [accessed: 23 Aug 2016].

Akamai (2016d). '{Open} – Get Started Building the Future of the Internet', available at: https://developer.akamai.com [accessed: 31 Aug 2016].

Akamai (2016e). 'State of the Internet - Internet Connection Speeds and Adoption Rates by Geography', available at: https://www.akamai.com/us/en/our-thinking/state-of-the-internet-report/state-of-the-internet-connectivity-visualization.jsp [accessed: 21 Sep 2016].

Amazon Web Services (2016). 'Amazon CloudFront', available at: https://aws.amazon.com/cloudfront/details/ [accessed: 20 Aug 2016].

Amsterdam Internet Exchange (AMS-IX) (2016). 'Connect Remotely', available at: https://ams-ix.net/connect-to-ams-ix/connect-remotely [accessed: 24 Aug 2016].

Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A., Lee, G., Patterson, D.A., Rabkin, A., Stoica, I. and M. Zaharia (2009). 'Above the Clouds: A Berkeley View of Cloud Computing', *UC Berkeley Technical Report EECS-2009-28*, February.

AT&T (2016). 'U.S. Network Latency', available at: http://ipnetwork.bgtmo.ip.att.net/pws/network_delay.html [accessed: 21 Sep 2016].

Body of European Regulators for Electronic Communications (BEREC) (2014). 'Monitoring quality of Internet access services in the context of net neutrality', *BoR,* (14)117.

Boorstin, J. (2015). ‘BitTorrent’s original content push’, *CNBC.com*, 11 February, available at: http://www.cnbc.com/2015/02/11/bittorrents-original-content-push.html [accessed: 23 Aug 2016].

Boyle, B. (2016). ‘Equinix closes Telecity Acquisition’, *Datacenterknowledge.com*, 18 January, available at: http://www.datacenterdynamics.com/content-tracks/colo-cloud/equinix-closes-telecity-acquisition/95540.fullarticle [accessed: 23 Aug 2016].

Briscoe, B. (2007). ‘Flow Rate Fairness: Dismantling a Religion’, *ACM SIGCOMM Computer Communication Review*, 37(2), 65-74.

Brynjolfsson, E., Hu, Y. and M.D. Smith (2003). ‘Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers’, *Management Science*, 49(11), pp. 1580-1596.

BusinessWire (2012). ‘EdgeCast Launches Licensed CDN Program Foloowing Global Deployment of Two Major Carrier CDNs’, 9 January, available at: http://www.businesswire.com/news/home/20120109006151/en/EdgeCast-Launches-Licensed-CDN-Program-Global-Deployment [accessed: 31 Aug 2016].

Calder, M., Fan, X., Hu, Z., Govindan, R., Heidemann, J. and E. Katz-Bassett (2013). ‘Mapping the Expansion of Google’s Serving Infrastructure’, *Proceedings of the 2013 conference on Internet measurement conference (IMC ’13)*, pp. 313-326.

Castro, J., Cardona, J.C., Gorinsky, S. and P. Francois (2014). ‘Remote Peering: More Peering without Internet Flattening’, *Proceedings of the 10th ACM Conference on Emerging Networking EXperiments and Technologies (CoNEXT ’14),* pp. 185-198.

Cedexis (2016). ‘Cedexis’, available at: http://www.cedexis.com [accessed: 20 Aug 2016].

Chatzis, N., Smaragdakis, G., Feldmann, A. and W. Willinger (2013). ‘There is more to IXPs than meets the Eye’, *ACM SIGCOMM Computer Communication Review*, 43(5), pp. 19-28.

Chatzis, N., Smaragdakis, G., Feldmann, A. and W. Willinger (2015). ‘Quo vadis Open-IX? Trying to boost public peering in the US’, *ACM SIGCOMM Computer Communication Review*, 45(1), pp. 12-18.

Chiu, Y.-C., Schlinker, B., Radhakrishnan, A.B., Katz-Bassett, E. and R. Govindan (2015). ’Are We One Hop Away from a Better Internet?’, *Proceedings of the 2015 ACM Conference on Internet Measurement Conference (IMC ’15)*, pp. 523-529.

Cisco (2001). ’IP Multicast Technology Overview’, available at: http://www.cisco.com/c/en/us/td/docs/ios/solutions_docs/ip_multicast/White_papers/mcst_ovr.html [accessed: 31 Aug 2016].

Cisco (2005). 'DiffServ – The Scalable End-to-End Quality of Service Model', *Whitepaper*, August, available at: http://www.cisco.com/en/US/technologies/tk543/tk766/technologies_white_paper09186a00800a3e2f.pdf [accessed: 31 Aug 2016].

Cisco (2016). Cisco Visual Networking Index: Forecast and Methodology, 2015-2020, June, available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html [accessed: 20 Aug 2016].

Clancy, M. (2016). 'Netflix moves all Global Traffic to Open Connect CDN', *RaidTVNews*, 19 March, available at http://www.rapidtvnews.com/2016031942170/netflix-moves-all-global-traffic-to-open-connect-cdn.html#axzz4IYunwH4Q [accessed: 28 Aug 2016].

Clark, D.D., Lehr, W.H., Bauer, S., Faratin, P., Sami, R. and J. Wroclawski (2006). 'Overlay Networks and the Future of the Internet', *Communications and Strategies*, 63(3), pp. 109–129.

Clark, D.D., Lehr, W.H. and S. Bauer (2011). 'Interconnection in the Internet: the Policy Challenge', *TPRC 39: The 39th Research Conference on Communication, Information and Internet Policy Paper*, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1992641 [accessed: 20 Aug 2016].

CloudFlare (2016). 'Plans', available at: https://www.cloudflare.com/plans/ [accessed: 30 Aug 2016].

DE-CIX (2016). 'The DE-CIX Blackholing Service', available at: https://www.de-cix.net/products-services/blackholing/ [accessed: 24 Aug 2016].

Dietzel, C., Feldmann, A. and T. King (2016). 'Blackholing at IXPs: On the Effectiveness of DDoS Mitigation in the Wild', in Karagiannis, T. and X. Dimitropoulos, eds, *Passive and Active Measurement: 17th International Conference PAM 2016 - Proceedings*, Cham et al.: Springer, pp. 319-332.

Dilley, J., Maggs, B., Parikh, J., Prokop, H., Sitaraman, R. and B. Weihl (2002). 'Globally Distributed Content Delivery', *IEEE Internet Computing*, 6(5), pp. 50–58.

Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., Zhan, J. and H. Zhang (2011). 'Understanding the Impact of Video Quality on User Engagement', *Proceedings of the ACM SIGCOMM 2011 conference*, pp. 362-373.

Durairajan, R., Barford, B., Sommers, J. and W. Willinger (2015). 'InterTubes: A Study of the US Long-haul Fiber-optic Infrastructure', *ACM SIGCOMM Computer Communication Review - SIGCOMM'15*, 45(4), pp. 565-578.

Economides, N. and J. Tåg (2012). 'Network Neutrality on the Internet: A Two-Sided Market Analysis', *Information Economics and Policy*, 24(2), pp. 91-104.

Equinix (2016). ‘Annual Report FY 2015’, available at: http://investor.equinix.com/phoenix.zhtml?c=122662&p=irol-IRHome [accessed: 31 Aug 2016].

Erman, J., Gerber, A., Hajiaghayi, M., Pei, D. and O. Spatscheck (2009). ‘Network-aware forward caching’. *Proceedings of the 18th international conference on World wide web (WWW ’09)*, pp. 291-300.

European Commission (2015). ‘Quality of Broadband Services in the EU’, published on October 22, 2015, available at: https://ec.europa.eu/digital-single-market/en/news/quality-broadband-services-eu [accessed: 21 Sep 2016].

European Telecommunications Standards Institute (ETSI) (2012). ‘Network Functions Virtualisation – An Introduction, Benefits, Enablers, Challengers & Call for Action’, *Introductory White paper*, available at: https://portal.etsi.org/nfv/nfv_white_paper.pdf [accessed: 20 Aug 2016].

Evans, D. S. and R. Schmalensee (2007). ‘The Industrial Organization of Markets with Two-Sided Platforms’, *Competition Policy International*, 3(1), pp. 151-179.

Faratin, P., Clark, D.D., Bauer, S., Lehr, W.H., Gilmore, P. and A. Berger (2008). ‘The Growing Complexity of Internet Interconnection’, *Communications & Strategies*, 72(4), 51-71.

Federal Communications Commission (FCC) (2015). ‘2015 Measuring Broadband America Fixed Broadband Report’, published on December 30, 2015, available at: https://www.fcc.gov/reports-research/reports/measuring-broadband-america/measuring-broadband-america-2015 [accessed: 22 Sep 2016].

Feldmann, A., Gladisch, A., Kind, M., Lange, C., Smaragdakis, G. and F.-J. Westphal (2010). ‘Energy Trade-offs among Content Delivery Architectures’, *9th Conference on Telecommunications Internet and Media Techno Economics (CTTE)*, pp. 1-6.

Flach, T., Dukkipati, N., Terzis, A., Raghavan, B., Cardwell, N., Cheng, Y., Jain, A., Hao, S., Katz-Bassett, E. and R. Govindan (2013). ‘Reducing Web Latency: the Virtue of Gentle Aggression’. *ACM SIGCOMM Computer Communication Review*, 43(4), pp. 159-170.

Frank, B., Poese, I., Lin, Y., Smaragdakis, G., Feldmann, A., Maggs, B., Rake, J., Uhlig, S. and R. Weber (2013). ‘Pushing CDN-ISP Collaboration to the Limit’, *ACM SIGCOMM Computer Communication Review*, 43(3), pp. 34-44.

Freedman, M.J., Vutukuru, M., Feamster, N. and H. Balakrishnan (2005). ‘Geographic Locality of IP Prefixes’, *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement (IMC ’05)*, pp. 153-158.

Galperin, H. (2016). ‘Localizing Internet Infrastructure: Cooperative Peering in Latin America’, *Telematics and Informatics*, 33(2), pp. 631- 640.

Gates, B. (1996). ‘Content is King’, *Microsoft Essay*, 1/3/96, available at: http://www.craigbailey.net/content-is-king-by-bill-gates/ [accessed: 30.08.2016].

Gerber, A. and R. Doverspike (2011). ‘Traffic Types and Growth in Backbone Networks’, *Optical Fiber Communication Conference and Exposition (OFC/NFOEC) and the National Fiber Optic Engineers Conference*, pp. 1-3.

Giotsas, V., Zhou, S., Luckie, M. and kc claffy (2013). ‘Inferring Multilateral Peering’, *Proceedings of the 9th ACM Conference on Emerging Networking EXperiments and Technologies (CoNEXT ’13)*, pp. 247-258.

Giotsas, V., Luckie, M., Huffaker, B. and kc claffy (2014). ‘Inferring Complex AS Relationships’, *Proceedings of the 2014 Conference on Internet Measurement Conference (IMC ’14)*, pp. 23-30.

Giotsas, V., Smaragdakis, G., Huffaker, B., Luckie, M. and kc claffy (2015). ‘Mapping Peering Interconnections to a Facility’, *ACM SIGCOMM Conference on emerging Networking EXperiments and Technologies (CoNEXT)*.

Google (2016). ‘Peering’, available at: https://peering.google.com/#/options/peering [accessed: 31 Aug 2016].

Greenberg, A., Hamilton, J., Maltz D.A. and P. Patel (2009). ‘The Cost of a Cloud: Research Problems in Data Center Networks’, *ACM SIGCOMM Computer Communication Review*, 39(1), pp. 68-73.

Gupta, A., Calder, M., Feamster, N., Chetty, M., Calandro, E. and E. Katz-Bassett (2014a). ‘Peering at the Internet’s Frontier: A First Look at ISP Interconnectivity in Africa’, in Faloutsos, M. and A. Kuzmanovic, eds, *Passive and Active Measurement: 15th International Conference PAM 2014 - Proceedings*, Cham et al.: Springer, pp. 204-213.

Gupta, A., Vanbever, L., Shahbaz, M., Donovan, S.P., Schlinker, B., Feamster, N., Rexford, J., Shenker, S., Clark, R. and E. Katz-Bassett (2014b). ‘SDX: A Software Defined Internet Exchange’, *ACM SIGCOMM Computer Communication Review – SIGCOMM ’14*, 44(4), pp. 551-562.

Hannigan, M. (2015). ‘Peering with Akamai’, *CaribNOG 10*, Belize City, Belize, November 2-6, available at: https://static1.squarespace.com/static/5138f2cae4b0f3422dd43abd/t/563a8549e4b0a7d393c4b674/1446675785424/CaribNOG10_Day_3_Peering_Martin_Hannigan_Akamai.pdf [accessed: 31 Aug 2016].

Hjembo, J. (2016). ‘Global Colocation Market Trends’, *Digital Realty/TeleGeography Joint Webinar*, April, available at: http://e27e3a335301d5434f3f-2ad8a26a73e5b573a6c0e67440b8c173.r11.cf5.rackcdn.com/slide-decks/2016/Telx-TGcolo%20FINAL.pdf [accessed: 31 Aug 2016].

Kohavi, R. Henne, R.M. and D. Sommerfield (2007). ‘Practical Guide to Controlled Experiments on the Web: Listen to Your Customers not to the HiPPO’, *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 959-967.

Krishnan, R., Madhyastha, H.V., Srinivasan, S., Jain, S., Krishnamurthy, A., Anderson, T. and J. Gao (2009). ‘Moving Beyond End-to-End Path Information to Optimize CDN Performance’, *Proceedings of the 2009 Conference on Internet Measurement Conference (IMC ’09)*, pp. 190-201.

Krishnan, S.S. and R.K. Sitaraman (2013). ‘Video Stream Quality Impacts Viewer Behavior: Inferring Causality using Quasi-Experimental Designs’, *IEEE/ACM Transactions on Networking*, 21(6), pp. 2001-2014.

Labovitz, C., Lekel-Johnson, S., McPherson, D., Oberheide, J. and F. Jahanian (2010). ‘Internet Inter-Domain Traffic’, *Proceedings of the ACM SIGCOMM 2010 conference*, pp. 75-86.

Laoutaris, N., Rodriguez, P. and L. Massoulie (2008). ‘ECHOS: Edge Capacity Hosting Overlays of Nano Data Centers’, *ACM SIGCOMM Computer Communication Review*, 38(1), pp. 51-54.

Leighton, T. (2009). ‘Improving Performance on the Internet’, *Communications of the ACM*, 52(2), pp. 44-51.

Lehr, W. and D. Sicker (2016). ‘Do you want your Internet with or without Video?’, *TPRC 44: The 44th Research Conference on Communication, Information and Internet Policy Paper*, available at SSRN: http://ssrn.com/abstract=2757374 or http://dx.doi.org/10.2139/ssrn.2757374.

Limelight Networks (2016). ‘ASN #22822’, available at: https://www.limelight.com/network/ [accessed: 20 Aug 2016].

Liu, X., Dobrian, F., Milner, H., Jiang, J., Sekar, V., Stoica, I. and H. Zhang (2012). ‘A Case for a Coordinated Internet Video Control Plane’, *Proceedings of the ACM SIGCOMM 2012 conference on Applications, technologies, architectures, and protocols for computer communication*, pp. 359-370.

Lodhi, A., Larson, N., Dhamdhere, A., Dovrolis, C. and kc Claffy (2014). ‘Using PeeringDB to Understand the Peering Ecosystem’, *ACM SIGCOMM Computer Communication Review*, 44(2), pp. 20-27.

Luckie, M., Dhamdhere, A., Clark, D.D., Huffaker, B. and kc claffy (2014). ‘Challenges in Inferring Internet Interdomain Congestion’, *Proceedings of the 2014 Conference on Internet Measurement Conference (IMC ’14)*, pp. 15-22.

Maggs, B. and R. K. Sitaraman (2015). ‘Algorithmic Nuggets in Content Delivery’. *ACM SIGCOMM Computer Communication Review*, 45(3), pp. 52-66.

McMillan, R. (2013). ‘Google handles 25 percent of North America’s Web Traffic’, *Wired Magazine*, 22 July, available at: http://www.wired.com/2013/07/google-internet-traffic/ [accessed: 28 Aug 2016]

Metzler, J. (2015). ‘The 2015 Guide to SDN and NFV’, January, available at: https://www.a10networks.com/sites/default/files/resource-files/2015Ebook-A10-all.pdf [accessed: 20 Aug 2016].

Miller, R. (2016). ‘Mapping Netflix: Content Delivery Network Spans 233 Sites’, *Data Center Frontier*, 19 August, available at http://datacenterfrontier.com/mapping-netflix-content-delivery-network/ [accessed: 28 Aug 2016].

Netflix (2016a). ‘Netflix Open Connect’, available at: https://signup.netflix.com/openconnect [accessed: 20 Aug 2016].

Netflix (2016b). ‘How Netflix Works with ISPs Around the Globe to Deliver a Great Viewing Experience’, 17 March, available at: https://media.netflix.com/en/company-blog/how-netflix-works-with-isps-around-the-globe-to-deliver-a-great-viewing-experience [accessed: 22 Aug 2016].

Netflix (2016c). ‘Can I Stream Netflix in Ultra HD?’, available at: https://help.netflix.com/en/node/13444 [accessed: 31 Aug 2016].

Niven-Jenkins, B., Le Faucheur, F. and Bitar, N. (2012). Content Distribution Network Interconnection (CDNI) Problem Statement, RFC 6707.

Norton, W.B. (2014). *The Internet Peering Playbook: Connecting to the Core of the Internet*, DrPeering Press.

Nygren, E., Sitaraman, R. K. and J. Sun (2010). ‘The Akamai Network: A Platform for High-performance Internet Applications’, *ACM SIGOPS Operating Systems Review*, 44(3), pp. 2-19.

Ofcom (2016). ‘UK Home Broadband Performance - The Performance of Fixed-line Broadband Delivered to UK Residential Consumers’, available at: http://stakeholders.ofcom.org.uk/market-data-research/other/telecoms-research/broadband-speeds/UK-home-broadband-performance-Nov-15 [accessed: 21 Sep 2016].

Poese, I., Frank, B., Ager, B., Smaragdakis, G. and A. Feldmann (2010). ‘Improving Content Delivery using Provider-Aided Distance Information’, *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, pp. 22-34.

Poese, I., Frank, B., Smaragdakis, G., Uhlig, S., Feldmann, A. and B. Maggs (2012). 'Enabling Content-aware Traffic Engineering'. *ACM SIGCOMM Computer Communication Review*, 42(5), pp. 21-28.

Popa, L., Ghodsi, A. and Stoica, I. (2010). 'HTTP as the Narrow Waist of the Future Internet', *Hotnets-IX Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks*, Article No. 6.

Pujol, E., Richter, P., Chandrasekaran, B., Smaragdakis, G., Feldmann, A., Maggs, B.M., and K.-C. Ng (2014). 'Back-Office Web Traffic on the Internet', *Proceedings of the 2014 ACM Conference on Internet Measurement Conference (IMC '14)*, pp. 257-270.

Qureshi, A., Weber, R., Balakrishnan, H., Guttag, J. and B. Maggs (2009). 'Cutting the Electric Bill for Internet-scale Systems', *ACM SIGCOMM Computer Communication Review*, 39(4), pp. 123-134.

Rayburn, D. (2014). 'Comcast launches Commercial CDN Service allowing Content Owners to Deliver Content via the Last Mile', 19 May, available at: http://blog.streamingmedia.com/2014/05/comcast-launches-commercial-cdn-service-allowing-content-owners-to-deliver- content-inside-the-last-mile.html [accessed: 20 Aug 2016].

Reed, D.P., Warbritton, D. and D. Sicker (2014). 'Current Trends and Controversies in Internet Peering and Transit: Implications for the Future Evolution of the Internet', *TPRC 42: The 42nd Research Conference on Communication, Information and Internet Policy Paper*, available at: http://ssrn.com/abstract=2418770 [accessed: 20 Aug 2016].

Richter, P., Smaragdakis, G., Feldmann, A., Chatzis, N., Boettger, J. and W. Willinger (2014). 'Peering at Peerings: On the Role of IXP Route Servers', *Proceedings of the 2014 Conference on Internet Measurement Conference (IMC '14)*, pp. 31-44.

Richter, P., Chatzis, N., Smaragdakis, G., Feldmann, A. and W. Willinger (2015). 'Distilling the Internet's Application Mix from Packet-Sampled Traffic', in Mirkovic, J. and Y. Liu, eds., *Passive and Active Measurement: 16th International Conference, PAM 2015, Proceedings*, Cham et al.: Springer, pp. 179-192.

Rochet, J.-C. and J. Tirole (2006). 'Two-Sided Markets: A Progress Report,' *RAND Journal of Economics*, 47(3), pp. 645-667.

Saltzer, J.H., Reed, D.P. and D.D. Clark (1984). 'End-to-end arguments in System Design', *ACM Transactions on Computer Systems*, 2(4), pp. 277-288.

Sandvine (2016). 'Global Internet phenomena: Latin America & North America', June.

Singla, A., Chandrasekaran, B., Godfrey, P.B. and B. Maggs (2014). 'The Internet at the Speed of Light', *Proceedings of the 13th ACM Workshop on Hot Topics in Networks*, October 27-28, Los Angeles, CA, USA.

Spring, N., Mahajan, R. and T. Anderson (2003). 'The Causes of Path Inflation', *SIGCOMM '03: Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, pp. 113-124.

Stocker, V. (2015). 'Interconnection and Capacity Allocation for All-IP Networks: Walled Gardens or Full Integration?', *TPRC 43: The 43rd Research Conference on Communication, Information and Internet Policy Paper*, available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2587833 [accessed: 20 Aug 2016].

Streibelt, F., Boettger, J., Chatzis, N., Smaragdakis, G. and A. Feldmann (2013). 'Exploring EDNS-Client-Subnet Adopters in your Free Time', *Proceedings of the 2013 conference on Internet measurement conference (IMC '13)*, pp. 305-312.

Stronge, T. (2015). 'International Network Supply, Traffic and Pricing', *NANOG On the Road*, Herndon, VA, USA, June 23, available at: https://www.nanog.org/sites/default/files/06-Stronge.pdf [accessed: 31 Aug 2016].

Sverdlik, Y. (2015). 'Digital Realty Closes $1.9B TELX Acquisition', *Datacenterknowledge.com*, 12 October, available at: http://www.datacenterknowledge.com/archives/2015/10/12/digital-realty-closes-1-9b-telx-acquisition/ [accessed: 23 Aug 2016].

Telefonica (2014). 'SDN-NFV: An Introduction', October, available at http://www.it.uc3m.es/fvalera/t2/5.pdf [accessed: 20 Aug 2016].

Telegeography (2016). 'Telecom Resources', available at: https://www.telegeography.com/ [accessed: 20 Aug 2016]

Temkin, D. (2016). 'The Real Cost of Public IXPs', *NANOG 67*, Chicago, Il, USA, June 12-15, available at: https://www.nanog.org/sites/default/files/Temkin_The_Real_Cost.pdf [accessed: 20 Aug 2016].

Valancius, V., Laoutaris, N., Massoulie, L., Diot, C. and P. Rodriguez (2009). 'Greening the Internet with Nano Data Centers', *Proceedings of the 5th ACM Conference on Emerging Networking EXperiments and Technologies (CoNEXT '09),* pp. 37-48.

Verizon (2016). 'Verizon ACCELERATE', available at: https://www.verizondigitalmedia.com/platform/accelerate/ [accessed: 31 Aug 2016].

Winstein, K. and H. Balakrishnan (2013). 'TCP ex Machina: Computer-Generated Congestion Control', *ACM SIGCOMM Computer Communication Review*, 43(4), pp. 123-134.

Zhao, M., Aditya, P., Chen, A., Lin, Y., Haeberlen, A., Druschel, P., Maggs, B., Wishon, B. and M. Ponec (2013). 'Peer-assisted Content Distribution in Akamai NetSession', *Proceedings of the 2013 conference on Internet measurement conference (IMC '13)*, pp. 31-42.