Volker Stocker, Georgios Smaragdakis, William Lehr, Steven Bauer

The growing complexity of content delivery networks: Challenges and implications for the Internet ecosystem

WISSEN IM ZENTRUM
UNIVERSITÄTSBIBLIOTHEK

Technische
Universität
Berlin

The Growing Complexity of Content Delivery Networks: Challenges and Implications for the Internet Ecosystem

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

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 their relationship to and implications for interconnection markets.

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

---

The Growing Complexity of Content Delivery Networks: Challenges and Implications for the Internet Ecosystem

1. Introduction

Content Delivery Networks (CDNs) emerged as overlay networks on the Internet in order to provide better support for delivering commercial content than was available using basic, "best-effort" Internet packet transport services. As the volume, complexity, and heterogeneity of Internet traffic has grown and evolved, so too have the ISPs and CDNs that provide the services used to deliver this traffic. The importance of CDNs within the Internet ecosystem has grown significantly over time – recent reports expect that CDNs will soon be handling over half of the global traffic on the Internet (cf. Cisco, 2016, p. 18).

In this paper, we describe the current CDN ecosystem and the forces that have driven its evolution. Section 2 provides an overview of the basic operation of CDNs and recounts how they have evolved since their market inception in the 1990s. Section 3 then presents a typology that sets forth the multiple types of CDNs and their relative strengths and weaknesses. Section 4 focuses on how CDN choices of where to locate their servers impact their operations, and Section 5 offers our speculations on where we see the CDN ecosystem going. Section 6 concludes and offers some thoughts and implications for the future Internet ecosystem.

2. On the Evolution of Content Delivery Networks

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. The basic protocols like the Transmission Control Protocol (TCP) that manage packet transmission on an end-to-end basis seek to provide decentralized congestion management and fair resource allocations across competing flows; however, these protocols fail to support the predictable, end-to-end Quality of Experience (QoE)[1] that is increasingly being demanded by commercial content and application providers.[2] Moreover, networks offering only a single service class are not well-suited for supporting the heterogeneous needs of traffic types with very different Quality of Service (QoS) requirements.

---

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

[2] For example, TCP performance degrades as the distance between communicating hosts increases and in the presence of packet losses along the end-to-end path (cf. Leighton, 2009). Also, TCP's flow-rate fairness principal may be exploited by applications (e.g., using swarming techniques to simultaneously operate multiple TCP connections) resulting in some applications capturing a disproportionate share of available capacity, while leaving other applications starved of capacity (cf. e.g., Briscoe, 2007).

While attempts have been made to expand the set of Internet protocols to enable better support for heterogeneous QoS requirements,[3] modifying the basic Internet infrastructure is difficult because it requires coordinating the widespread adoption of new protocols. In the Internet, control is decentralized and dispersed among many different and often competing entities. Although individual networks may deploy enhanced capabilities for active traffic management of on-net traffic, widespread inter-provider deployments are not generally available (cf. e.g., Stocker, 2015). Content Delivery Networks (CDNs) evolved to address the need for better support for differentiated content distribution without requiring modifications to the basic Internet architecture.

2.1. A Primer on CDN Service Provision

CDNs employ a scalable architecture of cache servers which are strategically distributed across the Internet and constitute an overlay "on top" of the Internet's basic packet transport infrastructure.[4] A typical CDN maintains a hierarchy of servers, with back-end servers ensuring the efficient intra-CDN distribution of content, and front-end servers at the edges used for handling user-server communications. These servers allow replicated copies of content to be stored at multiple locations across the Internet.

CDN providers employ complex software to match incoming requests for content to the "best" server for meeting each end-user request.[5] Requests flow from end-users to the selected front-end server, which delivers a copy of the requested content to the end-user, if a copy is already available on the server. If not, the front-end server passes the request further up the CDN-server hierarchy until a copy is located, which may entail pulling a copy from the content provider's

---

[3] Over time, the Internet suite of protocols have expanded to include better support for content delivery and other types of QoS differentiated packet transport services. For example, multicast enhances content delivery capabilities by allowing 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). Alternatively, newer protocols like DiffServ and IntServ provide support for better-than-best-effort packet delivery (cf. Cisco, 2005). Although such capabilities already exist in the Internet, they are neither uniformly available nor implemented in a standard way across ISPs which complicates efforts to make use of them.

[4] Cache servers are dedicated servers used to store content for subsequent delivery. In general, one can distinguish between three general types of cache servers involved in CDN service provisioning: front-end, back-end, and origin servers. The "origin server" refers to the server where the original or master copy of the content is stored. A key role for CDNs is to place additional copies of that content on other cache servers elsewhere in the Internet to improve content delivery services. The "front-end" servers are the ones that end-users communicate with; and these may communicate with appropriate "back-end" servers that are located within the CDN network and provide additional CDN functionality as we explain. For further details on CDN servers, see Leighton (2009), Nygren et al. (2010), or Flach et al. (2013).

[5] Determining the "best" server may depend on different metrics in the quality/cost-space in which different delivery features may be emphasized. As will be described in Section 4, the location of the server is important, and "location" may be defined in multiple ways. While in the case of CDNs with a network of dedicated servers, redirection decisions are typically done via the Domain Name System (DNS), distributed naming schemes are used in peer-to-peer CDNs.

origin server if a closer copy is not available (cf. Nygren et al., 2010). In this way, content is replicated and distributed across the CDN's footprint of servers. 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 (i.e., static versus dynamic),[6] the preferences of the content provider, end-user demand for the content (i.e., content popularity), what else is going on in the Internet (e.g., congested links), and the capabilities of the CDN provider. Further complications arise if origin servers are located at different content providers (e.g., as might be the case for advertisements and background text), or specialized security or stringent QoS requirements necessitate specialized routing treatment by the CDN provider.

Replicating content in multiple locations and jointly managing multiple servers in real-time, allows CDN providers to better balance server loads, which enhances the utilization efficiency of server capacity and capacity scaling, lowering content delivery resource costs and enabling CDNs to provide improved QoE performance for end-users. It helps CDNs better respond to flash crowds and denial of service attacks (cf. e.g., Nygren et al., 2010; Maggs and Sitaraman, 2015). When the content is available from multiple locations, single points of failure are eliminated, which improves reliability. The ability to match content requests with the "best" server for each request helps content providers reduce end-to-end delays and ensure a more consistent and higher QoE for their end-users (cf. Dilley et al., 2002; Chiu et al., 2015). CDN providers make use of proprietary protocols to distribute content across their server network, augmenting the basic Internet's capabilities. Content can be handled differentially on a per-request basis, without requiring special software or modifications to the underlying Internet infrastructure.[7]

Although CDNs can do much to enhance the performance of content delivery, their end-to-end control is limited by their dependence on the underlying ISPs for basic packet transport services. Most CDNs do not operate their own packet data networks or lease dedicated facilities to interconnect their servers, and rely on the ISPs to make packet routing decisions. The CDN overlay networks and the ISPs both have a role in determining the end-to-end QoE for content delivery, and the extent to which the CDN providers and ISPs coordinate their traffic management decisions can have a big impact on the overall QoE.

2.2. Innovation and the Cost of CDN Service Provision

When CDNs began to emerge in the mid to late 1990s, bandwidth prices for packet transit were expensive and accounted for a much larger share of the total costs of content delivery than they

---

[6] Dynamic content, as opposed to static content, needs to be frequently updated, and as a consequence may not be cacheable. For example, a telephone call or data associated with a highly interactive website (e.g., for gaming or providing real-time stock price data) may not be cacheable. Other content like magazine stories, movies, picture files, or software updates are static content that, once stored in a CDN's cache, does not require updating. For further discussion on the challenges of supporting dynamic content, see, for example, Leighton (2009, pp. 4f.) or Nygren et al. (2010, pp. 9ff).

[7] For example, CDNs can 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. See, for example, Nygren et al. (2010), or Winstein and Balakrishnan (2013).

do today. The costs of acquiring and maintaining servers (including energy to power and cool servers) and CDN personnel costs were less significant. Consequently, minimizing bandwidth costs for packet transit was the chief consideration for most CDN providers. This has changed over time. Technical innovation, changing market competition dynamics and large private and public investments in fiber networks and transatlantic links to expand transport capacity have contributed to steep reductions in the cost of bandwidth. Further, the growth and increasing importance of direct interconnections based on settlement-free (or, revenue-neutral) peering agreements (cf. e.g., Faratin et al., 2008; Clark et al., 2011) and the emergence and growth of Internet Exchange Points (IXPs) (cf. Chatzis et al. 2013) and other interconnection facilities (e.g., Equinix) which provide platforms for interconnecting multiple networks (cf. Labovitz et al., 2010) have helped CDNs and access ISPs avoid paying for transit services. These developments have contributed to a seismic shift in how the Internet is interconnected.

Although this has further amplified the general trend of declining unit prices for bandwidth around the world, significant price differences between regions persist. For example, 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). Finally, bandwidth pricing is typically subject to significant volume discounting. As a consequence, managing regional and volume-based bandwidth resources remains a key component of the CDN value proposition.

Due to other developments such as significant improvements in computing and communications technologies in accordance with Moore's Law and corresponding drops in the cost of servers and storage, the cost structure of CDNs has changed (cf. Armbrust et al., 2009). Over time, other cost components such as payments for hosting services (e.g., to the ISPs or IXPs providing physical locations where servers are hosted) and energy have become relatively more important CDN operating cost components. Depending on the CDN architecture, CDNs may be able to reduce costs by strategically locating services to take advantage of hosting and energy price differences across markets (cf. e.g., Qureshi et al., 2009; Greenberg et al., 2009).

At the same time that CDN costs have been changing, customer demand has become more differentiated. In addition to the continuing demand for the distribution of static content, content providers and their customers have growing needs to serve a broader array of more dynamic and interactive content. To meet this increasingly diverse 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. With the migration to everything-over-IP, the rise of new types of media is straining legacy notions of what constitutes traditional media content distribution. 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, concerts) and breaking news requires (near) real-time delivery. 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.

Meeting market demand for support for heterogeneous content with quite different QoS delivery requirements has important implications for CDN service provisioning and its costs. One-size-fits-all solutions are not cost-effective – either they fail to meet the requirements of the content with specialized QoS reqirements or over-provision capabilities for content that does not require those specialized services. Some CDNs are exploring niche strategies, focusing on specific types of content, specific architectures, or specific markets to better balance specialized requirements and costs with customized solutions. There are also general-purpose CDNs that offer a portfolio of differentiated services to address the needs of 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

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 in which end-users may wish to access 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, additional complementary services have become integral components in the service offerings of many commercial CDN providers. These include, for example, support for enhanced security, additional cloud functionalities (e.g., video-coding on the fly), and support for market metrics and analytics. These complementary services have become more important as the marketplace for basic content delivery services has become more competitive and diverse. 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). The growth in rich multimedia traffic, in particular entertainment video, has contributed significantly to the rise of CDNs, which in turn has helped propel further growth in over-the-top (OTT) entertainment services such as Spotify and Netflix, that drive still further demand for CDNs (cf. Lehr and Sicker, 2016). The resulting growth of the CDN ecosystem confronts content providers with a menu of options for how best to deliver content to their customers.

Some content providers with limited budgets, small audiences, or with content that can tolerate lower quality delivery performance may select approaches like centralized hosting, web-based delivery via platforms like Facebook, or free but limited CDN services for delivering their content. Larger and predominantly commercial content providers may choose to avail themselves of the services of general-purpose CDN providers like Akamai or Limelight, selecting service packages that are customized to their specific needs (cf. e.g., Akamai, 2016d). This middle-market of commercial content providers is the heart of the market for third-party CDN providers. Commercial entertainment providers confront intense competition for the attention (i.e., advertising supported) and discretionary spending (e.g., for paid entertainment content) of their target audiences and may thus consider faster and more reliable content delivery as a competitive

advantage. Many engage in "windowing" strategies by which they seek to segment their target audience in order to facilitate price discrimination to help them increase revenues.[8] At the upper end of the market, the larger content providers continually evaluate their make-vs-buy decisions, and the largest providers like Netflix and Google (YouTube) have outgrown the capabilities of third-party CDNs and have deployed their own specialized CDNs.

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 traditionally focused on generating revenues from selling bandwidth to support basic packet transport services have seen their transit revenues erode in the face of falling transit prices and restructured interconnection agreements (i.e., in particular the trend to direct peering). As a response, a number of ISPs are exploring their options for offering value-added services, and providing CDN services is one possibility. Many of the largest operators of full-service networks such as Telefonica, AT&T, and Telecom Italia, are investing heavily in Network Function Virtualization (NFV) and Software Defined Networking (SDN) technologies to "softwarize" their networks. These technologies enable finer-grained, flexible, and dynamic control over network resources. ISPs costs are lowered because operators can shift to commodity hardware; and because network functionality can be de-localized. For example, 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 co-located. Network infrastructure can be virtualized and then sliced into logical partitions that may support different levels of QoS and network 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).[9] With these new capabilities, these next generation ISPs are able to offer substitutes for much of the 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.

---

[8] Historically, the order of public release was used to implement price discrimination "windows". For example, 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).

[9] Providing wholesale access to basic computing, storage, and transport functionality is sometimes referred to as *Infrastructure-as-a-Service*; providing application or content hosting services is 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. For discussions of NFV and SDN, see Telefonica (2014), ETSI (2012), or Metzler (2015).

3. A Typology of CDN Architectures

As was illustrated in previous sections, the CDN ecosystem has grown increasingly complex and consists of a wide array of different CDN architectures and business models. In the following, we provide a typology of CDN architectures currently deployed in the Internet. Key features that distinguish the different CDN strategies include the strategy for deploying servers (e.g., the degree of distribution) and the business models employed (e.g., what performance characteristics are emphasized or what key applications are the market focus). Table 1 summarizes our findings. The first four types refer to multi-purpose CDNs based on different server deployment strategies, while the fifth type refers to specialized, single-purpose CDNs. The last three types describe new business models for delivering CDN services based on combining the resources of multiple market players.

[Insert Table 1 about here]

3.1. Datacenter-based CDNs

Some CDN operators concentrate large numbers of servers (i.e., server clusters or server farms) in a relatively small number of geographic locations. Limelight is one of the large CDNs that relies on this architecture. It maintains about 80 server clusters around the globe (cf. Limelight Networks, 2016). Smaller CDNs, such as CacheFly, Fastly, CloudFlare, and MaxCDN use this architecture with up to tens of server clusters. This strategy facilitates the realization of scale economies (e.g., volume discounting associated with bulk purchases of servers, power, and bandwidth capacity connections) and management efficiencies that can significantly lower operational costs. Instead of incurring the fixed (per-site) costs associated with having to secure, power, and manage a large number of geographically distributed server sites, datacenter-based CDN providers can focus on a smaller number of locations with a larger number of (potentially higher capacity) servers.

To be viable 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 peering locations.[10] For example, datacenters located close to IXPs are able to peer with multiple networks at a single location, thereby expanding the user population to which they can serve content. However, having content cached on servers located in only a few locations means that the distance data packets need to travel to end-users may be quite far. Also, such CDNs have limited end-to-end control over the QoE for packets that may have to traverse multiple networks, consuming additional transport resources and being more susceptible to Internet congestion than CDNs with more distributed server architectures. Consequently, these types of multi-purpose CDNs are best able to serve content for which the end-users are localized near datacenters and for which the content is delay-tolerant (and hence cacheable) or may be scheduled (e.g., to avoid peak congestion). Software updates and a lot of entertainment video may fit those requirements.

---

[10] As described in Section 2.2, key considerations in selecting the location of datacenters include the cost for energy and hosting, but also proximity to anchor tenants.

3.2. Highly Distributed CDNs

Akamai is the leading and best-known CDN provider with a highly distributed network of servers. It maintains about 220,000 servers deployed in more than 1,500 networks, and delivers more than 20% of global Web traffic (cf. Akamai, 2016a). These are strategically distributed at multiple peering locations and within access networks. The servers are organized into a hierarchy of edge-based servers hosted within ISP networks that manage front-end communications between the end-users and the server network, and back-end servers that tie together the distributed network of servers. Positioning edge servers closer to end-users – both in terms of network and physical distance – and employing optimized protocols allows Akamai's distributed CDN architecture to deliver substantial performance benefits (cf. Leighton, 2009; Maggs and Sitaraman, 2015). Such highly distributed CDNs rely on the deployment of multiple CDN modules, each offering a common set of general-purpose capabilities that may be customized by the CDN operator to support diverse types of content and applications (cf. Nygren et al., 2010). These may include support for static as well as dynamic and interactive Web or e-Commerce applications and content. Because the fixed (and sunk) costs of establishing and maintaining a distributed CDN comprised of CDN provider-owned servers is quite substantial, a large customer base is needed to render the architecture economically viable. A large size makes it possible to take advantage of economies of scale and volume 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 gain size-based bargaining power when negotiating interconnection and hosting arrangements with ISPs.

3.3. Peer-to-Peer CDNs

Peer-to-peer CDNs address the challenge of managing the high costs of providing content servers by shifting the costs for providing servers to end-users. One of the best-known peer-to-peer CDNs is BitTorrent. Most of the peer-to-peer CDNs have been non-profit or community-driven, with limited commercial support, and have focused either on specialized niche content distribution needs (i.e., specific to the community's shared interests), or are engaged in copyright infringement (where distributing ownership/control of the servers limits the CDNs vulnerability to copyright enforcement actions). Commercial peer-to-peer CDN services are offered by Resilio (formerly BitTorrent Sync), the (former) 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 (cf. Resilio, 2016).[11] Peer-to-peer CDNs work best when a large number of the member population is interested in the same collections of files, since the network members can rely on crowdsourcing to share desired content. 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). However, using CDN caches located at the end-users also means that the CDN has much less capability to manage end-to-end delivery performance (e.g., to control end-to-end latency) and that delays might be substantial as access networks have to be traversed twice along the path from sender to receiver. Consequently, peer-to-peer CDNs are

---

[11] Another provider of commercial BitTorrent-based peer-to-peer CDN services is Xunlei (cf. Dhungel et al., 2012).

most suitable for delay-tolerant bulk file distributions (e.g., audio, video, medical or scientific data).

3.4. Hybrid CDNs

A hybrid CDN approach seeks 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 that serve as CDN clients, providing additional capacity and increasing the number of distributed, replicated copies that the CDN may make available. Akamai's NetSession offering provides an example of such a hybrid CDN (cf. Zhao et al., 2013). The hybrid CDN provider combines the benefits of saving 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, the CDN needs to keep track of the state of replicated content cached on end-user servers as well as to maintain a content copy in an always-on server. Additionally, serving content from customer-maintained servers may suffer from the same sorts of performance issues as peer-to-peer CDNs (cf. Feldmann et al., 2010). Today, hybrid CDN architectures are typically used for large file downloads such as software updates.

3.5. Specialized CDNs

In recent years, a number of large content providers that formerly relied on third-party CDN services have built their own application-specific CDN platforms. For example, Google deployed their own caches (Google Global Caches) in more than 1,000 networks (cf. Calder et al., 2013; Streibelt et al., 2013) and Netflix introduced its Open Connect platform with servers in more than 233 locations worldwide (cf. e.g., Miller, 2016; Netflix, 2016a; Clancy, 2016). While Google's cache servers may have general-purpose capabilities, they are customized to support Google's content and applications. By some accounts, Google's CDN which includes YouTube traffic handles more than 25% of Web traffic (cf. McMillan, 2013). Meanwhile, Netflix is reported to be responsible for about one third of the peak Internet traffic on fixed networks in the U.S. (cf. Sandvine, 2016).

A major motivation for such vertical integration is to take advantage of the scale, scope, and specialization economies that can be realized by bringing the CDN functionality in-house. While this requires that traffic volumes are sufficiently large, specialization economies arise when the CDN architecture can be tailored to the application-specific (delivery) requirements of the content provider (e.g., by avoiding the need to support unnecessary general-purpose capabilities). Additionally, vertical integration may offer strategic benefits in the form of tighter control over how content is distributed. A particular complication arises when a single content provider serves traffic volumes sufficiently large to represent a significant share of the traffic delivered by an individual CDN. Hold-up risks on both sides may arise as neither the content provider nor the CDN can tolerate sudden shifts in traffic patterns. Resulting joint-management risks may be best managed through vertical integration.

3.6. Broker CDNs

CDN brokers have emerged that seek to take advantage of price differentials among CDN providers which may vary by region, allowing them to integrate and arbitrage offerings from multiple CDN providers to expand coverage or better customize services for their customers (cf. e.g, Mukerjee et al., 2016). Based on a meta-architecture for optimizing content delivery that relies on existing cache deployments, broker services 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 providers in the selection of CDNs. One or more CDNs may be used simultaneously, with the assignment of user requests to specific CDNs made dynamically. Typically, broker CDNs specialize in the provision of 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.

3.7. Licensed CDNs

ISPs and others interested in providing CDN services or functionality in-house may elect to license the CDN capabilities from an established CDN provider. Following this route reduces the upfront costs of establishing a CDN network and may allow sharing in the scale, scope, and learning economies realized by an established CDN provider. Although ISPs have an advantage in serving their own subscribers as they control the relevant last mile connections, their reach is limited. Consequently, licensing CDN capabilities can prove a useful strategy for meeting off-net coverage needs of content providers. For the established CDN providers, offering wholesale services to ISPs and others helps increase utilization and realize scale and scope economies for existing CDN networks. Providers like Akamai (cf. Akamai, 2016b) and Edgecast (cf. BusinessWire, 2012; Verizon, 2016) introduced off-the-self wholesale CDN solutions for ISPs. Two models for licensed CDNs are common. First, the set of ISP-owned CDN servers can be managed independently by ISPs seeking to operate standalone CDNs. A number of telco-based ISPs have chosen this model which is based on bare licensing of the CDN software platform, allowing the ISP's own servers to be upgraded to support CDN functionality.

Second, and the more commonly used approach, CDN providers may integrate ISP-owned servers into the CDN's (potentially global) network of servers. In this model, the CDN directly manages the ISPs' servers as front-end servers for the CDN. For example, Akamai is involved in a number of such collaborative licensing agreements with ISPs around the globe. Such collaborative licensing arrangements enable tighter integrated management of the CDN traffic, facilitating revenue sharing and making it easier to launch new applications. Additionally, ISPs and CDNs can more easily share information regarding access network conditions to facilitate joint optimization of end-to-end packet delivery. Improved performance and cost efficiency 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. e.g., Frank et al., 2013; Akamai, 2016b).

3.8. CDN Federations

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 individually. 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). ETSI (2013) and ATIS (2011) have issued standards related to the operation of such federated networks. Currently, the CDN provider Edgecast markets its OpenCDN as a commercial product that enables CDN federation.

4. The Multiple Facets of (Peering) Location

The location of CDN servers is a key determinant of CDN architecture, costs, and capabilities. In the following sub-sections we discuss different ways in which the "location" of servers impacts CDN operations.

4.1. Geographic Location

The geographic distribution of CDN servers plays an important role in the cost and performance of CDN services. Distributing CDN servers over a wide geographic area expands the range of options for serving content from a server that is geographically close to the end-user. Injecting traffic into a terminating access network in close proximity to end-users typically allows the packet to traverse fewer interconnection points that might be vulnerable to congestion. In addition, direct interconnections between CDNs and ISPs allow the traffic to bypass upstream, transit providers. In both cases, packet delivery consumes fewer network resources and typically experiences reduced packet delays.

Additionally, interconnection and hosting in multiple geographic locations enhances the resilience as well as the reliability of content delivery. Historically, CDNs relied on transit services purchased from large ISPs to deliver their traffic to end-users. Increasingly, in a trend that is also common among ISPs, CDNs are directly connecting to the ISPs to deliver their content (cf. Woodcock and Frigino, 2016, p. 10). For example, Akamai peers in more than 100 IXPs and private peering facilities (cf. Hannigan, 2015), while Google peers in more than 90 IXPs and more than 100 private peering facilities in a number of cities around the world (cf. Google, 2016). Maintaining a geographically distributed server architecture provides the CDN with a greater number of options for serving the content from a server that is close to the end-user, allows the CDN to route around "stroke points" (i.e., congested links or link outages), and enables the CDN to better balance server loads to enhance capacity utilization efficiencies in response to changing traffic patterns.

Although shorter paths often mitigate performance problems, the geographic distance between the servers may not correspond to the shortest or best network path. For example, in peer-to-peer CDNs or hybrid CDNs that serve content from end-user maintained caches, the path from one user to another may require traversing the access network twice. Alternatively, two CDN servers may be located in the same physical location (e.g., in the same carrier hotel), but one (e.g.,

datacenter-based CDN) may be interconnected on the upstream side of a peering link, while the other (e.g., licensed CDN) may be hosted by the ISP downstream on the network side. While the former server would be vulnerable to delays if the peering link is congested, the latter is not affected.[12]

Existing broadband access infrastructure and the connection speeds available vary significantly across geographic regions.[13] Faster connection speeds typically imply reduced delays and less of a benefit in locating content geographically close to the end-user. When access networks are slower or when interconnection links are likely to be congested, distributing caches closer to end-users on the access-side of interconnection links can deliver significant performance improvements. If latency in the access network dominates backbone latency, then those CDN architectures that heavily rely on the access network (e.g., peer-to-peer and hybrid CDN architectures) will generally perform worse than alternative CDN architectures.[14] However, if latency in the backbone network dominates, then architectures that locate servers inside the access network will typically provide better performance.

4.2. Virtual Location

Despite the importance of geographic location as described in the previous section, Internet routing protocols do not route packets based on the geographic (physical) location of the source or destination. Instead, routing is based on their virtual location as defined by their IP address which determines how packets are routed to and from the server. Servers with IP addresses managed by the same provider may be thousands of geographic miles apart (e.g., one in New York and the other in Los Angeles) (cf. Freedman et al., 2005); while two nodes in close geographic proximity that have IP addresses belonging to different administrative domains (i.e., autonomous systems or ASes) may be separated by network paths that require traversing thousands of miles, and may even involve crossing national borders (cf. e.g., Spring et al., 2003; Gupta et al., 2014a). Therefore, controlling the virtual location of CDN servers is also important (cf. e.g., Krishnan et al., 2009).

---

[12] According to recent studies, congestion at a peering link can increase delays by up to 40 msec (cf. Luckie et al., 2014).

[13] Studies on the state of residential access networks in the U.S. (cf. FCC, 2015b), the EU (cf. EC, 2015), and the UK (cf. Ofcom, 2016) illustrate significant differences in both upstream and download speeds across geographic regions and across different broadband access technologies. Although available connection speeds have increased during the last decade in most countries, Akamai (2016e) data on the connection speeds between users and their CDN servers across markets confirm that speed differences remain significant.

[14] 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 milliseconds and backbone latency 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).

In addition to traditional interconnection agreements, ISPs may exchange traffic with CDNs based on hosting services. CDN servers are strategically positioned and hosted within the ISP network and are assigned local IP addresses that belong to and are managed by the hosting ISP to help ensure packets are routed efficiently within the ISPs network, and avoid traversing inter-AS links unnecessarily. In addition to improving delivery performance, hosting servers on the ISPs network can reduce the load on interconnection links, improving the performance for other applications and allowing ISPs to postpone capacity upgrades that might otherwise be required.[15]

4.3. Colocation Hubs

Colocation hubs enable the interconnection of multiple ISPs and CDNs (cf. Lodhi et al., 2014). These are typically located at IXPs (cf. Chatzis et al., 2013) or other interconnection facilities for private interconnections (cf. Giotsas et al., 2015). Although most global Internet traffic is exchanged via private interconnections, more than 30 Tbps of traffic are exchanged at IXPs, with most of that activity taking place in Europe.[16] 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 interconnection facilities provides CDNs with multiple options for accessing a large number of networks and end-users while at the same time reducing both the geographic and virtual distance between source and destination (cf. Galperin, 2016). For example, Equinix maintains a number of interconnection facilities around the globe and in the United States, which they claim makes it possible to access 90% of the population in North America and Europe with only 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.[17] Large hubs are

---

[15] 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). 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).

[16] 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.

[17] 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).

typically 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. Resulting scale economies and other benefits of interconnecting at hubs may be further enhanced through mergers and acquisitions.[18] 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. CDNs that co-locate facilities at hubs have more opportunities for direct peering and hosting.

4.4. Innovation Hubs

A growing number of hubs offer support for a diverse array of innovative interconnection arrangements. These include differentiated types of partial transit and paid peering, as well as public peering via route servers that enable hundreds of interconnections 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 provided may include direct cross-connects, public peering, or tethering.[19] This diversity of mechanisms for physically interconnecting networks facilitates complex and nuanced traffic exchange arrangements (cf. Giotsas et al., 2014).

For example, remote peering enables a decoupling of a router's geographic location from the geographic location of the interconnection. A physical router presence is no longer necessary for networks to directly peer with other networks in many geographic locations. A CDN with a server cluster in a single location may thus 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. For example, more than 20% of the members in AMS-IX use remote peering (cf. Castro et al., 2014). Remote peering allows a CDN to expand its footprint at reduced cost compared with relying on transit services (cf. e.g., AMS-IX, 2016; Richter et al., 2014). Limelight and LinkedIn are two CDN providers that take advantage of this type of peering.

Additionally, innovation hubs typically offer a number of value-added services to their members. For example, several large IXPs such as DE-CIX, AMS-IX, and LINX, provide the latest black holing services that can be used to mitigate attacks and dispose of unwanted ingress or egress

---

[18] 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).

[19] How the CDN interconnects to customers and its servers can impact the CDN provider's ability to control the QoS of the services it offers. Direct cross-connects are based on dedicated circuit-switched connections between the peering networks and provide better QoS control for a CDN provider; whereas public peering relies on shared access to the IXPs switching fabric and may be more susceptible to QoS variations that are not under the CDN operator's control. Finally, tethering makes use of a virtual local area network (VLAN) on the IXP's shared switching fabric, and introduces additional potential sources of QoS variability that are not under the CDN provider's direct control (cf. Giotsas et al., 2015).

traffic.[20] Moreover, new technologies such as SDN enable even 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 makes these capabilities available to all networks using the hub for interconnection. Table 2 provides an overview of some of the capabilities that are available at such innovation hubs.

[Insert Table 2 about here]

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

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. Thus, they can compete with a broad portfolio of service offerings while at the same time taking advantage of cost economies. At the same time, the largest ISPs have been upgrading their networks through softwarization to expand their service portfolio and to lower network costs. Some ISPs are even 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 platforms and other wholesale services to vertically integrate into self-provisioning 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 or Limelight that own and manage a global network of servers and retain the backbone networking capabilities to allow them fine-grained control. In addition to the significant global scale and scope economies associated with managing such a network, Akamai benefits significantly from incumbency and its first-mover advantages associated with its success in acquiring a large installed base of customers. Although the potential market of commercial content providers, including entertainment media that want global-scale content delivery coverage is large, replicating the global reach of Akamai's dedicated server network would be extremely challenging for other third-party CDNs or even the largest ISPs. Akamai is able to spread the significant fixed and sunk costs of developing additional services over its existing and large customer base. Additionally, Akamai's control over traffic routing (i.e., where they inject CDN traffic into ISP networks) likely provides significant leverage in negotiating favorable colocation or interconnection terms with ISPs.

---

[20] Distributed Denial of Service (DDoS) attacks can result in large amounts of unwanted traffic that can cause congestion and service disruptions. In the course of a DDoS attack, large amounts of traffic are sent by a large numbers of distributed sources to overwhelm the network resources of the target network. For example, seemingly legitimate requests for information may be directed at a target server at a rate that exceeds the server's ability to respond, causing it to become overloaded and cease to function as designed. Blackholing describes the practice of redirecting unwanted traffic to nodes where such traffic can be harmlessly terminated and disposed of as if the traffic were being swallowed by a "black hole." A range of network operators and IXPs offer such services (cf. e.g., DE-CIX, 2016; Dietzel et al., 2016).

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.[21]

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 be capable of replicating Akamai's extensive capabilities and scale, it is feasible for such new CDN entrants to operate on much smaller margins since 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 (i.e., by traffic or customer type), 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 or a favorable location to content providers or Internet exchange facilities) to enter CDN services at low incremental 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, yet may appear to be marginally profitable for the would-be entrant.

Third, large content providers may find sufficient benefits from reducing content distribution 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 have lowered entry barriers into CDN services as already noted. 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 third-parties. Should they do so, they

---

[21] 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 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.

would confront potential channel conflict issues unless their content provider customers were offering complementary content.

Fourth, access ISPs that already play a critical role with respect to consumers QoE when accessing content and are feeling pressure as revenues from 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 subscribers, they may have additional advantages in determining how content might be cached or 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.[22]

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 based on horizontal integration. Such a strategy may make the most sense in markets like Europe – where there might be the need to establish Europe-wide 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. Corresponding strategies 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 and resource 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. In light of the significant growth in content that is mostly cacheable and entertainment-oriented (e.g., movie libraries), there are compelling mutual benefits to be realized for both CDN providers and ISPs from hosting CDN servers inside the ISP networks. 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 provide on-demand bandwidth and pop-up edge servers allowing

---

[22] 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.

CDNs to better dynamically scale their bandwidth and server needs. ISP-CDN coordination of cache updates and where content is injected into an ISP's network can help conserve interconnection capacity and reduce peak period congestion, enhancing the QoE of both content customers and users of other applications. Moreover, ISP-CDN collaborations may enable transparent caching, obviating the need for ISPs to deploy expensive middle-boxes.[23] Today, there are tens of active ISP-CDN collaborations (e.g., between Akamai and large ISPs such as AT&T, Deutsche Telekom, Telefonica, and Orange) that expand the number of available servers and enhance the efficiency of cache management.[24] This increases both the granularity by which multiple content copies can be cached close to end-users and the scope for traffic engineering (cf. Frank et al., 2013).

As the ecosystem continues to evolve towards next generation clouds and 5G networks, the need for and benefits of collaboration are likely to increase. Providing seamless support for mobile access to content is likely to call for new types of CDNs. For example, providing access across diverse wireless access networks, operating in different portions of the radio frequency spectrum, will call for better support for cross-layer protocol optimizations.[25] Moreover, since mobile devices are often more likely to be power constrained and/or have smaller screens, there may be a greater need for flexible and device-dependent coding that lowers the resolution by reducing the number of bits used to code the content or undertaking other energy saving strategies when delivering content to mobile devices. The industry's 5G vision specifies that networks should be able to support 1 msec latencies for services. Achieving this goal is likely to be impossible without much closer coupling of access and CDN networks.

Finally, regulatory policies are likely to have a significant impact on the evolution of the CDN ecosystem. For example, while CDN services offered by third-party providers are typically not regulated, the last-mile broadband access services often are. In a number of countries, policymakers have adopted so-called Network Neutrality (NN) regulations that are intended as a response to concerns that access ISPs may have market power that might be misused to discriminate against unaffiliated content or applications, thereby harming competition, increasing

---

[23] Middle-boxes are specialized, often proprietary, servers that are typically deployed by ISPs within their network to facilitate better traffic management and/or traffic shaping capabilities that can improve network utilization, implement differentiated services, and increase an ISP's control over QoS within its network. For further discussion, see Carpenter and Brim (2002).

[24] Corresponding strategies are also attractive for Tier 1 ISPs which may have developed a profound interest in diversifying their business strategies by participating in CDN service provision. For example, Level 3 operates its own proprietary CDN while AT&T is joining forces with Akamai based on an ISP-CDN collaboration.

[25] For example, TCP (transport layer protocol) often proves sub-optimal when used over wireless links (physical layer networks); and the performance of wireless links varies significantly across technologies and spectrum bands (e.g., Wi-Fi, LTE and satellite services operate in different frequency bands, with different protocols and network topologies). Cross-layer optimizations that jointly consider multiple network layers and make adjustments accordingly can improve performance.

costs, and limiting consumer choices.[26] Although the rules may differ by country and continue to evolve, NN regulations constrain the access ISP's scope and ability to actively manage data traffic within their networks and to offer price and QoS-differentiated content delivery services. Regardless of one's views as to the merits of NN regulation, it is reasonable to anticipate that such rules will create opportunities for regulatory arbitrage that will impact who provides CDN services and how those services are provided.

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 global general-purpose CDN providers like Akamai continuing to expand their range of services; 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.

6. Conclusions

In this paper, we reviewed how CDNs have evolved in conjunction with the broader trends driving changes in the Internet ecosystem. Over time, a complex landscape of CDN architectures and business models have emerged to flexibly address continuous 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. While some content providers do not need or cannot pay for additional support, and rely on basic Internet functionality to distribute their content, others are either so large or have needs sufficiently specialized that self-provisioning an in-house CDN offers the lowest cost and best strategic option. A wide range of third-party CDN providers serve the middle of the market. These range from incumbents to new entrants, regional to global providers, general-purpose to niche service providers, pursuing a range of business models. The various business models reflect each CDN's efforts to leverage its perceived competitive advantages. In the case of incumbents, this includes their first-mover advantages in establishing large customer bases and networks with global reach. Newer entrants often seek to leverage existing assets such as data centers or local ISP infrastructure, or to take advantage of the growing array of wholesale platform options for CDN or cloud services. Because entry costs into CDN services 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 (i.e., one to three) global providers like Akamai.

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

---

[26] In 2015, network neutrality regulations were introduced in the U.S. (cf. FCC, 2015a) and in Europe (cf. European Parliament and Council, 2015; BEREC, 2016). For a comparison of the U.S. and European frameworks, see Marcus (2016).

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, as well as the traditional division of labor between the two, is increasingly blurring.

While the benefits from increased integration of CDN and ISP functionality are significant, such integration poses both strategic and regulatory challenges. From a strategic perspective, both ISPs and CDNs may realize channel conflicts if they are vertically integrated. Non-affiliated content providers may allege network neutrality violations or confront challenges in negotiating last mile delivery services with competing ISPs. This may be less of an issue for ISPs in non-overlapping markets, or for CDNs that are self-provisioned by a single content or application provider. 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.

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 and the softwarization of networks discussed earlier, there will be expanded support for real-time traffic management. Figuring out where critical functionality and network intelligence should reside and who should control it (i.e., 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.

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].

Alliance for Telecommunications Industry Solutions (ATIS) (2011). *CDN Interconnection Use Case Specification and High Level Requirements*, Standard ATIS 0200003, available at: http://standards.globalspec.com/standards/detail?docid=1401215&familyid=JZTONEAAAA AAAAAA [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) (2016). 'BEREC Guidelines on the Implementation by National Regulators of European Net Neutrality Rules' *BoR (16) 127*, August.

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.

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.

Carpenter, B. and S. Brim (2002). 'Middelboxes: Taxonomy and Issues', RFC 3234.

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_paper09186a0080 0a3e2f.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. 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].

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

Dhungel, P., Ross, K.W., Steiner, M., Tian, Y., and X. Hei (2016). 'Xunlei: Peer-Assisted Download Acceleration on a Massive Scale', in Taft, N. and F. Ricciato, eds, *Passive and Active Measurement: 13th International Conference PAM 2012 - Proceedings*, Cham et al.: Springer, pp. 231-241.

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.

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.

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 (EC) (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 Parliament and Council (2015). Regulation (EU) 2015/2120 of the European Parliament and of the Council of 25 November 2015 laying down measures concerning open internet access and amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services and Regulation (EU) No 531/2012 on roaming on public mobile communications networks within the Union, 26.11.2015, OJ, L 310/1-18, Brussels.

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].

European Telecommunications Standards Institute (ETSI) (2013). 'CDN Interconnection Architecture', *Technical Specification*, ETSI TS 182 032 - V1.1.1, available at: http://www.etsi.org/deliver/etsi_TS/182000_182099/182032/01.01.01_60/ts_182032v010101 p.pdf [accessed: 20 Aug 2016].

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) (2015a). In the Matter of Protecting and Promoting the Open Internet, Report and Order on Remand, Declaratory Ruling, and Order, GN DOCKET No. 14-28, FCC 15-24, Adopted: 22 February 2015, Washington DC.

Federal Communications Commission (FCC) (2015b). '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.

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].

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.

Marcus, J.S. (2016). 'New Network Neutrality Rules in Europe: Comparisons to those in the U.S.', *Colorado Technology Law Journal*, 14(2), pp. 259-280.

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].

Mukerjee, M.K., Bozkurt, I.N., Maggs, B., Seshan, S. and H. Zhang (2016). 'The Impact of Brokers on the Future of Content Delivery', *HotNets 2016: Fifteenth ACM Workshop on Hot Topics in Networks*, November 9-10, Atlanta, GA, USA.

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].

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

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].

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].

Resilio (2016). 'Sync Everything', available at: https://www.resilio.com [accessed: 06 Jan 2017].

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.

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

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.

Woodcock, B. and M. Frigino (2016). '2016 Survey of Inter Carrier Interconnection Agreements', *BoR (16) 237*, November.

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.

Table 1: Typology 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	Multi-purpose, 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	Multi-purpose, no investment in dedicated infrastructure	File sharing, bulk transfers
Hybrid	Akamai NetSession	Dedicated servers combined with functionality at end-user equipment	Low	High	Multi-purpose, 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	Single-purpose, 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

Source: authors

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

Source: authors