analysys mason

Report for Korea Internet Corporations Association

IP interconnection on the internet: a white paper

Michael Kende, David Abecassis

May 21, 2020

Ref: 2020099-514

CONSULTING

Contents

The paper was sponsored by the Korea Internet Corporations Association and funded by its members. The Analysys Mason team produced this paper independently.

Analysys Mason Limited
North West Wing, Bush House
Aldwych
London WC2B 4PJ
UK
Tel: +44 (0)20 7395 9000
london@analysysmason.com
www.analysysmason.com
Registered in England and Wales No. 5177472

1 Executive summary

The internet is a global network of networks – including commercial, academic, research, and government networks. Interconnection is the glue that holds these networks together. Internet interconnection is commercially negotiated, rather than regulated, and has been since the internet was commercialized in the 1990s. Allowing these arrangements to be negotiated has enabled them to adapt to all the industry changes that have taken place over the past 20 years. In contrast with this established practice, Korean policy makers are proposing to impose charges on interconnection arrangements. In this paper we argue that these regulations can result in increased costs and latency for Korean end users when accessing their favorite content and services, including Korean content.

Internet interconnection arrangements have adapted themselves to the emergence of high-bandwidth video content delivery that predominates today. This has been possible because the arrangements are commercially negotiated, and thus able to evolve rapidly and organically. For example, content delivery networks (CDNs) emerged to deliver content closer to users. The first CDNs were independent companies that managed the content for third-party content providers. More recently, large content providers began to develop CDNs for delivering their own content.

Regardless of whether CDNs are independent or part of a content provider, they bring content closer to receiving internet service providers (ISPs), which lowers the costs for the ISPs and improves the quality of service for their end users. The development of CDNs involves installing caches to hold popular content (particularly video), points of presence (PoPs) where CDNs can peer directly with ISPs, and even building subsea cables between continents and countries, in order to deliver content to PoPs and caches. These investments are a response to commercial incentives to change business models and negotiate interconnection arrangements that improve access to content for end users.

In general, internet interconnection has not been regulated. In contrast with this approach, South Korea has been exploring regulated network charges between ISPs, which would require payments based on the volume of traffic delivered to an ISP. As discussed below, such a move could result in additional costs for content providers with a local presence, or they could decide to only make the content available outside the country, thus increasing the cost of accessing the traffic for South Korean ISPs. Either way, the possibility of regulation and ambiguity about how it will be imposed has had an impact on content providers' investments in South Korea. Furthermore, internet interconnection would be less efficient than the commercial outcomes in other countries, and it would impose increased costs and latency on South Korean end users with no mitigating benefit.

The internet has undergone significant changes since its commercial beginnings in the 1990s. In addition to remarkable growth rates, the internet has globalized from its historical roots in the US, the volume of traffic associated with content has grown exponentially, and content providers have begun to deploy their own networks to deliver the content to end users worldwide. The one constant has been that commercial negotiations, rather than regulation, have driven the changes in the interconnection arrangements, to the benefit of users of the advanced internet ecosystem. In South Korea, the need for regulation is unclear, the benefits so far not realised, and the costs all too visible in the form of a lower quality of experience for consumers.

2 Introduction

The internet is a global network of networks – including commercial, academic, research, and government networks. Interconnection serves as the glue that holds these networks together. The earliest interconnection arrangements have proven to be robust, and have adapted themselves to the emergence of high-bandwidth video content delivery that predominates today. This is possible because the arrangements are commercially negotiated, and so have been able to evolve as content providers developed new business models for delivering their content closer to users. This is also possible because these interconnection arrangements were not subject to any regulation that might have inhibited the innovation and investment that characterize all the changes over the past 20 years.

As the internet became commercial in the 1990s, interconnection arrangements were needed as a practical solution to enable networks to exchange traffic with one another, to the benefit of all, in the absence of a central coordinated backbone to govern the exchange of traffic. The voluntary arrangements that emerged, called 'peering', epitomized the attitude that drove the development and growth of the internet.

Early interconnection arrangements were driven by engineering considerations rather than by commercial factors or government mandate. The result was interconnection based on cooperation rather than competition or regulation. Despite exponential growth of the internet and the introduction of many new players and business models, including content delivery networks, this attitude – to cooperate to create the internet and then compete in downstream services – has remained intact. In the face of changes that may have been unimaginable to those developing the internet, and that would make it almost unrecognizable today, the one constant has been that internet interconnection remains a voluntary decision, rather than a regulatory mandate.

In contrast with international best practice, South Korea has been exploring regulated network usage fees between ISPs since 2016, requiring them to pay for volumes of traffic delivered to one another. One consequence, as we understand it, is that it provides a justification for ISPs to charge content providers a fee for delivering their traffic to other ISPs, as well as possibly the ISPs' own end users who requested that traffic. Although it is not unusual for there to be payments for exchanging traffic, it is not the norm for content providers to pay *termination* charges to bring traffic to end users,[1] and interconnection arrangements among ISPs are nearly always freely negotiated rather than regulated.

In this paper, we first introduce the basic concepts of internet interconnection (Section 3), including definitions of the different arrangements. We then explain the basis for unregulated internet interconnection arrangements and highlight what happened in the past when attempts were made to impose termination charges (Section 4). Finally, the paper concludes on the potential impacts of imposing a mandatory network usage charge on a market, in this case South Korea (Section 5). The annex provides an explanation of technical terms used in this paper.

1 Most of the more than 1.5 billion websites available on the internet pay for a connection to the entirety of the internet, not termination fees to the thousands of ISPs around the world.

3 Internet interconnection

The two main forms of internet interconnection are ‘peering’ and ‘transit’, which enable providers – including ISPs and content providers – to exchange traffic. Peering enables two providers to exchange their own traffic with each other, while transit enables one provider to pay another provider for the ability to exchange its traffic with the entire internet. The physical exchange of traffic can take place in private facilities or data centres, or at public points known as internet exchange points (IXPs). We define these in turn below, and provide an overview of how these arrangements arose.

The US government developed the internet (then called ARPANET), first to enable Department of Defense networks to connect, and then for academic and research networks (NSFNET). As commercial internet networks began to develop and expand in the 1990s, thanks in large part to the introduction of the World Wide Web, a fundamental question arose: how and where would these commercial networks exchange traffic?

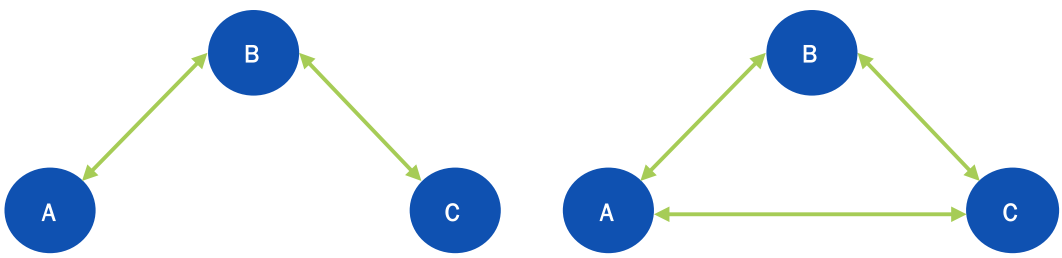
In the beginning, all the growing networks were small, and discovered a way to interconnect that was simple and effective and helped to grow the internet. These networks saw themselves as equals, or ‘peers’, and the form of ‘one-to-one’ interconnection that grew was called ‘peering’.

As the number of networks started to grow exponentially, with discovery mechanisms such as search helping users to find information in smaller and more-remote networks, the ability to exchange traffic on a ‘one-to-all’ basis across all networks became increasingly valuable. This was made possible through ‘transit’ arrangements, thanks to a hierarchy of interconnection relationships and the services being sold to an ISP’s customers.

3.1 Peering arrangements

A peering relationship between two internet providers enables them to exchange their own customers’ traffic with each another. On the left-hand side of Figure 1, Networks A and B exchange traffic with each another, and Networks B and C exchange traffic with each another, using peering arrangements. However, Network B does not exchange traffic between Network A and Network C. The simple reason for this is that Network B has no incentive to allow traffic from Network A’s customers to transit its network to reach the customers of Network C’s network – such an arrangement would cost resources and would not benefit any of Network B’s customers. In order for Networks A and C to communicate, they would have to peer with each another as well, as shown on the right-hand side of Figure 1.

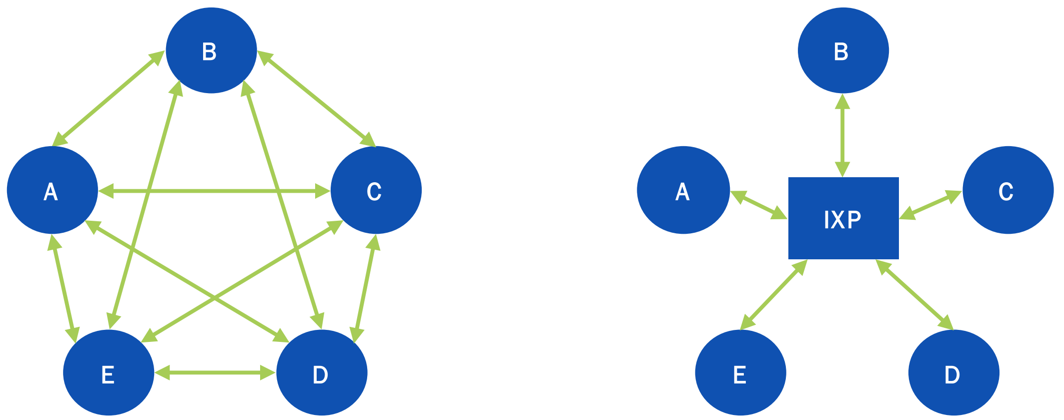
Figure 1: Peering arrangements [Source: Analysys Mason, 2020]



These peering agreements were initially all 'settlement free': that is, no money was exchanged to collect traffic originating from one network (Network A), and terminate it on the other network (Network B). Both Network A and Network B had to fulfill requests from their customers to reach users on the other network. As providers saw themselves as roughly equal, and were sending one another a similar amount of traffic, the cost to Network A of terminating traffic received from Network B was roughly the same as the cost to Network B of terminating traffic from Network A. Any payments would have effectively cancelled each other out, but it would cost money to measure the traffic and settle the accounts, so a settlement-free peering arrangement effectively saved money for all parties.

However, as more internet providers emerged the cost of the peering connections began to rise. For instance, if there were five providers (as shown on the left-hand side of Figure 2), then each would need connections with each of the other four providers to exchange peering traffic. The solution was to develop common points to exchange traffic, now known as internet exchange points (IXPs). As shown on the right-hand side of Figure 2, each provider needs one connection to the IXP, and then in the IXP traffic is efficiently exchanged among the peering partners.

Figure 2: The impact of an IXP on traffic flows between networks [Source: Analysys Mason, 2020]



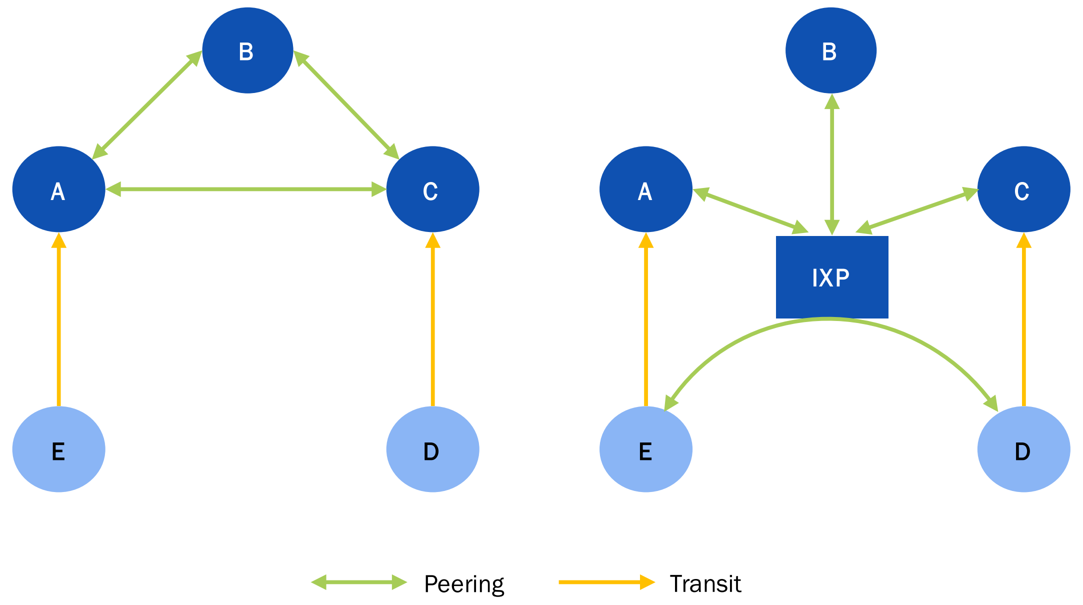
IXPs started in the US, and then they began to be established in Europe to help avoid a phenomenon known as '*tromboning*'. In the 1990s, all European internet providers had to have connections to the US, the historical home of the internet, in order to reach the largest base for content and users. European providers realised that it was cheaper to use those connections to exchange all of their domestic traffic as well, rather than connect with every other domestic provider (as on the left-hand side of Figure 2). This tromboning approach was not an efficient use of international connections, so eventually European providers started national IXPs, including the London Internet Exchange (LINX), and Amsterdam Internet Exchange (AMS-IX), to localize traffic exchange. These IXPs then grew into large regional IXPs.

3.2 Transit arrangements

The alternative to peering is known as transit. Transit is a paid service in which one provider becomes a customer of another, in return for access to all points on the internet. The transit provider establishes the connections to the rest of the internet through its other customers and its peering partners. Transit arose as a service for smaller networks – known as *internet service providers* (ISPs) – to deliver their traffic to and from the rest of the internet. The larger networks that sold transit became known as *backbone providers*.[2]

On the left-hand side of Figure 3, ISPs D and E purchase transit from Backbones A and C respectively. As a result, Backbone A will transit all the traffic from B and C to ISP E, which is a customer of Backbone A. Likewise, ISP D will receive all the traffic from Backbone C's peering partners. This enables the customers of ISP E to exchange traffic with ISP D, and vice versa.

Figure 3: Transit arrangements [Source: Analysys Mason, 2020]



With the growth and spread of IXPs, many ISPs, as well as content providers, became direct members of the IXP alongside backbones. This allows them to easily peer with one another, as shown by the curved line in the bottom right of Figure 3, which reduces their reliance on transit for those connections. There are normally criteria that networks impose on potential peers, such as a minimum scale or specific technical conditions.[3] If these criteria are not fulfilled, interconnection must still rely on transit despite the networks being present in the same location. We describe several such key criteria next.

2 In practice there is no definite distinction between backbones and ISPs. Some backbones have retail customers, and some ISPs have backbones. In this paper we use backbones to refer to providers that sell transit, and ISPs to refer to providers that purchase transit. In addition, while some backbones are domestic, we focus here on their role in delivering traffic between countries and continents. Likewise, ISPs may be international, but here we focus on their 'last-mile' role in domestic markets.

3 For example, Facebook's peering requirements specify: "a minimum of 50 Mbps of in-continent traffic destined to or through your network [...] up-to-date peeringdb entry for all public peering requests, including exchange information with properly formatted public fabric addresses, ASNs, and NOC/peering contact information".

3.3 Peering criteria

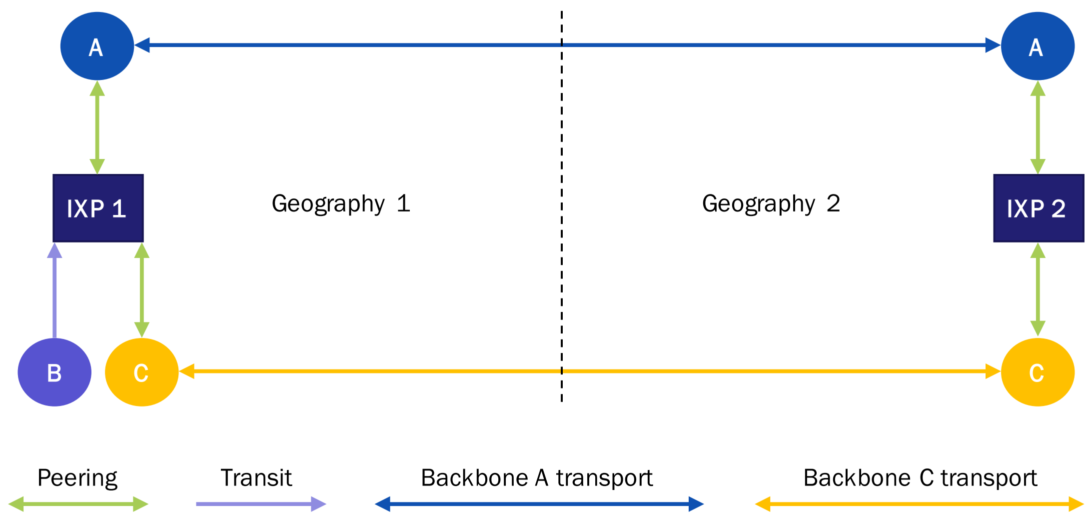
When two internet providers agree to exchange traffic with each other in a peering arrangement, each must devote network resources to terminating traffic from the other provider. These network resources have a cost, which providers seek to balance. In the early days, backbones were all connected to the same IXP, and had roughly the same mix of customers between end users and content providers. As a result there was a good balance of traffic, so costs were shared and settlement-free peering became the norm.

As the internet grew, two points of divergence eventually emerged: first, when some backbones grew across the US or internationally; and second, when some backbones began to focus more on delivering content. In some cases this changed the balance of cost between peering partners and, as a result, providers developed peering criteria to indicate the conditions under which they would peer with other providers. These peering criteria first emerged in the US during the late 1990s as a tool for self-regulation, to publicly signal how backbones would determine and evolve conditions of peering in light of changes taking place in the industry.

Differences in the geographic growth of backbones

As backbones began to grow around the world, they faced choices on where to peer and how to exchange traffic. In Figure 4, IXP 1 and IXP 2 could be on opposite sides of a country or in different countries. Backbones A and C span both geographies. Each is connected to both IXPs, and they also have their own transport capacity between the geographies. In contrast, provider B is a regional ISP only in Geography 1, and is only connected to IXP 1.

Figure 4: Peering criteria [Source: Analysys Mason, 2020]



When the similar Backbones A and C exchange traffic, they share the traffic load – all that is needed is an agreement about who carries the traffic in which direction. The common approach is known as ‘hot-potato routing’, where each backbone hands over the traffic at the common IXP nearest to where the traffic originated.[4] So, for instance, if a customer of Backbone A in Geography 1 asks for content that is delivered by Backbone C in Geography 2, the request is handed to Backbone C at IXP 1 for delivery to the other geography. In return, the content from Backbone C is handed over to Backbone A at IXP 2, and carried back to the requesting customer in Geography 1.

However, this approach is not possible for traffic exchange with ISP B. Backbone A would not find it cost effective to use its own transport to bring traffic to and from Geography 2 to ISP B under a settlement-free peering arrangement. That would require Backbone A to devote significantly more resources to the peering relationship than ISP B. Instead, ISP B can purchase transit from Backbone A or C to access networks in Geography 2 and beyond. Essentially, ISP B is purchasing access to both Backbone A's own customers and the customers of Backbone A's peers. ISP B must purchase this service in order to fulfil its obligation to provide internet access service to its customers, or else it would be required to build out its own backbone network.

Hot-potato routing allows similar backbones to broadly equalise the volume×distance of exchanged traffic that they carry. To ensure this is possible, backbones include geographic requirements in peering criteria, which specify the geographic spread across which peering connections must be made in order to keep this balance. For instance, Verizon Business requires a provider seeking peering in the US to have nodes in at least eight dispersed regions of the US.[5]

Differences in traffic balance

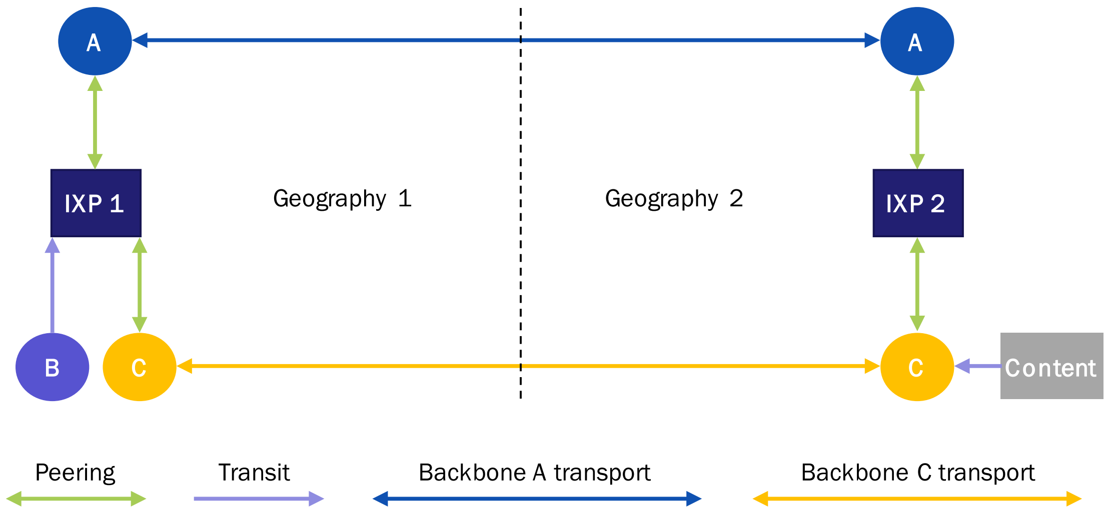
In addition to geographic coverage, the other concern that emerged over time was the relative amounts of traffic carried. As long as Backbones A and C had a roughly equal balance of downstream customers, including ISPs (which request content) and content providers (which deliver it), then the balance of traffic was likely to be the same.

This leads to the second point of recent divergence. With the growth of content providers, particularly those offering video, some internet backbones began to focus on content providers as clients, which started to shift the balance of traffic. For instance, in Figure 5 below, if Backbone C begins to originate more content in Geography 2, then Backbone A will have to carry significantly more traffic between IXP 2 and IXP 1. All the requests for content from Backbone A's customers in Geography 1 will be carried between the IXPs by Backbone C (under hot-potato routing), while in return all of the content will be carried back by Backbone A. The requests for content take very little bandwidth, while the content itself may take a significant amount of bandwidth.

4 In English, the term 'hot potato' likely originated from a game in which players pass around an object, which could be a potato. The object is not to be holding the potato when the game ends, and so the players pass the potato quickly, as if it were hot. Similarly, in the case of peering the providers pass the traffic on as soon as they can after receiving it.

5 See https://enterprise.verizon.com/terms/peering/

Figure 5: Peering criteria [Source: Analysys Mason, 2020]



As a result, backbones such as A began to include in their peering criteria that traffic exchange should keep within a certain ratio of sent-to-received traffic. For instance, Verizon Business requires the traffic exchanged with a peering partner not to exceed a ratio of 1.8:1 – that is, the partner cannot deliver more than 1.8 times what it receives.[6] Such criteria reflect that the *cost* of traffic exchange is higher for the company receiving more traffic, but it should be noted that this is not a measure of the *value* of the traffic exchanged. ISPs need content in order to attract end users, while content providers need ISPs in order to reach end users, and any arrangement based on value could only be arrived at through commercial negotiations.

The increase in content volume has had longer-term implications. Historically, when the traffic ratio exceeded the value specified in peering criteria, backbones such as A might ask to be paid for the excess traffic exchanged under the peering relationship, to compensate them for the cost of the additional traffic that they were carrying. Such '*paid peering*' is exactly the same as settlement-free peering from a technical perspective, but the party originating more traffic would pay the receiving party for carrying the content to its customers. However, it is worth bearing in mind that changes in the types of content traversing the internet came in response to end users making requests for the high-volume content. Thus, the changes to interconnection arrangements described below were made to help serve end users with their preferred content.

Commercial alternatives to paid peering soon began to emerge, which served the dual benefits of lowering cost and improving quality of service. The basis for these alternatives is for the provider with the content to reduce the distance over which the other backbone has to carry the traffic, to lower that backbone's costs. One solution is to keep all the same arrangements, but to switch from hot-potato routing to 'cold-potato routing'. This is a reversal of hot-potato routing, in which the provider with the content delivers it to the IXP closest to the provider whose customer requested the

6 See https://enterprise.verizon.com/terms/peering/

content, and thus the traffic load is taken away from the receiving backbone. Other solutions for helping to deliver content closer to the point where it was requested are discussed in Section 3.4.

The key issue is that the peering criteria described here have been developed to help provide guidelines for peering decisions on a commercial basis, not a regulatory basis. Likewise, decisions to move to paid peering or to deliver content closer to the receiving backbone also result from commercial decisions on the part of the provider with the content, and negotiations between the backbones.

We note here that the discussion involves backbones. In contrast, ISPs – particularly those that focus on residential end users – have asymmetric traffic patterns, which are built into their broadband technology. When an end user with a streaming video service clicks on the icon of a movie on their TV screen this sends a short request for the movie to the server, and the end user then receives perhaps two hours of high-quality video in return. As a result, since the very early days of ADSL (asymmetric DSL), broadband technology has had a built-in asymmetry. ISPs' customers want high-quality video, and that is what is delivered to their ISPs by content providers.

Over time, however, there has been a strong movement toward delivering content closer to the requesting backbone, not just to satisfy peering criteria, but more fundamentally to lower costs and increase the quality of service for end users. This movement has been strengthened by the emergence of a new business model for delivering content, which is described below.

3.4 Content delivery networks

As the volume of content traveling over longer distances grew, this not only added to the cost of delivery, but also increased the latency of the service, and affected the quality of service enjoyed by end users. To address these issues, content began to be distributed to multiple locations closer to end users' ISPs. A new business model emerged, called the content delivery network (CDN), to help content providers distribute their content efficiently.

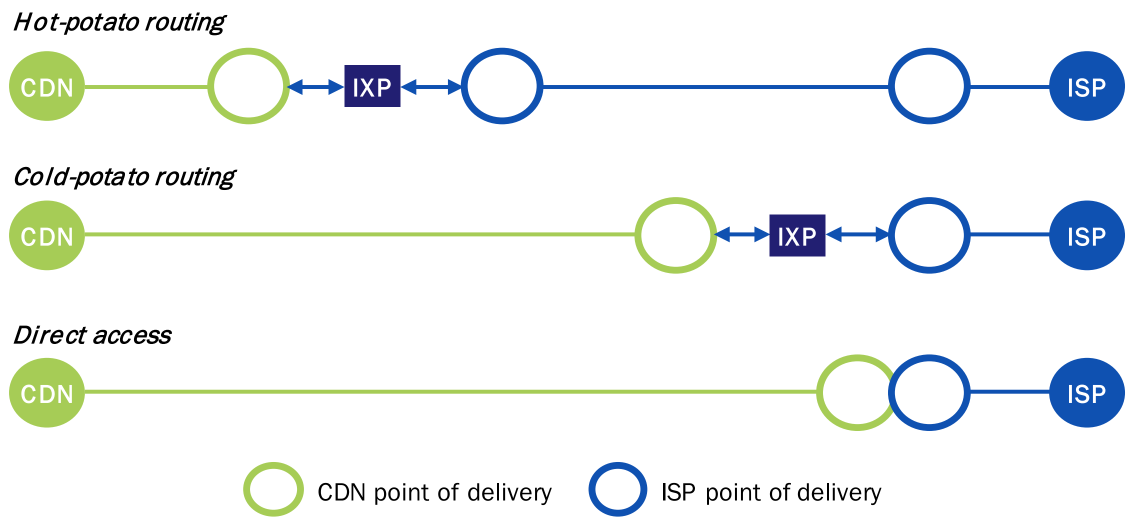
Broadly speaking, there are two types of content, each handled differently by a CDN:

- *Static content* is content that does not change, such as audio and video recordings, which can therefore be stored in multiple locations
- *Dynamic content* includes communications such as video calls or messages, and other interactions such as gaming which cannot be stored.

A CDN is a network that is engineered to help manage and distribute content around the world using different models. The simplest model is to host content in numerous distributed data centers, but still use hot-potato routing. A more-developed model is to place caches at IXPs for cold-potato routing, or place them directly 'on-net', in the networks of ISPs. See Figure 6 below for more details.

A cache holds popular static content, such as YouTube videos, where they can be delivered closer to users using cold-potato routing or on-net access. The use of caches saves ISPs a significant amount of money, because they do not have to use international transit every time someone requests the content – once it is in the cache they can deliver it using local transport. Given these savings from using a cache, ISPs do not typically charge the CDN for paid peering.

Figure 6: CDN content delivery points [Source: Analysys Mason, 2020]



A CDN can also set up points of presence (PoPs) in particular countries to deliver dynamic content. These PoPs can be used to peer with backbones and ISPs in multiple locations, and they can also be used to fill caches with static content. The CDN pays for the transport to deliver content to the PoPs, which further lowers the costs for ISPs that have access to the PoPs and caches. All of these actions both reduce the cost for the receiving network and improve the quality of service enjoyed by end users of the content.

The first CDNs were independent companies, established in regions and countries where their clients (third-party content providers) wished to deliver content. More recently, large content providers began to develop CDNs to deliver their own content, which required them to make significant investments throughout the content delivery value chain. Regardless of whether they are independent or part of a content provider, CDNs have shown that they can negotiate commercial arrangements with ISPs that both reduce costs for the ISPs and improve latency for end users.

The network developments of CDNs diminished the case for paid peering, as content could be delivered at the same IXP as it was requested, thus lowering the carriage cost for the receiving ISP. Today a number of content providers are making significant investments in network infrastructure to deliver content closer to end users.[7] This involves installing caches and PoPs, and even building subsea cables between continents and countries, in order to deliver content to PoPs and caches.

Increasingly, CDNs are putting equipment into ISP networks (also known as 'on-net'), to deliver the traffic directly. Data on interconnection arrangements published by the French communications regulator, Arcep, showed that on-net traffic made up 21% of the total volume in 2018, compared

7 David Abecassis and Richard Morgan, *Infrastructure Investment by Online Service Providers* (Analysys Mason, December 2018), http://www.analysysmason.com/Consulting/content/reports/Online-service-providers-Internet-infrastructure-Dec2018/

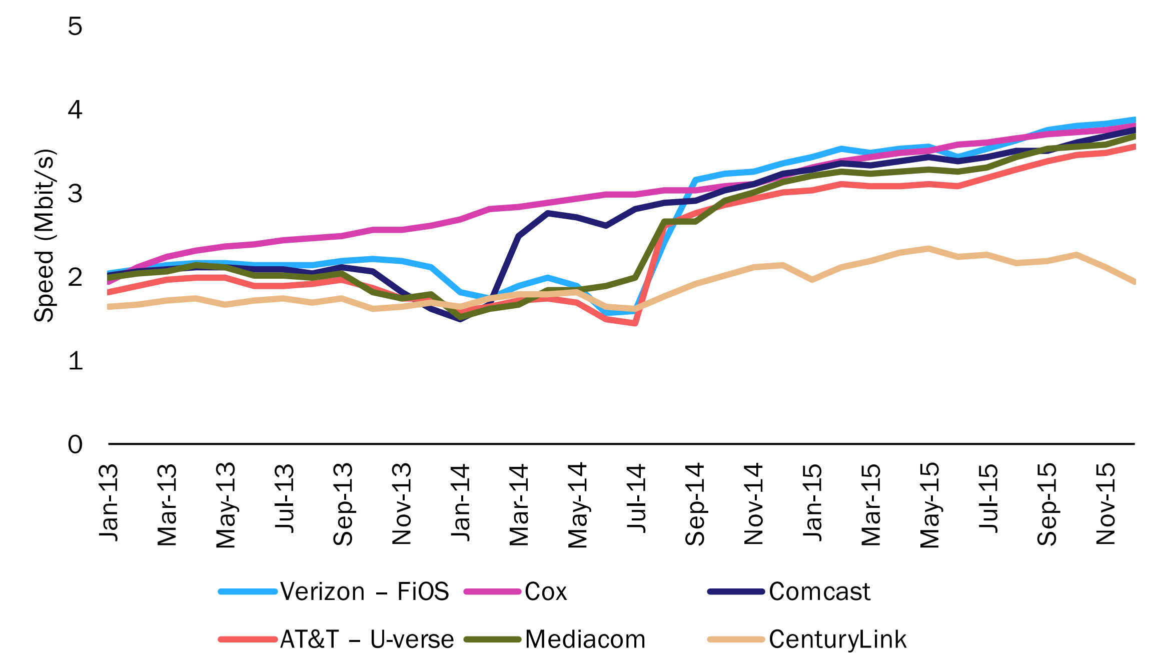
with 39% from transit and 40% from peering.[8] Although all traffic had risen since 2017, the proportion of on-net traffic in France had almost tripled, from 9% to 21%.

4 Unregulated interconnection

As CDNs began to grow and spread their content across and within an ever-increasing number of countries, content moved closer to end users and interconnection arrangements started to evolve. However, one constant was that the arrangements continued to be determined by commercial negotiations, rather than regulatory obligations. The CDNs continued efforts to build out their networks, which improved the quality of service for end users, and also lowered the cost for their ISPs. Some ISPs, on the other hand, sought to continue charging for peering. Typically, these efforts still involved commercial negotiations, but a few ISPs tried to lobby for regulatory intervention. As discussed below, these attempts to achieve regulated interconnection were not successful.

With regard to commercial negotiations, there were cases in which disputes over payments led to congested connections between content providers and ISPs, with end users experiencing slower speeds. Figure 7 illustrates how the speeds available to Netflix end users on certain ISPs slowed down owing to congestion in 2014, and then recovered as the issues were resolved, once ISPs accepted direct connections to Netflix's CDN.[9]

Figure 7: Netflix ISP Speed Index for US ISPs, 2013–15 [Source: ispspeedindex.netflix.com]



Another approach was for ISPs to seek regulatory intervention, and some lobbied regulators in favour of a variant of the 'calling party pays' model. This model is common in voice telephony,

[8] Arcep, *The state of the internet in France: 2019 report*, https://en.arcep.fr/uploads/tx_gspublication/report-state-internet-2019-eng-270619.pdf

[9] See https://groups.csail.mit.edu/ana/Measurement-and-Analysis-of-Internet-Interconnection-and-Congestion-September2014.pdf

where calls are passed between the operator of the calling party and the operator of the called party. The calling party pays for the call, and their operator pays a termination fee to the receiving operator for terminating the call. This has traditionally been justified on the basis of the benefit the calling party gains from being able to reach a target user on another network, and the peer-to-peer nature of voice communications. This model is used for the termination of international and national calls in many countries.[10]

For the internet, such arrangements are commonly referred to as 'sending party pays': the party sending IP traffic would pay for this traffic to be delivered, with flows of funds in the direction of the end user's ISP. These arrangements have largely been rejected by regulators, because they are not adapted to the situation on the internet. There are two main differences between telephony and the internet that are relevant in this context:

- First, there is a **significant regulatory difference between interconnection for telephony and for the internet**. Historically, telephone operators were regulated national monopolies in each country. The International Telecommunication Union (ITU) regulated interconnection between the operators for international calls on a calling-party-pays basis. As competition was introduced in individual countries, interconnection regulation was introduced to give competitors access to the networks of incumbent operators and allow the termination of calls between operators within a country on a calling-party-pays basis.[11]

  In contrast, backbones typically operated in a competitive environment from their very inception. They were able to access capacity from existing operators based on interconnection obligations, but did not have any regulations imposed on them.[12] This lack of regulations includes interconnection arrangements between internet providers, which could be undertaken on a commercial basis from the early days of the internet.[13] As discussed earlier, the result was a proliferation of peering relationships among backbones, which were largely settlement free.

- Second, in any case **traffic flows from a content provider to an end user are primarily initiated by the end user who wants to access the content** (as well as ancillary content that supports it, such as ads). This is not to say that the content provider does not gain any benefit from reaching end users, of course, but end users typically pay ISPs for the sole purpose of being able to request content from content providers anywhere on the internet. This demonstrates the

10 While calling party pays was the norm for fixed telephony, it was not always the case for mobile telephony. In some countries, such as the US and Singapore, the model was 'mobile party pays', in which each operator charged their own customer for making or receiving calls, and did not charge termination fees for incoming calls from other networks.

11 One significant difference between fixed and mobile is that mobile services typically began in a competitive environment, with equality between operators. As a result, mobile party pays (which is effectively like settlement-free peering between mobile operators) could be accepted by the mobile operators.

12 See, for instance, Jason Oxman, "The FCC and the Unregulation of the Internet," Office of Plans and Policy Working Paper Series (Federal Communications Commission, July 1999), https://transition.fcc.gov/Bureaus/OPP/working_papers/oppwp31.pdf

13 Michael Kende, "The Digital Handshake: Connecting Internet Backbones," Office of Plans and Policy Working Paper Series (Federal Communications Commission, September 1, 2000), https://www.fcc.gov/reports-research/working-papers/digital-handshake-connecting-internet-backbones

weakness of any analogy with voice telephony interconnection, where the calling party initiates the call and pays for it.

The resulting differences in interconnection arrangements between telephony and the internet could be quite striking. Many countries require the incumbent fixed operator to prepare an approved ‘Reference Interconnection Offer’ (RIO), which governs interconnection with competitive operators. These are complex and costly: for example, the RIO of STC, the incumbent operator in Saudi Arabia, is 37 pages long, contains 9 additional annexes, and requires many staff members within STC and the regulator to dedicate their time to preparing and administering it.[14]

In contrast, the Packet Clearing House (PCH) conducted a survey of internet networks in 2016 regarding their interconnection arrangements. Over 10 000 networks from 148 countries responded, detailing a total of 1 935 822 interconnection agreements. The results of the survey were striking. First, they revealed that 99.93% were simple ‘handshake agreements’ between the providers – a world apart from the complexity of telephony RIOs. Only 0.07% of the interconnection agreements (1347) involved a written contract.[15] Second, the survey revealed that the vast majority, 99.98%, were symmetric agreements, in which providers exchanged traffic without settlements or other requirements.[16] Only 0.02% (403 agreements) had asymmetric terms in them, including paid settlement terms.

Nonetheless, ISPs in several markets have long sought to get regulators to impose mandatory sending-party-pays charges for incoming traffic. The most notable example came in 2012, when the ITU’s International Telecommunication Regulations (ITRs) were being updated at the World Conference on International Telecommunications (WCIT). The European Telecommunications Network Operators Association (ETNO), the trade association representing large and incumbent telecom operators in Europe, put forward a proposal to include sending party pays in the new regulations for international internet traffic. The proposal was taken up by some countries, but not approved by the Member States of the ITU and thus not included in the ITRs.

Efforts to impose payments on the senders of traffic should be mindful of the dramatic changes that have occurred since the early days of the internet. At that time, content providers did not have a significant source of income, as users were not willing to purchase content, and payment mechanisms as well as advertising models were non-existent or immature. As ISPs did have a stable source of recurring revenue, content providers sought to share in that revenue by floating the argument that they were providing value to ISPs’ end-user customers, and thus deserved

14 See https://www.citc.gov.sa/en/RulesandSystems/RegulatoryDocuments/Interconnection/Pages/CITC_ReferenceInterconnectionOfferRIO.aspx

15 This is similar to an earlier survey by PCH involving 142 000 peering agreements, which found that 99.5% were handshake agreements, as reported in an OECD report by D. Weller and B. Woodcock (2013-01-29), “Internet Traffic Exchange: Market Developments and Policy Challenges”, *OECD Digital Economy Papers*, No. 207, OECD Publishing, Paris.

16 See https://www.pch.net/resources/Papers/peering-survey/PCH-Peering-Survey-2016/PCH-Peering-Survey-2016.pdf

compensation for driving demand for retail ISP services. Clearly, this was not accepted as a commercial basis for interconnection, either by ISPs or by any regulators.

Later, as content providers developed a successful revenue model, the arguments on revenue sharing flipped, as illustrated by discussions on sending party pays at the 2012 ITU WCIT. In this case, ISPs argued that content would have no value if it could not reach end users, and thus that content providers should compensate them through sending party pays.[17] As noted earlier, this argument was unsuccessful at the WCIT, and the change in perceived value flows demonstrates the danger of a regulatory solution. If the original revenue-sharing proposal put forward by content providers had been adopted, it would have been difficult to remove or reverse. In contrast, commercial negotiations and business models have been able to accommodate these changes.

5 Impact of interconnection regulation in South Korea

Interconnection arrangements between internet providers have evolved to deliver overlapping benefits – including lowering the cost of traffic delivery, reducing or avoiding congestion, and optimising latency. Settlement-free peering was adopted as a means to directly exchange traffic without termination charges. IXPs were then introduced as a means to interconnect more efficiently, in part to avoid costly and slow 'tromboning' arrangements. These benefits derived from interconnection arrangements that were freely negotiated rather than regulated – but regulation of interconnection in South Korea means these benefits for South Korean end users are in jeopardy.

Content providers have made significant investments to ensure improvements in traffic delivery continue apace. CDNs (both independent and those that belong to content providers) deliver traffic closer to ISPs, often as members of IXPs, or may even put equipment into ISPs' networks to deliver content. These CDNs reduce the ISPs' dependence on expensive backhaul capacity. Content providers have also begun to invest directly in submarine cables, to deliver traffic around the globe using their own capacity. Again this improves the latency of content delivery, lowers costs for ISPs, and provides significant benefits for end users.

All of these trends are a response to changes in the nature of content and evolving business models. Introducing mandatory regulated charges between ISPs for delivering traffic threatens to disrupt the evolution of interconnection arrangements, and prevent the full benefits reaching consumers. This is true irrespective of whether the charge is for each bit, or takes account of inbound/outbound ratios (as was recently announced in South Korea). Such a move could increase costs for content providers, and thus affect content delivery in a variety of ways.

Local content providers (e.g. Naver, Kakao) typically host their content in South Korea itself, and pay for connectivity through an ISP. If that ISP has to pay to deliver the content to other ISPs, it will pass those charges onto the content providers. The effect will be to increase the content providers' costs compared to those of local content providers in unregulated markets. In particular, it will be difficult for smaller content providers or start-ups to enter the market and grow. An alternative

17 Geoff Huston, "Carriage vs. Content," *The ISP Column* (blog), July 2012, http://www.potaroo.net/ispcol/2012-07/carriagevcontent.html

solution is for the content provider to provide caches to all the ISPs, but that is not feasible for all content providers, and certainly not start-ups.

The impact of interconnection regulation on international content providers is different, as they already operate a network outside South Korea. Overall, as discussed earlier, large content providers are actively seeking to build out their networks closer to end users. This improves the quality, latency, and availability of traffic delivery to end users, at no extra cost to ISPs. However, if an international content provider wants to build its network out to South Korea, it has to deliver content to end users through local ISPs.

International content providers have two options. First, they can interconnect directly with each ISP in the country through PoPs in their network, to avoid sending-party-pays charges among the ISPs. Second, they can decide to only make content available outside the country, where South Korean ISPs would have to pick it up. This would prevent South Korean ISPs from sending the content between them, and thus avoid them paying any content-related charges. However, smaller and larger ISPs appear to have different incentives at this point:

- *Smaller ISPs* generally benefit most from content providers' networks being present in South Korea, and they appear to have agreed to collect traffic from content providers without seeking compensation in the form of a fee; however, the content providers may not wish to provide a cache to each small ISP.
- *Larger ISPs* are actively seeking to negotiate the payment of a fee by content providers, arguing that if ISPs have to pay one another to exchange traffic, then so should content providers with a network PoP in South Korea.

As a result, some international content providers have decided to only exchange traffic outside South Korea, either elsewhere in Asia or in the US. This suppresses content providers' investment in South Korea, both the installation of edge network elements (including caches and PoPs), and possibly subsea cables and other forms of infrastructure. As well as missing out on the direct benefits of such investment – which are instead going to other markets that are more open – South Korea will also be foregoing indirect economic benefits.[18]

This imposition of regulated payments increases the cost of traffic for ISPs, as well as the latency of traffic exchanges in many cases – both of which are detrimental to end users. Figure 8 shows where ISPs in South Korea are exchanging traffic offshore, including likely with CDNs that might otherwise exchange traffic in South Korea.

18 For an estimate of the economic benefits of Facebook connectivity initiatives, including network infrastructure, see: David Abecassis, Michael Kende, Elena Korsukova, Richard Morgan, Sviat Novik, *The impact of Facebook's connectivity initiatives in the ASEAN region* (Analysys Mason, May 2020), https://www.analysysmason.com/Consulting/content/reports/impact_of_facebook_connectivity_asean

Public Peering Exchange Points

Filter

Exchange / ASN	IPv4 / IPv6	Speed / RS Peer
AMS-IX 4766	80.249.209.129 2001:7f8:1::a500:4766:1	10G ○
DE-CIX Frankfurt DE-CIX Frankfurt Peering LAN 4766	80.81.192.170 2001:7f8::129e:0:1	10G ○
Equinix Los Angeles 4766	206.223.123.44 2001:504:0:3::4766:1	20G ○
Equinix Palo Alto 4766	198.32.176.102 2001:504:d::66	20G ○
LINX LON1 Main 4766	195.66.224.147 2001:7f8:4::129e:1	10G ○
NYIIX 4766	198.32.160.49 2001:504:1::a500:4766:1	20G ○
SIX Seattle MTU 1500 4766	206.81.80.131 2001:504:16::129e	30G ○

Private Peering Facilities

Filter

Facility / ASN	Country / City
Equinix Los Angeles (LA1) 4766	United States of America Los Angeles
Equinix Palo Alto (SV8) 4766	United States of America Palo Alto
Equinix Seattle (SE2/SE3) 4766	United States of America Seattle
Equinix Singapore (SG1) 4766	Singapore Singapore
Flexential - Portland/Hillsboro 2 (PDX02) 4766	United States of America Hillsboro
HKCOLO - Sino Favour Centre 4766	Hong Kong Hong Kong
MEGA-i (iAdvantage Hong Kong) 4766	Hong Kong Hong Kong

Figure 8: Korea Telecom peering points [Source: PeeringDB, 2019]

These arrangements incur significant transit costs for ISPs. These 'offshore' peering relationships provide a non-regulated avenue for South Korean ISPs to collect the traffic that their customers request from content providers, outside the reach of South Korean regulations.

In addition, the current situation impedes the development of a stronger IXP in South Korea, where more traffic exchange could take place. An IXP would provide an efficient location for content providers to connect to ISPs, in one of several ways.

- A CDN could attach a cache directly to the IXP in order to distribute content to connected ISPs, or it could connect indirectly through an ISP
- A CDN could establish a PoP at the IXP, again to peer with ISPs for traffic exchange
- IXP nodes are often hosted in data centers, where content providers could establish a direct connection with an ISP, sometimes known as a private network interconnect (PNI).

All of these options would lower the costs for ISPs to access content, as well as improving the latency. However, with content providers facing the possibility of having to compensate ISPs for receiving traffic, these connections will not widely develop around an IXP.

Although some of South Korea's smaller ISPs have installed caches inside their networks, which they fill through connections to content provider networks outside the country, this only partly mitigates the downside of the current situation. We understand that larger ISPs have resisted installing caches at various points in time, to try to force content providers to pay for their traffic to be delivered in South Korea (as they then suffer from more-constrained or lower-quality bandwidth for their content). The result is a lose–lose situation for South Korean stakeholders: ISPs have to pay to backhaul traffic from abroad, while content providers may see less engagement with their services in South Korea if backhaul capacity is insufficient.

Finally, the protracted ongoing discussions on the imposition of content charges, as well as ambiguity about how they would be imposed, creates uncertainty for all providers. This uncertainty can affect investment decisions, irrespective of how, or even whether, the underlying regulations are implemented. Again, commercial negotiations – as relied on elsewhere – would deliver better results, by finding the most efficient solution in terms of both cost and latency, and helping to stimulate investment in South Korea.

6 Conclusion

The internet has undergone significant changes since its commercial beginnings in the 1990s. In addition to remarkable growth rates, it has globalized from its historical roots in the US, the volume of traffic associated with content has grown exponentially, and content providers have begun to deploy their own networks to deliver the content to end users worldwide. The one constant has been that commercial negotiations, rather than regulation, have driven changes in interconnection arrangements.

In general, internet interconnection arrangements have not been regulated. At a global level, an attempt to impose 'sending party pays' through ITU regulations was not adopted. At a national level, the situation in South Korea helps to demonstrate why: regulating internet interconnection either imposes additional costs on content providers (which have to build out their network to multiple ISPs in South Korea and possibly pay them to receive the content), or results in content being made available abroad, thus increasing the costs for South Korean ISPs to access content requested by their end users. Either way, the possibility of regulation and ambiguity about how it will be imposed has had an impact on investment. In addition, the resulting interconnection arrangements are less efficient than the commercial outcomes in other countries, and impose increased costs and latency on South Korean end users with no mitigating benefit.

Unlike in traditional telephony, there is no clear economic rationale for imposing termination charges. End users pay their ISPs for internet access, for which they expect to get access to all of the internet. Since the very start of the internet, commercial negotiations have proven effective at maximizing the incentives for all providers to carry traffic efficiently, while adapting to changes in traffic patterns. Regulators and policy makers around the world have consistently rejected the idea that content providers should pay mandatory usage fees to ISPs. Developments to date in South Korea demonstrate that deviating from this established practice brings no benefit, and imposes costs on South Korean consumers, businesses and content providers.

Policy makers in South Korea and elsewhere define policy and regulations by assessing the evidence available to them. In this case, the need for regulation is unclear and its impact appears overwhelmingly negative.

Annex: terminology

Backbone	An internet backbone provider has a network that carries traffic within regions, across countries, or around the world. Its customers include large enterprises, ISPs, and content providers, and it exchanges traffic between them and with other backbone providers. It uses *peering* arrangements with other backbone providers, and sells *transit* to its customers.
Cache	A cache is used to store static content such as videos for a CDN. Popular content can be 'pushed' into the cache by the CDN, or it can be 'pulled' into the cache by an end user when they ask for the content. Once in the cache, content is closer to the end user so that it can be accessed with less latency, and it lowers costs for ISPs (as they do not have to use expensive international transit to access the content each time it is requested).
Content delivery network (CDN)	A CDN is a company that distributes content on behalf of a content provider, using caches and PoPs. The CDN may be independent of the content providers, or it may be developed by a content provider to distribute its own content.
Dynamic content	Dynamic content is content that is personal to the user and changes over time. Emails, customized webpages, and video conferences are all examples of dynamic content.
Internet exchange point (IXP)	An IXP is a location where internet providers, including ISPs, backbones, CDNs, and enterprises, can meet and efficiently exchange traffic, using peering or transit arrangements. Instead of having to arrange separate circuits for each provider, each provider only needs one connection to the IXP to exchange traffic with all the other providers. An IXP can also help to attract content, by providing an efficient way to distribute the content.
Internet service provider (ISP)	An ISP provides internet access to end users, using mobile or fixed broadband connections, and can also provide access to enterprises.
Latency	Latency is the amount of delay involved in sending traffic between points on the internet. It is the delay experienced while downloading a webpage or streaming a video. It is also the delay experienced by end users communicating with one another. Latency can cause a video to pause while it is being viewed, or it can make it difficult to conduct an internet phone call.
Peering	Peering is an arrangement between two internet providers to exchange their own traffic. For instance, two backbones can use peering to exchange their ISP customers' traffic, or an ISP can peer with a content provider to deliver the content requested by its end users. Peering partners will not allow peering traffic to *transit* their networks – that is they will not deliver content from one

	peering partner to another. Peering is typically 'settlement free', meaning that neither party pays or is paid to exchange traffic with the other.
Point of presence (PoP)	A point of presence is a location where a CDN can deliver dynamic content to a backbone or ISP.
Static content	Static content does not change depending on the user or over time. Static content can be copied and stored in caches all over the world, to be delivered to end users. Examples include songs and videos.
Transit	In a transit arrangement, a backbone will sell access to the entire internet to another backbone or ISP. This includes access to the backbone's own customers as well as the backbone's peering partners.
Tromboning	In a country without an IXP, ISPs will often exchange all of their traffic abroad using international circuits, including traffic that is destined for a user in the same country. This is more efficient than establishing separate links to each other ISP. It is called tromboning because the traffic will leave and re-enter the country in a shape similar to that of the musical instrument.