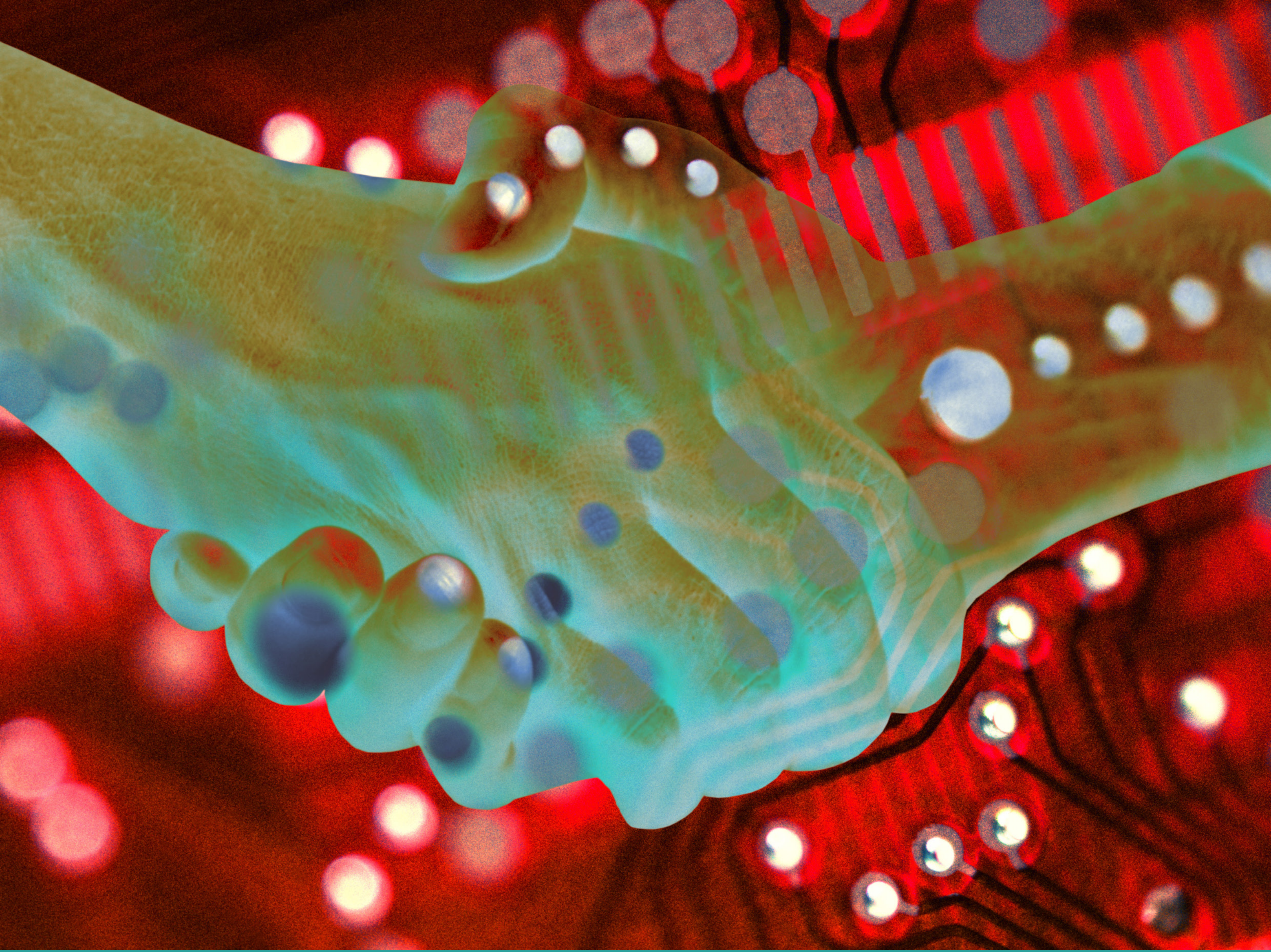


The Future of the Internet

Innovation and Investment in IP Interconnection

May 2014

Arthur D Little

LIBERTY GLOBAL

Content

Foreword

Scope of this study: "The Internet" and "IP Interconnection"

The future of the Internet is a widely debated public policy theme all over the world. Questions are raised on how to preserve the public "best-effort" Internet as an "open" platform for innovation and competition, and how to combine economic- & societal value creation and sustainable returns on investment. Although so far the history of the Internet has been an incredible success in organically developing a self-adapting complex of network business relations, concerns are raised about how the Internet will be able to sustain an adequate quality of experience for the end-user in the future. This ability may come under pressure by a spectacular boom in Internet traffic volumes in the coming years, resulting in unprecedented demand for reliable, ubiquitous Internet access and mass uptake of bandwidth-intensive services and applications. To illustrate this point: by 2020, more than 50% of the world's population will be online. This means an increase from 2.7 billion users in 2014 to 5.0 billion users by 2020. By 2025, "The Internet of Things" will comprise around 50 billion connected devices. By 2030, machine-to-machine ("M2M") communication is expected to constitute more than 50% of IP traffic.

The question therefore seems justified as to whether the Internet can cope with this evolution, and who and what is needed for the Internet to evolve and adjust to these changing circumstances.

One part of the answer lies in capacity, quality and traffic management in the Internet access network (fixed or mobile), or so-called "last mile", owned or operated by an Internet Access Provider over which end-users access the Internet. This part is the subject of "net neutrality" discussions and mainly covers the front end, consumer-facing side of the Internet.

The other part of the answer lies in the so-called "up-stream" side of the Internet. This is where the Internet access networks connect with (i) each other, (ii) bulk IP traffic transportation networks and undersea cables connecting continents and (iii) content & application server parks located across the globe. This "IP Interconnection" part of the Internet solely consists of wholesale agreements, which determine the technical & economic conditions under which IP traffic is delivered from the originating party (for example, a content & application provider or an ISP). This is done via several exchangeable delivery networks of multiple Internet connectivity providers (often used in parallel) to the residential Internet access networks of terminating ISPs, and vice versa.

IP Interconnection is, and has been, an essential building block for the quality and functionality of the Internet as ultimately experienced by the end-user, despite the fact that the end-user is no party to IP-Interconnection arrangements. IP Interconnection models have adapted to changes in Internet usage and traffic patterns caused by disruptive applications or technologies (for example, predominantly digital distribution technology moving from decentralized peer-to-peer to centralized streaming) and facilitated IP content delivery accordingly. In many ways, the extent to which the IP Interconnection sector is able to innovate itself defines the scope of evolution of the Internet as a platform for future applications.

Therefore, IP Interconnection developments have an impact on overarching objectives in Internet public policy debates that, essentially, focus on warranting end-users' quality of experience over the public Internet, or "Best-Effort" Internet.

With this report, we want to unravel some of the complexity in IP Interconnection and identify the main drivers of change in IP Interconnection. We also analyze the effects of investment and innovation by the IP Interconnection players on the future capability of the public Internet, end-user quality of experience, competition and the scope for new Internet applications.

In order to achieve this, the study first analyzes the latest stage in the evolution of the Internet, and then the corresponding trends in IP Interconnection reflecting this evolution, asking if the IP Interconnection players will be able to continue to reach innovative interconnection business models and participate in the Internet value chain. Finally, the study analyzes how different IP Interconnection scenarios may affect the future of the public "Best-Effort" Internet, which is linked to unlocking the potential of innovation and new, exciting application paradigms such as the "Internet of Things" and the "Internet of Humans".

Sincerely,

Gregory Pankert
Partner
TIME* practice
Arthur D. Little

Andrea Faggiano
Principal, Head of SASCAR**
TIME* practice
Arthur D. Little

Karim Taga
Global Practice Head
TIME* practice
Arthur D. Little

* TIME: Telecommunication, Information Technology, Media and Electronics
**SASCAR: Strategic Advisory Services for Competition and Regulation

Key Messages

The Internet is vital, continuously evolving and has developed into a new media platform

1. The Internet has been transformed into a new media platform, as the nature of Internet traffic has changed from static data & text file transfer to streaming interactive media content.
2. The Internet has become mission critical for most Content & Application Providers. Minor disturbances in the quality of delivery directly impact the willingness of end-users and advertisers to pay for online services.
3. The future development of the Internet as a media platform is impacted by increasing global connectivity, proliferation of smart devices and streaming media services which cause spectacularly higher traffic volumes, greater imbalances in traffic flow and changing traffic patterns.
4. IP Interconnection is an essential building block for the quality & functionality of the Internet as ultimately experienced by the end-user, despite the fact that the end-user is no party to business-to-business IP Interconnection arrangements.

IP Interconnection, so far, adapted well to support the changing nature of the Internet, and remains dynamic and competitive

5. The IP Interconnection value chain converges, but remains dynamic and competitive. Proliferation of Content Delivery Networks and Internet Exchanges, commoditization of IP transit and CDN prices challenge existing interconnection models and enable new ones.
6. From the early days of "IP transit" and "Peering", a genuine mix of viable application/content delivery strategies is accessible to all players seeking connectivity.
7. Content & Application Providers and ISPs are setting the pace and determining the nature of IP Interconnection innovation by vertically integrating and interconnecting directly, which disintermediates pure Internet connectivity providers to some extent.
8. Changes in the IP Interconnection ecosystem lead to tension between IP Interconnection players. However, disputes concern less than 1% of all IP Interconnection agreements and are solved without regulatory intervention in more than half of these cases.
9. End-users have not been substantially or structurally affected by IP Interconnection disputes.

Future applications require IP Interconnection models to evolve towards providing higher-quality assurances, which will impact the current "best-effort" Internet

10. Innovation in IP Interconnection is needed to support further development of the Internet and accelerate take-up of next-generation applications (Internet of Things, Internet of Humans) that require IP Interconnection Quality of Service (latency, jitter, packet loss) extended with new parameters (e.g. security, data protection).
11. Variants of Paid Peering, Deep Caching, Assured Delivery or Secure M2M are among the innovative IP Interconnection business models that could lay the foundation for an advanced Internet platform, based on assured end-to-end Quality of Service Internet Platform – complementary to "Best Effort".
12. "Best-effort" Internet is and will no doubt continue to be essential in the future, and there is early evidence to indicate that it can continue to improve and coexist with complementary end-to-end Quality of Service platforms if properly monitored.
13. Private investment in IP Interconnection has led to structurally improved conditions for the future development of the public Internet. Content comes closer to end-users (by direct interconnection and local content caching), Internet performance is improved by adoption of new application technologies (e.g. "adaptive streaming") and IP network resources are abundant (e.g. higher capacity in the "last mile").

Executive Summary

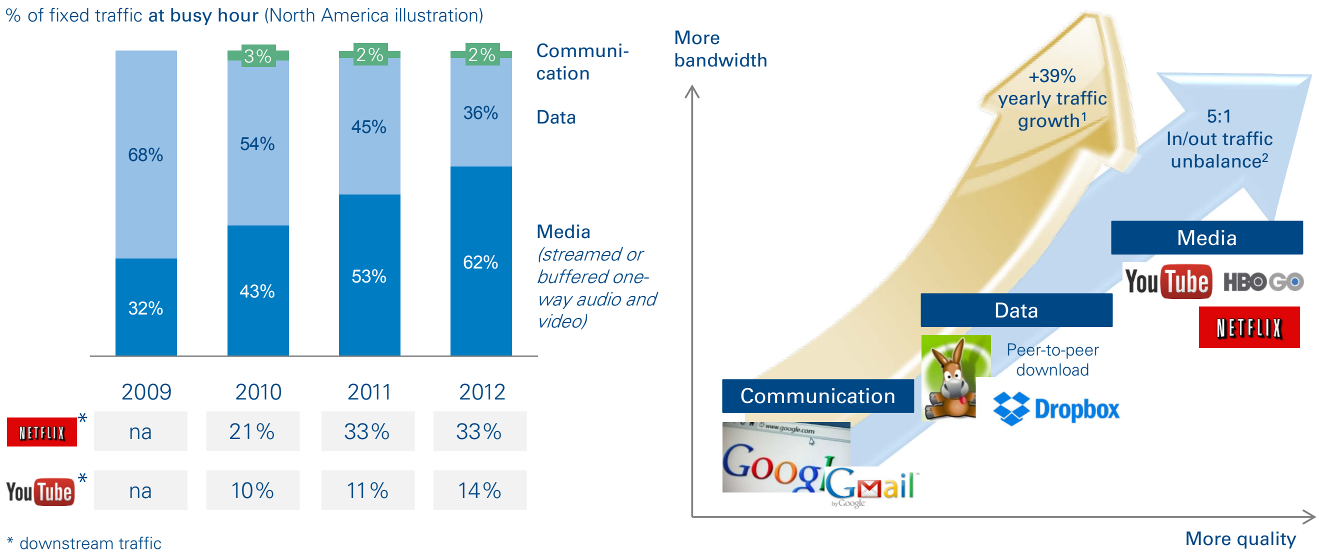
The Internet is vital, continuously evolving and has developed into a new media platform

Over the last decades, the nature of Internet traffic has changed from static data & text to interactive media content, effectively transforming the Internet into a new media platform as its usage shifted to richer types of content, particularly streaming video.

The future development of the Internet as a media platform is nowadays challenged by increasing global connectivity, proliferation of smart devices and streaming media services, which cause spectacularly higher traffic volumes, greater traffic imbalances and changing traffic patterns. Internet traffic doubles almost every two years, and traffic patterns have changed as a result of real-time streaming overtaking peer-to-peer as the predominant form of digital distribution. Internet access networks experience significant in the order of 5:1 imbalances between incoming and outgoing traffic because of the media-related nature of traffic, which mainly flows one way, from content providers to end-users.

In 2014, the Internet reached over 2.7 billion individuals and has become mission critical for most Content and Application Providers. Indeed, minor disturbances in the quality of delivery directly impact the willingness of end-users and advertisers to pay for online services.

Figure A - The changing nature of the Internet and the new capacity/quality requirements



* downstream traffic
COMMUNICATION: services provided by application (Skype, WhatsApp, iMessage, FaceTime,etc.); DATA: file sharing (Bit Torrent, eDonkey, etc), web browsing, social networking, email, etc.; MEDIA: streamed or buffered audio and video (Netflix, non-linear TV services;
1. 2009-2012 CAGR; 2. Interviews. Source: ITU, Sandvine; Arthur D. Little analysis

IP Interconnection, so far, adapted well to support the changing nature of the Internet, and remains dynamic and competitive

IP Interconnection is an essential building block for the quality & functionality of the Internet as ultimately experienced by the end-user, despite the fact that the end-user is no party to business-to-business IP-Interconnection arrangements. IP-Interconnection models adapt to changes in Internet traffic patterns caused by disruptive applications, or technologies, and facilitate IP content delivery accordingly. In many ways, the extent to which the IP Interconnection sector is able to innovate itself defines the scope of evolution of the Internet as a platform for future applications.

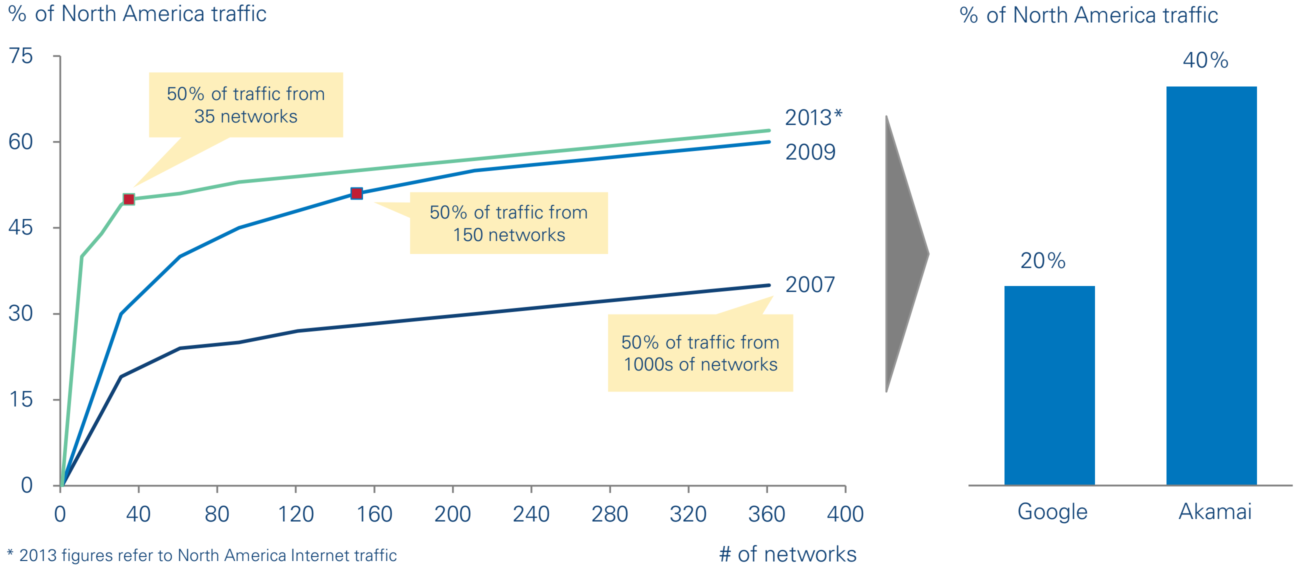
The Internet and the underlying IP Interconnection ecosystem demonstrated an organic ability to evolve and adapt. Alternative business models (such as peering and Content Delivery Networks) challenged existing ones and improved the overall efficiency of IP Interconnection, leading to a cost reduction of around 30% per annum over the last decade.

From the early days of "IP transit" and "Peering", a genuine mix of IP Interconnection models is currently available to both ISPs and Content & Application Providers (CAPs) seeking connectivity. This is the result of three major developments:

1. **Decentralization of the Internet:** the emergence of national and regional Internet Exchanges facilitates private peering arrangements by increasing number of ISPs' access networks edges in one central location.
2. **Commoditization of IP Interconnect prices** (falling IP transit, CDN or router costs) led to substitutability of IP Interconnection products and countervailing powers in the IP Interconnection value chain.
3. **Proliferation of Content Delivery Networks:** Content & Application Providers leverage the increased value of their Internet content by building proprietary caching server parks or, alternatively, using independent, commercial CDN services that are located close to the ISPs' access networks.

Still, the majority of Internet traffic is progressively being concentrated to a limited number of large Content & Application Providers and a few wholesale carriers. In 2013, 35 networks carried 50% of all Internet traffic in North America, down from 150 networks in 2009. The concentration of IP traffic is a major evolution in the IP Interconnection value chain, and has the potential to influence the negotiating power among connectivity stakeholders and affect the current equilibrium in the Internet ecosystem.

Figure B – Traffic concentration trends (2007, 2009, 2013)

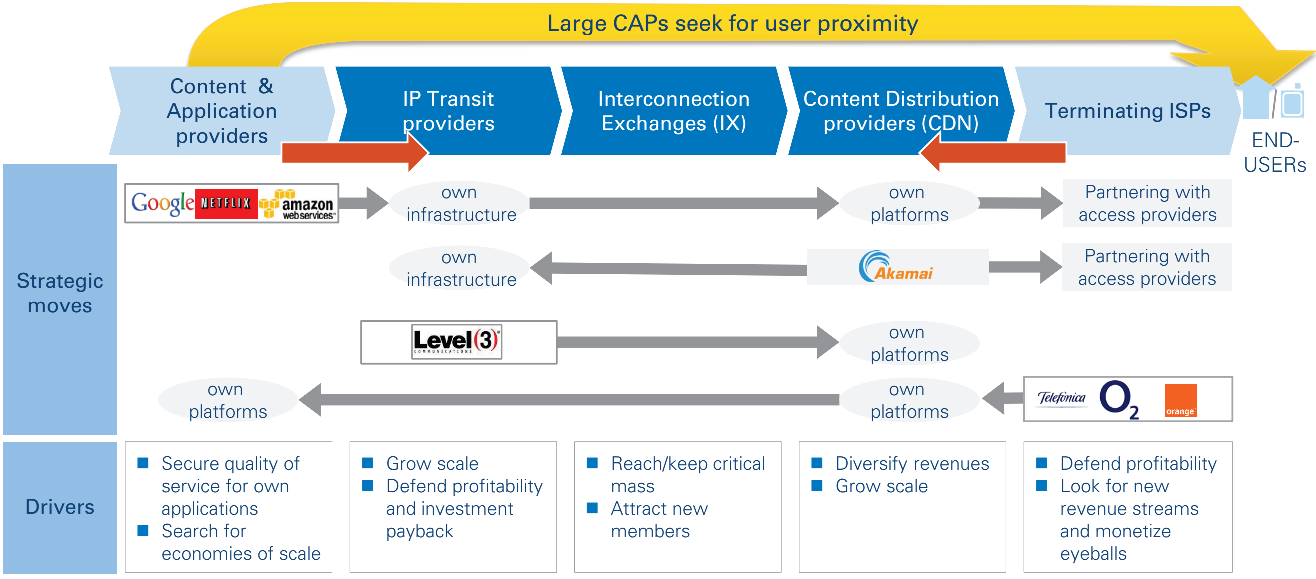


* 2013 figures refer to North America Internet traffic
Source: Deepfield; Arthur D. Little analysis

In the last years, the largest Content & Application Providers and Internet Service Providers have been setting the pace and determining the nature of IP Interconnection innovation through vertical integration. Content & Application Providers seek end-user proximity and are increasingly investing in proprietary Content Delivery Networks or relying on third-party CDNs. ISPs invest in network-based content delivery platforms ("deep caching") for internal purposes and as a service to third-party Content & Application Providers.

As a result, Content & Application Providers and Internet Service Providers increasingly interconnect directly, disintermediating pure Internet connectivity providers to some extent. Improving control on the quality of delivery over the Internet is the main motivation. This is true not only for Internet-based CAPs, but also increasingly for the video-streaming strategies of traditional broadcasters (e.g. BBC iPlayer's average daily unique users grew 33% year on year since 2009). The equilibrium in the IP Interconnection value chain has subsequently changed, and traditional IP Interconnection players have adapted to maintain their competitiveness. Internet

Figure C – Trends over the IP Interconnection value chain



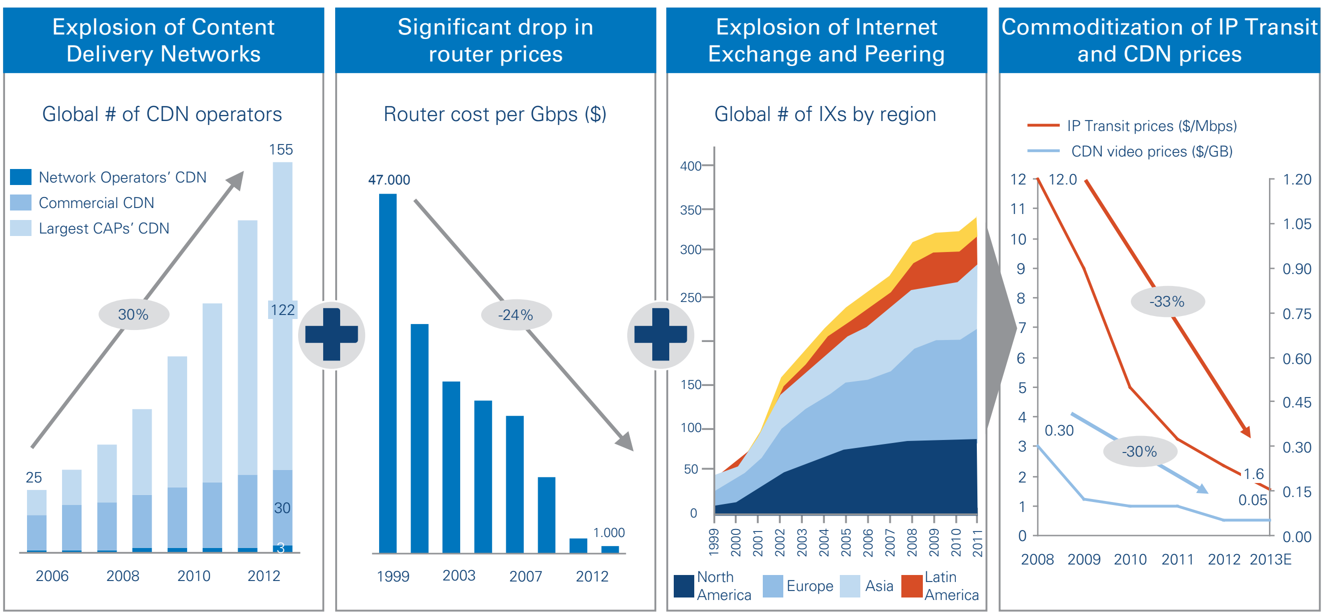
Source: Arthur D. Little analysis

connectivity providers such as IP Transit providers, independent CDN providers and Internet Exchanges are under pressure to innovate and diversify their service offerings (e.g. offering "partial transit", commercial open CDNs or web security) or attract a critical mass of traffic through consolidation (e.g. international carrier Level 3 acquiring its competitor, Global Crossing).

Changes in the IP Interconnection ecosystem meant that tensions between IP Interconnection players intensified. However, disputes concern less than 1% of all IP Interconnection agreements, and are solved without regulatory intervention in more than half of these cases. There are a number of reasons for this:

- IP Interconnection (including upgrade) costs account for just a marginal share, i.e. less than 1% of the overall connectivity costs.
- Countervailing powers emerged by changing the IP Interconnection economics that keep the value chain in balance:

Figure D – Economic drivers of IP Interconnect evolution and countervailing powers



Source: ITU, Informa, Packet Clearing House, Dr. Peering, Cisco, streamingmedia.com, Web sites, Arthur D. Little analysis

a) IP Transit and Peering have become substitutes in terms of cost
b) Falling IP Transit and CDN prices make high-quality transport and CDN strategies accessible to smaller CAPs
c) Falling IP Transit prices balance against paid-peering cost pressures
d) Strong retail competition prevents market foreclosure by ISPs

End-users have not been substantially or structurally affected by IP Interconnection disputes. The commercial interest of parties prevailed, and mutually acceptable solutions were found. Associated Interconnection costs have not proven to be prohibitive to core business models so far.

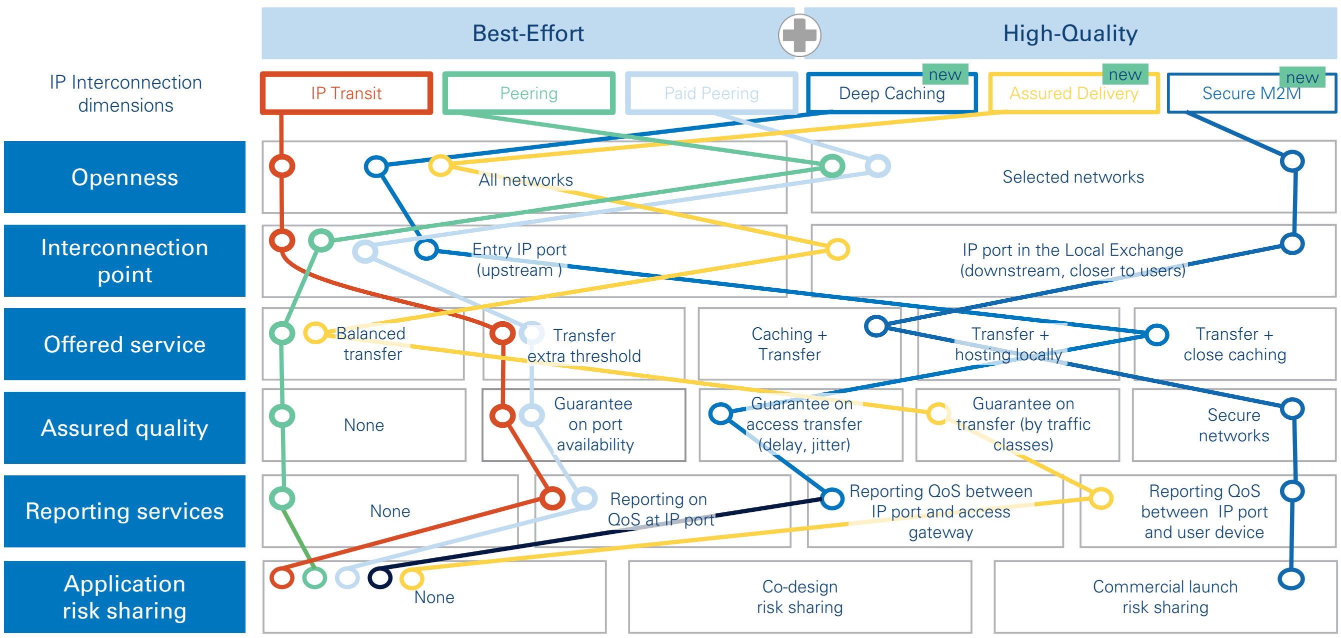
Future applications require IP Interconnection models to evolve towards providing higher-quality assurances, which will impact the current "best-effort" Internet

In addition to continuous improvement of connectivity between Content & Application Providers and access networks, innovation in IP Interconnection can support further development of the Internet and accelerate the take-up of next-generation applications that require uncompromised quality. In particular, the Internet of Things and the Internet of Humans application landscapes can unlock an economic value potential in the range of trillions of euros by 2020. However, advanced Internet platforms, i.e. beyond Best-Effort, may be required for next-generation applications, which could bring an Internet of Things and an Internet of Humans to life.

IP Interconnection Quality of Service needs to be extended to new parameters (e.g. latency, jitter, packet loss, security, and data protection). As the Internet evolves from nice-to-have services to mission-critical services, next-generation applications for sectors such as the Financial Services industry, the Electronic Payment sector, high-security Governmental Bodies (police, military, emergency services, etc.) will generate a demand for new IP Interconnection requirements going well beyond additional throughput capacity. It will expand to delivery features relevant for streaming video, such latency reduction, availability, jitter control and packet-loss limitation. Security and data protection deserve special attention as they play a critical role in the safe use of next-generation applications, especially in scenarios foreseeing the Internet of Things leading to M2M applications such as connected cars with remote-start features.

Variants of Paid Peering, Deep Caching, Assured Delivery and Secure M2M are among the innovative IP Interconnection business models that could lay the foundation for an advanced Internet platform, based on an assured end-to-end Quality of Service Internet platform – complementary to Best-Effort.

Figure E – Innovation in IP Interconnect business models



Source: Arthur D. Little analysis

Still in the advent of new IP Interconnection business models, the Best-Effort Internet is, and will no doubt, continue to be essential in the future, and there is early evidence to indicate that it can continue to improve and co-exist with complementary end-to-end quality of service platforms if properly monitored. Best-Effort has long co-existed with business-to-business IP managed services, as well as with ISPs' managed IPTV platforms, and its average and peak connection speeds still increased by respectively 12% and 23% since 2007, increasing to 21% and 26% since 2011.

Private investments in IP Interconnection resulted in a number of trends with structurally improved conditions for the future development of the public Internet:

a) Content comes closer to end-users: Direct interconnection with fewer Internet connectivity providers allows for better structural conditions on Quality of Service - i.e. lower latency, lower risk of packet loss and jitter.

b) New application technologies improve performance: This includes codecs, adaptive streaming and content distribution algorithms.

c) Abundance of IP network resources: Capacity in the "last mile" became larger. IP transit and Content Delivery services were being commoditized, and new quality delivery opportunities were created, with network-based "deep-caching" technologies.

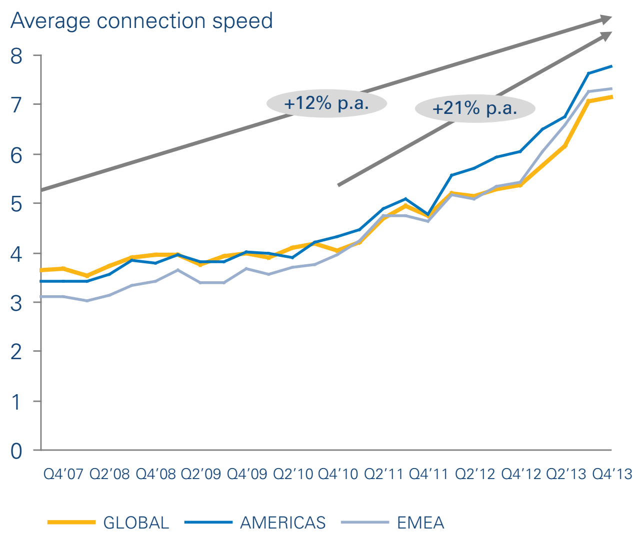
d) High dynamism in the Internet ecosystem: The high value at stake for all stakeholders allows new disputes to be resolved quickly and create new business relationships.

Figure F – IP Interconnection structurally contributing to Internet's improvement

IP Interconnection trends: better structural conditions

- **Content comes closer to end-users**
  (direct interconnection with fewer intermediaries, allowing for better structural conditions on QoS: lower latency, lower risk of packet loss and jitter)
- **New application technologies improve performance**
  (codecs, adaptive streaming, content distribution algorithms)
- **Abundance of IP network resources**
  (larger pipes in the last mile, IP Transit and Content Delivery services being commoditized, new opportunities with deep caching technologies)
- **High dynamism in the ecosystem**
  (the high value at stake for all players allows the new disputes to be resolved rapidly and new business relationships to be created)

Best-Effort Internet is continuously improving



Source: Akamai, Arthur D. Little analysis

The public Internet will stand to benefit mostly from private investments in IP Interconnect architecture aimed at shortening the distance that Internet traffic needs to travel before it reaches the "last-mile" Internet access networks. This is accomplished by storing popular content/applications in local servers that form part of proprietary or commercial CDNs, or in network-based "deep-caching" servers. Shorter travel distances for IP content with fewer intermediaries implies less chances for "bumps" in the road. It also increases the prospects for more manageable end-to-end controls, leading to an overall higher quality of experience for the end-user.

1. The Internet is Vital and Continuously Mutating

1.1. The Internet is now mission critical commercially, and calls for quality delivery

The Internet rapidly and dynamically evolved from an experimental network to a mass-market interaction platform

The story of the Internet began in 1957, when the USA responded to the USSR, launching Sputnik into space with the creation of the Advanced Research Projects Agency (ARPA). The Agency's mission was to become the leading force in science and new technologies. Around 1994 the Internet was used for the first time as an open commercial platform, enabling the launch of the first Internet ordering system (Pizza Hut) and the first Internet Bank (First Virtual). From that point on, all the major Internet companies emerged: Google was launched in 1998, MySpace in 1999, Apple's iTunes Store in 2003, Facebook in 2004, YouTube in 2005 and Twitter in 2006. Netflix (created in 1998) started to offer online streaming services in 2008. Soon, their rapid growth and potential success would bring these Internet companies to become listed companies, Google showing the lead in 2004. Facebook (2012) and Twitter (2013) made remarkable entries to the stock market.

From its birth in 1969 to nowadays, the Internet has evolved and mutated in many different ways:

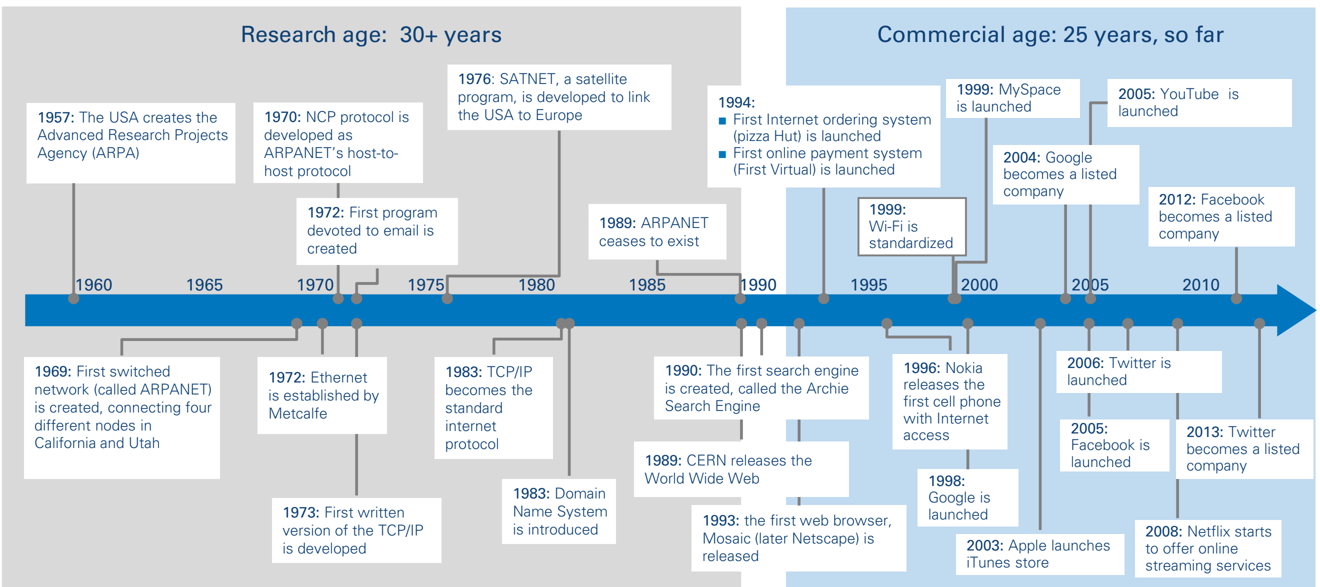
- Internet usage ceased to be a US-centric phenomenon and achieved a global reach;
- Traffic boomed and regional poles emerged, while the nature of Internet content progressively changed from static text and simple data to interactive media and entertainment;
- The way of accessing the Internet shifted from dialing in via fixed networks to an always-on mobile experience
- New network requirements emerged, and quality of delivery – which, in the early stages was not that important – became mission critical.

The popularity of the Internet grew substantially, and the penetration of Internet users within the world's population strongly increased.

Over the 2005-2013 period, the number of individuals using the Internet[1] grew yearly at 13%, from 1 billion to almost 3 billion, mainly outside the US.

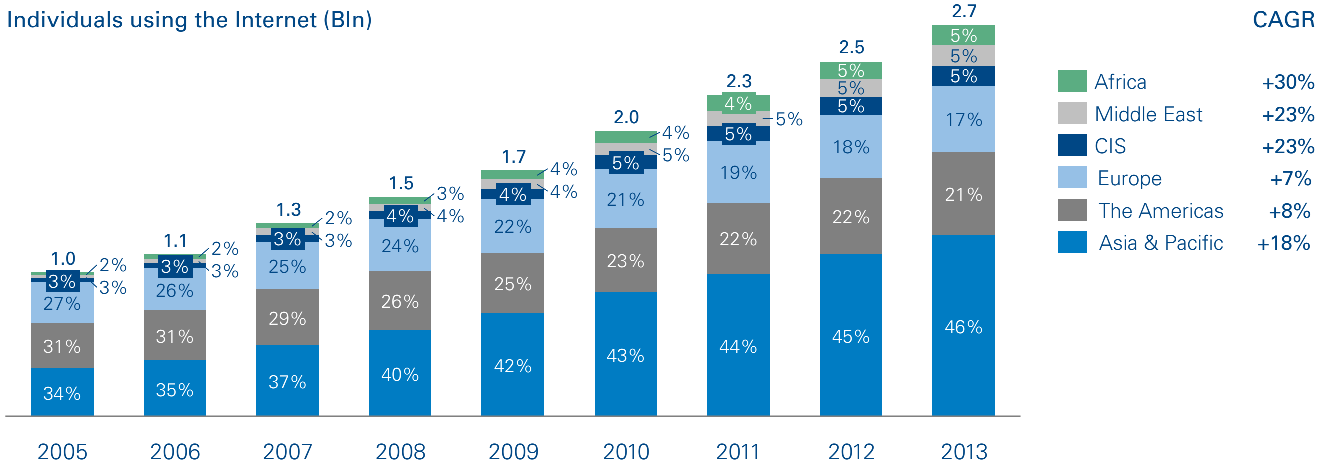
1 Source: ITU; an Internet user is someone aged 2 years old and above who went online in the past 30 days

Figure 1: The Internet timeline



Source: Arthur D. Little

Figure 2: Individuals using the Internet



Source: ITU, Arthur D. Little analysis

Regional data from ITU reveals that this growth is consistent across all world regions. Europe's and America's relatively low growth during these recent years (7% and 8%, respectively) is explained by their earlier take-up of the Internet, and both regions together now account for only 38% of global Internet users.

However, Internet traffic origination is not equally distributed across global regions: most of the IP traffic originates in North America (34%), followed by Asia & Pacific with 33%, Europe with 24%, Latin America with 8% and the Middle East & Africa with only 2%.

CISCO indicates that IP traffic will continue to grow, but at lower growth rates: highest growth rates are expected from 2013 to 2017 in the Middle East & Africa (35%) and Asia & Pacific (24%).

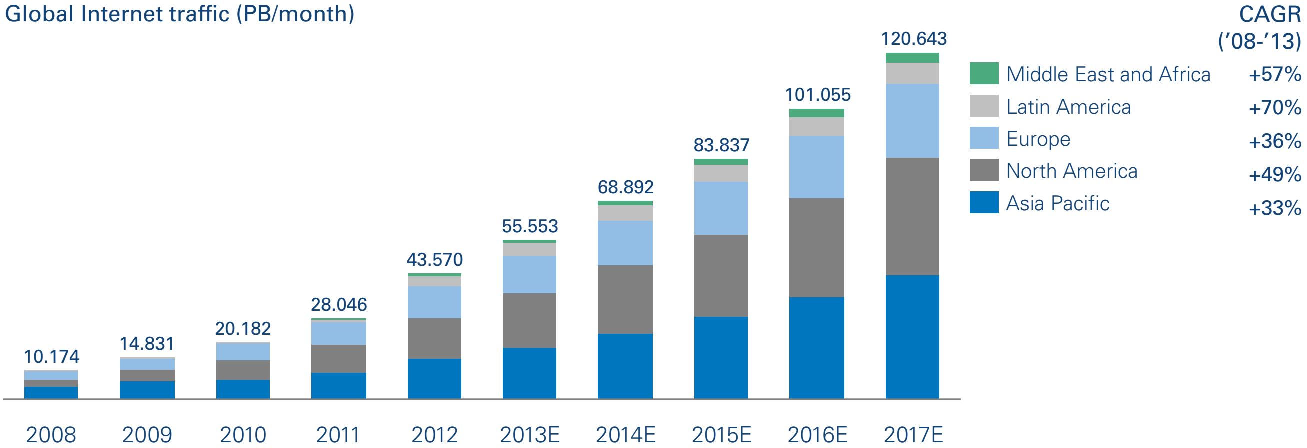
In the beginning, Internet access relied mainly on fixed-network infrastructure and end-users experienced the Internet through desktop computers. Over the years, a vast array of connected devices emerged, such as smartphones and tablets, through which end users could experience the Internet in mobility. These events significantly increased Internet's accessibility and boosted its penetration in the global population.

The number of broadband (BB) subscriptions has grown over 2007-2013 at a growth rate of 28.7 %, from 614 million to 2.8 billion, but when looking closer at their composition, it is clear how the mobile broadband development did become the main driver of this significant growth trend, especially in developing countries[2].

Between 2007 and 2013, the share of mobile broadband access increased from 43% of total broadband subscription to 75%. As the number of Internet users and broadband subscriptions has increased, the volumes of IP traffic carried over the Internet

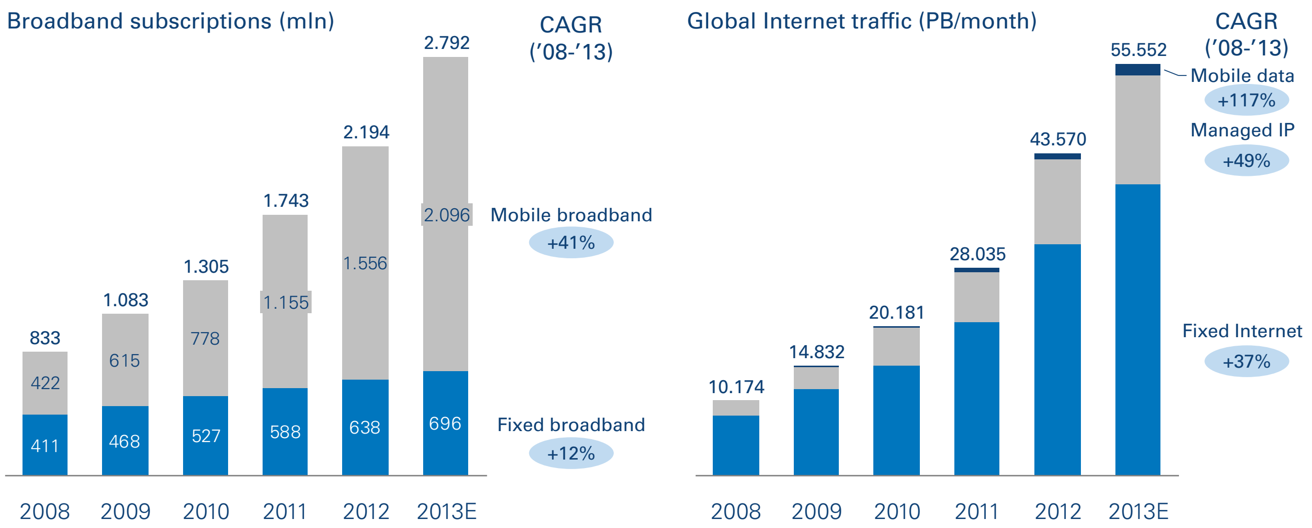
2 Classification of developing and developed country is available at http://www.itu.int/ITU-D/ict/definitions/regions/index.html, and has been made according to UN M49

Figure 3: Global IP traffic by region



Source: CISCO, Arthur D. Little analysis

Figure 4: Internet Broadband Subscription and Global IP traffic by type



Source: ITU, Arthur D. Little analysis

have grown tremendously. Cisco Visual Networking Index shows that traffic has risen from 10.1 exabytes (1018 bytes, or 1 billion gigabytes) per month in 2007 to 55.5 EB/month in 2013, with a 40% CAGR, and is expect to reach more than 120 EB/month in 2017 (with a growth rate of 21% from 2013 to 2017).

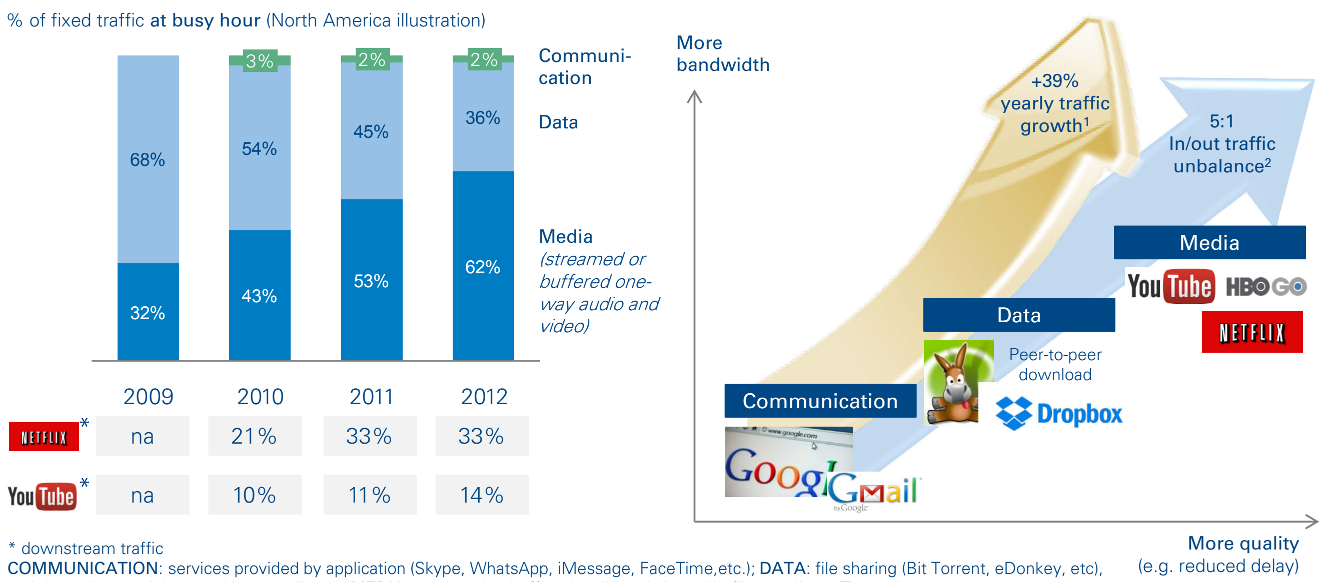
Fixed-Internet traffic accounts for about 70% of Global IP traffic, and has grown at a CAGR of 37%, reaching 39 EB/month. It is expected to reach 82 EB/month in 2017, counting for roughly 68% of the total IP traffic.

Managed IP services such as IP Virtual Private Networks (IP VPN), historically applied in business-to-business environments, have also grown significantly from 2007, at a CAGR of 49%. These have reached 15 EB/month in 2013, accounting for 26% of total IP traffic; CISCO projections in 2013 indicate that IP managed services traffic could reach 27 EB/month by 2017, slightly decreasing its share over total Internet traffic (23%).

Global mobile traffic has grown at notable rates (CAGR of 117%), reaching 1.6 EB/month in 2013, but it still accounts for only 3% of total IP traffic; however, CISCO expects it to rise to 11 EB/month by 2017, accounting for more than 9% of total IP traffic.

This growth of IP traffic volumes – data consumption per user grew from 12 gigabytes per user per month in 2008 to roughly

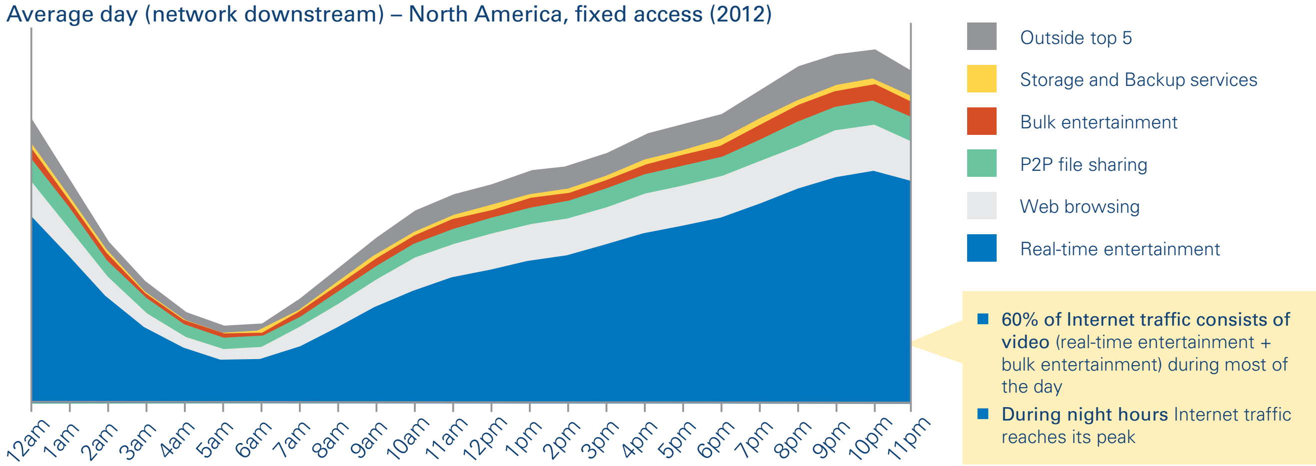
Figure 5: The changing nature of the Internet and the new requirements



* downstream traffic
COMMUNICATION: services provided by application (Skype, WhatsApp, iMessage, FaceTime,etc.); **DATA**: file sharing (Bit Torrent, eDonkey, etc), web browsing, social networking, email, etc.; **MEDIA**: streamed or buffered audio and video (Netflix, non-linear TV services;
1. 2009-2012 CAGR; 2. Interviews. Source: ITU, Sandvine; Arthur D. Little analysis

Figure 6: Daily Internet traffic by content type

Average day (network downstream) – North America, fixed access (2012)



Note: Storage and Back-up Services: PDBox, Netfolder, Rapidshare, etc..; Bulk Entertainment: iTunes, movie download services; P2P Filesharing: BitTorrent, eDonkey, Gnutella, Ares, etc…; web Browsing: HTTP, WAP browsing; Real-time Entertainment: streamed or buffered audio and video, peercasting, specific streaming sites service (Netflix, Hulu, YouTube, Spotify,..);
Source: Sandvine, Arthur D. Little analysis

20 gigabytes per user per month in 2013 – is not only driven by the increase in the number of broadband subscribers, but is also, and most importantly, linked to the change in nature of the Internet traffic itself.

The nature of Internet traffic changed from static data and text to interactive media content

The latest measures available show that the bulk of total usage growth comes from real-time streaming devices; the Internet has transformed itself from a data- and file-transfer platform into a new-media platform, and its usage has shifted to richer types of content, particularly video.

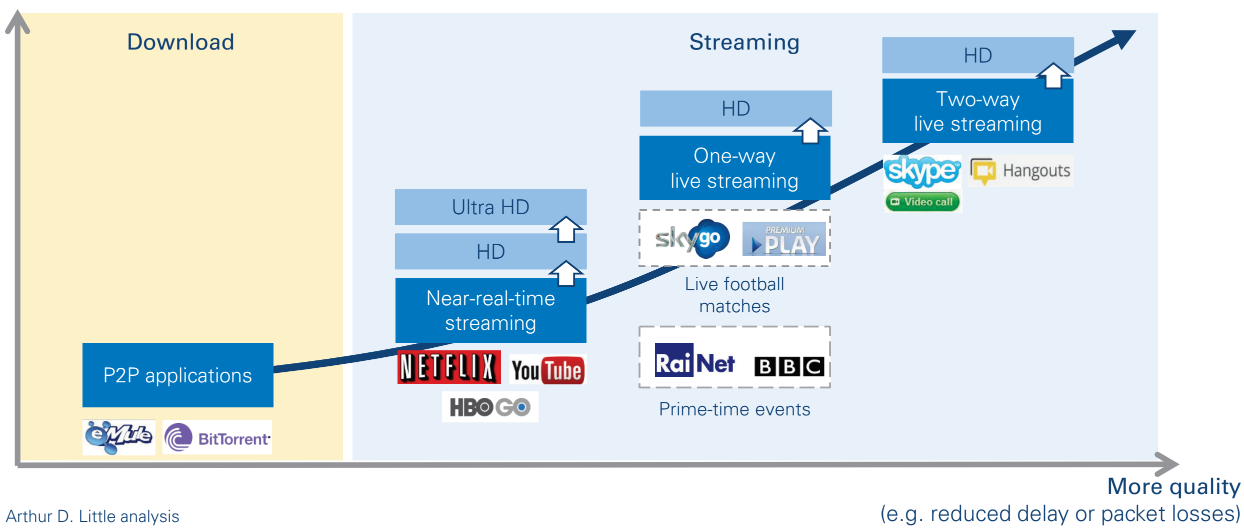
Today more than 60% of Internet traffic in the US is media related and, seemingly, such share is expected to grow further in the coming years

The consequences of such a shift materialized into increased demand for a higher bit rate and delivery quality. During the early ages of the Internet, communication interactions were established through sequential (asynchronous) applications, and on-time delivery was not important (e.g. emails); nowadays, higher throughput and reduced delivery time are essential for a good quality of experience (see figure 5 overleaf).

Furthermore, traffic kept doubling almost every two years, and traffic patterns changed as a result of real-time streaming

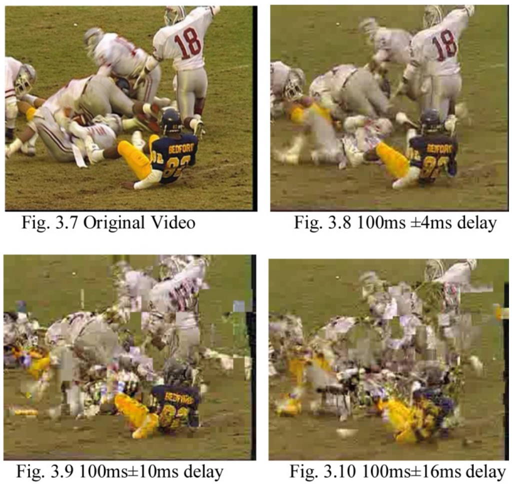
Figure 7: New connectivity requirements

Uncontended capacity



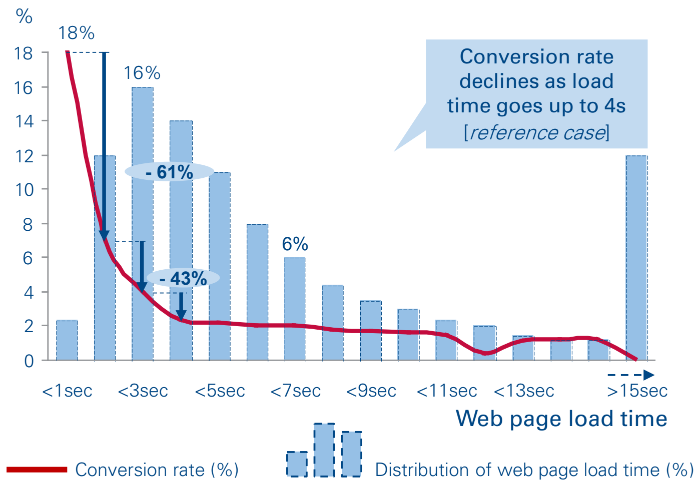
Source: Arthur D. Little analysis

Figure 8: Impact of delay on Media Customer Experience and on an eCommerce site performance



Source: Akamai, ReissRomoli-Telecom Italia, Arthur D. Little analysis

eCommerce conversion rate and web page load time distribution



http://www.bth.se/com/mscee.nsf/attachments/5297_Effect_of_Delay_Delay_Variation_on_QoE_in_Video_Streaming_pdf/$file/5297_Effect_of_Delay_Delay_Variation_on_QoE_in_Video_Streaming.pdf

becoming the predominant form of digital distribution. Access networks experienced significant imbalances (in the order of 5 to 1) on average between incoming and outgoing traffic just because the nature of traffic today is media related and streaming, and therefore mainly flows one way from content providers to end-users.

Although Internet consumption changes during the day and peaks between 9 and 10 pm, streaming media and real-time applications account for the majority of the traffic at any time of the day.

This evolution of the role and of the type traffic carried through the Internet pushes the amount of uncontended bandwidth effectively needed for each individual user to new levels. But more importantly, it also increases the quality level required for delivery of content.

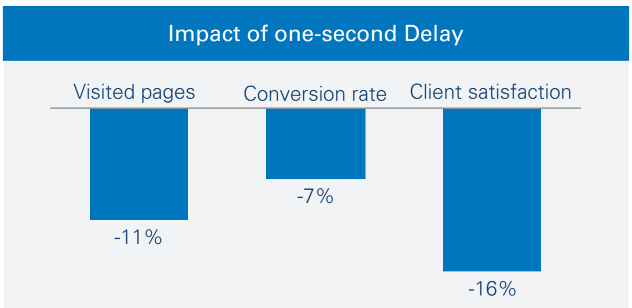
As High-Definition media becomes increasingly popular, reduction of delay and packet loss is becoming critical for the newer applications. But it becomes essential as a new class of applications may spread, such as live streaming (e.g. streaming sport events across popular video platforms) or two-way live streaming videos (e.g. Skype video-calls and "Hangouts").

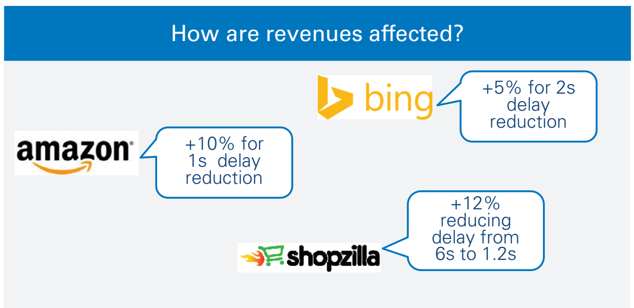
Today, Internet content delivery quality matters and means money

Nowadays the Internet is a mission-critical platform used by many companies to sell their own services and products to end-users. Consequently, and not surprisingly, the quality of the end-user's experience has become increasingly important because it directly affects end-users' purchasing decisions, with financial implications that can be quantified.

For example, variation of delay of just 16 milliseconds (on top of the foreseen average delay of 100 milliseconds) can damage the end-users' experience while watching a video stream, and therefore undermine his willingness to buy video content on

Figure 9: Business economic value of Internet quality





Source: Aberdeen Group, Joshua Bixby, Company reports, Arthur D. Little estimates the Internet. Along the same lines, the conversion rate (i.e. the proportion of visitors who actually buy) of a popular eCommerce web site can drop by a factor of 10 as the average load time for a web page increases from 1 to 4 seconds.

Many companies providing products and services via the Internet are pointing out the impact of a higher web page load time on several key performance indicators, such as the number of visited pages, the conversion rate or client satisfaction. For instance, eCommerce leader Amazon estimates that an increase of 10% in its revenues would be achievable through a delay reduction of 1 second. Similarly, Bing, a web search engine, found that reducing the delay by two seconds would result in a 5% revenue increase.

1.2. The IP Interconnection evolves and follows the evolution of the Internet

The Internet is made of IP Interconnection at the bottom of a complex stack of service layers

As the Internet went through the various evolutions described above, the (interconnection) practices that enable the Internet to function followed a continual process of reinvention and innovation.

The capacity of the networks that support Internet traffic has been continuously upgraded while the various market forces have adapted the way they interface and interact. The terms of those interactions, referred to as "IP Interconnection agreements", have been updated in line with new requirements imposed by new demand paradigms.

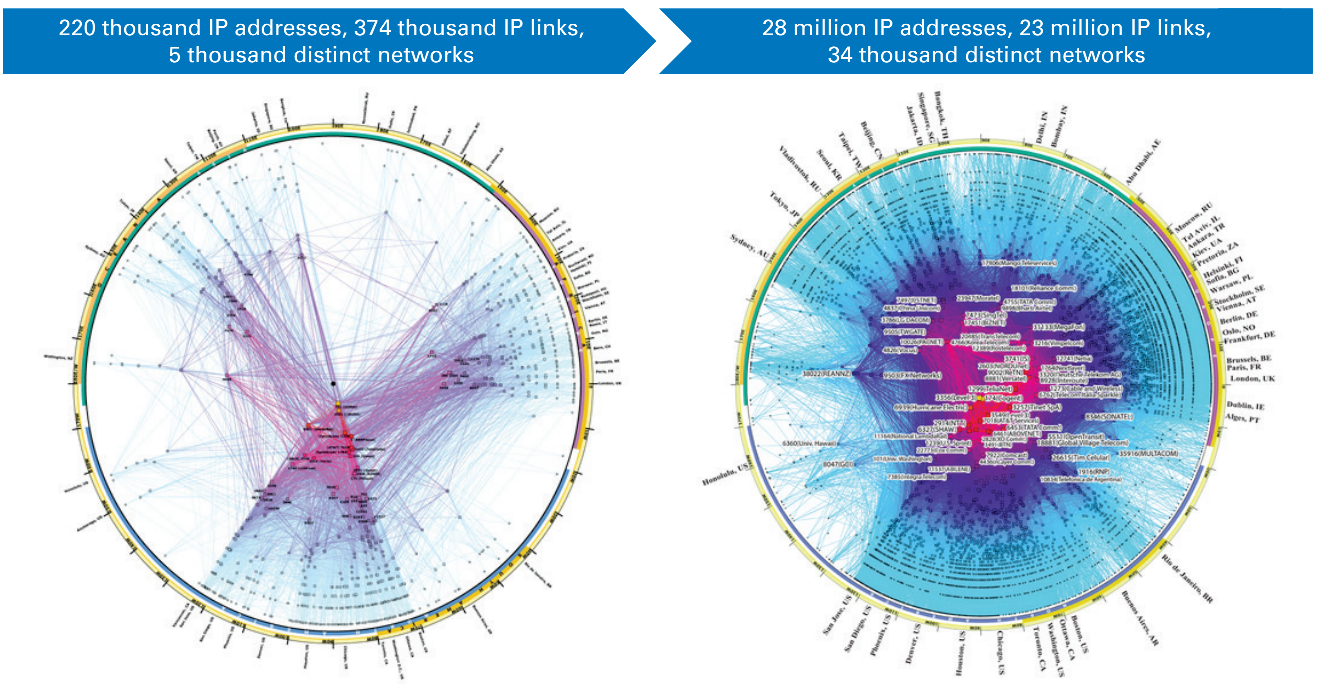
The Internet is made up of so-called "IP Interconnections", i.e. IP links that connect some tens of thousands of so-called "Autonomous Systems" (distinct IP networks: public, private, academic or other types of networks).

All the interconnected networks in the Global Internet use the Internet Protocol Suite (TCP/IP) to exchange information with each other. The Internet Protocol (IP) is the primary communication protocol for delivering packets of information from the source to the destination, without any form of central coordination that tracks or maintains the state of the network. The Transfer Control Protocol (TCP) adds some control mechanisms on top of it.

This Internet suite provides many unique features.

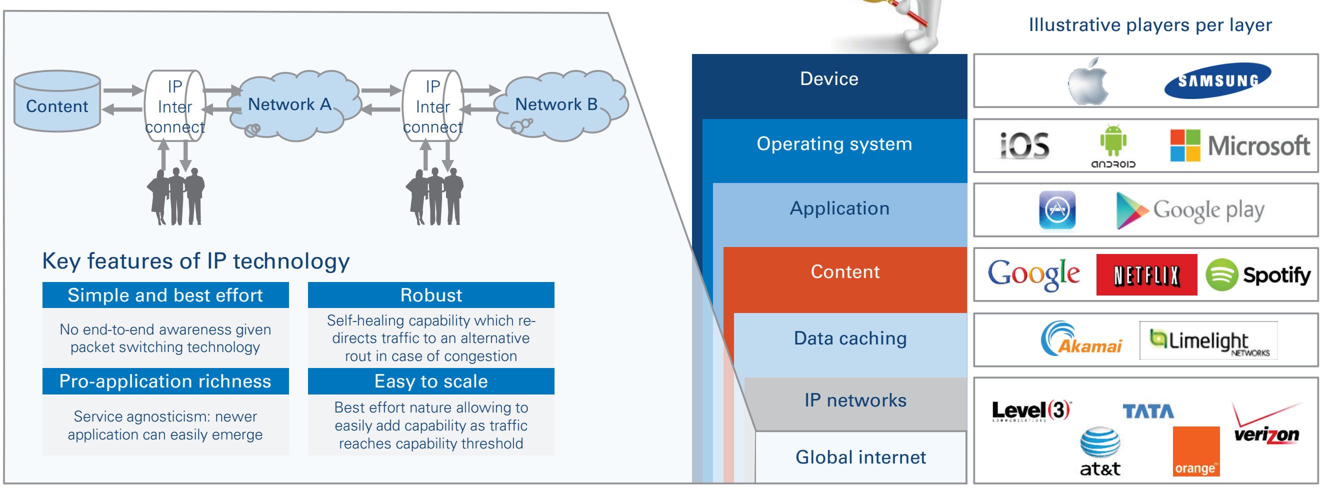
First of all, the IP confines a best effort service nature to all traffic sent. Indeed, Internet routers are programmed to do "their best" to deliver packets to the requested destination,

Figure 10: Representation of the Global Internet in polar coordinates: January 2000 vs January 2013



Source: CAIDA, Ark, Copyright UC Regent; Arthur D. Little analysis

Figure 11: The IP technology and the Internet stack



Source: Arthur D. Little analysis

thereby routing them to the "next (best) hop". The Internet Protocol offers no guarantees that packets will not be lost, delayed, corrupted or duplicated. With this Best-Effort mechanism, all users are served with an unspecified variable bit rate and delivery delay, solely depending on the current traffic load.

The lack of central intelligence makes the Internet an easy-to-scale network: capacity can be added progressively as traffic reaches capacity thresholds at each individual node of the global interconnected networks.

Furthermore, the Internet protocols provide robustness and self-healing capabilities: in fact, whenever a congestion issue is revealed, packet delivery is slowed down and/or packets are redirected and find alternative routes to reach their destinations. For the benefit of reducing network complexity and keeping the core of the Internet as simple as possible, the error-correction intelligence is located in the end nodes (the last node before reaching the destination, i.e. routers or computers) of each data transmission.

Another unique feature is that the Internet Protocol treats all packets independently of their content. They are individually addressed or routed according to the information carried by each unit, making it an application-agnostic transport layer. This feature is often considered the magic ingredient that enabled newer applications to easily and virally emerge.

Nevertheless, the Internet as we experience it encompasses many other layers, and the Global Internet is at the very bottom of a complex stack – the Internet stack. The end-users experience the Internet through a vast set of connected devices, operating systems, applications and online content, but are often not aware of the complex underlying interconnected structure of physical, data link, network and transport layers that make it possible to use all the fascinating services.

At the very bottom, IP networks and the global system of IP Interconnection layers perform several tasks:

- They provide the hardware with a means of sending and receiving data via the physical network (composed of cables, antennas, cards and other physical components);
- They convey the bit stream (electrical impulse, light or radio signal) through the network at the electrical and mechanical level;
- They manage the switching and routing technologies, creating logical paths for transmitting data from node to node;
- They adapt the speed of data streams in order to prevent or recover from congestion and re-route data streams when specific routes fatally fail;
- They ensure transparent transfer of data between end systems, or hosts, and are responsible for end-to-end error recovery and ensuring complete data transfer.

Content and applications make a long journey before we can consume them

When an end-user (an individual or a company) wants to access a particular video, website or other application (e.g. games), he/she connects to a server of a Content or Application Provider (CAP), which sends him/her the data packets that compose his/her video, webpage, etc.

Typically the end-user will use a software application (e.g. a video player) to display the content on his/her device (PC, tablet, smartphone, TV, etc.). The device will require an operating system to run the above-mentioned application (e.g. Windows, iOS, Android as the most widespread operating systems). His/her device will then connect to the network of his/her Internet Service Provider (ISP, which operates the local-access fixed (DSL, Cable, Fiber) or mobile network. From there the data packets will travel through one or more interconnected networks (Autonomous Systems) from the end-user to the CAP and vice-versa. International carriers provide the long-haul connectivity to foreign networks.In some cases, the content is already stored locally by Content Delivery Networks (CDN), shortening the digital content journey.

The number of networks through which the data packets will travel depends on the network congestion at that time and each individual packet will take the best route available at the instant that it is send out or reaches a new point where two networks interconnect. The CAP's location, its content delivery strategy, occasional circumstances or just the physical distance from the end-user influence which of the four different types of routes data packets take on their journey.

In summary, the quality of experience for the end-user will mainly depend on the efficiency of all network interconnections, as well as the performance of the device, operating system and application. Though they are also critical in the quality of experience, the last three are independent of the way the networks interface, and will not be further addressed in this report.

1.3. Introducing the actors in the IP Interconnection

The IP Interconnection value chain is composed of those actors that enable the delivery of traffic from a sourcing point A to a terminating point B. The sourcing point A is generally an application or content server, and the terminating point B is the end-user consuming that particular application or content.

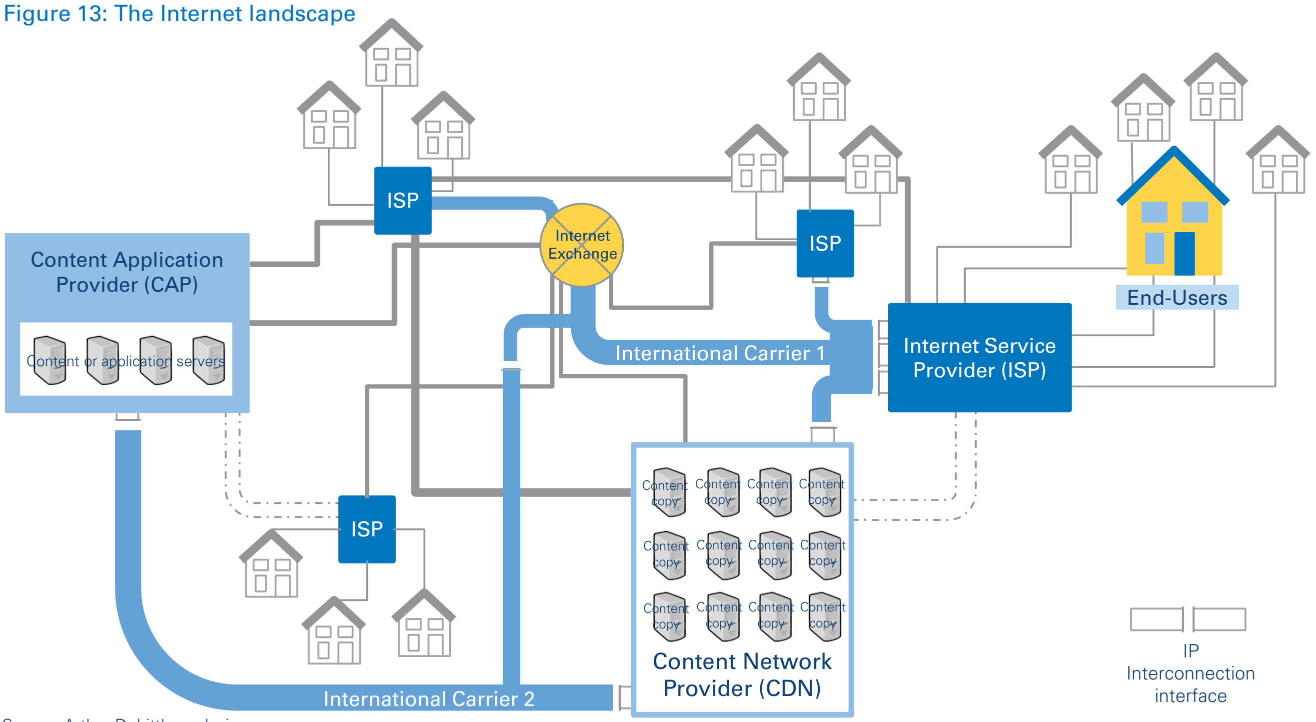
In the IP Interconnection value chain, five types of actors can be identified:

1. **Content/Application Providers (CAPs):** These players provide applications or content to be consumed by end-users by means of application servers located in one or multiple locations across the globe. A CAP can connect to the Internet through an Internet Service Provider, a Content Delivery Network or an IP Transit Provider, or directly connect to an Internet Traffic Exchange using its own proprietary infrastructures.
2. **IP Transit providers:** These players provide international connectivity, i.e. access to all possible originating and terminating countries around the globe. No player is really covering all possible locations (from both a coverage and capacity perspective); therefore, global connectivity is provided by interconnecting of several networks. (These networks take the form of regional networks, undersea cables, bilateral international terrestrial cables, satellite links covering some islands or remote locations, etc.)
3. **Internet Traffic Exchange providers (IEXs):** These players provide the locations where two or more parties can interconnect reciprocally. These locations are provided by consortia or private players, and are sparse, limited to the main global metropolitan areas or strategically convenient geographic locations (e.g. industrial districts, city centers, ports).
4. **Content Delivery Networks (CDNs):** These players provide networks of server farms around the globe, in which it is possible to store non-real-time content for local consumption by end-users (for efficiency and quality reasons). These players provide an overlay infrastructure built on top of the global IP interconnectivity.
5. **Internet Service Providers (ISPs):** These players provide local-access networks through which end-users (both individuals and enterprises) are attached. An ISP can act as both originating and terminating ISP according to the specific relation – i.e. the originating ISP is the one to which the CAP is connected, while the terminating ISP is the one connecting the end-user.

Figure 12: The main actors of the Internet landscape

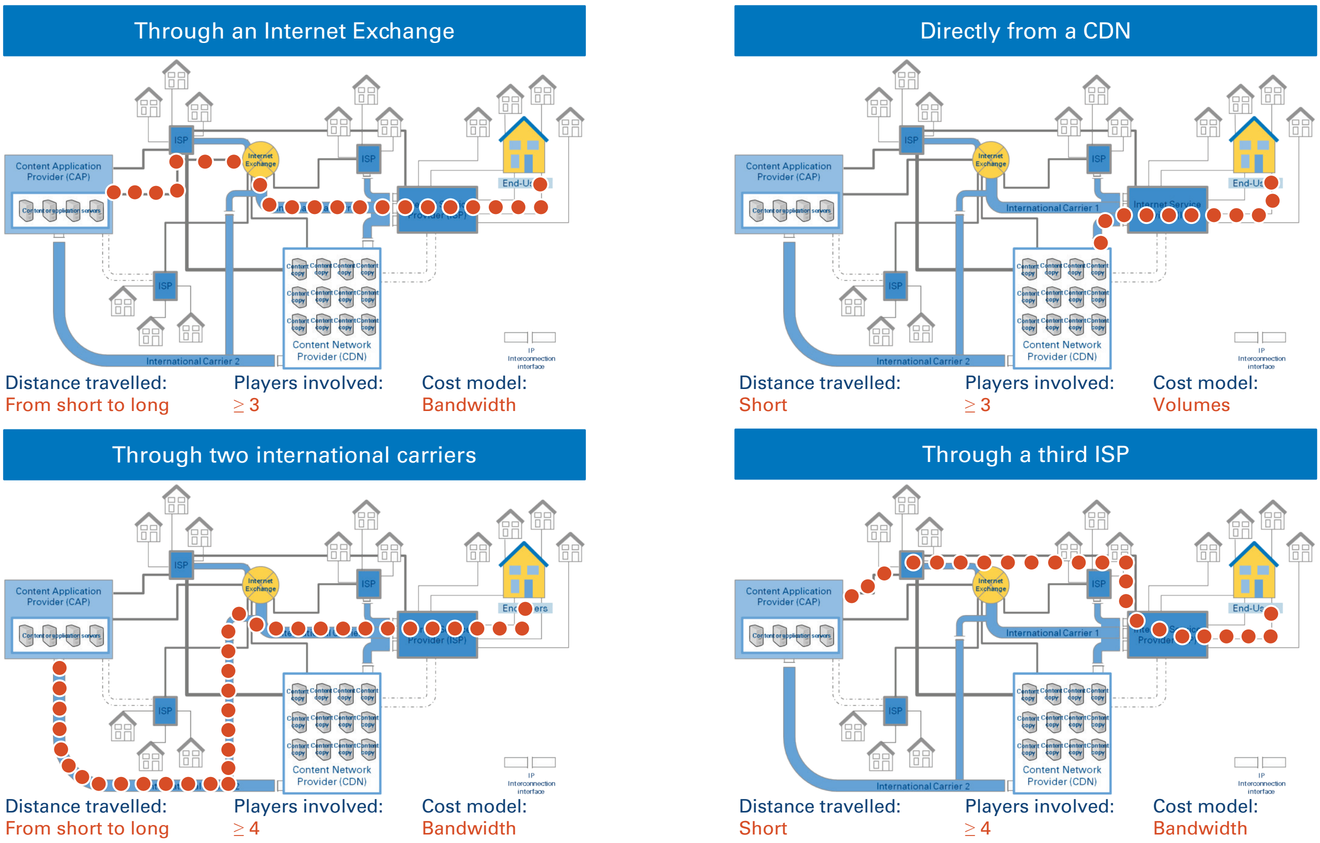
Source: Arthur D. Little analysis

Figure 13: The Internet landscape



Source: Arthur D. Little analysis

Figure 14: The predominant ways content or applications come to end-users



Source: Arthur D. Little analysis

1.4. Description of IP Interconnection business models

IP Interconnection has traditionally been established via two main commercial models: Transit and Peering

IP Interconnection arrangements are historically established via two commercial and non-regulated models selected by interconnecting parties on the basis of a "make-or-buy" decision: *IP Transit* and *Peering*.

When two players want to interconnect (or upgrade their interconnection), they need to decide whether to buy transit or to opt for a peering agreement.

Both of these practices are based on a private decision between the two interconnecting parties.

IP Transit is a bilateral agreement in which a carrier provides connectivity to all global IP destinations reached by its network to another (access seeker) ISP or Content/Application Provider, and is responsible for delivering the incoming and outgoing traffic to or from third parties. Typically, smaller ISPs buy transit from global carriers to reach the entire Internet. Transit is usually priced per capacity (Mbps), and customers are often required to commit to a minimum volume of bandwidth.

Carriers capable of global connectivity are labeled "Tier-1". Tier-1 networks do not need to buy "transit", so they do not pay another company to accept its traffic and distribute it to all networks connected to the Internet. (Tier-1 networks tend to have peering agreements with another.) It is interesting to note that no Tier-1 carrier can currently, with its own network, interconnect with all Autonomous Systems present in the global Internet. In general, global international carriers with more than 20% reach of IP addresses achieve the status of Tier-1 carrier. (Such status has been achieved by few players, such as Level3, AT&T, TeliaSonera and TI Sparkle).

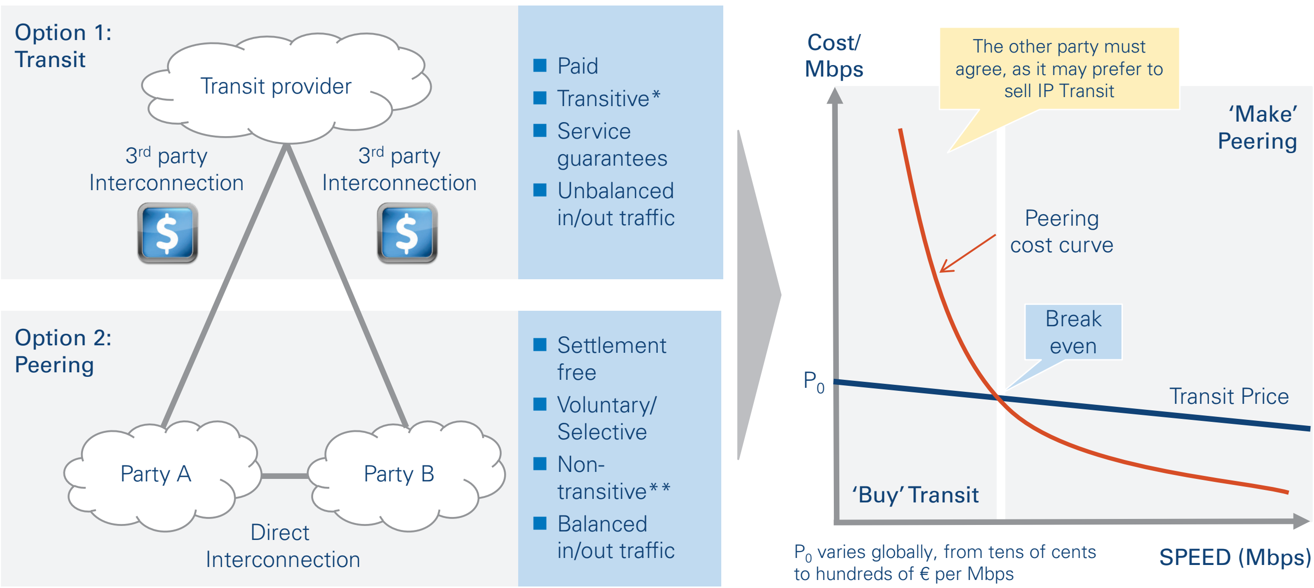
Peering is the alternative model for IP Interconnection. With the globalization of the Internet and the growing Internet traffic volumes, reciprocity to interconnect cost-efficiently is increased in self-interest. Therefore, ISPs wanting to avoid costly and volume-sensitive IP transit services often decide to directly interconnect – i.e. to "peer".

Through a peering agreement, two parties commit to exchanging traffic between one another. However, unlike Transit, Peering does not provide access to the full Internet, i.e. peering parties exchange only IP addresses connected directly to their networks (non-transitive).

Initially, Peering agreements were closed via simple hand-shakes, and did not imply any type of payment (settlement free). It is nevertheless important to note that peering does not come without a cost to both peering partners: the establishment of a peering interconnection (even if settlement-free) requires the peers to bear set-up and maintenance costs – more specifically:

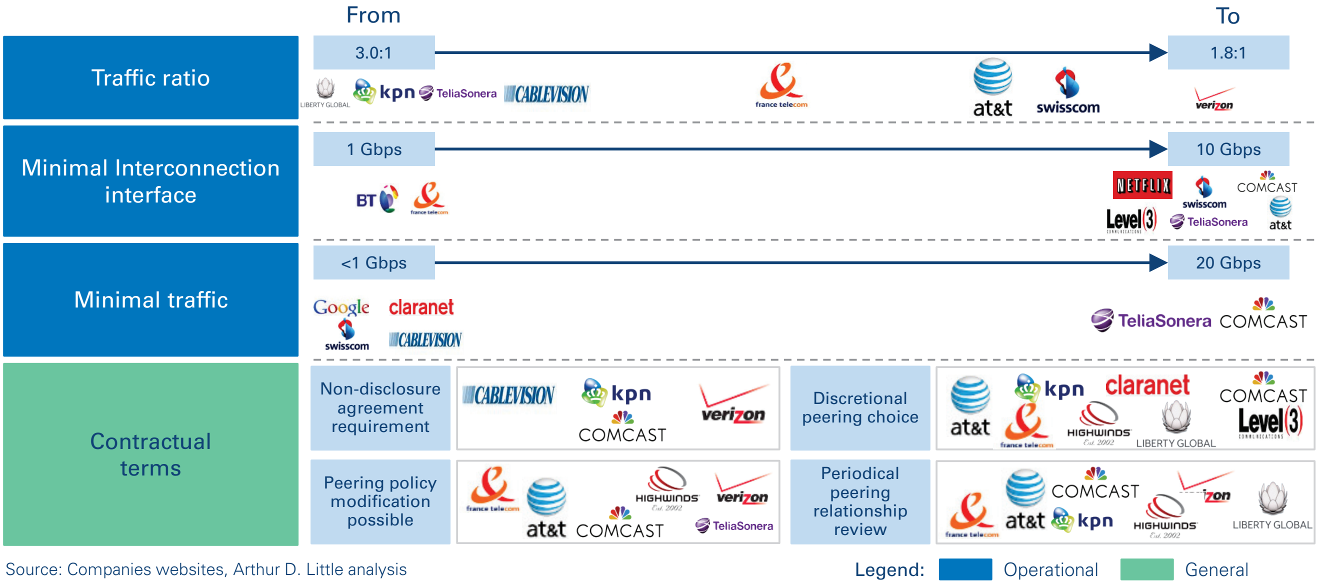
- Transmission costs to route the traffic to the peering location;

Figure 15: IP Interconnection options: Transit versus Peering



* Transitive: all possible IP destinations accessed by interconnected parties are made available;
** Non Transitive: only IP destinations accessed directly by the two peering parties are made available.
Source: Arthur D. Little analysis; William Norton (Dr Peering.com)

Figure 16: Overview of Peering policy features



Source: Companies websites, Arthur D. Little analysis

- Colocation fees: operating costs, mainly related to the space and power, and network equipment in a physical location providing access to the peering switches;
- Equipment costs (e.g. routers/switches);
- Peering port fees (when peering occurs in public locations and the Internet Exchange consortium offers access to shared routing fabrics).

In the early days of peering, most applications using the Internet were transferring limited and fairly balanced data amounts between Internet users. Hence, the underlying prerequisite for those Peering agreements was a balanced volume of Internet traffic to be exchanged between the two parties. Several economic and technical reasons supported the widespread use of peering agreements:

- Avoiding the higher cost of transit services;
- Increasing the IP network robustness by means of redundant routes that reduced dependence on one or few transit providers;
- Increasing the routing control capability thanks to the availability of alternative routes to reach a given destination;
- Reducing the delivery delay (latency) via direct interconnections.

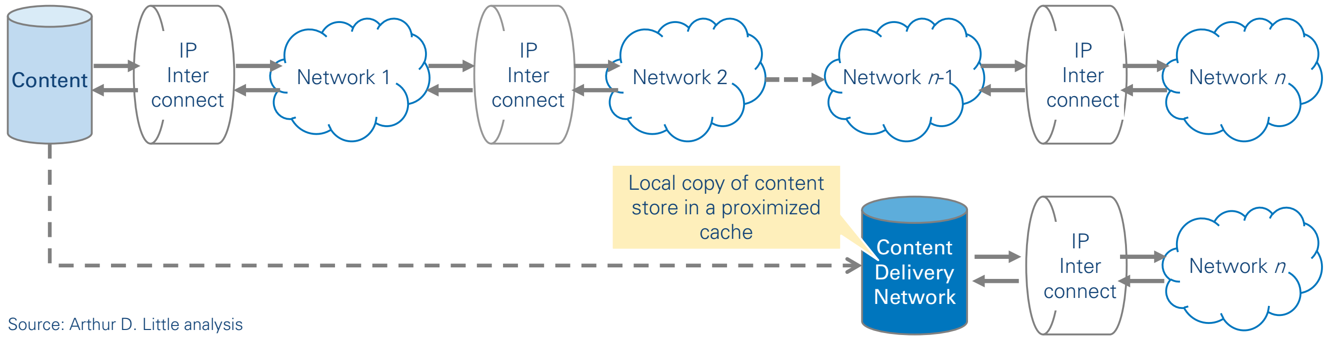
A survey by OECD, published in early 2013, reported that 99.51% of the 142.210 surveyed Peering agreements were "handshake agreements" in which the parties agreed to commonly understood terms without creating a written document.

Transparency in Peering

As the traffic has boomed and the number of Peering agreements risen, ISPs started to structure their Peering relationships with the help of specific Peering policies, which are usually published on their web sites. A review of 18 companies' public Peering policies reveals the following recurring features:

- *Balanced Traffic Ratio:* Peering players require the traffic exchanged to be balanced, imposing in/out ratio thresholds (between 3:1 and 1.8:1). It is worth noting that CAPs or CDNs do not specify any traffic ratio requirements in their peering policies, as their in/out traffic ratio is unbalanced by definition;
- *Minimum Interconnection Interface:* Often there is required minimum threshold of interconnection capacity links at Peering locations (1 Gbps and 10 Gbps as standard interfaces);

Figure 17: Basic working principle of Content Delivery Networks



Source: Arthur D. Little analysis

- *Minimum traffic:* some operators, wanting to ensure that their investment is adequately leveraged, ask interconnection links to be fed with a minimum traffic load, often measured in Gbps (using the 95th percentile measure, average measure, peak traffic, etc.). These required minimum thresholds vary significanlty and have been found to range between 25 Mbps and 20 Gbps;
- *Other contractual terms:* Other contractual terms generally refer to the possibility of modifying the Peering arrangements at anytime. They cover items such as the right to periodically review the Peering relationship or to choose whether or not to peer at either party's discretion.

A crucial part of Peering agreements is upgrading connections when they start to fill, because congested ports slow down the rate at which bits flow between networks. Usually, when connections reach about 50% utilization, the Peering partners agree to upgrade them in a timely manner.

The impact of proliferation of public Internet Exchanges on Peering

The rapid widespread use of Peering agreements was accelerated by the establishment of public locations at which to peer, the so called Internet Exchange Points (IXPs). The rising number of worldwide IXPs provides evidence that the Internet decentralization trend leads to a shift from global transit agreements to more regional Peering-based interconnections. Following the initial success of several IXPs in the USA and Europe, the expansion of IXPs has accelerated, becoming a reality in Asia and developing countries as well.

An IXP is a large data center where operators can collocate their servers and/or connect to one another using direct cross-connections. The biggest European IXPs are based in Amsterdam, London and Frankfurt.

So-called **"Public Peering"** at IXPs involves a large number of operators connecting, through a shared Peering fabric (an Ethernet switch) at public IXPs. Public Peering mutualizes the peering investments and operating costs, thereby reducing the Private Peering costs (when two operators exchange traffic across a dedicated link).

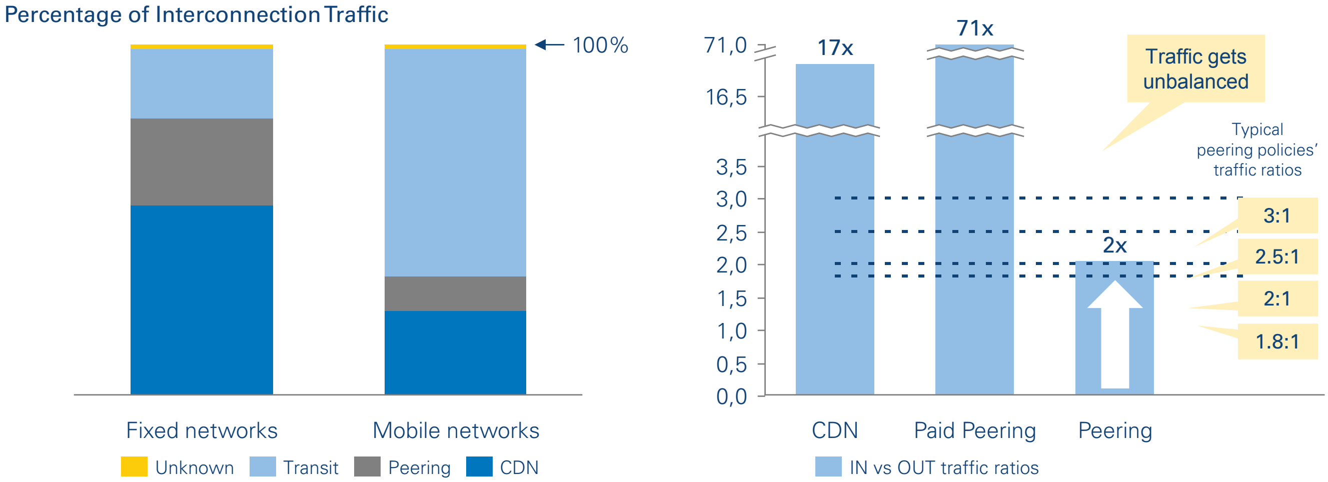
Private Peering is performed by creating a direct physical connection (usually consisting of one or more 10GE fibers) between two networks. The connection is made from only one network to another, for which a set fee is paid to the owner of the infrastructure (such as a datacenter).

Still, Private Peering remains a preferable choice when the amount of traffic exchanged between the two ISPs is high and peers want more flexibility to agree on capacity upgrades needed to avoid congestion.

Peering Variants appeared over time following new interconnection strategies and tactics aimed at better circumventing IP Transit fees or the limitations of Peering policies. Here are few examples:

- Content Peering refers to a Peering relationship in which the in/out traffic exchanged is unbalanced. Such situations typically occur when an ISP interconnects with a Content or Application Provider;
- Paid Peering foresees a paid settlement for a structurally unbalanced in/out traffic flow, and when the imbalance of traffic between service providers rises above a certain traffic ratio. Paid Peering is often applied to a Content Peering arrangement;
- Multi-Homing Peering is an approach, or rather a tactic, that uses interconnections with more than one ISP in order to improve redundancy and/or avoid in/out traffic unbalances;
- Partial Transit refers to a practice, or rather a tactic, in which an ISP resells to smaller ISPs the access to the interconnection routes secured via a Peering relationship, e.g. at an IXP. The corresponding price is slightly higher than the price of transporting traffic to the IXP, but generally less than IP Transit fees.

Figure 18: CDN relevance (UK case) and examples of in/out traffic imbalances



Source: Case example, Arthur D. Little interviews

Content Delivery Networks (CDNs) emerged in response to requirementsfor higher quality delivery

The explosion of media content and the mass adoption of commercial practices over the Internet imposed new quality requirements that traditional Transit and Peering interconnection services could not meet.

First, with traditional interconnection practices, content packets, looking for the best available route, tend to pass through several hops, with the risk of introducing excessive delivery delay or packet loss.

Second, a large part of the content requested by end-users is not provided in real-time, and consequently does not need to be constantly refreshed.

Third, physical distance negatively affects the final throughput (as further described in this report)

Therefore, Content Delivery Networks (CDN) were set up, and hosted, copies of the most popular content and/or applications in a given region in local caches. The proximity, gained by shorter paths to the final users, secure Quality of Experience at the entry of the local-access network while decreasing interconnection costs related to transit.

The mainly one-way nature of the traffic related to the media content implies that CDN interfaces, similarly to Paid Peering interfaces, show structural in/out traffic imbalances.

Pure CDN players (such as Akamai, Limelight and Edgecast) and Carrier-CDN players (such as Level3) emerged and grew by providing content aggregation and caching services. Traffic delivered through CDN players progressively increased, and now accounts for about 50% of global Internet traffic.

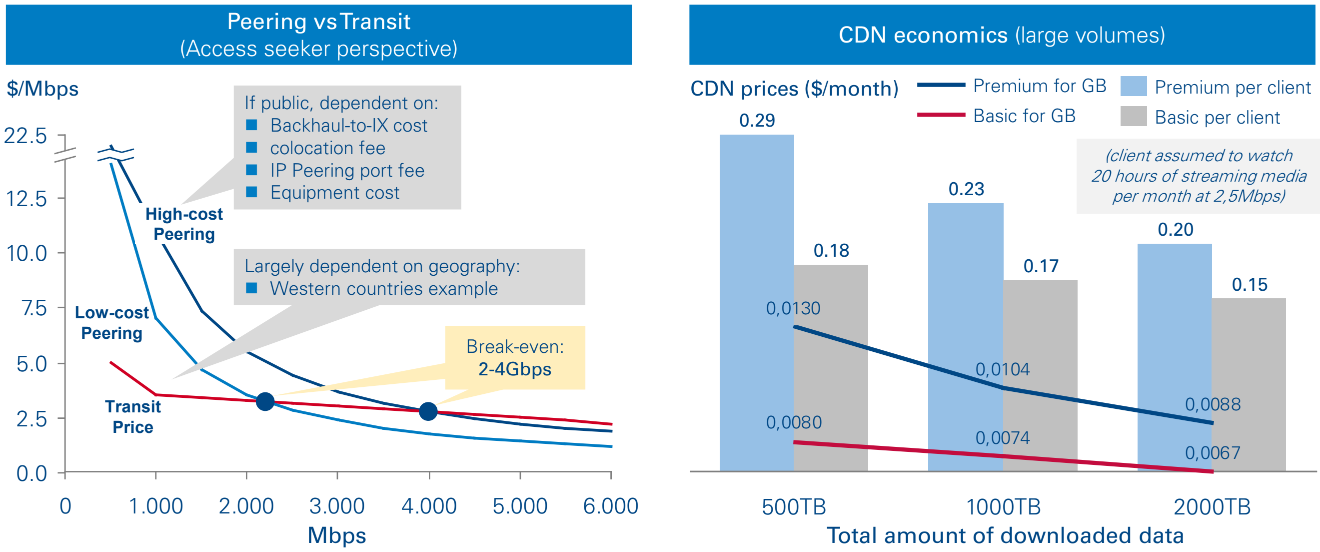
Statistics regarding the traffic split at IP Interconnection interfaces are scarce, but recent analysis by several regulators (e.g. Ofcom) has revealed that in the United Kingdom CDNs account for 55% of the traffic on fixed networks and 25% on mobile networks.

The variance in traffic mix on fixed and mobile network interfaces largely reflects the difference in the nature and volumes of content and applications consumed over those networks. As a matter of fact, fixed networks have traditionally carried much more video content and larger volumes than mobile networks.

The development of CDNs also contributed to traffic regionalization:

- Looking at international bandwidth evolution by route, it can be observed how in recent years the International capacity has grown less than overall IP traffic;
- With the development of larger regional networks, the share of international capacity connected to the United States and Canada has reduced in all regions, except for Latin America. Observations show that North America's largest IP interconnection partners, Europe and Asia, have seen their share of international bandwidth connected to the US decline from 24% in 2007 to 16% in 2011, and from 58% to 42%, respectively.

Figure 19: Illustrative economic evaluations between interconnection practices



Source : Dr Peering, Arthur D. Little analysis

The selection of the IP Interconnection mix requires an arbitration exercise

The selection of the most adequate IP Interconnection service(s) (Peering, Transit and/or CDNs) implies a delicate arbitration exercise involving many factors, as there is no solution fitting all purposes.

On one hand, IP Transit agreements bring costs for delivering traffic (the higher the traffic volumes, the higher the transit fee), but they also offer:

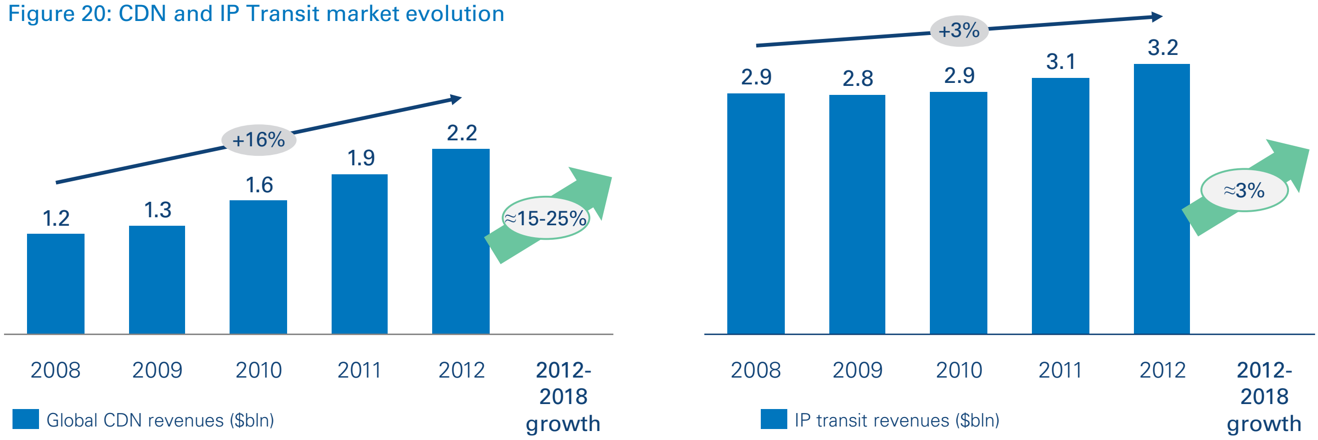
- Flexibility and redundancy in managing interconnection interfaces;
- Professionalism in the complex job of routing IP traffic and assuring quality of delivery in case of unexpected incidents.

Opting for IP Transit also saves some investments and operating costs required for the establishment of Peering.

On the other hand, Peering is, in general, more convenient when volumes (hence the bandwidth required) between two parties are large enough for unit costs to be lower than transit costs. In principle, Peering also brings the convenience to eliminate redundant hops in the delivery of traffic, thus reducing delivery delay (latency) and the risk of packet losses.

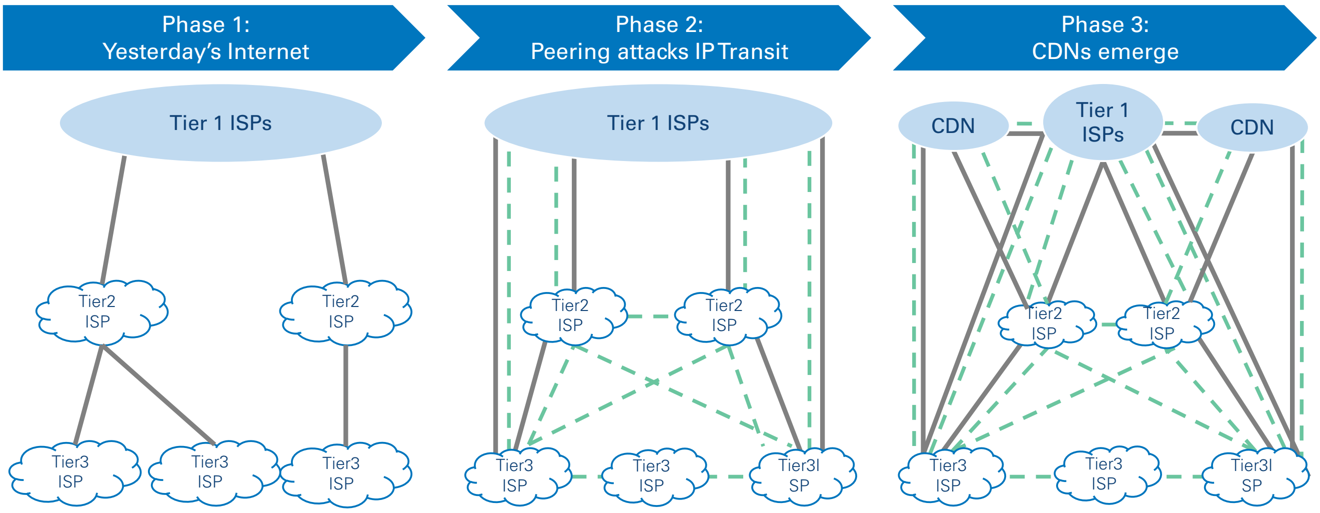
Eventually, interconnection to a Content Delivery Network allows the content provider to distribute its own content through a direct interconnection that enables savings versus transit and assures better quality (thanks to proximity and a dedicated cache fabric tailored to end-users' demands).

Figure 20: CDN and IP Transit market evolution



Source: ITU, Arthur D. Little analysis

Figure 21: Evolution of the Internet architecture into a robust and flat mesh



Source: Arthur D. Little analysis

Although the price dynamism in the sector and the nature of the service make static comparative assessments difficult (especially when comparing Transit to CDN), it is possible to conclude that Peering and CDN services are cost-effective choices for access seekers, in particular content providers, when capacity at the interconnection interface increases beyond 2-4 Gbps and/or downloaded volumes are massive.

Content Delivery Network economics differ from IP Transit and Peering, as their main drivers relate to low latency (generally, linked to the number of data centers among which content is distributed) and downloaded volumes (drivers are GB instead of Mbps). CDN pricing can double according to the service level requested – i.e. the number of data centers involved.

The combined revenues of the IP Transit and CDN segments reach around $6Bln, well below 1% of the total telecom landscape.

Notwithstanding the increased competition among the different interconnection practices, CDN and IP Transit business kept on growing, although the former at a higher rate (16% CAGR) and the latter close to stagnation (3% CAGR).

The CDN market is forecast to continue growing signficantly in the coming years, while the expected market growth for IP Transit is weak or moderate according to the geographies.

The dynamism in IP Interconnection practices brought robustness and efficiency to the Internet

Competition between the three IP Interconnection forms (Peering, Transit and CDN) drove the evolution of the architecture and topology of the Internet Interconnection landscape. The landscape moved from an initial hierarchical structure to a flatter structure.

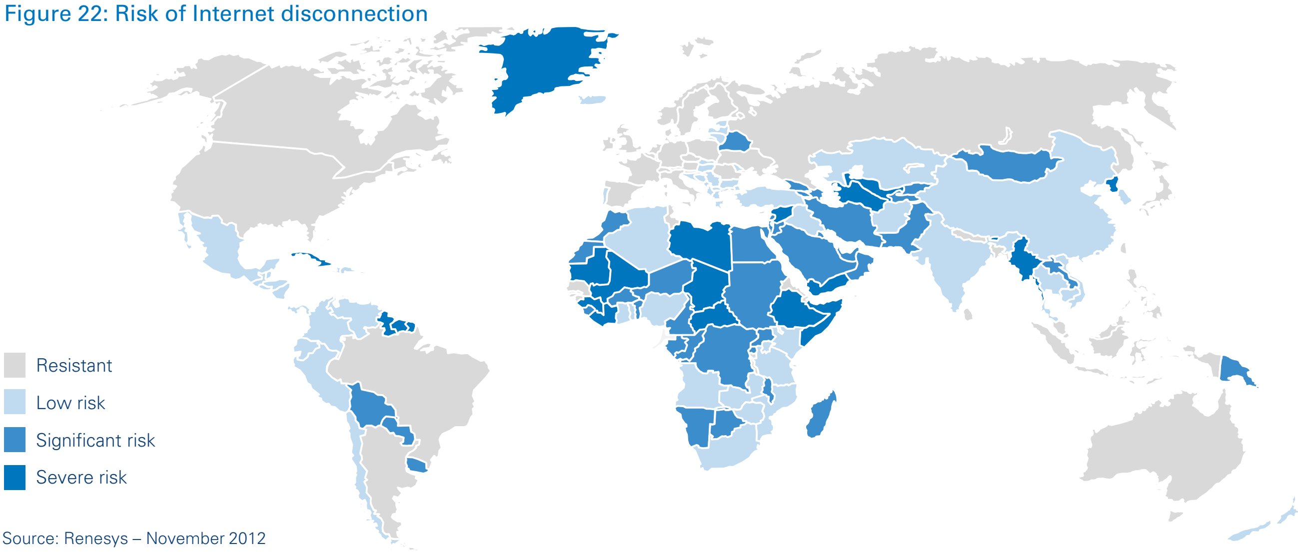
In the beginning, the topology followed a hierrachical structure, with a few Tier-1 global ISPs at the top of the pyramid. The access to global Internet interconnection was then cascaded down via wholesale transit services to regional ISPs (Tier-2 ISPs) and local players (Tier-3 ISPs), which were in charge of collecting and delivering the content or application. Hence, Tier-3 local players relied on interconnection agreements with Tier-2 and Tier-1 operators to provide Internet services to end customers.

With the rapid emergence of Peering and CDNs, direct inter-connection links between any type of player (Tier-1, Tier-2, and Tier-3) were enabled, thereby creating a substantially more robust topology with significantly less points of failure.

Most countries, in particular among the developed countries, now have up to 40 interconnection points to the global Internet. This significantly reduces the risk of losing Internet connectivity. The global Internet disconnection risk can be mapped (according to Renesys) as function of the number of providers of international Internet connectivity, as follows:

- Severe, with just one or two providers
- Significant, with three up to nine providers
- Low, with 10 to 39 providers
- Resistant, with 40+ providers

Figure 22: Risk of Internet disconnection



Source: Renesys – November 2012

Next to robustness, efficiency also increased thanks to the development of photonic and computing technologies, which drove Transit and Peering interconnection costs down substantially in the last 15 years (at an overall CAGR of -24%).

Note that notwithstanding the significant interconnection price cut for both Transit and CDNs, the increase in Internet traffic explains the growing revenue forecasts for CDNs and Transit.

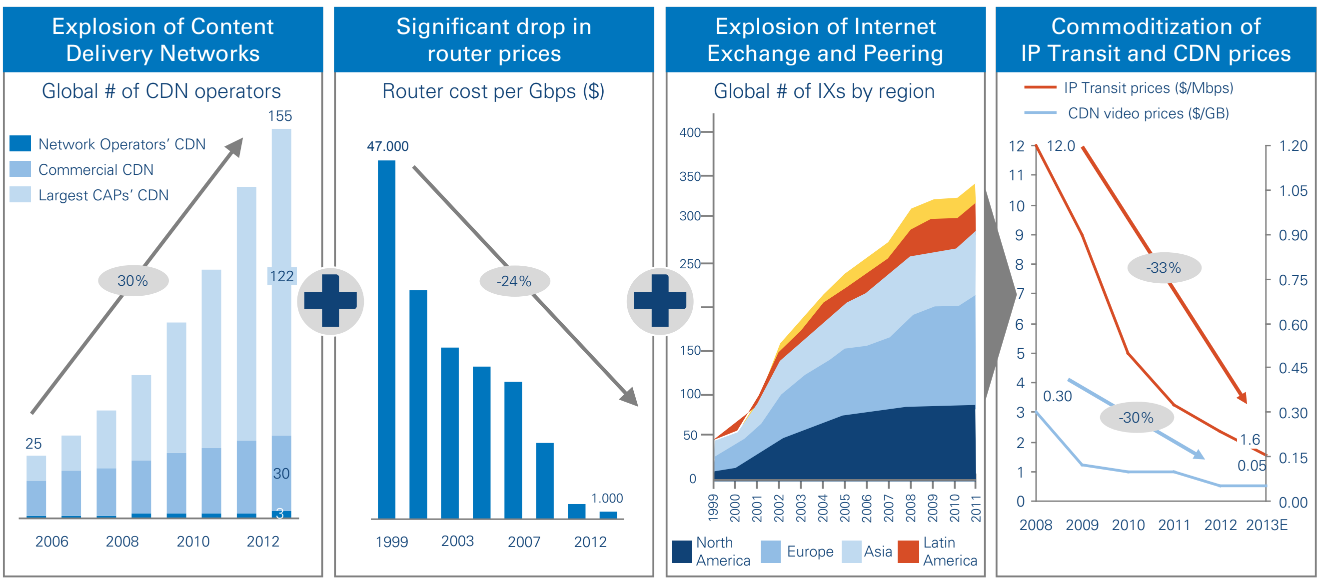
In summary, the combined effect of the growing adoption of CDNs, substantial decreases in interconnection costs and the explosion of Internet Exchange Points resulted in significant efficiency improvement by reducing IP Interconnection prices by more than 30% annually.

This implies that, so far, market forces have reacted well to the explosion of Internet traffic and the morphological change in the nature of the Internet.

Application/Content Providers have many alternatives for interconnecting, depending on their size

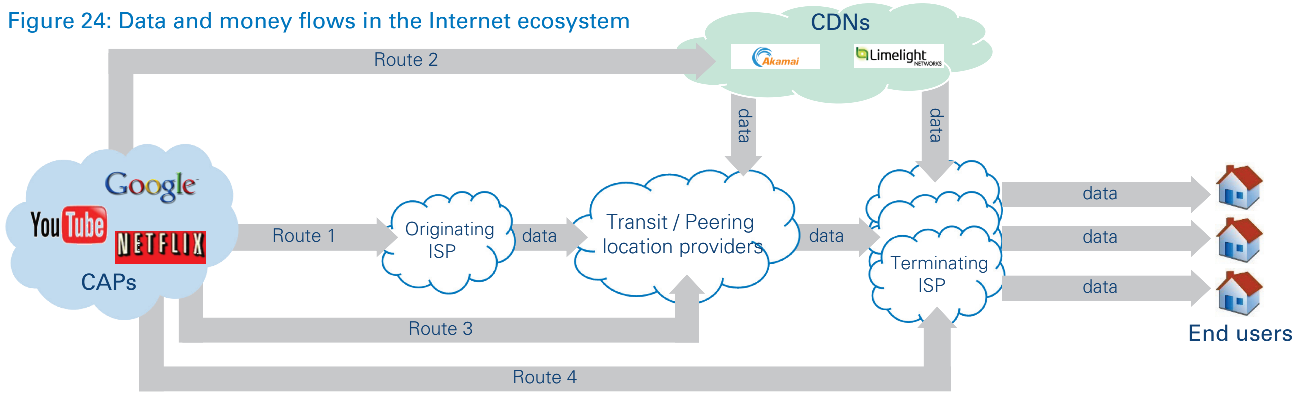
Depending on the nature and volume of their traffic and the available financial and infrastructure resources, several routes are available to Content and Application Providers to bring their services to end-users. They can:

Figure 23: Drivers of IP Transit and CDN price drops



Source: ITU, Informa, Packet Clearing House, Dr. Peering, Cisco, streamingmedia.com, Web Sites, Arthur D. Little analysis

Figure 24: Data and money flows in the Internet ecosystem



Source: Arthur D. Little analysis

- Connect to a local ISP (route 1), the most popular option for smaller CAPs;
- Directly connect to an independent CDN (route 2), an option that is mainly used for applications with stricter delivery requirements;
- Directly connect to a Transit Provider or peer with a Tier-1 or Tier-2 ISP (route 3); this option is generally addressed by larger CAPs that offer content to IP Transit Providers in exchange for a better global connectivity service;
- Connect directly to the terminating ISP (route 4): large CAPs have the opportunity by means of propietary infrastructures to connect directly to the most important terminating ISPs at global level.

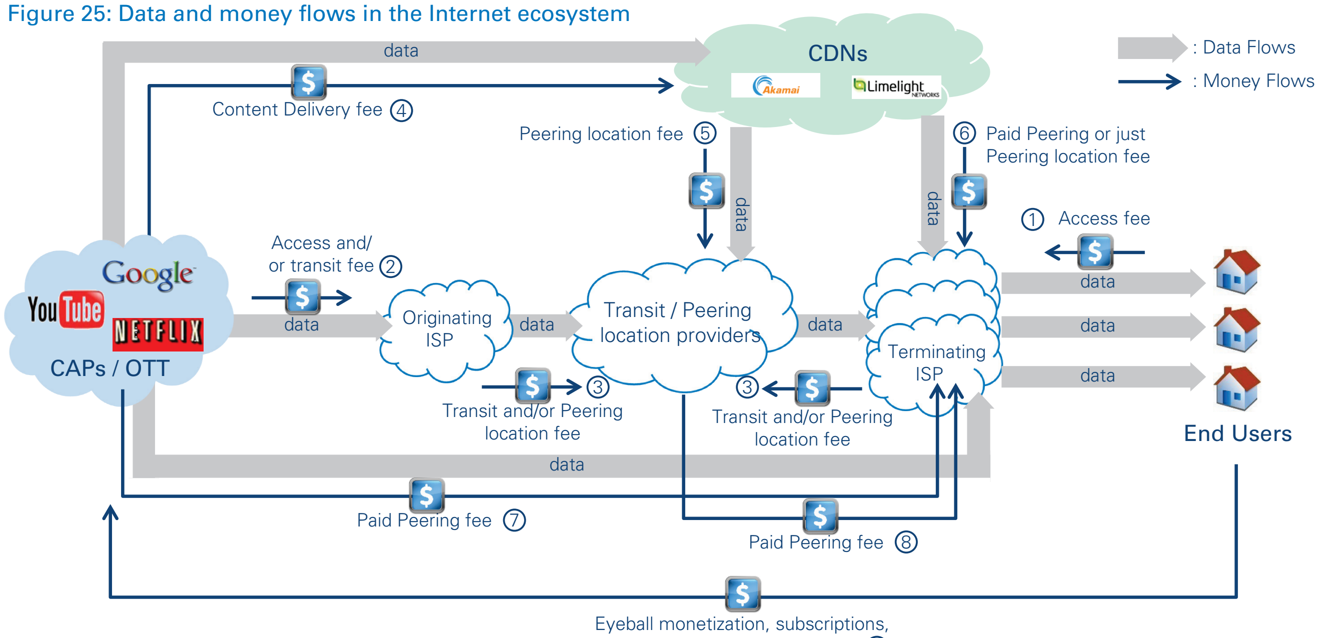
On the Internet, data traffic means money

The Internet is all but not for free and data traffic determines money flows among involved parties in a complex way.

End-users buy connectivity services from Internet Service Providers (ISPs) in order to be granted access to the services and content provided or sold by Content and Application Providers (CAPs). Traditionally, fixed and/or mobile telecom operators and cable operators act as local access providers or terminating ISPs Terminating ISPs assure their access to the global Internet by paying a transit provider and/or investing in Peering capacity.

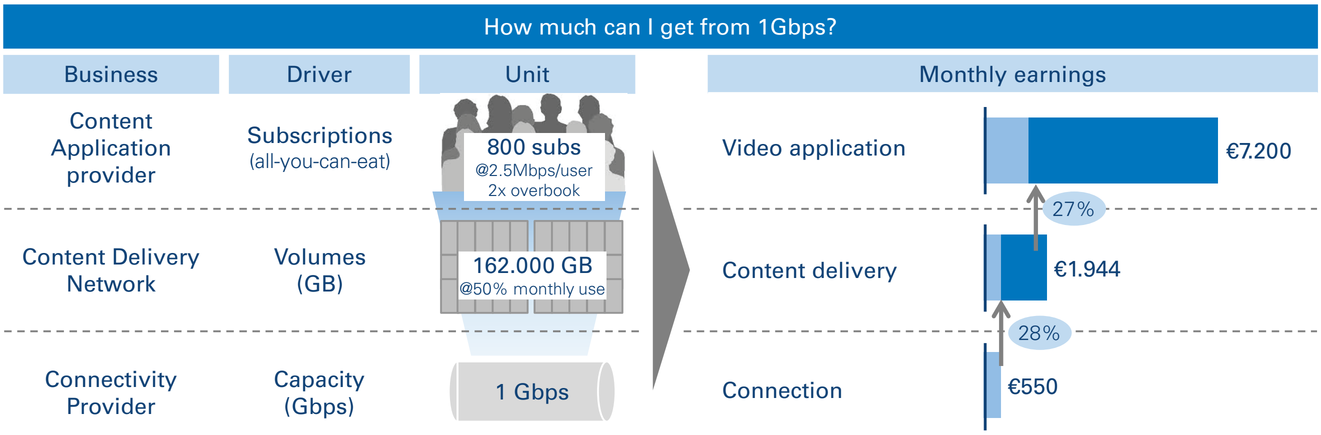
At the other end, CAPs also need to connect to the Internet by paying an access and/or transit fee to the local-access network operators or to global ISPs specializing in Internet transit services. CAPs may alternatively opt to buy content delivery

Figure 25: Data and money flows in the Internet ecosystem



Source: Arthur D. Little analysis

Figure 26: Illustrative calculation of how much value can be extracted from 1 Gbps of IP Interconnectivity



Note: 1) Subs are estimated assuming average 2,5Mbps per streaming user in busy hour and 2x overbooking; GB (Gigabytes) are calculated assuming 50% capacity utilization per day. 2) Earnings are calculated assuming 0.55€/Mbps/month for Western countries local connectivity and 0,013€/GB for CDN service below 500TB and 9€/user/month for a Internet video subscription.
Source: Arthur D. Little analysis

services and caching capacity from commercial, independent CDN providers, or even invest in and roll out their own Content Delivery Networks. E.g. Netflix's Open Connect, Google Global Caches, Amazon's CloudFront. Reportedly, Apple is also building its own CDN to manage its growing iCloud service usage, as well as hosting and delivering content from the iTunes and App Stores, both streamed and downloaded.

Besides, CAPs generate revenues from distributing content to end-users over the Internet ("Over-The-Top"). They therefore set up different types of business models, such as:

- Eyeball monetization, i.e. having advertisers pay to reach end-users;
- Directing end-users' subscriptions or pay-per-use for content or services (e.g. subscriptions to video services or information sites);
- Intermediation of transactions between online merchants and final customers through an online marketplace (e.g. eBay, amazon, iTunes);
- Any mix of the above.

Furthermore, Content Delivery Network operators must interconnect at public or private network interfaces and may need to pay private Peering colocation or paid peering fees.

Finally, in the recent years, Global ISPs and CAPs directly interconnecting with terminating ISPs have been requested to pay for peering when traffic volumes exceed peering policies.

Overall, the Internet ecosystem results in a complex mix of interlacing business models that are built upon the global availability of IP connectivity:

- ISPs provide and sell connectivity, for access at local or international level. Their access business model is currently driven by a capacity measure, i.e. gigabit per second (Gbps);
- Content Delivery Network operators sell caching and web acceleration services. Their business model is mainly driven by volumes (and sometimes by server throughput capacity also called egress capacity), i.e. Gigabyte (GB) or Terabyte (TB);
- CAPs sell services and/or content. CAPs can apply any mix of the abovementioned business models driven by any proxy of traffic volumes – i.e. web clicks, page/video views, unique visitors, downloads, transactions, paid events or subscriptions.

We illustrate in figure 26 how different business models pile up, assuming a video application delivered via a CDN generating 1Gbps connectivity. In our illustrative model, each layer is an essential input to the layer above. According to our estimates, in principle, connectivity costs can represent around 28% of the content delivery revenue and, correspondingly, Content Delivery Network costs represent around 27% of the video application provider's revenues.

This suggests that, in order to achieve cost savings and protect their margins, CAPs will seriously consider integrating vertically towards deploying their own IP Interconnection networks if their volumes explode.

Connectivity costs can represent around 28% of content delivery revenue and, correspondingly, content delivery network costs represent around 27% of the video application provider's revenues

Key messages

- The Internet has been transformed into a new media platform, as the nature of Internet traffic has changed from static data & text file transfer to streaming interactive media content.
- The Internet has become mission critical for most Content & Application Providers. Minor disturbances in the quality of delivery directly impact the willingness of end-users or advertisers to pay for online services.
- The future development of the Internet as a media platform is impacted by increasing global connectivity, proliferation of smart devices and streaming media service,s causing spectacularly high traffic volumes, greater imbalances in traffic flows and changing traffic patterns.
- IP Interconnection is an essential building block for the quality & functionality of the Internet as ultimately experienced by the end-user, despite the fact that the end-user is no party to business-to-business IP Interconnection arrangements.
- The IP Interconnection value chain remains dynamic and competitive. Proliferation of Content Delivery Networks and Internet Exchanges, commoditization of IP Transit and CDN prices challenge existing interconnection models and enable new ones.
- From the early days of "IP Transit" and "Peering", a genuine mix of viable application/content-delivery strategies has been accessible to all players seeking connectivity.

2. So far, the IP Interconnection Value Chain has Adapted Well

Changes in the Internet environment generated new balances, which triggered actions and reactions and limited friction among the different stakeholders. Consequently, regulators started to look into the main developing trends in this particular segment of the telecom market.

Nevertheless, players in the ecosystem have been able to solve their disputes in an effective and timely manner (given the commercial relevance of their mutual interests at stake in the short term). Accordingly, it is not a surprise that most regulators have so far decided not to intervene in order to avoid altering the productive equilibrium that is developing so well in the Internet ecosystem.

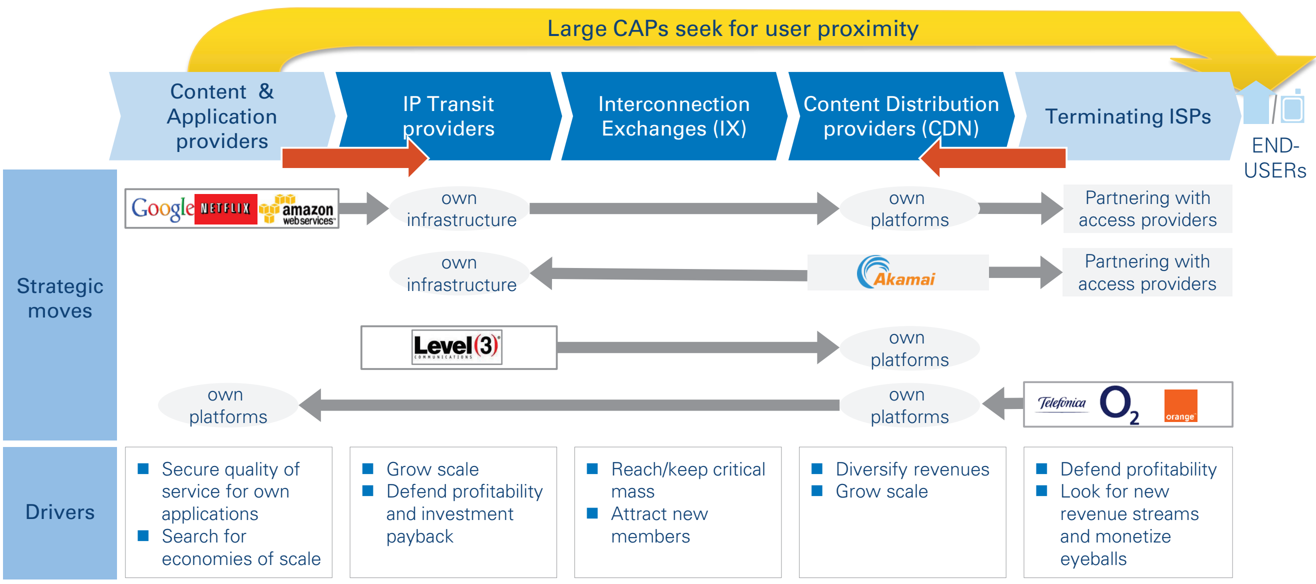
2.1. Content and Application Providers and Terminating ISPs are setting the pace of IP Interconnection innovation

The IP Interconnection value chain reveals an ongoing repositioning of the two major groups of actors: the Content and Application Providers and Terminating ISPs. As they are aiming to promote their own respective interests, they are currently setting the evolutionary pace of IP Interconnection:

1. Content and Application Providers seek end-user proximity and, when possible, try to integrate vertically;
2. Terminating ISPs try to enlarge their wholesale service portfolios by offering services directly to Content and Application Providers.

Larger Content and Application Providers have an incentive to integrate vertically and secure proximity with their end-users. It provides them with higher operational control over the Quality of Experience as well as economies of scale. Still their strategy goes beyond achieving a cost saving on IP Interconnection. As Quality of Experience becomes mission critical in differentiating from other Content and Application Providers, the end-to-end performance of IP Interconnections can be a source of strategic advantage. Indeed, investments in proprietary IP Interconnection solutions such as Content Delivery Networks, helps them to bypass many of the limitations of the public Best-Effort Internet. Furthermore, it improves the negotiating power towards Terminating ISPs. Of course, end-user proximity strategies through vertical integration are implementable only by the largest Content and Application Providers; smaller CAPs will opt for outsourcing such tasks to independent CDN providers.

Figure 27: Trends over the IP Interconnection value chain



Source: Arthur D. Little analysis

On the other side of the value chain, Terminating ISPs continue to invest in network-based content and application delivery platforms in order to defend their profitability and search for new revenue streams to compensate for their declining distribution revenues. Such platforms enable delivery of both third-party content and applications (i.e. from Internet-based Content and Application Providers) and their own content and applications (i.e. in direct competition with the CAP).

IP Transit providers, independent CDN providers and even Internet Exchanges are increasingly under pressure to innovate and diversify their service offerings (e.g. offering "partial transit") and attract a critical mass of traffic through consolidation, as Content and Application Providers and Terminating ISPs are interacting directly more and more, and thereby re-balance the entire value chain. A progressive commoditization of international IP Transit bandwidth and content delivery services is ongoing and has a significant impact on prices and volumes. Therefore, IP Transit providers and independent CDN players try to diversify their revenues by enlarging their service portfolios in other value-added segments such as web security and application acceleration platforms.

As Quality of Experience becomes mission critical in differentiating from other Content and Application Providers, the end-to-end performance of IP Interconnections can be a source of strategic advantage.

Countervailing powers keep the IP Interconnection Value Chain in balance

IP Transit and Peering become substitutes:

Internet Exchanges (IXs) allow ISPs to connect to each other directly via peering rather than through third-party network providers. They allow data to be accessed and transmitted locally, rather than using an upstream IP Transit provider. In this way, the boom of Internet Exchanges enabled the diffusion of peering that becomes a substitute for IP Transit in some circumstances.

Falling IP Transit and CDN prices also make high-quality transport and CDN strategies accessible to smaller CAPs:

Falling prices of IP Transit and CDN services are occurring in all geographies (ranging from -10 to -35% YoY, depending on Internet market maturity and infrastructure availability), and offer benefits on average to all players. Some information asymmetry and scale effect may still influence the commercial negotiation, but possible imbalances can be recovered fast (in less than one year, if contractual terms allow it). The resulting commoditization of IP Transit and CDN services makes high-quality transport (e.g. small latency routes) and CDN strategies affordable for all players in all geographies, irrespective of their size.

Falling IP Transit prices balance against Paid-Peering, resulting in limited business for terminating ISPs:

IP Transit and CDN prices are strictly correlated and influenced by the Peering opportunity cost. Paid Peering naturally follows the same pricing laws, and is hence influenced by the falling IP Transit prices. As IP Transit prices are falling well below one Euro for high volumes, Paid Peering prices are dropping to around one Euro. For instance, a European incumbent is known to be offering €1.5 Mbps per month, presumably to be discounted against large volumes.

Thus, falling IP Transit prices ensure that Paid Peering agreements stay in line with the competitive market pricing for IP Interconnection connectivity.

It should be noted that Paid Peering only represents marginal revenues for Terminating ISP. For instance, a large CAP, accounting for 40% of the IP traffic and paying €1.5 Mbps per month to a large European incumbent whose IP Interconnection interface is in the order of 600-1000Gbps, would only generate a modest €4-7 million per year.

Healthy retail competition disciplines and prevents market foreclosure by ISPs:

Retail markets in Europe are typically composed of two to five ISPs (two or more fixed ISPs and three or more mobile ISPs, acknowledging the fact that fixed and mobile actors tend to converge). Leaving aside the legal arguments, healthy competition among those ISPs alone is sufficient to prevent market foreclosure. In the hypothetical case that a CAP would be subject to foreclosure by a Terminating ISP (for example, if the ISP decides not to enter into a Peering relationship with a given web service's CAP), there are still multiple Internet connectivity providers that could be used by that CAP to access that ISP's Autonomous System indirectly. In addition, the CAP would keep access to other Terminating ISPs' networks, and the affected end-users, likely to be highly attuned to issues such as ISP service levels, broadband speeds and net neutrality, would be motivated to change to different ISPs that would provide access to the CAP's services and/or sufficient quality of delivery.

Healthy retail markets therefore act as an efficient lever to prevent market foreclosure.

2.2. Internet Content and Application Providers look for quality control

In recent years, Content and Application Providers have increasingly been paying attention to improving their control of the quality of delivery over the Internet. This is true not only for Internet-based CAPs, but also for the video-streaming strategies of traditional broadcasters. A prominent example is provided when reading the Internet distribution guidelines issued by the BBC during of the preparation of the 2012 summer Olympics: clearly, the British public broadcaster was seeking to assure high streaming quality and higher control.

Overall, as the popularity of streaming video grows, leading Content and Application Providers feel increasing pressure to assure quality of experience and have direct control over this. In addition to implementing streaming technologies that give CAPs greater control (such as "adaptive streaming"), several Internet delivery strategies can be developed, depending on the CAP's strategic objectives, capabilities and size; but there is a tendency for bigger CAPs to opt for in-house approaches when the end-user's Quality of Experience becomes mission critical.

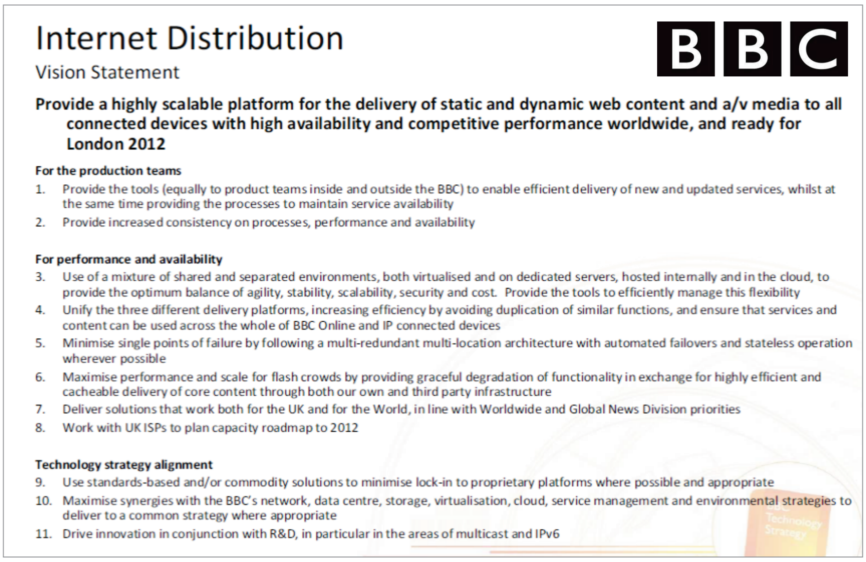
An illustration of such an in-house Internet delivery strategy is provided by Netflix and Google, which developed and rolled-out their own proprietary Content Delivery Networks and deep-caching platforms (i.e. content caches implemented within the ISP's access network, not at its edge, and, thus, closer to the end-users). The main rationale is to lower costs of delivery while substantially increasing end-users' Quality of Experience.

Furthermore, Netflix offered ISPs the possibility of peering directly at one of its eight settlement-free peering exchanges, or to install Netflix's proprietary CDN platform, Open Connect, into the ISPs' networks. Proprietary deep caching solutions are a clear cut for CAPs aiming at controlling the quality of content delivered while saving costs. Indeed, deep caching also creates strategic advantages against new players and enhanced negotiation power towards independent commercial CDN and IP Transit providers. Some commentators argue that the development of proprietary infrastructures may, given the Best-Effort nature of the open Internet, generate an indisputable strategic advantage against smaller competitors.

For instance, Netflix's Open Connect infrastructure strongly supports the launch of Ultra High-Definition (4K) services. This move certainly allowed Netflix to keep both IP Transit and content-delivery costs under control, while simultaneously offering direct control over the delivery quality promised to its subscribers.

Besides, it is remarkable how, in 2012, the launch of Open Connect immediately impacted the company's existing relationships with CDN providers. Prior to the launch of Open Connect, Netflix's traffic was delivered through the three main global CDN players (Level 3, Akamai, Limelight) in almost equal proportions. However, six months later, almost 78% of Netflix's traffic was redirected to its proprietary Content-Delivery infrastructure, leaving only 22% of its total traffic to the company's traditional CDN suppliers.

Figure 28: BBC's online technology strategy



Elements of Internet Distribution strategy:

- Combine shared and dedicated environments to reach the optimum balance in terms of agility, scalability, security and cost
- Minimize the number of single points of failure with a multi-redundant, multi-location architecture
- Manage flash crowds by graceful degradation for highly efficient and cacheable delivery through own and third-party infrastructure
- Use standard and commodity solutions where appropriate
- Maximize synergies with own network, data center, storage, virtualization, cloud and service management

Source: BBC, Arthur D. Little analysis

Figure 29: Content delivery strategies go in-house when size becomes critical

	Outsourcing	Hybrid	In-house
Description	■ Major services outsourced with a turnkey contract	■ Sourcing of managed cloud services vs. hybrid environments with internal dedicated/ cloud resources	■ Delivery of core services through internal resources
Pros	■ limited internal hardware, software and human resources required ■ Low investment level / low entry barriers	■ Competence center for delivery of a subset of services ■ Moderate investment level / moderate entry barriers - sustainable as needed	■ Full control over QoE and management of demand variability and service hurdles ■ Competence centers for core services (specialized ICT for delivery services vs. standard for corporate services)
Cons	■ Control over QoE and management of demand variability and service hurdles ■ Limited internal competencies (for managing providers and SLAs)	■ Partial control over QoE and management of demand variability and service hurdles	■ High investment level / high entry barriers ■ Internal hardware, software and human resources required
Best fit	■ Video delivery is not core ■ Media volumes are low ■ Low quality of service/performance required to secure end-user's quality of experience	■ Video delivery is relevant, but not mission critical ■ Media volumes are not yet above the critical threshold for investments ■ Performances and QoE are relevant	■ Video delivery is mission critical ■ Media volumes are above the critical threshold for investments ■ Excellence in media delivery and QoE is pursued

– Control and web delivery maturity +

Leading publishing web property	MEDIASET	BBC	NETFLIX

Source: Arthur D. Little interviews

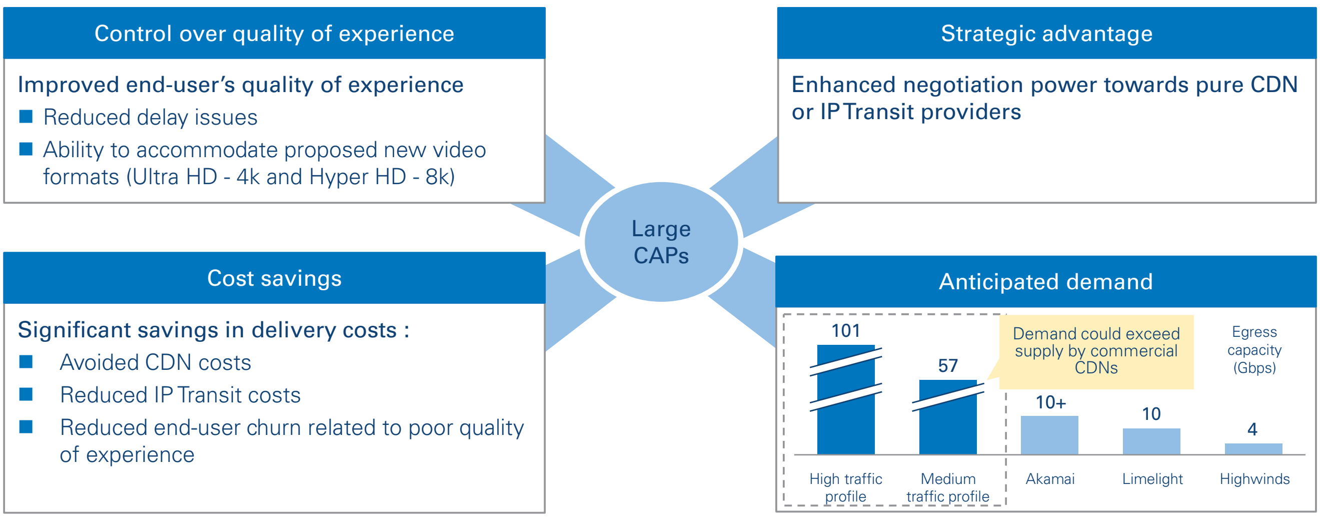
Yet, the falling prices of Content Delivery Network services tend to support smaller CAPs in developing their own CDN-based delivery strategies if it is via commercial independent CDN service providers.

Moreover, ISPs can also provide an alternative solution: they are developing their own commercial caching solutions as an alternative to deep caching by Content and Application Providers, as for security and liability reasons (i.e. illegal or harmful content), they are reluctant to accept proprietary third-party equipment in their networks.

Still, no Internet delivery strategy can fit all possible needs. We expect that Content and Application Providers will be "multi-homing", i.e. combining internal solutions with third-party services in order to balance the investments required and the level of Quality of Experience delivered. Such a strategy will create a robust environment of shared and dedicated resources, allowing:

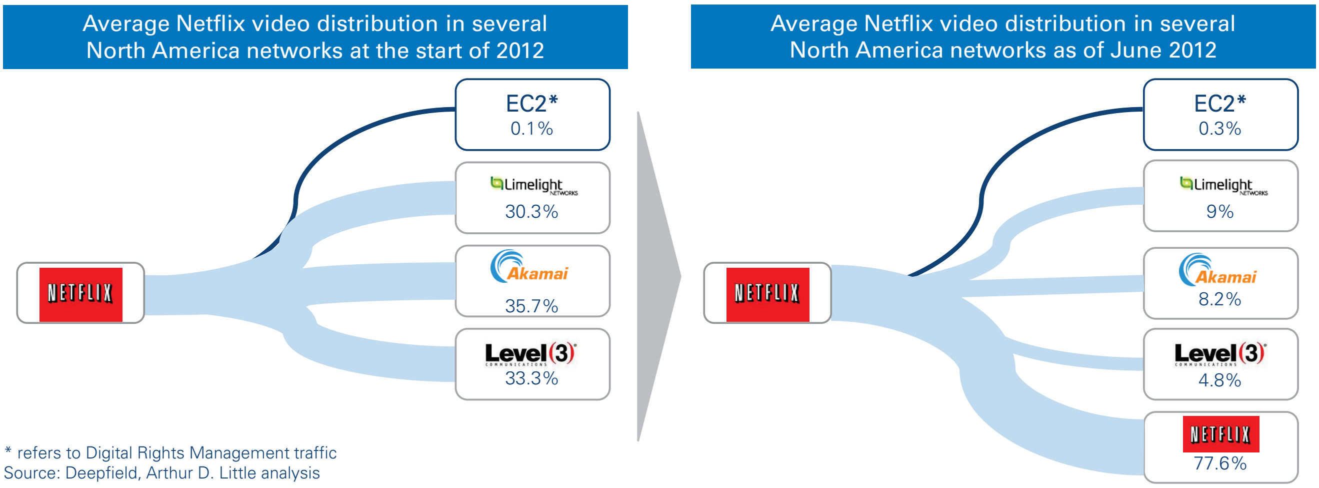
- Optimal balancing agility, scalability, security and cost;
- Minimization of single points of failures via a multi-redundant, multi-location architecture;
- Maximization of synergies with own network, data center, storage, virtualization, cloud and service management.

Figure 30: Rationale for proprietary content-delivery solutions



Source: Companies websites, Arthur D. Little analysis

Figure 31: Netflix Open Connect impact on content distribution model



* refers to Digital Rights Management traffic
Source: Deepfield, Arthur D. Little analysis

2.3. A new power emerged: the arrival of the Internet Global CAPs

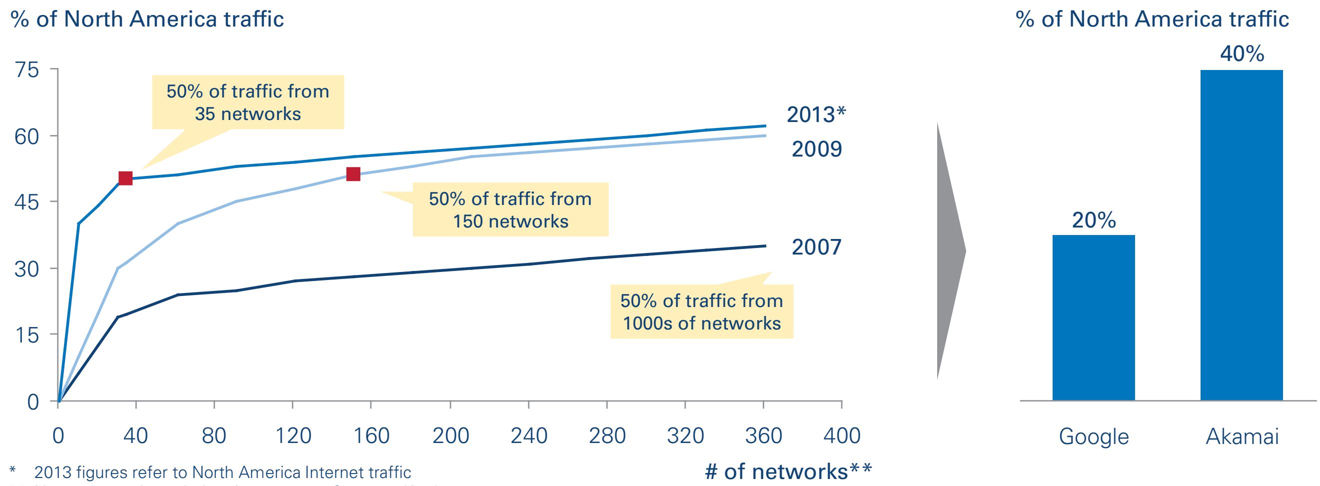
Internet Global Content & Application Providers have emerged and progressively concentrated the majority of traffic

Over a decade ago, content on the Internet was supplied by myriad end-users and CAPs in a balanced way over thousands of networks. But this situation changed radically around 2007, to the point that today less than a few dozen networks manage half the total Internet traffic. In North America, fewer than 35 Network Autonomous Systems account for more than 50% of Internet traffic.

Internet traffic was progressively concentrated by a limited number of large companies, defined as "Hyper-Giants" by some European regulators, such as Ofcom. These are the leading Content and Application Providers (e.g. Google, Amazon, Netflix) and the leading independent Content Delivery operators (e.g. Akamai), which are rapidly becoming the main traffic conveyors on the Internet.

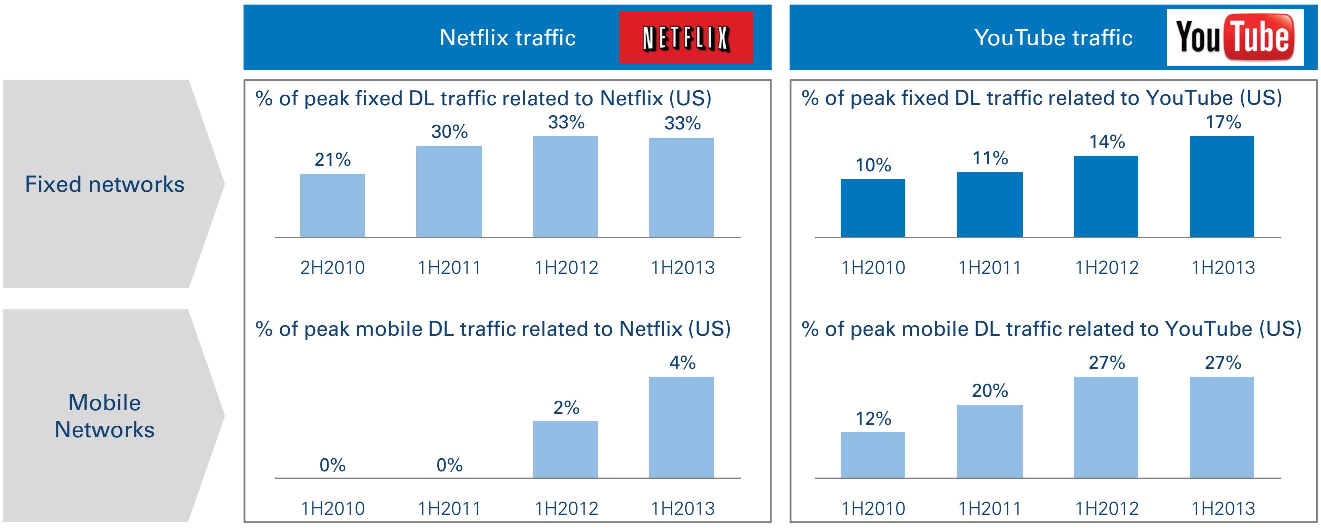
In the first half of 2013, YouTube and Netflix accumulated 50% of total download traffic during the peak time of the US fixed networks, and 30% of that of the mobile networks.

Figure 32: Traffic concentration by Autonomous Systems (2007, 2009, 2013)



* 2013 figures refer to North America Internet traffic
** Networks are intended as Autonomous Systems (Ass)
Source: Deepfield; Arthur D. Little analysis

Figure 33: Netflix & YouTube share of Internet traffic



Source: Sandvine, Arthur D. Little analysis

At the global level, Google's Autonomous System carried about 20% of the global IP traffic in 2013 (+14pp with respect to 2009, and +18,6pp versus 2007), making Google's Internet presence bigger than that of Facebook, Netflix and Twitter combined.

A large proportion of Google's extensive Internet presence is driven by data-heavy YouTube videos on fixed networks. Furthermore, the heavy usage of YouTube and (location-based) searches on mobile devices only confirm the trend.

The emergence of Internet global CAPs is a major evolution in the value chain, and has the potential to influence the negotiating power among stakeholders and therefore the current equilibrium in the ecosystem:

- Global CAPs are in a position to leverage their size to negotiate reduced unitary prices (per GB or per Mbps) in exchange for massive volumes at the IP Interconnection level;
- They are also in a position to leverage their exclusive premium content and applications, and thereby potentially steer end-users' decisions regarding selection of an Internet Service Provider.

2.4. Investment strategies for future IP Interconnection

The changing nature of the Internet from a data platform into a media-delivery platform is accompanied by an explosion of traffic at IP Interconnection nodes – i.e. network in and out ports. There is also a high degree of imbalance between incoming and outgoing traffic flows. Apart from the need for greater capacity, the demand for higher transmission quality by new (subscription based, advertising based and transactional) content & application business models needs to be satisfied.

Capacity-related investment strategies

Capacity at IP Interconnection nodes needs to be continuously upgraded throughout the interconnected networks passing Content and Application Providers' traffic to end-users, in particular at the edges of the ISPs' access networks.

The investments required for such capacity upgrades need to be put in perspective with the ongoing competitive dynamics in the market of content and application delivery services. With the possibility of delivering content and applications over the Internet, Content and Applications Providers started to provide their services "Over-the-top" (OTT) – i.e. over the open Internet instead of the close content distribution platforms of Terminating ISPs – thereby cannibalizing part of the Terminating ISP's revenues.

As voice and SMS services become commodity services and generate declining revenues, the competitive pressure has increased even further among Terminating ISPs (fixed telecom operators, mobile telecom operators and cable operators).

ISPs are challenged to manage any incremental costs of investments in IP Interconnection capacity upgrades and relate any additional investments with new revenue flows.

The symbiotic relationship between content and application services provided over the Internet and Internet broadband access services is challenged as traffic continues to grow considerably and subscriber growth and/or price increases become insufficient to stabilize the flattening – or even decreasing – revenues of the telecom sector. This is due to both commoditization of voice and SMS services and cannibalization by Over-the-top competition.

ISPs seek to adjust existing practices designed for a symbiotic growth scenario and implement new IP Interconnection models to support innovative, quality, value-added applications. For this purpose, ISPs started to take various initiatives at IP Interconnection level:

a) Eliminate inefficiencies or redundant intermediation (e.g. disconnect from redundant IP Transit or Content Delivery Network operators);
b) "De-Peering": Rationalize the number of Interconnection interfaces where the volumes involved are below the break-even point (e.g. Peering relationships with interfaces below 2-4Gbps);
c) Avoid investments or cost increases beneficial only to one party (e.g. capacity increases at asymmetric Peering ports);
d) Propose new commercial arrangements (e.g. Paid Peering when traffic is heavily asymmetric);
e) Open their interactive platforms: Monetize the opportunity cost of hosting potential OTT competitors (e.g. hosting a video OTT provider on own IPTV platform);
f) Vertically integrate: Address upcoming quality of delivery needs with new value-added delivery solutions (e.g. offering their own network caching-, CRM- or data center services directly to Content and Application Providers;
g) Make greater use of, or introduce innovative delivery technologies for offering managed quality over IP networks (such as MPLS and IPX.).

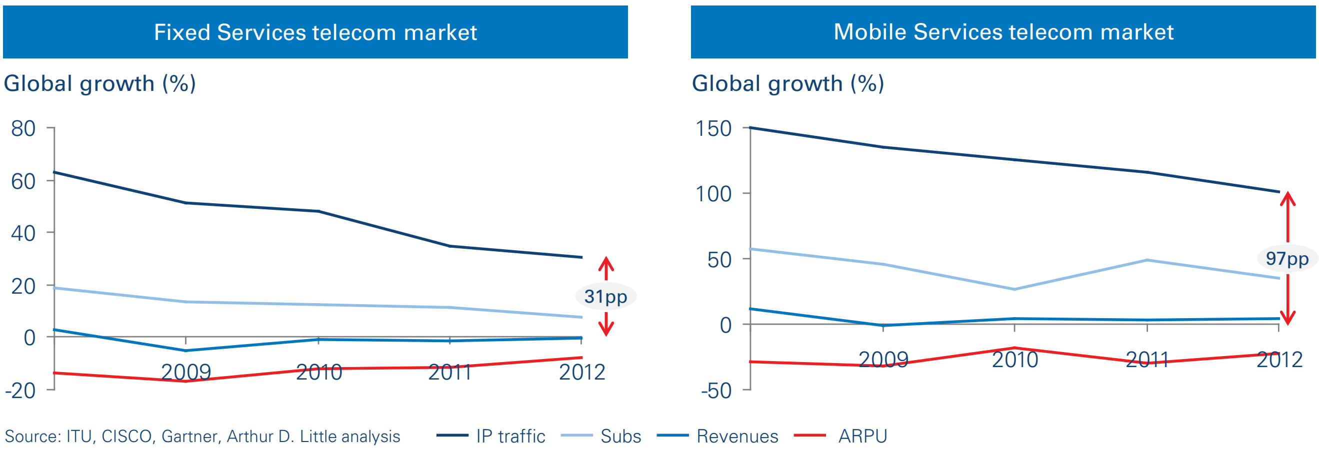
These initiatives support the overall trend to move IP Interconnection from legacy relationships driven merely by volumes and number of interfaces to quality relationships in which specific levels of services are negotiated and delivered.

Quality-related investment strategies

It is noteworthy that, in addition to investing in building proprietary CDNs, CAPs are actively investing in new bandwidth-saving technologies to make the end-user's Internet experience, especially for video consumption, smoother, and reduce imbalances at IP Interconnection interfaces. In particular, Google is known to be active in developing:

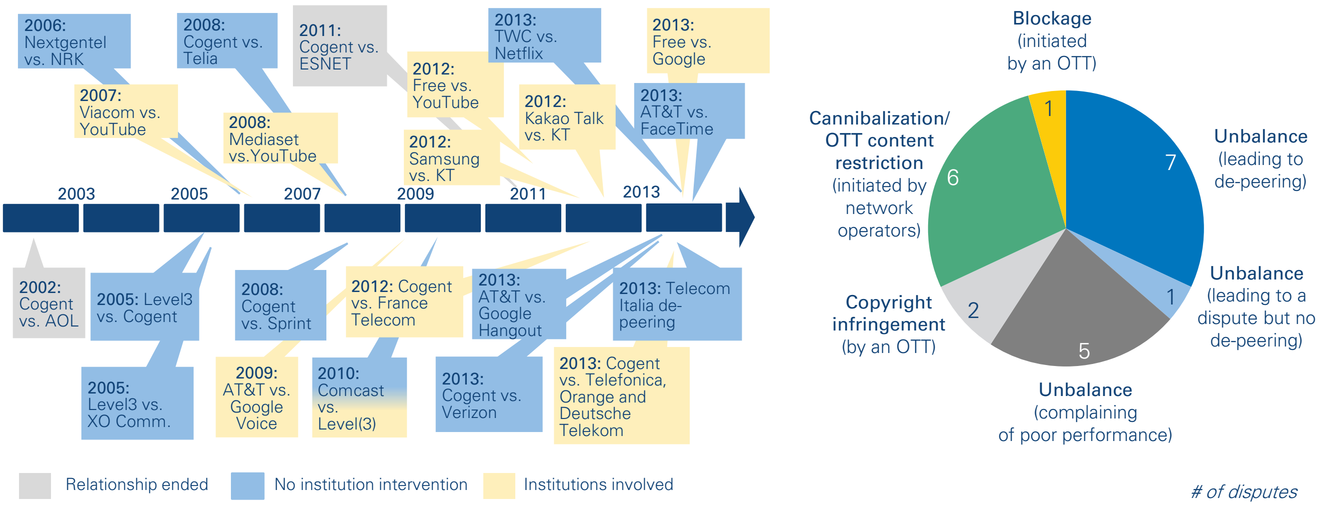
- Adaptive streaming technologies: these technologies can break down individual videos into multiple segments according to each available video definition. The video player can then adjust the video definition (e.g. from 720p down to 480p) to account for temporary fluctuations in bandwidth or congestion;
- Codecs for video compression: video data can be efficiently transmitted across the Internet while using compression technology. New video formats such as WebM, also known as VP9, aim to halve the bandwidth required for HD video through an open-source platform.

Figure 34: Growth comparison between traffic volumes vs. revenues, subscribers and ARPU



Source: ITU, CISCO, Gartner, Arthur D. Little analysis

Figure 35: Selected disputes at IP Interconnection level



Source: Websites; Arthur D. Little analysis; Woodcock, OECD

2.5. Friction occurs in the IP Interconnection value chain, but are (so far) quickly resolved

Since 2008, disputes in the ecosystem have arisen and intensified, yet their resolution is usually quick.

Disputes only concern less than 1% of all IP Interconnection agreements, and are solved by commercial agreement in more than 50% of cases.

The changes in the IP Interconnection value chain and the search for economically and technically sustainable business models to support interconnection models in the future (i.e. satisfying the cost and quality requirements of next-generation Internet content and applications) bring tension between the different players, which sometimes translates into disputes.

An analysis of the major disputes starting from 2003 onwards reveals that most disputes are triggered by two recurring causes:

- Traffic imbalances;
- Cannibalization of ISPs' services by OTT services.

Figure 36: EU regulator approach to IP Interconnection disputes

EU regulatory attention points

1. Degradation or interconnection failure due to failed negotiations between players

→ Impossibility for users to access, distribute or use the application and services of their choice

2. Anticompetitive behavior

→ Interconnection could be used to engage in anticompetitive behavior towards the source, the destination, or the content of the information being conveyed

EU current state of regulatory decisions

No stringent regulatory framework
…given the current state of the market and calls for caution with any binding measure of regulation whose effects could be potentially harmful

Focus on two principles: non-discrimination and transparency

No linkage between non-discriminatory conditions of interconnection and net-neutrality debate

Acknowledgment that IP Interconnection regards a general issue on the financing of networks and the economic balance between operators and users

Interest in gathering information from the players
…monitoring market trends and the QoS of Internet access in order to be ready to settle disputes

Source: BEREC, Arthur D. Little analysis

Other causes relate to various other types of imbalances or blockages.

Still, it is worth noting that these disputes only relate to a limited number of existing IP Interconnection agreements (less than 1% according to a study by the OECD) and in more than 50% of cases, the involved parties have been able to solve them without third-party intervention. Since 2008, only two dispute cases led to the end of the interconnection relationships.

Certainly, there have been numerous other disputes that never achieved a public dimension and were settled privately.

So far, Regulatory Authorities have generally preferred a non-interventionist approach and focused their attention on the principles of non-discrimination and transparency.

In particular, European Regulators focused their monitoring activities on two main concerns:

1. The risk that failure in negotiations between Internet players may result in Internet service degradation for end-users;
2. The risk of anticompetitive strategies that may be implemented by biggest players.

Accordingly, they opted for a non-stringent regulatory framework that allows the market to freely find its equilibrium, and innovate its business models. This acknowledges the IP Interconnection mechanisms' place at the core of the financing of required network investments, as well as the economic balance between actors of the Internet value chain and end-users.

Key messages

- The majority of Internet traffic is progressively being concentrated to a limited number of large Content & Application Providers, the so-called "Hyper-Giants".
- The emergence of Internet global Content & Application Providers is a major evolution in the IP Interconnection value chain and has the potential to influence the negotiating power among connectivity stakeholders and affect the equilibrium in the Internet ecosystem.
- The IP Interconnection value chain converges. Content & Application Providers and ISPs are setting the pace and determining the nature of IP Interconnection innovation by vertically integrating and by interconnecting directly, disintermediating pure Internet connectivity providers to some extent.
- Changes in the IP Interconnection ecosystem led to tension between IP Interconnection players. However, disputes concern less than 1% of all IP Interconnection agreement,s and are solved without regulatory intervention in more than half of these cases.
- Countervailing powers emerged by changing IP Interconnection economics that keep the value chain in balance, and end-users have not been substantially or structurally affected by IP Interconnection disputes.

3. Revolutionary Future Applications that Require New Delivery Features

3.1. The future Internet driven by the Internet of Things and the Internet of Humans?

The Internet is continuously mutating, and new Internet applications emerge in an unpredictable way, very often exceeding our imagination.

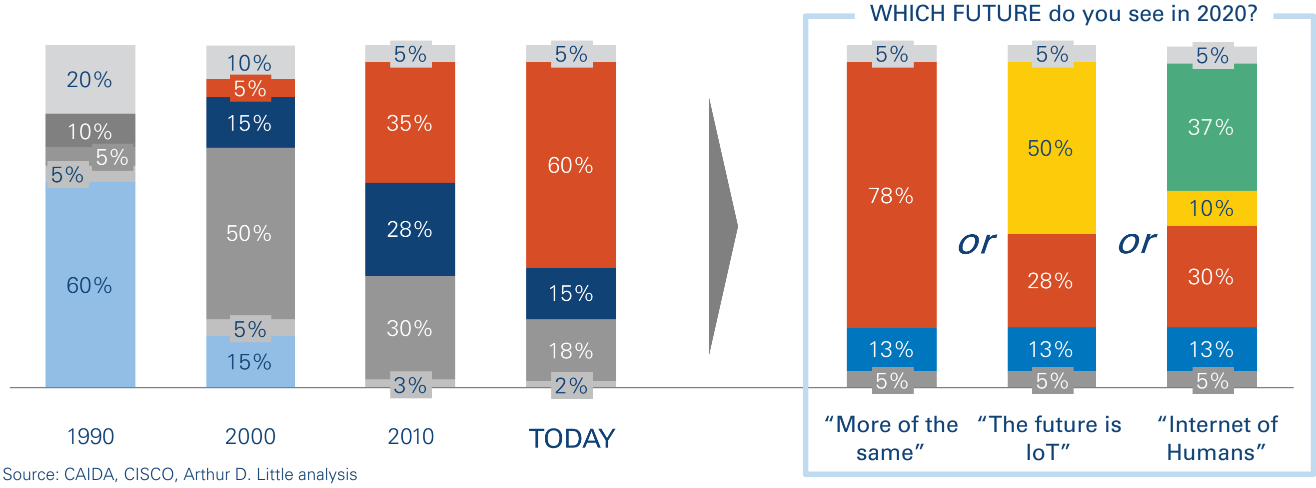
The history of the Internet application landscape is rich in changes and disruptions. In the early 90s, the "File Transfer Protocol" (FTP) application accounted for more than 50% of the total traffic. A decade later, the share of FTP traffic was substantially reduced, being replaced by growing web browsing. Similarly, web browsing accounted for the lion's share for 10 years, before being supplanted by media content (particularly one-way video downloading and streaming), which cumulated in over 30% of total traffic in 2010. At the same time, peer-to-peer exchange of media content between end-users reached a similar share. Nowadays, one-way video traffic has grown to 60% of total Internet volumes and peer-to-peer traffic is losing ground relative to this.

This teaches us how unpredictable the evolution of the Internet application landscape can be, raising the question of whether one can reasonably forecast the future of Internet.

Looking forward, we can only work with scenarios. For the next decade, most stakeholders agree with three likely outcomes by 2020:

- *More of the same:* For the next decade, one-way video streaming traffic may remain the main source of traffic. Most of the content consumed would be generic, hence limiting the need for real-time connection – i.e. the content could easily be cached. Traffic volumes will continue to increase as the number of end-users grows, their individual usage increases and higher-definition standards are brought to the market;
- *The Future is the Internet of Things (IoT):* Alternatively, we could experience a mass adoption of mission-critical machine-to-machine (M2M) applications, cumulating in up to 50% of traffic, such as remote health monitoring & care, driverless connected vehicles, smart grid and smart traffic control. Besides this, governments may seek to increase the efficiency of their administration and push for the development of secured e-Administration applications.
- *The Internet of Humans:* We could also see the mass-adoption of two-way, real-time high-definition video applications, which would enable rich and remote human interactions. Such online human interactions would be driven by the emergence of advanced collaboration in the context of telemedicine, online crowd-working, etc.

Figure 37: Internet traffic split by applications: yesterday, today, tomorrow



Source: CAIDA, CISCO, Arthur D. Little analysis

FTP | E-mail | TELNET | DNS | WEB | P2P | Data | 1-Way Video | IoT M2M | 2-Way Realtime Video | Other

Nonetheless, it appears clear that the advent of next-generation applications may substantially change the Internet eco-system and push the current performance requirements to new levels. This may imply new network architectures, IP Interconnection business models, definition of Quality of Service and Quality of Experience, Internet Governance, etc.

As the IP Interconnection eco-system that we know today may have little to do with these future requirements, a closer look at envisioned promising next-generation applications is needed.

Great services are emerging that could revolutionize society and our interactions within it

Every day, the Internet community (software developers, Content and Application Providers, professional service providers, equipment manufacturers, telecommunication operators, etc.) is releasing an increasing number of fascinating concepts of future applications that show how the everyday interactions in our professional and personal lives could evolve.

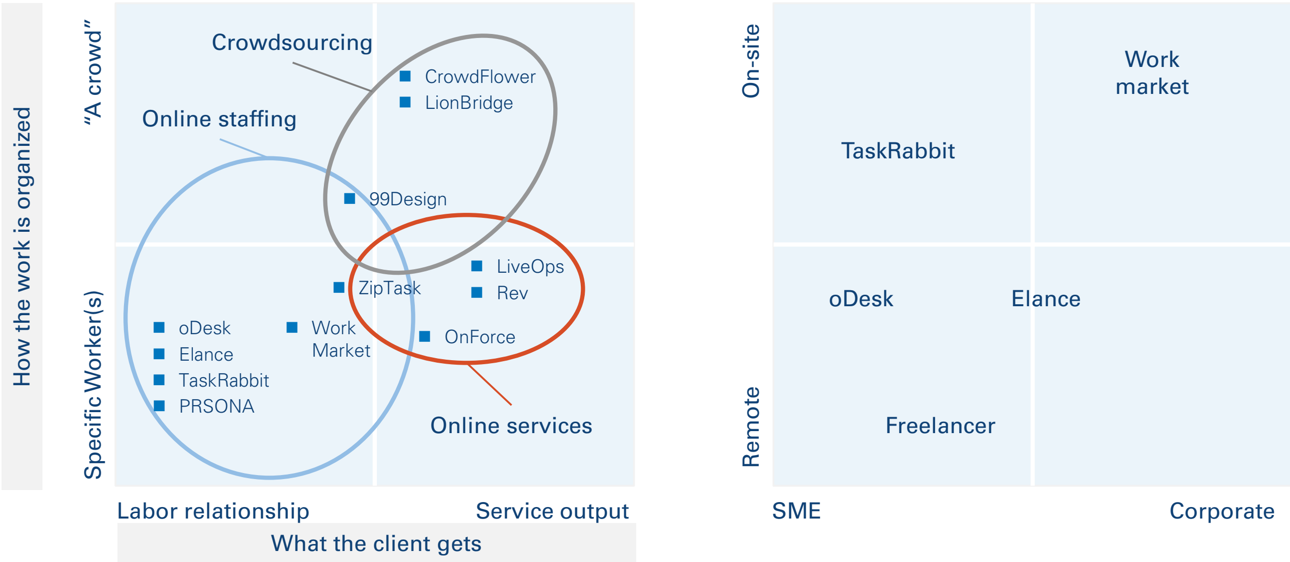
Hereafter, we illustrate how applications promise to revolutionize human-to-human interactions through Ultra-high-definition real-time video communication:

- **Ambient presence** would enable remote human interaction through wall-sized screens for, as an example, a telemedicine service encompassing diagnosis, treatment, monitoring, and, patient education, and providing convenient, site-independent access to expert advice and patient information;
- **Secure home delivery** would allow remote control of access to our homes and offices by remotely operating door-locking systems, monitoring the identity of the person knocking at the door and by ensuring that the package (e.g. online purchase) or service (e.g. plumber) delivery is finalized in a secure environment;
- **Remote care** would allow patients affected by chronic diseases to interact in real-time with remote relatives and access on-site medical equipment that provides critical information regarding vital functions;
- **Online personal training** would enable online training, coaching or professional advice via real-time video communications;
- **Ubiquitous HD videoconferencing** would enable all working desks to access any other working desk(s) in real-time and with a high-definition standard for improved collaboration;
- **Advanced work collaboration** would integrate ubiquitous HD videoconferencing with interactive boards for immersive collaboration experiences.

Figure 38: Internet of Humans - examples of next-generation applications

Source: Corning, Microsoft, Sprint, Arthur D. Little analysis
1. http://www.corning.com/adaymadeofglass/videos/index.aspx
2. http://tomorrowawards.com/showcase/1222/sprint-autobahn-thirty-six
3. http://www.itproportal.com/2013/09/02/ifa-2013-live-breaking-news-photos-and-analysis-from-berlin/
4. http://tomorrowawards.com/showcase/1222/sprint-autobahn-thirty-six
5. https://www.microsoft.com/office/vision/
6. http://www.corning.com/adaymadeofglass/videos/index.aspx

Figure 39: Example of companies in the Workforce-as-a-Service (WaaS) market



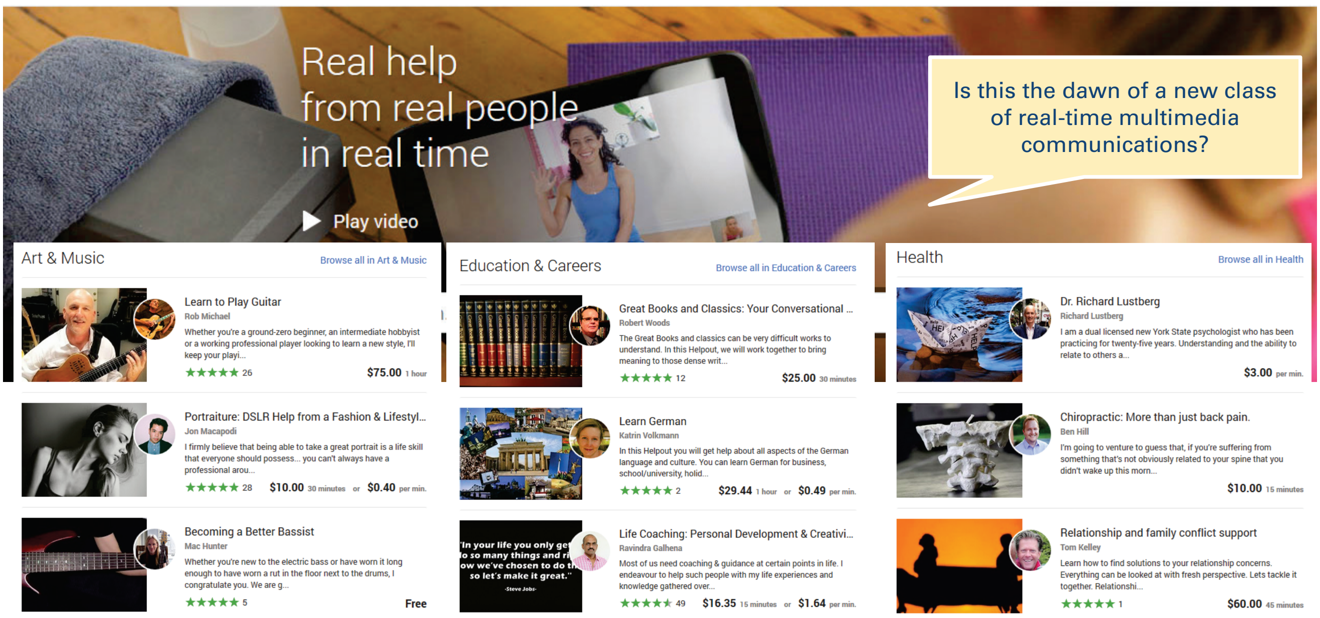
Source: WorkMarket, Arthur D. Little analysis

The common factors among these applications are stringent requirement in terms of seamless reachability and access, Quality of Experience, connectivity availability and reliability, security and privacy of the information involved. These seem distant from the current Best-Effort Internet services, which certainly offer open access to any application but cannot guarantee any Quality of Experience, especially when more than one network is involved and a network-to-network interface is required.

Next-generation applications are already reality!

The above-mentioned next-generation applications concepts may be less far-fetched than it could appear at first hand. Many Internet players are heavily investing in application platforms that provide crowd-working services, and some actors believe that in the future a large portion of work will be carried out by freelancers engaged remotely through dedicated job marketplaces.

Figure 40: Google Helpouts: the platform enabling people to help people in real-time



Source: Google Helpouts website

As WorkMarket puts it, "the traditional relationship between employer and employee is evolving as technology, globalization, and regulations disrupt the labor market. The most dramatic shift has been the rise of the freelance economy, which currently includes 17.7 million independent workers, and is expected to grow to 24 million by 2018 in the United States alone (MBO Partners, The State of Independence in America). With a new pool of talent available on an on-demand basis, businesses are rapidly shifting to an extended workforce model as freelancers continue to augment full-time employees."

Consequently, many start-ups are positioning themselves along the multiple dimensions of the Workforce-as-a-Service (WaaS) market.

In November 2013, Google launched its new platform Google Helpouts, a service platform that organizes a marketplace to bring together end-users and professional service providers or subject-matter experts. End-users pay for live video sessions with experts who provide, via online video support any type of service ranging from step-by-step instructions on how to cook a turkey on Thanksgiving Day to marriage counseling. (Figure 40)

Meanwhile, TeleCure, which provides high-quality medical care via telephone and video consultations, announced that it would use Google Helpouts' platform to deliver its services. Similarly, TakeLessons, an online marketplace that provides students with the best instructors, announced that it would launch on Helpouts by bringing a top-notch selection of music teachers, performing-arts instructors and tutors to the platform.

Such initiatives are expected to have a huge impact on the Internet application landscape, as they are promoting a new class of applications, which differ greatly from currently watching a video over YouTube or Netflix. They involve two parties interacting in real-time through a video communication, whose quality must be good enough to make the service worth paying a financial transaction.

3.2. Today's Internet is a Best-Effort and finite (yet-not-scarce) resource

Most people have already experienced what they would define as inconsistent Internet delivery quality, usually reflected by:

- Buffering video applications or web pages that require several seconds to load;
- Poor quality of video streaming or video communications due to high jitter;
- Heavily fluctuant instantaneous download speeds, which are often lower than the peak speed offered by a Terminating ISP.

Such situations are the consequence of the today's Internet being a Best-Effort and finite (yet-not-scarce) resource. The Internet we know today is affected by various limitations due, for example to physical constraints when long, end-to-end distances apply, or to the intrinsic features of the control layer governing the IP protocol (the Transfer Control Protocol, TCP).

While an exhaustive technical analysis would require diving into a much broader set of considerations and details, the key high-level characteristics of the current Internet's available performance can be highlighted.

Best-Effort prevails

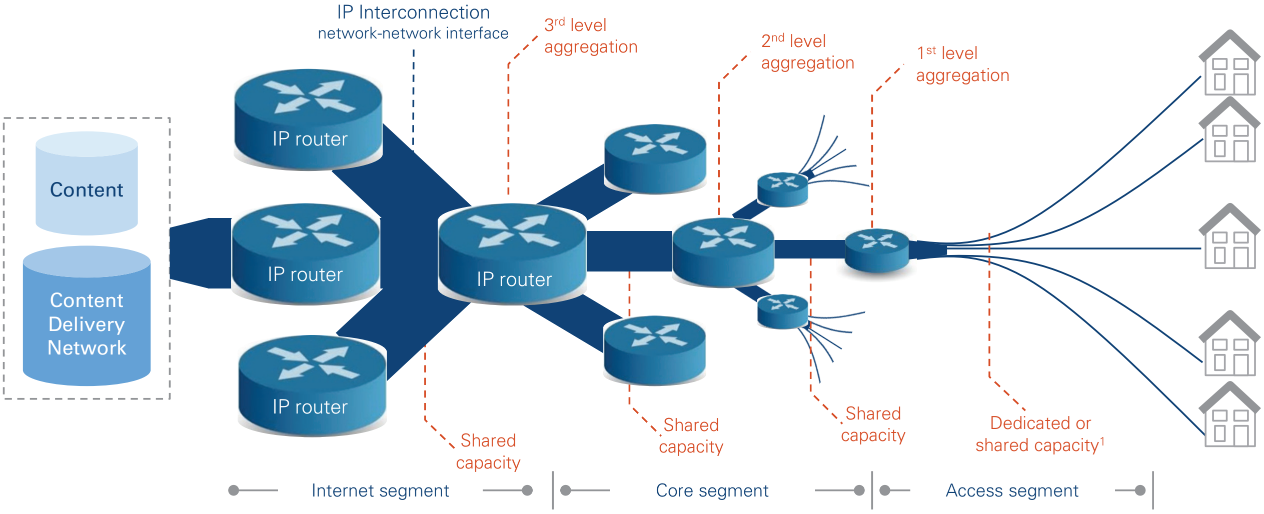
As introduced in Chapter 1, the Best-Effort feature is simultaneously one of the magic ingredients that enabled the fast expansion of the IP technology and its intrinsic constraint.

Indeed, by its very nature, the Internet Protocol cannot, similarly to the traditional circuit-switched telephony system, establish a circuit with dedicated logical resources that interconnect the originating and terminating parties. The IP is a packet communication technology.

This implies that after each aggregation node, the available capacity is shared among the different outgoing and incoming traffic flows – hence, among the different end-users. Congestion can only be mitigated with abundant installed resources or undiscriminating traffic management. Though these mitigation techniques can assure an average throughput, they can only partially assure other delivery features such as the variance of packet arrival (jitter) which greatly affects video delivery.

Jitter becomes even more critical when data packets (whether a file, email or video communication) must travel many hops from the CAP to the end-user. In this journey, many resources are involved – from the originating server to the middle router to the terminating access network, until the processing CPU of the receiving device (PC, smartphone, tablet, etc.) – and their improper dimensioning or the instantaneous traffic load

Figure 41: Sketching an IP network



Source: Arthur D. Little analysis

at each of the involved nodes may affect the overall quality of experience.

Out of the new network protocols and technologies that are currently being investigated, only a few (e.g. MPLS technology, Software Defined Networks and IPX protocol) would be capable of establishing upon request the release of proper logical resources at network interfaces.

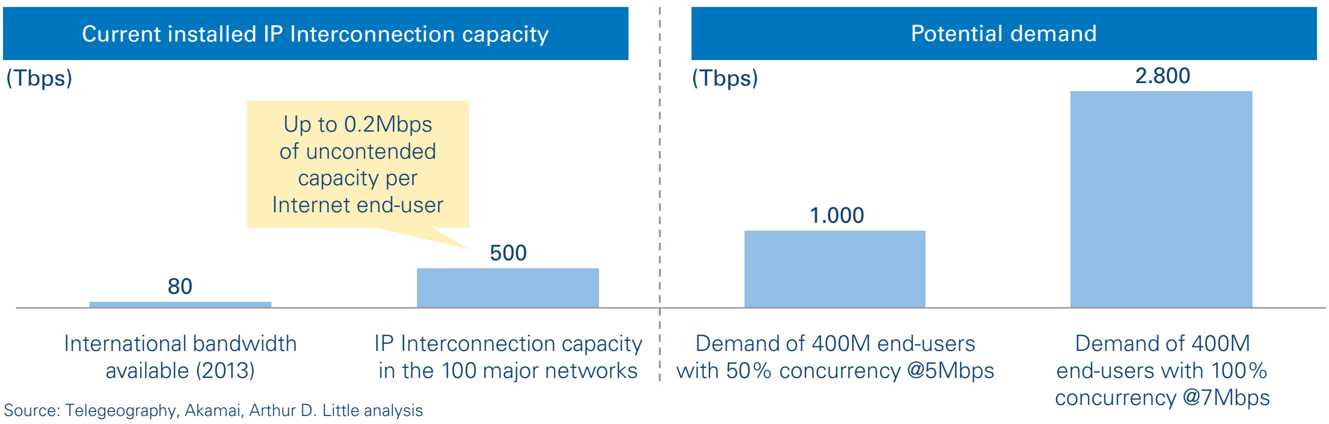
Still, the Best-Effort nature of IP technology does not necessarily imply low performances. On the contrary, the simplicity of adding capacity at interconnection nodes enabled its wide and fast propagation.

Many believe that the quality achieved is "Good Enough" for the vast majority of applications.

Internet resources are finite (yet not scarce)

For example, contrary to wireless spectrum, the IP Interconnection capacity is not a scarce resource because it is possible to easily add capacity at each interconnection. But it is finite, in the sense that each interconnection node can count on a finite amount of routing capacity before new investments are made. The fact that each interconnection node is shared among all traffic flows implies that each user benefits from just a small portion of the uncontended capacity.

Figure 42: Current IP Interconnection capacity and potential demand



Source: Telegeography, Akamai, Arthur D. Little analysis

The IP Interconnection capacity is not a scarce resource because it is possible to easily add capacity at each interconnection. But it is finite, in the sense that each interconnection node can count on a finite amount of routing capacity before new investments are made.

Given that all end-users do not use the Internet simultaneously, they do not request IP Interconnection capacity concurrently, and IP Interconnection nodes are therefore dimensioned according to statistical overbooking rules (i.e. the foreseen capacity per end-user is only a fraction of the capacity available in his local access network).

However, the new nature of the Internet as a media platform is challenging current assumptions; especially during primetime. It is not rare that concurrency can exceed 50% for special live events such as sport games or concerts.

Figures relating to the current capacity supplies at IP Interconnection levels are disclosed with parsimony. Still, Telegeography suggests that 80Tbps of international capacity were installed by 2013, and Akamai estimates that 500Tbps are available at the IP Interconnection interfaces of the 100 major global networks.

500Tbps made available to the 2.7 billion global Internet end-users implies that there would be roughly 0.2Mbps of uncontended capacity available at the IP Interconnection level for each end-user. Such an estimate resonates well with assumptions by the French Regulatory Authority ARCEP, which assessed the average peak-hour capacity consumption in fixed networks between 0.1Mbps and 0.3Mpbs.

The finite nature of shared resources at IP Interconnection level becomes even more obvious when comparing the current installed capacity with future demand scenarios in which hundreds of millions or even billions of end-users concurrently request access during primetime.

Yet, this challenge is not beyond achievement, as the required costs related to IP Interconnection capacity increases are relatively low. According to ARCEP, the upgrade costs to double the uncontended capacity for each end-user would be around €0.15.

The real challenges are that this installed IP Interconnection capacity remains shared and that higher uncontended capacity cannot be allocated upon request to specific applications (i.e. maintaining a risk that high-quality, demanding applications will still not receive sufficient logical resources). Furthermore, it remains inefficient to increase IP Interconnection capacity when more developed caching solutions could significantly avoid repetitive transmission of generic content.

The Internet is subject to latency and packet loss

The physical distance between origination and termination substantially impacts the final transmission throughput. This causality mainly relates to the Transport Control Protocol's (TCP) working mechanism that governs the speed of transmission over IP networks.

By measuring the response time from the receiving party, the TCP tests the connection capability between the origination and termination points and progressively allows the origination server to send greater volumes of data packets. In the event that the response time increases or packet losses trigger retransmissions, the TCP decreases the throughput window at the origination point.

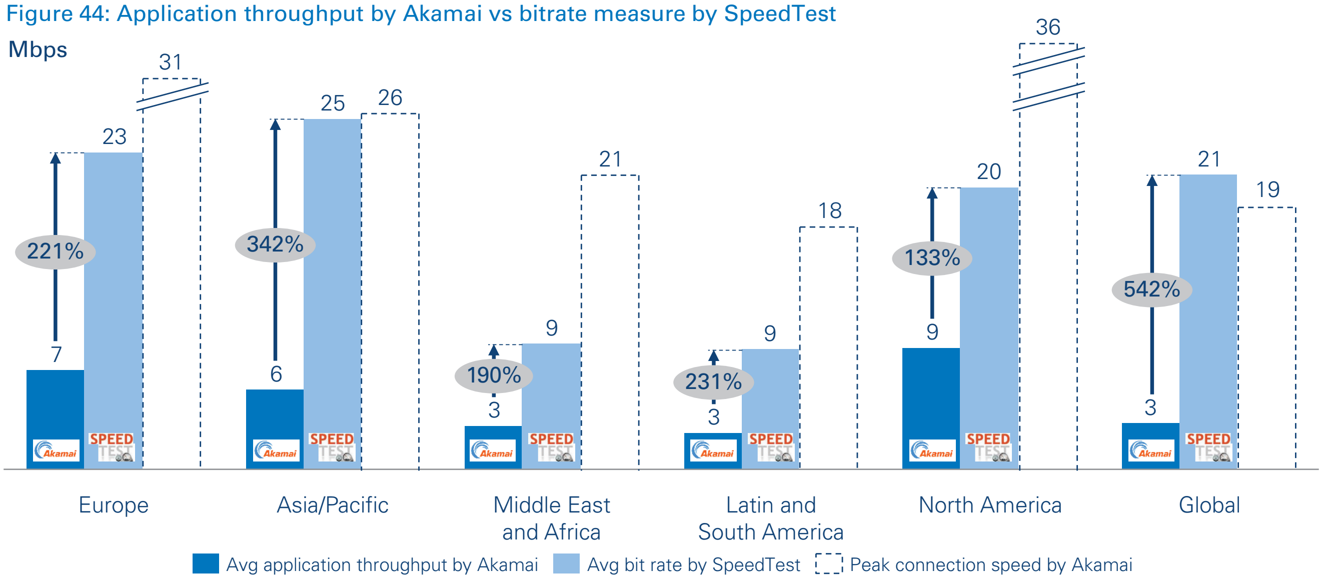
With this mechanism, the TCP can only progressively increase the throughput, unless special techniques circumventing such limitations apply (e.g. web acceleration).

Figure 43: Relationship between distance (miles), latency (milliseconds) and throughput (Mbps)

Distance from content/application server to end-user	Network latency	Packet loss	Throughput	Download time
Local (<160 km)	1.6 ms	0.6%	44 Mbps	12.2 min
Regional (800-1.600 km)	16 ms	0.7%	4 Mbps	2.2 hrs
Intra-continental (<5.000 km)	48 ms	1.0%	1 Mbps	8.16 hrs
Intercontinental (<9.000 km)	96 ms	1.4%	0.4 Mbps	20 hrs

Source: Akamai

Figure 44: Application throughput by Akamai vs bitrate measure by SpeedTest



Source: Akamai, SpeedTest measures collected by Netindex.com, Arthur D. Little analysis

Akamai performed several studies on this phenomenon, and reveals the link between distance, latency, packet loss and throughput: throughput and time to download decrease with increasing latency and packet losses.

Although loss is often insignificant in frame relay networks (less than .01% on average), it is very significant in IP VPN networks that go into and out of certain markets such as China, where loss rates commonly exceed 5%, and are often much higher. In the latter scenario, high loss rates can have a catastrophic effect on performance.

The Internet's quality is hard to measure

The analysis above demonstrates that the Internet quality of experience is a function of many factors, and its fluctuations depend on the number of concurrent end-users requesting online content and applications.

The Best-Effort IP interconnected network delivers data according to the available resources, without any assurance of reliability, delay bounds or throughput requirements. As a result, the performance is highly variable, never guaranteed and always below its nominal capacity.

It is often difficult, if not impossible, to identify the limiting bottleneck, which can be caused by congestion at an IP Interconnection or aggregation node, at the CAP's server farm or even at the processing unit of the end-user's devices.

Although this study does not intend to cover the review of all factors impacting the Internet Quality of Experience, it should be noted that other technical parameters such as link asymmetry, protocols tuning, front-end optimization, client elaboration time and policy management. also play a role in the underlying Quality of Service of IP Interconnection services.

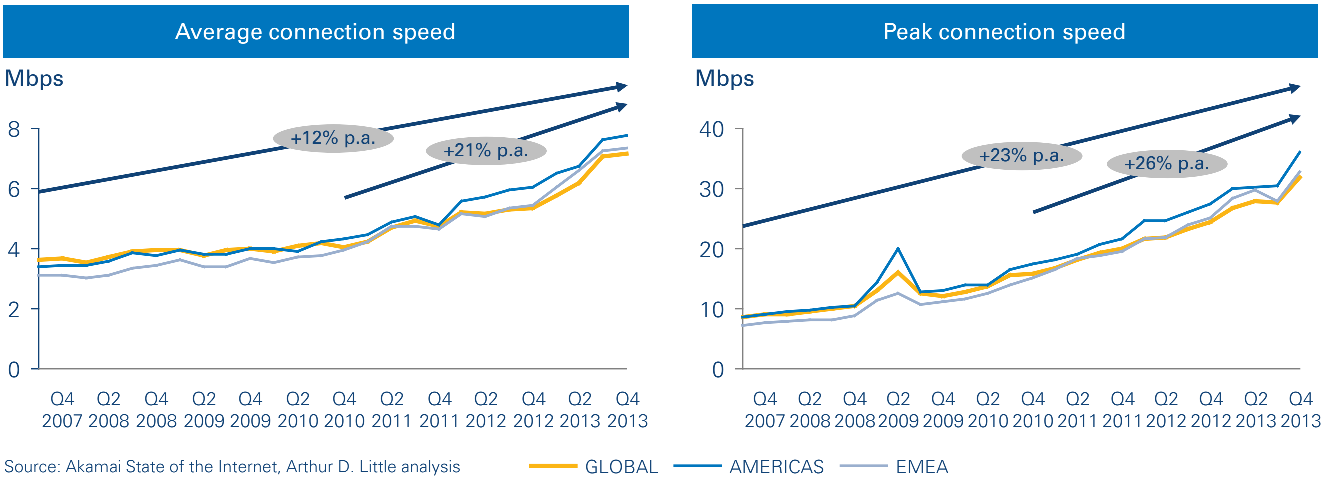
The variability of the Internet's performance is delicate to monitor and interpret. There have been several attempts to track the progress of the Internet Quality of Experience, and a number of initiatives can be listed:

- **Content and Application Providers** such as Netflix (Netflix ISP Speed Index) and Google (Google Video Quality Report – *"YouTube HD Verified"*) track (video) delivery capability of ISPs;
- **Regulatory Authorities and Governments** push end-users to adopt speed tests to measure the performance of fixed and mobile access networks;
- **Some equipment vendors**, especially Content Delivery Networks, track the average values of server-to-client connection performances.

Akamai's measures of peak and average speeds reflect a sum of factors, including the different deployment and mix of access technologies (xDSL, FTTX, cable, mobile) as well as the various network management policies and resources.

Consequently, the reader must pay particular attention when going through such results, as measurement methodologies, target measuring parameters and statistical techniques lead to very different results. This is illustrated by the wide variance in results found when comparing, for example, Akamai's measures of application throughput and SpeedTest's measures of access bit rate.

Figure 45: Evolution of Internet speeds



Source: Akamai State of the Internet, Arthur D. Little analysis

Can capacity upgrades accommodate higher-quality requirements?

The Best-Effort Internet ecosystem continuously achieved, year after year, improvement in the average speed that end-users enjoy. So far, the quality of delivery, based on an average throughput provided, was sufficient to allow the Internet to proliferate.

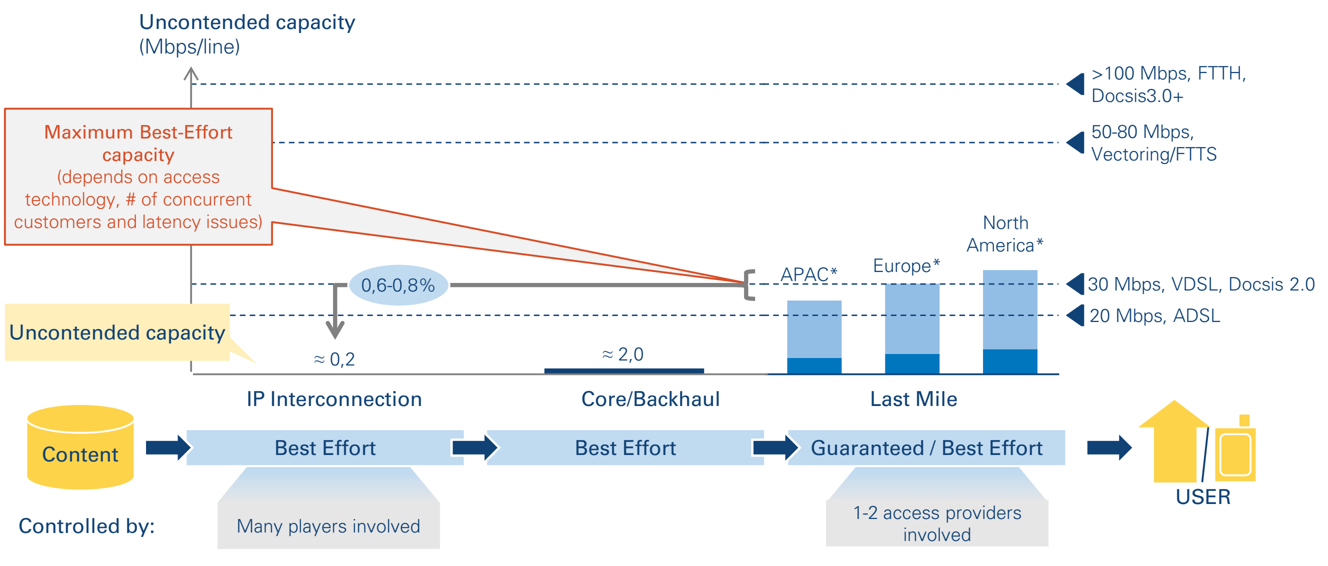
As demand for Internet content and applications will rise, it is likely that more IP Interconnection capacity will be deployed, whether by adding resources or leveraging technological progress.

In particular, as demand for concurrent consumption of content and applications will increase, the statistical overbooking ratios will be reviewed and investments will be made at IP Interconnection interfaces.

However, even when substantially increasing the statistical overbooking ratios, the risk remains that, at a given instant, the overall number of end-users and the nature of their content and applications will generate traffic that exceeds the dimensioning of IP Interconnection nodes, thereby resulting in a degraded Quality of Experience.

The performance of applications on globally interconnected networks is affected by a large number of factors, in addition to bandwidth. The notion that bandwidth solves all, or even most, application problems is, simply put, a myth. At the network level,

Figure 46: The gap between capacity in the last mile and in the other segments (2013)



Source: Akamai, Arthur D. Little analysis; * 2013 data application performance is limited by high latency, jitter, packet loss and congestion.

As an example, Google advocates[3] the greater relevance of latency versus capacity on throughput. Upgrading connectivity from 1Mbps to 2 Mbps halves the page load time (PLT), but upgrading from 5Mbps to 10Mbps results in a mere 5% improvement. At the opposite end, the latency graph tells an entirely different story. For every 20ms improvement in latency, there is a proportional improvement in page-loading times.

There are many good reasons for this: an average page is composed of many small resources that require many connections, and throughput is closely tied to the time required by our browsers to communicate with the originating content server. Technically three round-trip-times – RTTs – are required before a connection is established. Hence, an end-user paying for an Internet connection with a nominal download bandwidth (e.g. 30Mbps) that theoretically enables him to view High-Definition, real-time video streaming, could still be subject to traffic jams at IP Interconnection nodes and unable to fully enjoy his video.

Noteworthy, given the non-discriminatory nature of the Internet Protocol, is that it is very likely that a substantial share of the traffic relates to content and applications that do not have the same quality requirements (latency, jitter, security, etc.) as the requested real-time video, and is prevented from being more efficiently prioritized without degrading the end-user's Quality of Experience.

For the abovementioned reasons, it is not surprising that pioneering applications, such as Cisco's Telepresence and Telemedicine, have, so far, been rolled out via dedicated networks (i.e. static Virtual Private Networks, outside the open Internet), in which higher costs and operational complexity represent a barrier to the mass adoption of such applications.

Eventually, the question is raised as to what extent the Best-Effort nature of the Internet can actually be a constraint to the take-up of next-generation applications and whether the industry should look for an alternative solution capable of providing on-demand, guaranteed capacity.

3.3. New requirements are emerging, beyond bandwidth

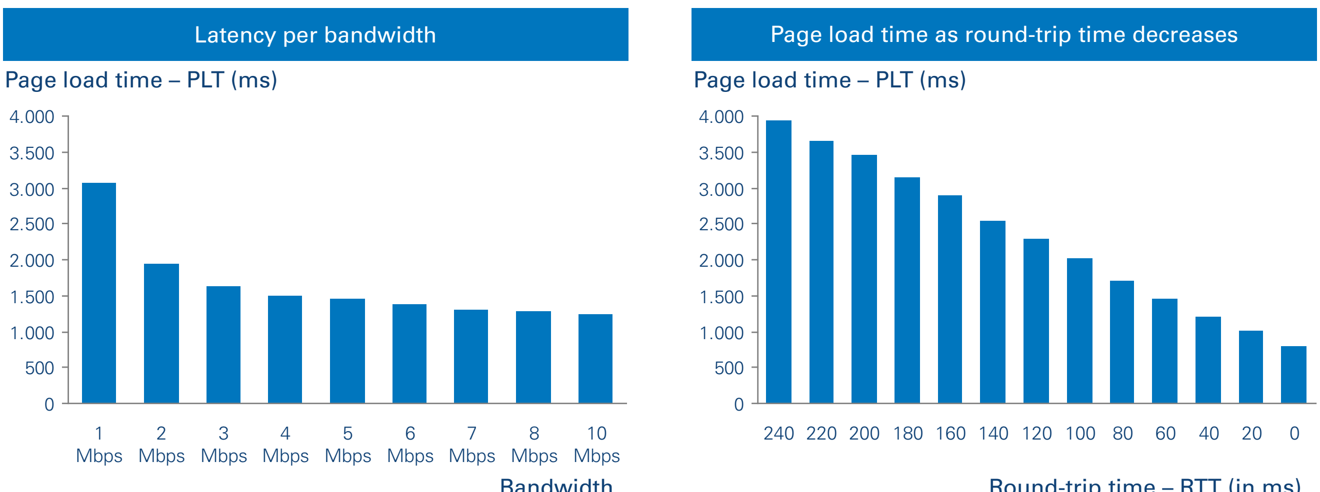
Next-generation applications demand many new requirements, much more than just bandwidth.

In recent years, much of the attention in the industry and among regulatory bodies has been focused on the development and deployment of new network technologies capable of higher bandwidth the access networks. Accordingly, we observed the progressive roll-out of FTTx and DOCSIS 3.x fixed networks and 3G/4G mobile networks.

Looking forward, we expect that next-generation applications will generate a demand for new IP Interconnection requirements going well beyond additional throughput capacity, expanding to delivery features such as latency reduction, availability, jitter control, packet-loss limitation and security.

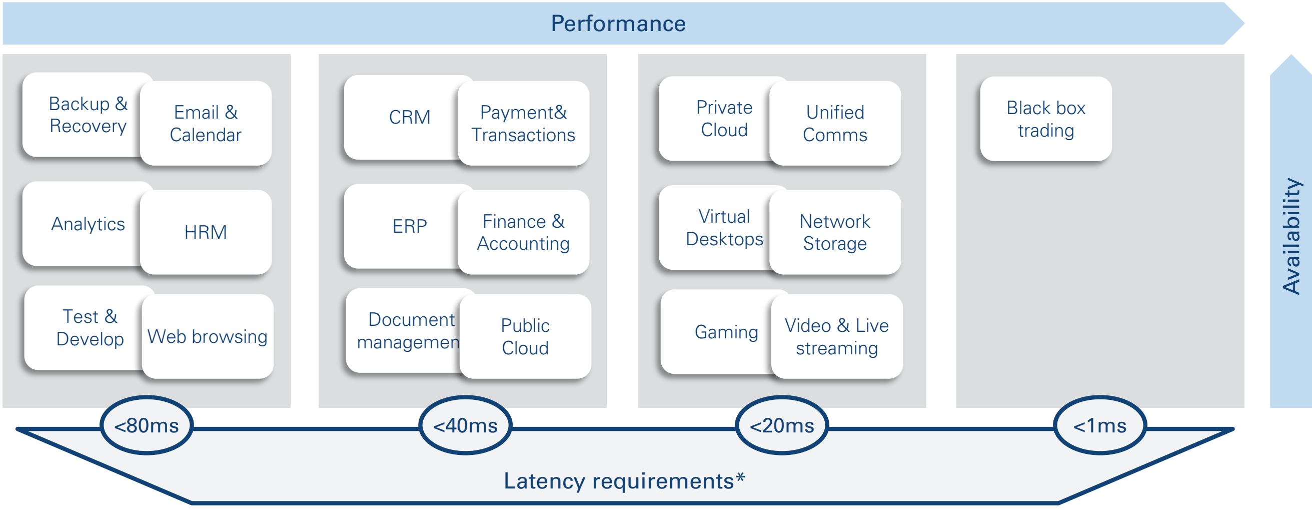
3 http://www.igvita.com/2012/07/19/latency-the-new-web-performance-bottleneck/

Figure 47: Latency versus bandwidth as a performance driver



Source: Google, Arthur D. Little analysis

Figure 48: Internet services traditional and emerging transport features



Source: Interxion, Arthur D. Little analysis; * round trip delay

Actually, such demand for requirements other than mere bit rate has existed for over a decade in Business-to-Business (B2B) segments.

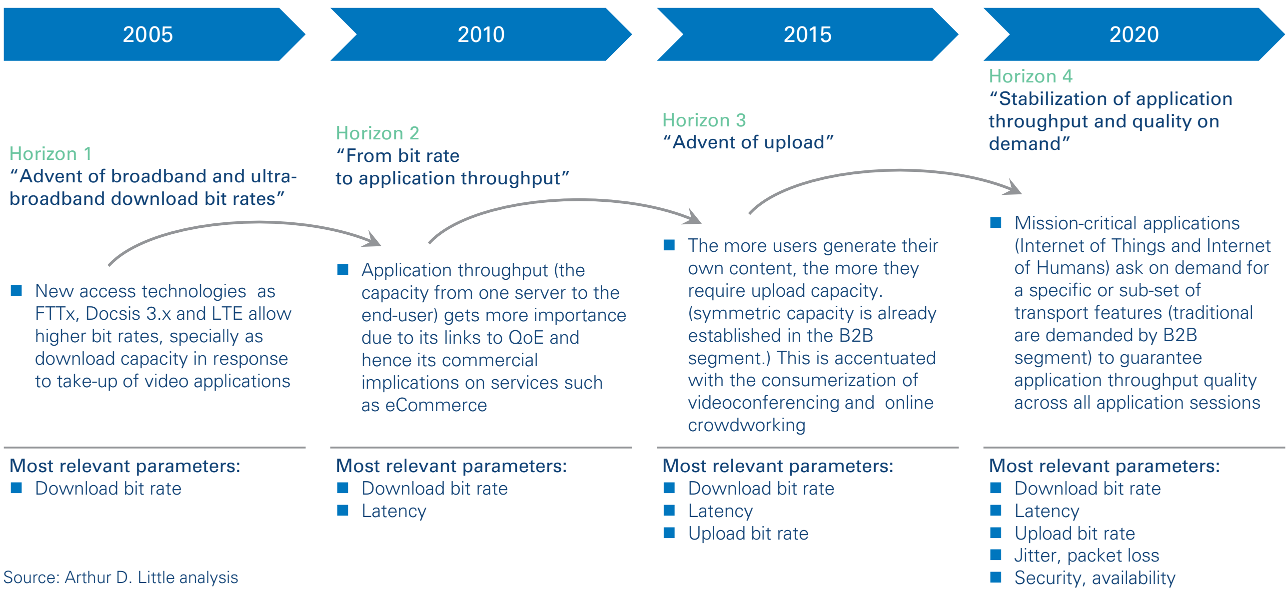
Sectors such as the Financial Services industry, the Electronic Payment sector, high-security Governmental Bodies (including police, military, emergency services, etc.), the Broadcast sector, etc. have all made ample requests for "advanced" services.

Indeed, key data transmissions in the B2B segment rely on dedicated networks that require customized design and significant efforts to be developed. However, given the high price tag associated with such dedicated networks, it seems unfitting to promote the global mass adoption of next-generation applications over an open-connectivity platform.

Among all new requirements, security and data protection deserve special attention as they play a critical role in the safe use of next-generation applications, especially in scenarios foreseeing the Internet of Things.

In such scenarios, the number of connected devices is expected to increase exponentially (Ericsson and Cisco anticipated 50 billion devices connected to the Internet by

Figure 49: Internet services traditional and emerging transport features



Source: Arthur D. Little analysis

2020). Highly heterogeneous systems (e.g. houses and office access systems, appliances and vehicles) become increasingly connected and extremely vulnerable to cybercrime, turning security into the main priority.

The level of concern rises even more if we consider that only a small portion of the population is currently aware of the risks incurred in extreme connected scenarios, in which privacy and security cannot be guaranteed.

Particular attention must be paid to M2M applications. These applications imply that many devices connect and communicate with no or low human intervention of control. A survey conducted by Beecham Research highlighted that data loss and corruption, access intrusion and distributed denials of service (DDoS) are growing.

Imagine a connected car with remote-start features being turned off by an unauthorized individual or a burglar accessing home-sensor information to determine the best moment to break in.

Cisco argues that future security techniques need to focus not only on protecting end-points and network boundaries but the fact that the network itself will become a central point of attention in future security programs. Other analysts even speculate that with so many connected devices, it will be impossible to secure the end-points; therefore, security will need to be implemented on the sole element that these devices have in common: the network.

According to a study from Arbor Networks ("Worldwide Infrastructure Security Report, Volume IX"), DDoS attacks are one of the top concerns for today's organizations, and together with large-scale malware, can severely compromise an ISP's core equipment, resources and business-critical IP services. Over recent years, Arbor Networks conducted global surveys among service providers to determine their experiences with security threats. According to the data received, the size of the largest DDoS attack gradually increased between 2002 and 2012, and in 2013, participants reported attacks ranging from 100 Gbps to an alarming 309 Gbps.

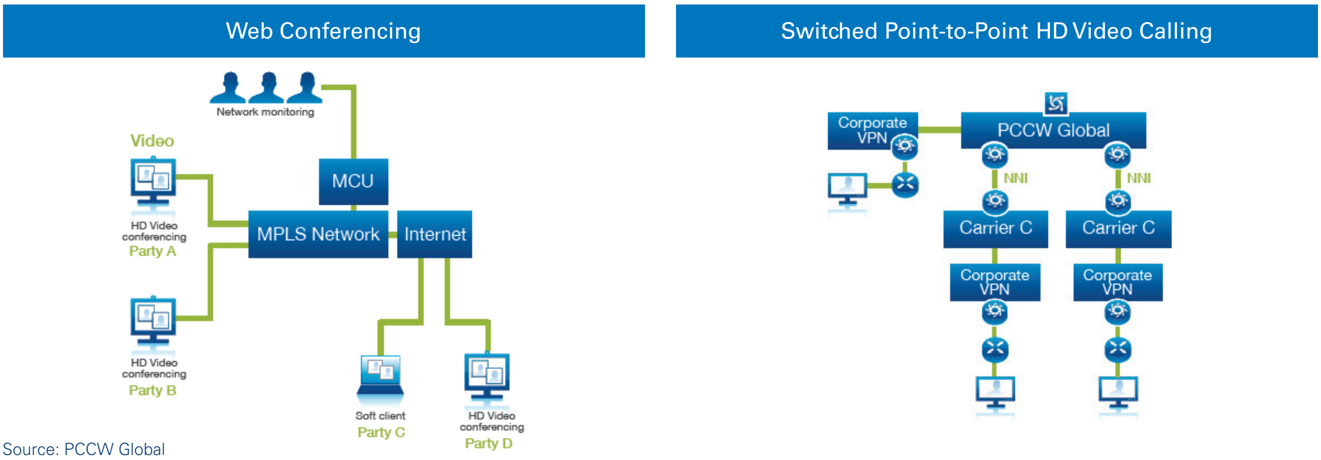
One of the possible solutions, suggested by Cisco in its *Annual Security Report* (2014), is to rely on the networks to continuously monitor and analyze data and thereby identify malicious behavior as soon as it appears.

New solutions and technologies on the horizon

Several technologies and solutions are currently being investigated to improve the quality control of IP Interconnectivity and better support new requirements in IP networks. The most notorious innovative solutions and technologies are worth mentioning here:

- **Transparent Caching: content delivery techniques extended to all web content.**
  While Content Delivery Networks provide web content acceleration exclusively to those Content and Application Providers that have requested a managed service and subscribed to a commercial agreement, Transparent Caching extends web content acceleration to all generic content that is "frequently requested" by end-users. Such popular content is transparently and automatically "captured" by caches installed within the ISP networks, and then retransmitted from a location in the network that is closer to end-users. According to initial estimates, Transparent Caching could provide web content acceleration to at least 50-60% of the total Internet traffic. Next to the benefits to end-users, Terminating ISPs would save costs by avoiding redundant retransmissions of IP traffic.
- **MPLS-like circuits governed by software-defined paradigms: differentiating IP flows dynamically inside the network.**
  Multi Packet Label Switching (MPLS) is a technology capable of "coloring" IP traffic flows within a network with certain "labels", with the aim of routing them, for example, according to their priority levels. MPLS technology has long been used in core networks in order to manage indiscriminate traffic flows or Virtual Private Networks (VPNs), in particular for business customers. Still, up to now, it remained confined to the B2B perimeter and the core network. A broader adoption of MPLS technology, i.e. to the very edge of the network and extended to consumer networks, would enable a complete new set of guaranteed delivery services. This opportunity is expected to materialize in the context of a significant evolution of networking technologies that foresees the introduction of programmable and remotely controllable network elements. Such innovative concepts may come with Software Defined Network (SDN) and/or Network Function Virtualization (NFV) technologies.
- **IPX paradigms beyond mobile voice: managing IP quality while passing from network to network.**
  IP Packet Exchange (IPX) is a technology that aims to inter-connect IP-based networks by implementing a standard cascading mechanism for passing and enforcing Quality of Service (QoS) at the IP Interconnection interface. The IPX concept has been around for years, and many believe that only now that mobile operators, deploying their LTE networks, are moving to full IP transition, IPX will at last really take off. Historically, IPX has been focused on voice applications, and up to now, it has not really addressed the general problem of IP Interconnection. Still, it is gradually moving to a multi-service platform and looking to simulate the way voice circuits worked on legacy networks for everything that will eventually pass over an IP network.

Figure 50: HD Video switched Point-to-Point concept by PCCW Global



Source: PCCW Global

Specifically, new features covered will enhance quality (i.e. jitter, latency and packet-loss, depending on what quality a specific network-to-network interface can handle and security.

- **Switched point-to-point interconnection: a dedicated solution for HD videoconferencing.**
  With the growth of the video-conferencing market and reflecting the online migration of human interactions, the industry asks itself how to accelerate the path towards enhanced and mass-market solutions. Many have realized that HD video-conferencing needs to step away from the closed "enterprise-as-an-island" environments and address the increasing requests for greater interoperability and interconnectivity. Acknowledging this trend, PCCW Global proposes to introduce standardized IP Interconnection interfaces that would enable placing of video-calls via different networks through a mechanism similar to the one used in the legacy phone system. Each user would have an identifier to call to, and the standard network-to-network interfaces would take the responsibility to forward the call and maintain the quality during the entire session. Similarly to the traditional phone service, the calling party is delivered through a managed interconnection to the receiving party whenever the calling party wants to start the video call. According to PCCW Global, this service would differ from many current videoconferencing solutions where end-points are all bridged to a central server that interconnects them (see attached figure). The key challenges of such solution reside in
  - the critical mass of interconnections to be reached among global carriers (i.e. such managed interconnections do not exist today and the application take-up depends on the global coverage), and
  - a current absence, though being investigated, of a solution for identifying the two end-points of the video-call (e.g. numbering, domain service, etc.).

Numerous industry bodies are working on frameworks and standards for the adoption of HD video-calling. For instance, the i3 Forum has established a working group looking at high-definition video communications issues within the IPX domain. Another example is the GSMA: the global association of GSM providers set HD video-calling on its agenda and launched, for this purpose, the Open Visual Communications Consortium (OVCC) to help shape an industry environment that is ready to deliver these services.

3.4. New IP Interconnection business models are being developed

The evolution of applications, competitive repositioning over the value chain and progress of routing technologies so far described, are profoundly re-shaping the IP Interconnection ecosystem and preparing the ground for the future Internet platform.

Business incentives, technologies and potential solutions materialize into something tangible which can be described as sustainable business models. These clarify the underlying value propositions and mechanisms that have been adopted to share the value created by the interaction of different players in the ecosystems.

Based on interviews and comparative analysis performed by Arthur D. Little, a morphological box can describe how IP Interconnection business models may develop. This analysis has no intention of being exhaustive, but gives a framework that helps to understand how new IP Interconnection business models can address new classes of IP Interconnection requirements.

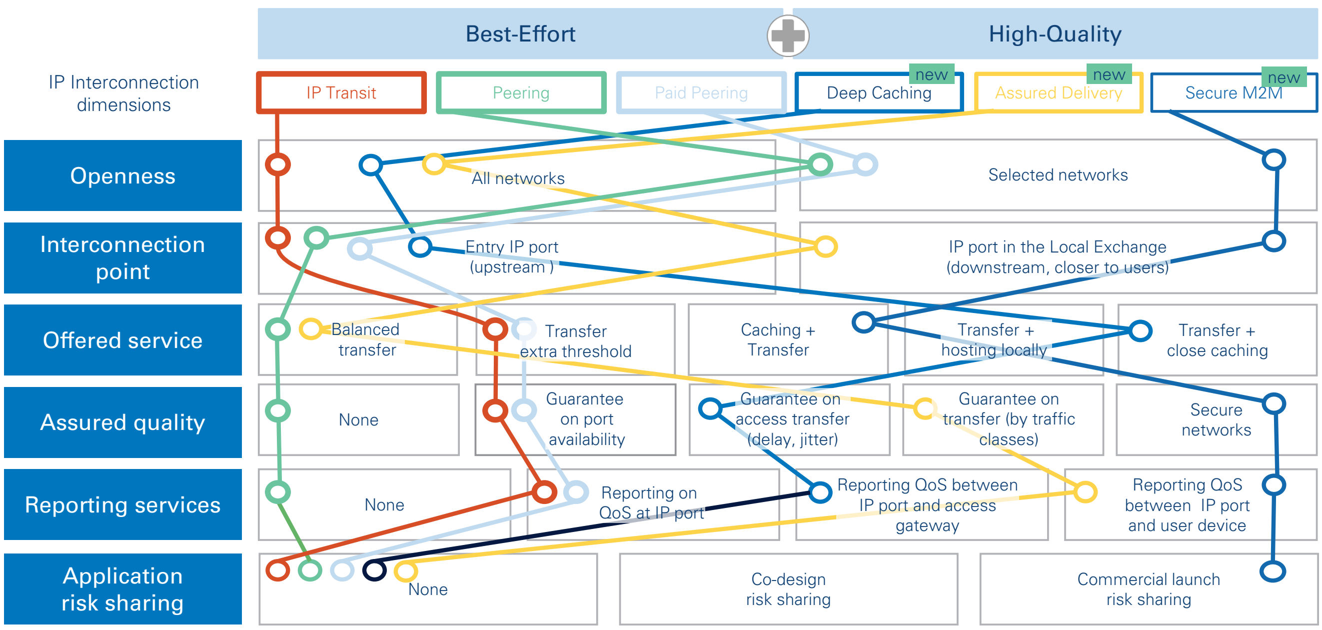
Our analysis also highlights that traditional IP Interconnection business models, i.e. Peering and Transit, will be, to a greater extent, complemented by new business models that are likely to constitute the basis of the future Internet platform. This process is naturally driven by the market forces in presence.

For the sake of illustration, IP Interconnection arrangements have been characterized according to six – not necessarily exhaustive – dimensions:

1. **Openness:** The degree of availability of such arrangements to access seekers (e.g. accessible to all or only selected networks);
2. **Interconnection point:** Location of the IP Interconnection interface (e.g. entry-point vs. deeper points into the network);
3. **Services offered:** Portfolio of services offered by the specific arrangement (e.g. balanced traffic transfer, allowance for extra-thresholds traffic, traffic transfer plus caching, transfer plus hosting);
4. **Assured quality:** The nature, location and level of Quality of Service offered (e.g. none, guaranteed on port availability, guaranteed on traffic transfer, secure networks);
5. **Reporting services:** Specific reporting services as essential complements to Service-Level Agreements (e.g. reporting QoS at IP port, reporting QoS between IP port and access gateway);
6. **Application risks' sharing:** E.g. none, co-design risk sharing, commercial launch risk sharing.

From the combination of the alternative options available for each dimension, more than 100 possible outcomes exist.

Figure 51: IP Interconnection business models



Source: Arthur D. Little analysis

Only three interconnection solutions still constitute the majority of interconnection arrangements today.

This evidence suggests that a greater number of interconnection business models can be imagined and are already likely to be a subject of discussion among the IP Interconnection actors.

Beyond the novelty of the already-known Paid Peering model, which allows direct interconnection between a Content and Application Provider and an access network with an unbalanced traffic ratio, at least three new IP Interconnection business models have been identified that have the potential to lay the foundation for a complementary assured end-to-end Quality-of-Service Internet platform for mass market use:

1. **Deep Caching** is an improvement on CDNs that provides better delivery quality by further reducing the physical distance between content and end-users. Managed or transparent caches, installed deep into Terminating ISPs' networks, are offered to interconnection seekers. They provide reduced latency, and consequently enable further improvement to throughput and packet loss. Strictly speaking, it is not an innovative delivery method, but rather an efficient and effective work-around for services such as on-demand premium video.
2. **Assured Delivery** is a more advanced networking option that foresees the dynamic establishment of dedicated routing for specific applications by applying the innovative concepts of software-defined networks (SDNs), network functions virtualization (NFV) and massive use of MPLS technology. Ad hoc network capacity is released dynamically and upon request. The interconnection seeker requests, as a permanent rule or only on demand, the establishment of managed capacity across two points (the interconnection interface and terminating interface, being either an end-user or another network), with a defined service-level agreement thereby attached (e.g. capacity vs latency vs. jitter).
3. **Secure M2M:** This networking option is similar to the assured delivery, but security features are added at network edges, and some parameters (e.g. latency or availability) are stressed for mission-critical applications (e.g. micro-electronic payments).

The new business models are already being discussed, or even commercially offered

Dedicated arrangements through dedicated services or even networks, have always existed for the business segment, although SDN/NFV technologies are expected to push networking possibilities to the next level in the mid-term. At the opposite, many mass-market applications cannot access these dedicated services, as they are not available over the public Internet. Still, as the demand for stricter networking requirements is increasing, new services are expected.

Figure 52: Google and Netflix in-house Content Delivery solutions

Google – Global Caches

Proposed approaches:
- **In the core:** global network to connect Google's datacenters and transport traffic to Internet aggregation ports
- **At the edge:** peer directly with access network operators and ISPs
- **Beyond the edge:** development of Google Global Cache platform in access providers' networks to deliver content even closer to users and reduce transport costs for operators

Deep cache platform offered in access networks to store contents even closer to users

Netflix – Open Connect

Proposed approach:
- **At the edge:** connect directly Open Connect to access networks with free peering at common Internet Exchanges
- **Beyond the edge:** save even more transit costs by putting Open Connect in or near access networks*
- **Open source:** availability to share hardware design and open source software components of the servers for co-design and improvements

Deep cache platform offered in access networks to store contents even closer to users

amazon webservices – Direct Connect

Proposed approach:
- **Dedicated network connection**
- **Better experience:** Reduced network costs, increase throughput and a more consistent network experience than Internet-based connections (no latency variations)
- **Compatibility:** The same connection to access public resources while maintaining network separation between the public and private environments
- **Expandability:** virtual interfaces reconfigured at any time to meet changing needs

Deep cache platform offered in access networks to store contents even closer to users

Source: Google, Netflix, AWS, Arthur D. Little analysis; * Netflix indicates that more than 11 access networks worldwide accepted

Evidence of real demand for new services is provided by largest CAPs' ambition to "move beyond the edge" – i.e. beyond the IP Interconnection interfaces – and get closer to end-users.

As an example, Netflix and Google both launched their own Content Delivery Networks with an offer to ISPs to install their proprietary solutions "within" terminating ISP's networks (not just interconnecting at the traditional interconnection interface).

This illustrates that the Deep-Caching business models are presumably settled for strong take-up in the short term while advanced networking option such as Assured or Secured Delivery may need to wait for the full deployment of SDN/NFV technologies, although Amazon Web Services Direct Connect already represents a good example of latent demand to have "a more consistent network experience than Internet-based connections", as stated on Amazon Web Services' web site.

Key messages

- Innovation in IP Interconnection can support further development of the Internet and accelerate the take-up of next-generation applications that require uncompromised quality.
- Today's Internet is a Best-Effort and finite (yet-not-scarce) resource
- The Best-Effort nature of IP technology does not necessarily imply low performance; average and peak connection speeds have increased by 12-14% since 2007, with acceleration to 19-21% since 2011.
- The Internet is subject to latency and packet loss and Advanced Internet platforms – i.e. beyond Best-Effort – may be required for next-generation applications that could bring an Internet of Things and an Internet of Humans to life.
- IP Interconnection Quality of Service needs to be extended to new parameters (e.g. latency, jitter, packet loss, security, data protection).
- Variants of Paid Peering, Deep Caching, Assured Delivery and Secure M2M are among the innovative IP Interconnection business models, which could lay the foundation for an advanced Internet platform based on assured end-to-end Quality-of-Service Internet Platform – complementary to Best Effort.

4. New Business Models Could Accelerate Innovation and Value Creation

As stated in Chapter 3, the Internet application landscape is hard to predict; still, some future application landscape scenarios could be supported or accelerated by new IP Interconnection business models.

The potential value creation generated by next-generation applications could be substantial, but will require creating the conditions – i.e. the Internet technological platform – to enable those next-generation applications to take up.

4.1. The most advanced application landscapes could generate substantial economic value creation

Forecasting the value associated with the next-generation application landscapes is a dedicated exercise based on numerous assumptions and educated estimates. Although the forecasts of the value at stake greatly vary among analysts in the order of 10 – most analyses converge towards ranges of economic contribution in the range of trillions of dollars.

Gartner, a research company in the ICT industry, indicated that although in 2015 the combined IT and telecom market will amount to nearly $4 trillion, the incremental revenue generated by Internet-of-Things suppliers (i.e. hardware, embedded software, connectivity services, information services) is estimated to reach close to $300 billion per year by 2020. Gartner predicts that the Internet of Things will create greater economic value for all organizations, and the total economic value contribution from the Internet of Things will be around $1.9 trillion in 2020 – of which manufacturing (15%), healthcare (15%) and insurance (11%) are to be the largest beneficiaries.

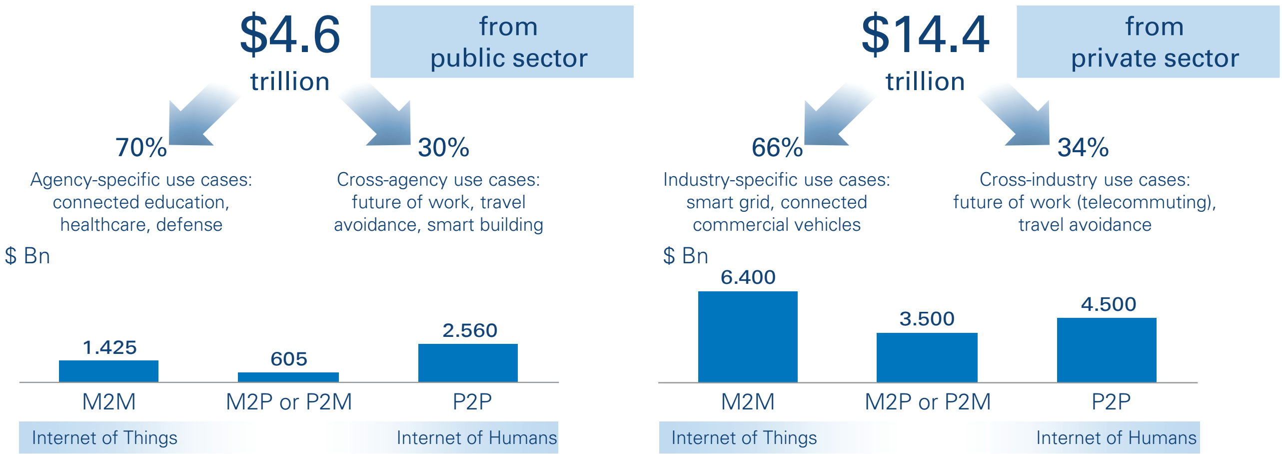
Among the several studies addressing the topic of economic value creation in the hyper-connected world, Cisco puts forward more detailed research, revealing significant value creation estimates by 2022 for what it terms as the "Internet of Everything". According to Cisco, the advent of the Internet of Everything would be associated with a value at stake worth of around $19 trillion when combining the value impact from both the public and the private sectors:

- In the public sector, the Internet of Everything could generate around $4.6 trillion in value at stake over the next decade. In the public sector, Cisco defines the Value at Stake as the potential value that can be created by public sector organizations in terms of lower costs and societal benefits from greater efficiency. Benefits from programs such as connected transportation, smart roads, social care, and education accrue as reductions in overall costs, especially through better targeting and control of resource usage. Other programs have indirect benefits for government — economic, social, or environmental — but direct benefits for citizens and businesses in terms of reduced transactional costs and time saved, or in terms of external benefits such as better quality of life. The value creation would be mainly driven by:
  - improving labor effectiveness for new and existing services;
  - improving effectiveness in police forces through greater situational awareness and connected command centers, vehicles, and supplies;
  - improving labor efficiency and capital-expense utilization, leading to reduced operational costs;
  - shortening "search" times, improving the environment and producing better health outcomes;
  - improving the ability to match supply with demand, while also enhancing monitoring and compliance.

  Major application examples include smart parking, water management, gas monitoring, chronic disease management, road pricing, telework, connected learning and connected militarized defense

- On the other side, the Internet of Everything is forecasted to create around $14.4 trillion in value at stake (higher revenues and lower costs) in the private sector by:
  - reducing selling, general, and administrative (SG&A) expenses and cost of goods sold (CoGS) through improvement of business process execution and capital efficiency;
  - creating labor efficiencies that result in fewer or more productive man-hours;
  - eliminating waste and improving process efficiencies;
  - increasing customer lifetime value and growing market share by adding more customers;

Figure 53: Value at stake associated by Cisco to Internet of Everything scenario by 2022



Source: CISCO, Arthur D. Little analysis

– increasing the return on R&D investments, reducing time to market, and creating additional revenue streams from new business models and opportunities.

Major application examples include smart factories, connected marketing and advertising, smart grid, connected gaming and entertainment, smart buildings, connected commercial ground vehicles, connected healthcare and patient monitoring and connected private college education.

From the more detailed analysis provided by Cisco, the following values at stake can be linked to the emergence of future application landscapes:

- Internet of Humans (i.e. Person-to-Person communication for collaboration purposes): around $7 trillion;
- Internet of Things (Machine-to-Machine communication for automation purposes): around $8 trillion;
- The remaining $4 trillion relating to hybrid Machine-to-Person or Person-to-Machine communication typically linked to analytics purposes – e.g. connected healthcare.

4.2. The acceleration of the most advanced application landscapes requires guaranteeing a link with the most suitable Internet platform option

Cisco's impressive estimates on the value at stake associated with the Internet of Everything nonetheless imply the fulfillment of at least three prerequisites:

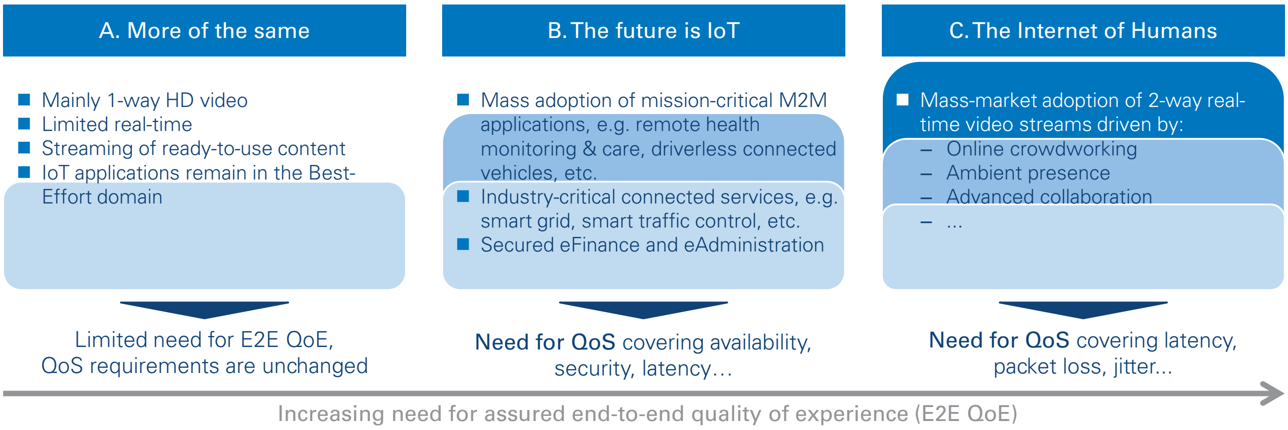
1. Plans for policymakers, governments and companies to target given application landscapes;
2. Commitment to embrace the key hyper-connected paradigm implied by the application landscape, and subsequent commitment of public administrations and companies to align their processes and operating models;
3. Contribution to the development of a robust and secure Internet technology platform capable of enabling and providing the required connectivity services.

Therefore, although it is a key ingredient in the advent of various future application landscape scenarios, the future Internet Platform is derived from (as opposed to driving) the ambition for and commitment to materializing a specific application landscape.

The clarification of the preferred application landscape is therefore essential to indicate which specific functionalities need to be developed within the enabling Internet technology platform (not the contrary).

The more the ambition moves towards the most advanced application scenarios, the more we see a need to assure network quality control.

Figure 54: Internet Application landscape scenarios



Source: Arthur D. Little analysis

As an example, in the Internet of Humans scenario, the underlying platform should be capable of seamlessly delivering a significant amount of conversational video services i.e. two-way video-streaming services for collaboration purposes among people. Such services are proven to be extremely sensitive to upload bandwidth and more sophisticated network parameters such as latency, jitter and packet loss.

These quality requirements are difficult to control in the current Best-Effort Internet, and only specialized IP-VPN services, widely applied in the business-to-business segment, can guarantee fulfillment of those more stringent IP Interconnection requirements.

In a very advanced application landscape scenario, conversational services could even require establishing, on demand, logical circuits that cannot currently be provided over open IP networks. Indeed, the current Best-Effort IP networks are limited by their native packet-based nature and the absence of standards guaranteeing the quality of delivery at IP Interconnection nodes.

Conversely, IP Interconnection requirements are definitely less restrictive in application landscape scenarios, such as "More of the Same", in which the dominant applications related mainly to one-way video streaming or download of generic content (e.g. audio/video streaming, video-on-demand download) and/or basic interactive services (e.g. web search and peer-to-peer data transfer.)

University research studies show that, next to the end-user's willingness to pay for the service, the take-up of advanced ICT applications is subject to achievement of a service quality level beyond what psychologists call the "cognitive absorption

Figure 55: Quality of Service versus Quality of Experience

	Quality of service (QoS)	Quality of experience (QoE)
Definition	**Guarantee of technical parameters** intended to measure network performances (e.g. availability, minimum bandwidth, average bandwidth, peak bandwidth, latency, etc.)	**Holistic experience** as perceived by end-users when using the combination of applications, devices, software and network (i.e. through the whole Internet stack)
Perspective	**Inside-out,** from the network to the end-user (technical conditions driven)	**Outside-in,** from the end-user to the network (commercial value drives)
Relevance	**Contractual** as minimum threshold applicable to a set of technical parameters (selected in function of end-users' demand) and determinant of the final price	**Drives the overall user satisfaction** and, ultimately, the take-up of specific or innovative applications and end-users' willingness to pay
Limitations	**Partial view of the overall quality** provided to end-users because it is focused just on a subset of parameters and/or lacking the end-to-end control capability of the service	**Difficult to monitor** as it requires the capability to control quality end-to-end, i.e. from a content or application server to an end-user's device, or from an end-user's device to another end-user's device

Source: Arthur D. Little analysis

threshold". This threshold refers to the quality level necessary to make the ICT platform "transparent", or seamless to the end users – i.e. no perceptible interruption or degradation of service.

From Quality of Service to Quality of Experience

Hence, the strict observance of quality requirements becomes crucial to assuring a high take-up of next-generation applications.

In the recent years, the demand for Quality-of-Service at the network level has consistently and progressively expanded its focus to cover additional parameters that would better capture the quality perceived by end users (i.e. Quality of Experience). BEREC defines Quality of Experience as the relationship between the performance expected from a specific service and the subjective perception obtained after the use of the service, which largely depends on the Quality-of-Service parameters.

The most advanced application landscapes (Internet of Things and Internet of Humans) are expected to require the effective and efficient management of a larger set of QoS parameters (including latency, jitter and/or packet loss) for the benefit of an improved QoE. Furthermore, the quality control should apply "End-to-End" irrespectively of the originating and terminating networks.

Thus, the ambitioned future application landscape is crucial in defining the requirements of the future Internet Platform and industry actors, regulatory authorities and policy-makers should clarify their vision for the future Internet and application landscape before taking positions on Internet governance.

As mentioned earlier, End-to-End Quality of Service and End-to-End Quality of Experience have been driving the specifications of connectivity services in the business-to-business segments for a while already. Assuming that the paradigms of Internet openness and application agnosticism are preserved, the extension of such capabilities to the "open" Internet platform (as opposed to the specialized IP-Virtual Private Networks) could be a lever for democratization and mass adoption of the most advanced next-generation applications.

Key Messages

- The Internet-of-Things and the Internet-of-Humans application landscapes can unlock an economic value potential in the range of trillions of euros by 2020.
- The future Internet Platform is derived from (as opposed to driving) the ambition of and commitment to materializing a specific application landscape.
- The strict observance of quality requirements becomes crucial to assuring a high take-up of next-generation applications.
- Most advanced application landscapes (Internet of Things and Internet of Humans) are expected to require the effective and efficient management of a larger set of QoS parameters (including latency, jitter and/or packet loss) for the benefit of an improved QoE.
- Quality control should apply "End-to-End" irrespectively of the originating and terminating networks.
- End-to-End Quality of Service and End-to-End Quality of Experience have been driving the specifications of connectivity services in the business-to-business segments for a while already.
- Assuming that the paradigms of Internet openness and application agnosticism are preserved, the extension of such capabilities to the "open" Internet platform (as opposed to the specialized IP-Virtual Private Networks) could be a lever for democratization and mass adoption of the most advanced next-generation applications.

5. Three core assumptions can drive the evolution of the future Internet Platform

5.1. The future Internet Platform will, as always, organically grow from stakeholders' various interests

The partially complementary and partially conflicting interests of key stakeholders allowed the Internet to grow

The evolution of the Internet platform entails the most diverse interests and a high number of stakeholders; it is therefore essential to understand the stakeholders' map in order to ensure a constructive contribution to the ongoing debate on the evolution of the Internet platform.

In our taxonomy, the stakeholders' map includes six groups of stakeholders: end users, ISPs, CAPs, International Carriers, CDN providers, Regulators & Institutions.

The different or even divergent interests that can apply within each group – e.g. large vs small ISPs or large vs. small CAPs – reflect intra-sector competitive issues. Therefore, ISPs and CAPs are usually split into two groups covering respectively the larger actors and the smaller ones to better mirror the diverging interests they have. Interests obviously also differ and/or diverge between stakeholders' groups and illustrate the competitive issues over the IP Interconnection value chain.

Stakeholders' interests are centered on end-users who ultimately decide on application take-up and service subscriptions via their willingness to pay set prices for services.

The divergent forces can contribute to the positive creation of new services, economic value and consumers' surplus. However, this requires that actors step down from polarized positions, and that their business relationships are not influenced by unfitting (regulatory) rules.

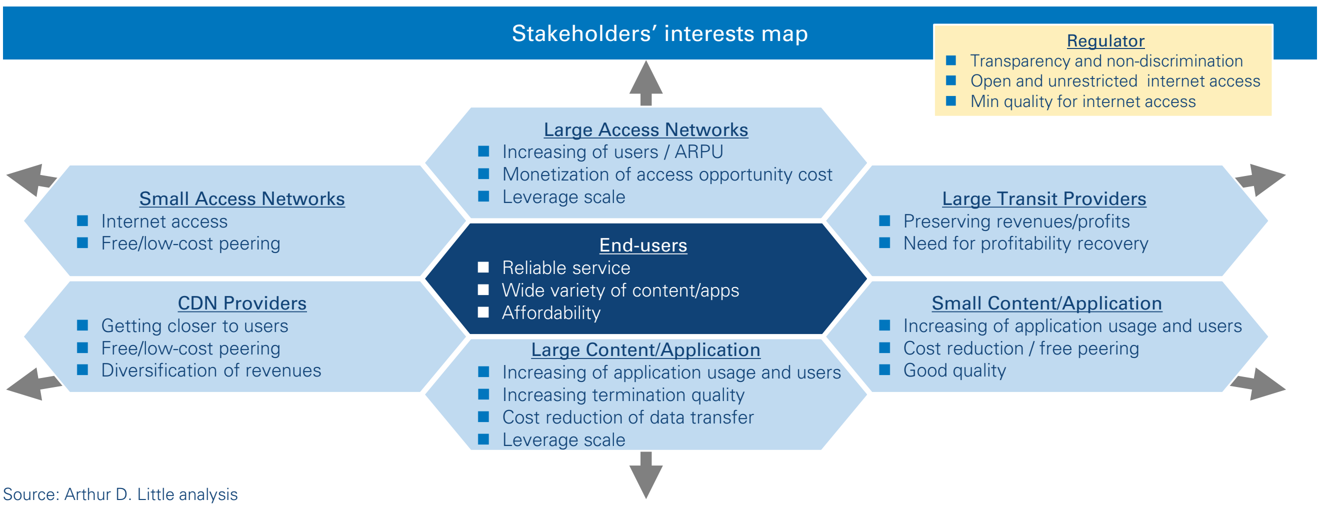
Finding the optimal equilibrium requires balancing the universality, affordability and access openness of the Internet with safeguarding the industry's operational efficiency and supporting the innovative process of freely inventing and developing new value propositions based on new IP Interconnection services and/or business relationships.

While the different positions fuel the debate, the Internet ecosystem continues to evolve through the creation of new business relationships, as witnessed by some recent public moves:

a) Terminating ISPs partner with CDN providers

- Orange and Akamai entered into a global content delivery alliance in late 2012;
- Telefonica and Akamai entered into a global content delivery alliance in March 2014.

Figure 56: Internet stakeholders' interests map



Source: Arthur D. Little analysis

b) Large CAPs partner with Terminating ISPs

- Google and Orange signed a special transit deal in France in February 2013;
- Comcast signed an IP Interconnection deal with Netflix early in February 2014;
- Orange agreed to host and serve Netflix in its network in preparation for Netflix's service launch in France in late 2014;
- Verizon signed an IP Interconnection deal with Netflix in late April 2014.

5.2. Three core assumptions have the potential to shape the future Internet platform

The debate regarding the Internet evolution is complex to illustrate concisely, but our analysis identified, among the recurrent themes, the few core assumptions that have the potential to shape the future Internet Platform – i.e. invalidated they lead to different scenarios.

We found at least three key assumptions with such potential:

1. **Assumption #1** - Best-Effort Internet is insufficient to support next-generation applications;
2. **Assumption #2** - Best-Effort Internet can co-exist with Guaranteed-Quality Internet Services;
3. **Assumption #3** - New IP Interconnection business models are needed to accelerate innovation.

Assumption #1 - Best-Effort Internet is insufficient to support next-generation applications

As illustrated at the end of Chapter 3, the Best-Effort nature of Internet and IP Interconnection interfaces has fundamental implications:

On one side, although the Best-Effort Internet is providing a satisfactory answer to today's Internet needs, it raises questions as to its ability to enable wide adoption of next-generation applications that have more stringent requirements; such reduced latency, limited latency variation (limited jitter) and reduced packet loss:

- Advanced application landscapes will increase the focus on End-to-End Quality of Experience and cause greater attention to be paid to a larger set of Quality-of-Service parameters in the open Internet Platform;
- Adding substantial and relatively cheap IP Interconnection capacity will statistically reduce the risk on QoS on those new parameters, but will the achieved levels be sufficient to enable next-generation applications?
- The rapid emergence of technologies such as caching, deep caching, adaptive streaming and video compression tends to illustrate that adding unmanaged capacity at IP Interconnection nodes is not, alone, sufficient to guarantee E2E QoE.

On the other side, the Best-Effort nature of the Internet with its easy IP Interconnection setup, scalability and application agnosticism, promoted the fast growth of the today's Internet.

A consistent proportion of IP Interconnection stakeholders are reluctant to consider such innovative practices, as they fear the current configuration of Internet standard offers insufficient reassurance in the way of non-discrimination at the network layer for still quite vague up-side potential.

Figure 57: Core assumptions having the potential to shape the future Internet platform

Assumptions	
Assumption 1	→ **Best Effort Internet quality is insufficient for next-generation applications** (e.g. the availability of new interconnectivity features such as latency, jitter, packet loss, availability for next-generation applications such as Internet of Humans – ambient presence, online crowd working, advanced collaboration, etc. – and Internet of Things)
Assumption 2	→ **Best-Effort and High-Quality Internet can coexist, without cannibalization** (e.g. the availability of complementary application-agnostic interconnection services and the assurance to not degrade the quality of the Internet as we know it)
Assumption 3	→ **New IP Interconnection business models are needed to accelerate innovation** (e.g. the availability of new interconnection services such as Deep Caching, Assured Delivery, Secure M2M to complement the currently available IP Transit, Peering or Paid Peering)

Source: Arthur D. Little analysis

Those opinions are well summarized by BEREC:

- Best-effort Internet results, in most cases, in a high-quality of experience for users, even for delay-sensitive applications such as VoIP or Video streaming, as demonstrated by popularity of applications such Skype, Viber, YouTube, BBC iPlayer, iTunes Video, Netflix, etc. Therefore, while not providing a guaranteed delivery of data, Best-Effort Internet does not imply low performances (i.e. low speed, high latency or high jitter).
- A guaranteed end-to-end QoE is unrealistic over the Internet; i.e. neither commercially nor technically realistic:
  - QoS is an end-to-end concept that is not natively supported in IP networks. A new standard language would need to be developed across interconnected networks;
  - Traffic exchanged at interconnection points is highly aggregated and thus statistically stable, i.e. there is no significant traffic peak load variation over time; consequently, there is no need to implement traffic management mechanisms at interconnection points;
  - It takes time for any network to reach the critical size (in number of connected destinations); innovative practices would similarly not create substantial value before long;
  - Negotiating one-to-one QoS agreements with each involved party would be cumbersome (e.g. mobile roaming interconnection agreements);
  - Quality of Experience is also affected by end-users' devices, operating systems, the intrinsic quality of the content or application, and is therefore unlikely to reach the end-user (who pays for it);
- Alternative mechanisms are available for improving end-to-end performance:
  - E.g. end-point-based congestion control for reduction of the traffic load;
  - Internet Exchange Points and Peering increase the number of alternative routes, and thereby reduce congestion;
  - CDNs are available to improve the user's perception of an application's quality (QoE).
- Whether implementing end-to-end QoS across IP networks is economically a viable and future-proof strategy is largely affected by the costs of adding IP Interconnection bandwidth. The significantly decreasing trend of cost in the core and backhaul IP networks would enable IP capacity abundance strategies.

Stakeholders who believe that Best Effort "is not sufficient" tend to emphasize the opportunities associated with the most advanced application landscapes scenarios and the need to address new practices, while the other party tends to follow a reactive and pragmatic approach of "So far so good; why change it?"

Assumption #2 – Best-Effort Internet can coexist with Guaranteed Quality Services

We have seen that managed services offering a guaranteed Quality-of-Service and Quality-of-Experience have been applied for a long time in the business-to-business segments. They typically rely on the allocation of dedicated capacity and use of traffic management techniques to comply with the required service features. Quality-Guaranteed Services are operated within dedicated IP networks, working (and possibly interworking with the open Internet) through an admission-control gateway, often optimized for specific applications.

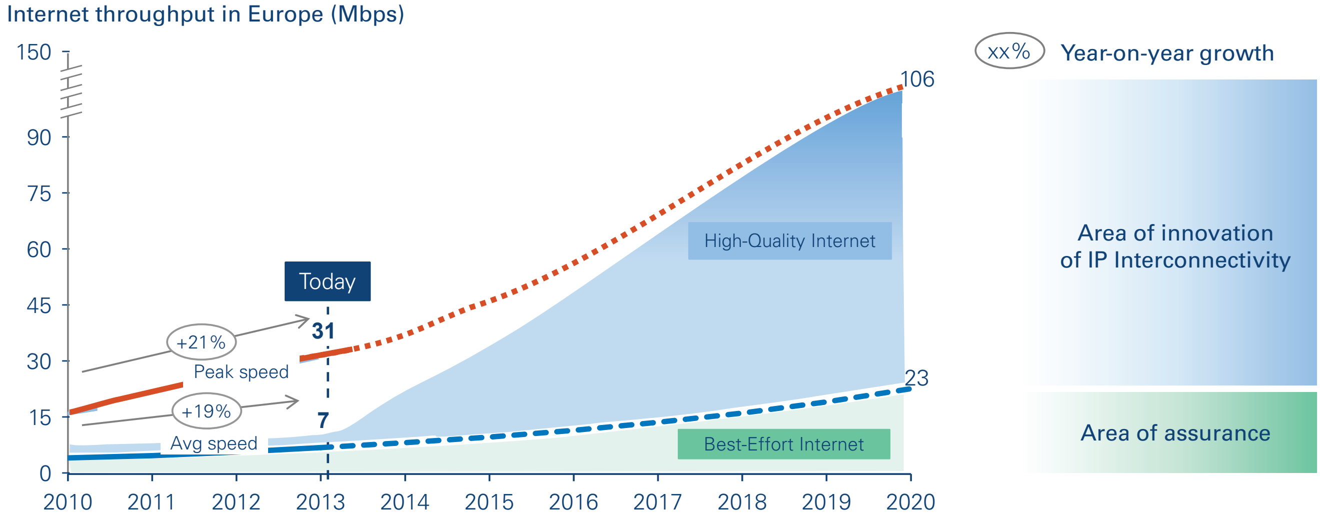
There is a concern that a mass extension of Guaranteed Quality Services from the IP Interconnection business-to-business segment to the residential Internet access networks in the form of "specialized services" may come at the expense of the "Best-Effort" open Internet access-services. Traffic prioritization would enable more advanced applications and attract investments, thereby reducing the incentive for IP Interconnection actors to invest further in improving the performance of the Best-Effort Internet. Therefore the sustainable co-existence of both services (i.e. Best-Effort Internet and Guaranteed Quality Services) is questioned.

Guaranteed Quality services have been traditionally offered to the business segment as IP Managed Services, accounting for around 20% of the global IP traffic, and to consumer segments through vertically integrated IPTV platforms (accounting for an even larger share of access network traffic).

Akamai's historical data on the State of The Internet demonstrate that both the average and the peak bandwidth available per user has improved over the years, concluding that Guaranteed Quality Services did not prevent the improvement of Best-Effort Internet services' performance. These indications lead us to presume that Guaranteed Quality Services can be further built additively and non-discriminatorily on top of the Best-Effort Internet.

In principle, the concern that ISPs could be incentivized to focus on managed IP services, with the consequence of slowing down investments in Best-Effort IP Interconnection capacity, is understandable. Therefore, in a scenario of mutual growth for Best-Effort Internet and Guaranteed Quality Services, it would appear legitimate to introduce a framework and tools to

Figure 58: Evolution scenario: Best-Effort Internet Access Services and Quality-Guaranteed Services



Source: Arthur D. Little elaboration on Akamai 'State of Internet report Q3 2013'

monitor and steer the continuous development of Best-Effort performance.

In such a scenario, public Internet Access Services will remain an area of assurance, and should keep growing in terms of both peak bandwidth and average bandwidth, while Guaranteed Quality Services may come as innovative and investment areas to fully exploit the capacity available in the last mile (subject to the available access network technology).

Should Guaranteed Quality Services be implemented additively and openly on top of Best-Effort services, the Internet Platform would result in a richer service portfolio.

Assumption #3: New IP Interconnection business models are needed to accelerate innovation

So far the Internet has consistently allowed for innovation to progress at a steady pace and new applications have emerged regularly (see Chapter 1).

However, earlier sections have extensively illustrated that the nature of Internet traffic is changing and the nature of the underlying quality features is evolving as well (i.e. from download and upload speed to QoE). This is illustrated by the demand for innovative interconnection arrangements nowadays, and more is expected in the coming years.

Figure 59: Debate space around the three identified core assumptions

Debate space	AGREE	DISAGREE
Assumption 1 ➔ Best-Effort Internet quality is **insufficient** for next-generation applications	■ Best Effort is a finite resource ■ Future applications require more than unmanaged capacity ■ Best Effort cannot guarantee new features on demand	■ Best Effort is net neutral and scales efficiently ■ IP capacity is abundant ■ If just adding bandwidth is cheap, Best Effort is the choice
Assumption 2 ➔ High-Quality and Best-Effort Internet **can coexist,** without cannibalization	■ High-Quality Internet can build up additively ■ Best Effort can enjoy the take-up of new technical practices ■ Both can grow together	■ Best Effort is not low quality ■ Prioritization introduces an incentive to degrade Best-Effort ■ There are no tools to measure quality and to safeguard
Assumption 3 ➔ New IP Interconnection business models are needed to **accelerate innovation**	■ New business models enrich the space of opportunities for Content & Application Providers ■ New business models are already emerging	■ Foreclosure and market predation would be incentivized ■ Innovation is already happening at steady pace

Source: Arthur D. Little analysis

Already in the short term, (ultra) high-definition conversational services will require assured end-to-end Quality of Experience. In the medium-term application landscape, secure network features will be requested to support the take-up of advanced Machine-to-Machine or Internet-of-Things applications such as eHealth & eCare, Home Security, vehicle telematics, etc.

Chapter 3 has shown that a number of new business models are being investigated by IP Interconnection actors. Traditional IP Transit and Peering practices are now complemented, though still to a limited extent, by Paid Peering practices. Innovation in the IP Interconnection ecosystem is also materializing at a steady pace. ISPs are investing in innovative delivery solutions in their core networks. CAPs are applying new streaming techniques and proposing to move beyond the "network edge", while Content Delivery platforms are being complemented by new security services.

5.3. Three options can be foreseen: Best-Effort 2.0, Quality-Guaranteed Services and Both Worlds

Sound arguments can be provided in favor of, or against each of the three core assumptions and sketching out the available options to IP Interconnection stakeholders, with regard to the future IP Interconnection Platform will help develop a more holistic and balanced view.

Next to inventing a new revolutionary technology or technique that would improve IP connectivity beyond any issues currently identified, the main drivers to improve IP Interconnections are:

- Investing in capacity upgrades at IP Interconnection nodes, i.e. providing an abundance of IP Interconnection resources;
- Extending the capabilities for managing E2E QoS to the open Internet.

Accordingly, three options can be identified to drive the development of the Internet Platform in the immediate future by combining those two drivers:

1. "Best-Effort 2.0" relying on unmanaged capacity upgrades in IP Interconnection;
2. "Quality-Guaranteed Internet Services" opting for dedicated resources and traffic-management techniques over the public platform to improve end-users' Quality of Experience of Over-the-top content and applications;
3. "Both Worlds", investing simultaneously in both drivers.

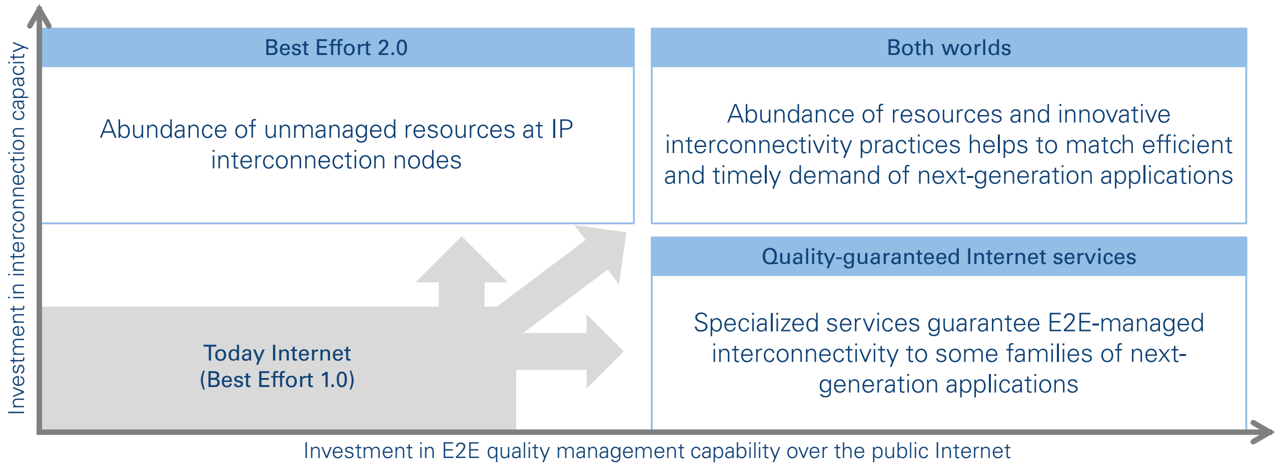
This analytical framework intends to serve as a tool to nurture the ongoing debate and contribute to a constructive collaboration among stakeholders. Hereafter, the potential implications in terms of strategic moves, new IP Interconnection business models and innovation are highlighted for each option.

Option 1: Best-Effort 2.0

The "Best-Effort 2.0" option implies that Quality-of-Service and Quality-of-Experience issues are solved by adding extra capacity at IP Interconnection interfaces. Therefore, the "Best-Effort 2.0" option could result in the emergence of a dual Internet composed of an open Best-Effort Internet while E2E proprietary networks develop around the largest CAPs.

The largest CAPs, for which QoE becomes essential in the short term, are likely to continue to invest in their own IP Interconnection infrastructures, possibly up to the local-access network (e.g. proprietary deep-caching). In practice, they would be deploying their own almost end-to-end IP networks focused on delivering their own content and applications to Terminating ISPs.

Figure 60: Possible options for the future Internet Platform



Source: Arthur D. Little analysis

IP Transit providers and CDN providers would come under strong pressure as they would mainly serve as a back-up route for the largest CAPs, and therefore risk encountering difficulties in maintaining/reaching critical mass and economies of scale.

On the other side, Terminating ISPs would reluctantly invest in Quality-of-Service platforms. Indeed, the ruling Best-Effort principle at IP interfaces would annihilate the business case for investing in new solutions that would require extra financial coverage by end-users in exchange for a higher level of Quality of Experience. Still, resources at IP Interconnection interfaces would keep increasing in line with the evolution of access networks.

Smaller CAPs would benefit from the falling IP Transit and Peering costs to design and implement viable delivery strategies. However, the number of involved actors and the Best-Effort principle would limit them as far as the End-to-End Quality-of-Service level they could access, and therefore cap their innovation potential in the context of the Internet of Things and Internet of Humans.

With Best-Effort 2.0, no major new IP Interconnection business model would be expected, as no fundamental assumption has changed compared to the current situation. Consequently, the emergence of next-generation applications is likely to be delayed as QoE is not widely and affordably guaranteed end-to-end.

Still, next-generation applications could emerge when directly developed or supported by the largest Content and Application Providers, which thanks to their proprietary infrastructure, would be able to offer the highest Quality of Service until the access network.

Option 2: Quality-Guaranteed Internet Services

The "Quality-Guaranteed Internet Services" option assumes that the increasing need for extended End-to-End Quality-of-Service is addressed by enabling new business models that rely on some form of active differentiating traffic management. This option could focus on newer Business-to-Business or Business-to-Business-to-Consumer services (e.g. Home Delivery services), but mainly would serve as a complement to Best-Effort interconnections.

With such an option, new IP Interconnection business models would emerge around selected families of applications, most likely focusing on premium video, security and reliability features required for Internet-of-Things applications (accounting for another estimated 10-20% of global IP traffic).

Without substantially increasing the IP Interconnection capacity, not-yet-existing conversational applications that would generate substantial traffic and require high levels of E2E QoE are unlikely to emerge in the public Internet, especially for the consumer segment, and the largest CAPs will continue to invest in their own proprietary infrastructures.

Alternative business models built around selected applications risk lacking critical scale, and prices of high QoE could remain expensive (as often occurs for IP-VPN services). Besides, the pressure will be strong on IP Transit and CDN providers, which are likely to develop niche positions.

Over the medium term and thanks to their own infrastructures, Quality-Guaranteed Internet Services dedicated to new families of applications will become affordable for the largest CAPs.

The emergence of new applications is expected to be mainly related to premium video and IoT applications. Their higher cost is likely to limit a rapid mass-adoption, and thereby a significant macro-economic impact. The most advanced applications, which cannot access Assured E2E QoS services over the public Internet, will only emerge if developed and supported by the largest CAPs (which can offer the highest level of QoS until the access network).

Option 3: Both Worlds

The "Both Worlds" option has the potential to tap into the demand of next-generation services to the greatest extent, e.g. the combination of the Internet of Things and Internet of Humans, by leveraging on the volume driver.

First, many new (IP Interconnection) business models would be available affordably and non-discriminatorily to different Internet players, and this would compensate the necessity to deploy own E2E infrastructures aimed to by-pass the limitations of the current Internet platform. Consequently, various IP Interconnection business models would emerge non-discriminatorily according to the needs of all existing and new players in the Internet eco-system (not solely the largest CAPs). Moreover, given the multitude of interconnection possibilities and related IP Interconnection business models, the emergence of competition would solve possible inefficiencies.

On the one hand, the largest CAPs could focus on innovation and investments in development of Content and Applications. On the other hand, Terminating ISPs could focus their investment resources in the most efficient way, i.e. adding more resources, possibly as a function of the overall QoE (no more restricted to the delivered download and upload speed) that they want to provide to their end-users, with an impact on their retail revenues per end-user.

This option would result in a substantial acceleration of time-to-market for new generations of applications, and allow tapping into the potential for accelerated macro-economic growth.

The take-up of a new generation of applications is also expected to contribute to increased overall Internet usage and Internet penetration in all social and age classes.

Figure 61: Key characteristics of Internet Platform options

	Best Effort 2.0	Quality-guaranteed Internet services	Both worlds
Improved speeds (i.e. more capacity)	◕	◔	◕
Improved QoS (i.e. extended to latency, jitter, packet loss, etc.)	◔	◑	◕
Acceleration of innovation (i.e. enabling new application landscapes)	◔	◑	◕
Unlocked value creation potential	◔	◔	◕
Certainty and timeliness of value creation impact	●	◑	◑

Source: Arthur D. Little analysis

Key Messages

- The future Internet Platform will, as always, organically grow from stakeholders' various interests
- Three core assumptions have the potential to shape the future Internet platform:
  1. Best-Effort Internet quality is insufficient for next-generation applications;
  2. Best-Effort can coexist with Guaranteed Quality Services;
  3. New IP Interconnection business models are needed to accelerate innovation.
- "Best-effort" Internet is and will no doubt continue to be essential in the future, and there is early evidence to indicate that it can co-exist with complementary end-to-end Quality-of-Service platforms if properly monitored.
- Private investment in IP Interconnection has led to structurally improved conditions for the future development of the public Internet. Content comes closer to end-users (by direct interconnection and local content caching), Internet performance is improved by adoption of new application technologies (e.g. "adaptive streaming"), IP network resources are abundant (e.g. higher capacity in the "last mile").
- The public Internet will stand to benefit mostly from private investments in IP interconnection architecture aimed at shortening the distance that Internet traffic needs to travel.

Glossary

Adaptive streaming	A technology that allows adaption the video definition and the application bit rate to the user's connection conditions
Aggregation	A point in networks where multiple traffic flows coming from multiple upstream sources are combined for subsequent downstream delivery
Application accelerator	A technical solution that eases access to a given Internet application (or family of) and boosts its performances
Application landscape	Used here as the whole set of applications that can run smoothly over the Internet
AS, Autonomous System	A network identified by a set of IP addresses (IP-routing prefixes) under the control of one or more organizations. Most Autonomous Systems are physically composed of proprietary and leased infrastructure
Availability	Used here as the degree to which a certain connection or resources are available for use, it is expressed in percentage (the ratio of the total time a connection or resource is capable of being used during a given interval to the length of the interval)
Bandwidth	Used here as synonymous of capacity with a given connection because the transmission capacity of each medium can be expressed in terms of spectrum available for that particular medium
BEREC	The Body of European Regulators for Electronic Communications
Bit rate	Used here as a measure of the data transmission speed as the number of bits that can be sent per each second
Caching	Used here as the storage of content in places located closer to the Internet users who consume it
CAGR	Compound Annual Growth Rate
CAP	Content and Application Providers: players providing content and applications through the Internet
Capacity	Used here as a measure of a given link or connection to transmit data
CDN	Content Delivery Network/ Players providing caching services
Concurrency	Used here as a rate of use of a given resource (connection or application) by a multiplicity of requesters
CPU	Central Processing Unit; it refers to the component in electronic devices (e.g. computers) in charge of elaborating computational instructions
Deep-Caching	A CDN variant in which content caches are installed within networks (not at their borders or edges) and closer to final users. This would allow a shorter path between the content server and the final user
DNS	Domain Name Server; it is the system of associated structured names (e.g. www.wikipedia.com) to IP numbers (which would be difficult to remember)
DOCSIS	Data Over Cable Service Interface Specificatio;, it is a technology that allows transmission of broadband data on hybrid fiber-coaxial cables (those used for TV signal transmission). The latest version of this technology is DOCSIS 3.1
E2E	End to End, e.g. from an application server to the final user
EB	Exabyte
Eyeballs	In media, users who consume content
FTP	File Transfer Protocol, a legacy application used to transfer data over the Internet
FTTx	Fiber to the X, the 'X' can have many alternatives such as Home (FTTH), Building (FTTB), Cabinet (FTTC), Street (FTTS), etc.

GB, Gbps	Gigabyte, Gigabit per second
GE	Gigabit Ethernet, a networking solution mainly in-building
Hop	An end-to-end connection; may be composed by several intermediate segments which are also referred as 'hops'
Host	A computer/server attached to the Internet
Hosting	The service to run an application on a third-party's servers
IAP	Internet Access Provider, often synonymous with ISP
ICT	Information and Communication Technology
IEX	Internet Exchange, synonymous with IX
In/out traffic	Traffic data exchanged through the Internet can have two directions in principle: the traffic arriving to the host is said 'in' while the traffic leaving the host is said 'out'. This also applies to any network interface
Interconnection	The interworking of two distinct networks
Interface	Here used as the point of interconnection of two networks or subsequent network segments
Internet Delivery	The service to transfer data from the application server to the final user
Internet Demand	Used here as traffic generated by users needing to access Internet applications
Internet Platform	The whole set of hardware, application and protocols that enable the Internet delivery service
IoE	Internet of Everything, as defined by Cisco
IoH	Internet of Humans
IoT	Internet of Things
IP	Internet Protocol
IP Transit	An IP Interconnection regime
IP VPN	IP Virtual Private Network, a logical channel that allow the data transmission with certain Service-Level Agreement (speed, availability, security, etc.)
IPTV	A technology that allows the provision of TV service over IP networks; it is often realized through the use of dedicated resources or networks
IPX	IP Exchange, a set of technologies and protocols – still under investigation – that would enable quality-capable interconnection interfaces
ISP	Internet Service Provider
IX	Internet Exchange
Jitter	Used here as the variance in time of the arrival of data packets
Last mile	Used generally in fixed networks, the last segment that connects homes and offices to the serving network
Latency	The time required for a data packet to travel from source to destination
M2M	Machine to Machine, an application through which two machines communicate
M2P	Machine to Person, an application through which a machine communicates with a person
MB, Mbps	Megabyte (refers to capacity), Megabit (refers to speed) per second
MPLS	Multi-Protocol Label Switching, a networking technology
Network Edge	From a CAP's point of view, the interconnection point to an access network
NFV	Network Function Virtualization, an advanced technology for implementing and operating networks
OTT	Over The Top, it generally refers to applications developed and run over IP networks
P2M	Person to Machine, an application through which a person communicates with a machine
Packet loss	The occurrence losing data packets during transmission to the detriment of Quality-of-Experience

PB	Petabyte
Peak traffic	The maximum volume of traffic per second achieved in a given period (day, month, etc.) in a given direction (downlink or uplink)
Peering	The IP Interconnection regime under which two parties directly interconnect and exchange traffic
Peer-to-Peer	A technology that allows two Internet hosts to exchange data by acting as both suppliers and consumers of content data
PP	Percentage Point
QoE	Quality of Experience (a qualitative term describing how the end-user experiences the Internet)
QoS	Quality of Service (a technical engineering term)
Round Trip Time, RTT	The time required by a packet to go from the source to the destination and return
Routing Technologies	The technology used for switching data packets
SDN	Software-Defined Network
Speed	Used here as synonymous with bit rate
Streaming (application)	An application that doesn't require the complete download of a video file for consuming
TB	Terabyte
TCP	Transfer Control Protocol, a protocol of the TCP/IP suite at the basis of the Internet. It can control the speed of data packet delivery and its resending in case of congestion over IP networks
Throughput	Used here as the successful data delivery over a communication channel, measured in bits per second.
Tier 1, Tier 2, Tier 3 ISPs	Used here as the status to describe an international carrier on the basis of its size and global coverage of IP addresses. Tier 1 operators are so big that they never need to buy transit from other operators to get full connectivity to the Web, but instead, may interconnect just through peering relationships
Traffic management	A technique to maximize the use of a communication channel in the presence of concurrent requests
Traffic Patterns	Used here as the way Internet consumption is manifested through the day, expressed as volumes or capacity consumed per hour
Transit Providers	International carriers selling IP Transit
Uncontended capacity	The amount of connection capacity that is reserved for a specific user (synonymous with minimum guaranteed capacity)
Web acceleration	A technical solution that eases access to Internet applications and boosts their performances
WebM /VP9	A technology for compressing video files
xDSL	Digital Subscriber Line, a technology that allows transmitting broadband data over copper cables (those used for phone service). The 'X' stands for the main variants, e.g. ADSL (up to 20 Mbps download) or VDSL (up to 70 Mbps download)

Notes

Notes

Notes

Notes

Notes

Authors

If you would like more information or to arrange an informal discussion on the issues raised here and how they affect your business, please contact:



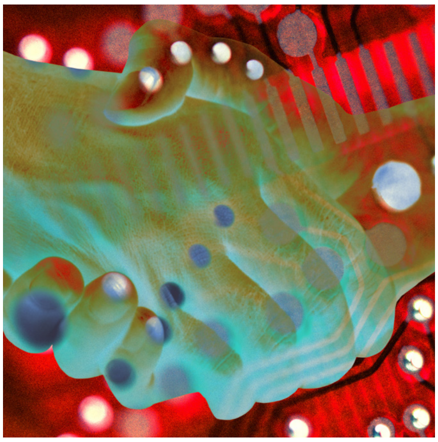
Gregory Pankert
Partner
Telecommunication, Information Technology, Media and Electronics practice
Arthur D. Little, Brussels
Mobile +32.473.99.72.27
pankert.gregory@adlittle.com

Andrea Faggiano
Principal
Head of Strategic Advisory Services for Competition and Regulation - TIME
Arthur D. Little, Rome
Mobile +39.335.784.4995
faggiano.andrea@adlittle.com

Karim Taga
Global Practice Head
Telecommunication, Information Technology, Media and Electronics practice
Arthur D. Little, Vienna
Mobile +43.664.230.7189
taga.karim@adlittle.com

Acknowledgements
We want to thank all industry managers and policy makers who constructively contributed to our study as well as Giulia Strusi and Francesco Sabatini, of the Arthur D. Little's staff team, who spent determination and passion for its finalization.

Liberty Global commissioned Arthur D. Little to author a study on the topic of Innovation and Investments in IP Interconnection. The objective of this work is to contribute to a debate currently high on the agenda of industry, policy and regulation with objective analysis of the key trends, empirical evidence and a holistic perspective. The study reflects Arthur D. Little's thoughts on the topic of IP Interconnection, supported by industry analyses as well as case studies and company examples based on publicly available information. In the process of writing the study, over 30 Global industry managers and policy makers were interviewed whose expert contribution is reflected in this work. The study provides a basis for discussion for key stakeholders across public and private sectors on a broad set of topics related to IP Interconnection developments and future strategic, policy and regulatory priorities.



Arthur D. Little

As the world's first consultancy, Arthur D. Little has been at the forefront of innovation for more than 125 years. We are acknowledged as a thought leader in linking strategy, technology and innovation. Our consultants consistently develop enduring next generation solutions to master our clients' business complexity and to deliver sustainable results suited to the economic reality of each of our clients.

Arthur D. Little has offices in the most important business cities around the world. We are proud to serve many of the Fortune 500 companies globally, in addition to other leading firms and public sector organizations.

For further information please visit **www.adl.com**

www.adl.com/FutureOfTheInternet