Current business models and services; scenarios for the future; high-level requirements - How can the future Internet look like?

Deliverable/Milestone: D2.1

Date: 31/05/2010

Version: 2.0

Editor:	María Ángeles Callejo, Carlos Ralli (TID)
Deliverable nature:	R
Dissemination level: (Confidentiality)	PU
Contractual Delivery Date:	30/04/2010
Actual Delivery Date	17/05/2010
Suggested Readers:	All public, specially OTTs
Total number of pages:	119
Keywords:	QoS, business models, guarantees, ecosystem

Version notes:	This version has been agreed by the consortium

Abstract

This deliverable reflects on the main trends in the current Internet in terms of recent evolution of new services and expected traffic growth and usage patterns. This deliverable identifies the main deadlocks or stumbling blocks looking at today's Internet business. These deadlocks or stumbling blocks are hindering or will delay the deployment of QoS based solutions in the networks and the offering of services with assured quality level. Furthermore, the deliverable analyses the current state of the art of network services, the related business models and technical solutions, to identify important limitations. The deliverable identifies and categorizes key factors and options that have significant impact on the future options for interconnect strategies and multi-NSP collaboration. Key options or models for innovative and dynamic multi-carrier contracts and markets are presented, as well as some reflection about the need for a bootstrapping case that is sufficiently attractive among NSPs to kick-off and lead the deployment of QoS solutions and services, such as those suggested by ETICS. On this basis service oriented scenarios are identified and described in which flexible end-to-end Assured Service Quality (ASQ) is required to match the users' QoS expectations in terms of ubiquity, quality and high capacity. The selected scenarios have been described considering the current trends and the possible evolutions of the Internet. An initial set of high level business and technical requirements is directly derived from the identified scenarios, to be further elaborated in deliverable D2.2, to guide the ETICS activities in WP3 and WP4.

EXECUTIVE SUMMARY

This deliverable analyses the current state of the art of network services, the related business models and technical solutions, to identify important limitations. On this basis service oriented scenarios are identified and described in which flexible end-to-end Assured Service Quality (ASQ) is needed to match the users' QoS expectations in terms of ubiquity, quality and high capacity. This analysis considers the recent evolution of the networking market, as well as trends that could be considered as sources of future opportunities for the QoS deployment in the networks. These opportunities could mainly come from:

- The increasing demands for new services such as video and real time communications mainly due to the social dimension that is characterizing the recent Internet usage. Progresses on terminal equipments such as HD/3D TVs and the more and more "always-on/connected" users will not only increase the network usage but also push for consuming contents with better quality.
- The important traffic explosion in mobile networks that has led the mobile operators to specify fair usage policies, like the traffic volume caps usually specified in the mobile customers contracts.
- The evolution of connectivity and other valued services in the mobile networks that require additional capabilities to ensure quality because of the shared mobile access resources.
- The roaming needs for the different services.
- The evolution of - services provided to corporate and SME customers and users. This is an important point since the deliverable focuses on *future networks* and not only on Internet related aspects as we know Internet today.

To complement the analysis and to further provide drivers and directions for further work within ETICS, this deliverable identifies the main deadlocks or stumbling blocks looking into today's Internet business. These deadlocks or stumbling blocks that are delaying or will delay the deployment of QoS based solutions in the networks. A central point discussed here is **why service differentiation over the global Internet has not happend**? Over-provisioning has often been the simplest (and faster) solution to provide end-users with a certain level of quality; as a side effect, the best-effort Internet is currently perceived by end-users as "good enough" for most of the services. Even Over-The-Top (OTT) providers, like Google, Akamai, etc., have been relying on this "perception" to develop their new services in full overlay. However, the growing popularity of YouTube-like applications for the consumption of more and higher resolution contents is putting new strict constraints on the networks, and the limitations of the best-effort approach are rapidly emerging. Therefore, a new question is arising about **the long term sustainability of the over-provisioning approach, mostly because of the less clear return on investments for operators in the current network ecosystem.** This question is particularly relevant for mobile networks, where the resources are scarce and must be shared among the end users in the same cell, and over-provisioning is intrinsically limited by the radio spectrum and its regulation. In this mobile framework, policies for assuring the fair usage of network resources are just implemented; therefore, the introduction of new policies to assure the quality of service

to specific inelastic traffic should not be revolutionary. Similar scenarios for a "natural" QoS deployment may also be sketched for fixed networks, particularly in the access side.

In pure technical terms, several building blocks are already available or are emerging for the different network sections and the wired/wireless technologies, like BGP IP/MPLS VPNs, PCE, Congestion control, etc.. These tools are not sufficiently integrated in multi-carrier environments and cannot support dynamic interconnection agreements (yet). A further limitation is their lack of integration between them and the business and operation layers/systems adopted by the operators for the network services deployment. The ETICS project is working towards a solution for all these issues, with a particular focus on the inter-carrier dynamics.

To analyse the current situation, this deliverable describes the business models adopted in the current networks, considering some important services like:

- VoIP and IPTV services (including VoD & Gaming and more personalized services), which still suffer from quality issues to really replace the traditional voice and TV services;
- Internet access services, which have so far been a tremendous success partially because of the flat rate model, but now suffer the lack of service guarantees particularly during the busy hours;
- VPN services, mostly for corporate users, usually perfectly assured while deployed intra-domain, but still suffering from the lack of real QoS guarantees in multi-domain scenarios. The increase ratio of tele-workers and mobile devices (smartphones, netbooks and laptops) with professional services will further push for QoS improvements in the near future;
- Interconnection with IP Exchange (IPX), which is certainly more advanced with respect to QoS provisioning, but is still not fully deployed and/or accepted as considered as too costly.

The deliverable also discusses, identifies and categorizes key factors and options that have significant impact on the future options for interconnect strategies and multi-NSP collaboration. It provides a preliminary reflection on some future high-level Telco options and what may characterize the future competitive landscape. Key options or models for innovative and dynamic multi-carrier contracts and markets are presented, as well as some reflection about the need for a bootstrapping case that is sufficiently attractive among an initial set of NSPs to kick-off and lead the deployment of QoS solutions to enable end-to-end ASQ and attractive services such as those suggested by ETICS.

Finally, a set of future Internet services scenarios is described considering different perspectives:

- Users: Real time communications for end users (including unified communications on mobiles), Real time social networking with good quality interactive Video services, Remote Access presentation (for home automation and multimedia content sharing), Gaming as a Service (in particular with simple set top boxes at home and the intelligence and rendering located in the network) and virtual drives (either for personal usages or for backend services);
- Business: enhanced VPN offers with a simpler provisioning of L2 and L3 VPNs among interconnected operator networks, the integration of value-added services and application awareness, the creation of "Virtual Private Services" extending VPNs with other IT/communication services, Telematic services distributed among multiple ISPs (for disaster prevention and relief,

remote health services, etc.), advanced tele-presence services (to go beyond current communication services, allowing for real remote interaction in business environment and reduce expenses and the ecological impact of business travels);

- Wholesale: Inter-provider Multi-cast streaming services, creating real opportunities for citizens or companies to broadcast high quality video/voice services to defined groups of users, teaming up with the Telcos to access their customer database and having access to multi-cast tunnels as high-speed and low delay network services, carrier-driven CDN benefiting from a better collaboration between the content distribution and the actual network capabilities, Dynamic assured QoS connectivity services controllable with an API open to third parties.

These scenarios have been used to derive the initial set of requirements reported in this document. This set will be further elaborated in WP2 and will also guide the definition of Business models in WP3 and of the overall ETICS architecture in WP4.

TABLE OF CONTENTS

1. INTRODUCTION

One of the main ETICS objectives is to create (and develop the technical solutions to support) a new ecosystem of innovative QoS-enabled interconnection models between Network Service Providers (NSP). The ETICS ecosystem will allow for a fair distribution of revenue shares among all the actors of the service delivery value-chain. As part of this main goal, WP2 has to identify, describe and analyse the technical and business requirements to define the capabilities and boundaries of the ETICS solutions.

The analysis of the current state of the art including the business models and the technical solutions is an important starting point to identify the most important service oriented scenarios in which a flexible end-to-end QoS in network services is needed. Since all these scenarios should be built considering current business models, trends and limitations as well as possible evolutions, it is mandatory also to identify a general multi-actor reference model where the interaction between the different stakeholders are described. For instance, to better understand current limitations of business models and technical solutions it is important to be aware of the main deadlocks or stumbling blocks looking at today's Internet business. This knowledge and awareness is important when considering interconnect options and suggesting solutions and services for the future.

Once these service oriented scenarios are identified, an initial set of the high level business and technical requirements is provided, which will be further elaborated in deliverable D2.2.

In order to achieve these goals the document is structured as follows:

- Section 2 provides an overview of the market, technology and business trends. This section is focused on a short-medium timescale. It considers the recent evolution of the market as well as the trends that could be considered as sources of future opportunities. It also offers a discussion about the current deadlocks and stumbling blocks that could prevent the deployment of QoS based solutions. Finally, even though this deliverable does not aim to progress on the technical solutions to implement the proposed scenarios, section 2 provides a short overview of some technologies that represent good candidates to the deployment of QoS solutions in an inter-carrier environment.
- Section 3 gives an overview of the state of the art focusing on core Telco services, business models, value chains and technical solutions, including main limitations. This section is essential to understand the next steps to be followed in order to define new scenarios that also consider migration requirements.
- Section 4 identifies and categorizes key factors and options that have significant impact on the future options regarding interconnect strategies and multi-NSP collaboration. It provides a preliminary reflection on some future high-level Telco options (analysing the main characteristics that identify each one); then, it identifies the key elements that should be analysed in order to choose specific options and technical solutions (e.g. evolution of the traffic); it also provides models for multi-carriers contracts and markets, and it finally provides some reflection about the need for a bootstrapping case that could lead the deployment of ETICS based solutions.

- Section 5 describes the initial set of scenarios that have been identified by the different partners. These scenarios will be further elaborated in WP3 from the business point of view and in WP4, where architectural solution for their implementation will be studied.
- Section 6 infers the high level technical and business requirements that will be further elaborated in D2.2 and it identifies the next steps to be followed in WP3 and WP4 in order to address the stated requirements.

In addition to the core contents in the document, the following annexes are provided at the end of the document: Annex A provides a deeper analysis of the state of the art of some technical solutions that are briefly described in the main document. Annex B provides an overview of the Bandwidth on Demand services considering both technical and business aspects. Finally, Annex C presents the template that has been used by the consortium in order to identify and analyze the scenarios that are presented in this document.

2. MARKET, TECHNOLOGY AND BUSINESS TRENDS

This section first provides an overview of the current status of the market, considering the evolution of the traffic in the Internet and how the traffic growth is having a direct impact on the operators' margins. This can make the QoS (Quality of Service) management an important topic for an efficient management of the network resources and an opportunity for having new sources of revenue providing incentives for new investments of operators in their network. This short overview does aim to address all the scenarios that have been identified by ETICS as candidates to increase the deployment of QoS capabilities in the networks but it also introduces some well known problems in the current Internet that could motivate the deployment of QoS features.

Next, we discuss what could be the problems derived from just providing best effort services and why QoS provisioning is still not available in today's networks.

Finally, even though this document does not aim to discuss the technical implementation of the ETICS solutions, an overview of the concepts, technologies and the techniques enabling QoS is provided with a special focus on their limitations and possible utility.

2.1. OVERVIEW OF CURRENT INTERNET USAGE AND CHALLENGES

As the first step to build the Internet of the future, it is essential to analyze the expected evolution of the Internet users' behaviour. One of the main characteristics of present day network planning is the high uncertainty of the users' demand evolution. Nowadays, the main characteristic of end users is their diversity: there are multiple applications (Peer-to-Peer, streaming, voice over IP, blogs, social networks, chats, gaming, etc.) with heterogeneous requirements, that can be accessed from multiple devices (mobile devices, PCs, game consoles, tablet PCs, etc), and using different types of connectivity (wireless of different types, fixed by different media).

There have been multiple attempts to evaluate how the different Internet users behave, but, probably, a classification of the end users' behaviour according to their age could be the best approach to foresee what is expected in the future from the end users perspective. In this sense, a new generation called Generation Z or digital natives (born in 90s and 21st century) is used to the technical changes and has much more knowledge about the technology as it is reported in [GENZ06]. Effectively, due to the continuous technical evolution they have lived, they are much used to adapt to the new interfaces, services, etc.; and they have, in fact, a very important knowledge about the usage of the different technologies. This generation is highly connected to Internet and makes a lifelong use of the communication.

It is clear that in a short time span, different patterns of usage of the network have arisen, due to the ever richer offer of services and to the lower entry age to the Internet. Therefore, regarding Internet connectivity services, many new requirements must be considered in the evolution of the Future Internet and Future Networks: future networks must support new traffic demands, reliability to allow the end users to trust the availability of their connections, ubiquity of the access, security in the service usage, flexibility

to adapt to the different requirements, neutrality and openness to allow the development of new services, and ability to provide advanced services which combine all these characteristics.

In the coming years, the evolution of traffic will have huge impact on dimensioning in future networks. As stated by Cisco in [CIS09] and represented in Figure 1, it is clear that we are witnessing a huge increment of the demand of new multimedia applications that require better network performance or better Quality of Service (QoS), such as on-line gaming, video streaming or videoconference.

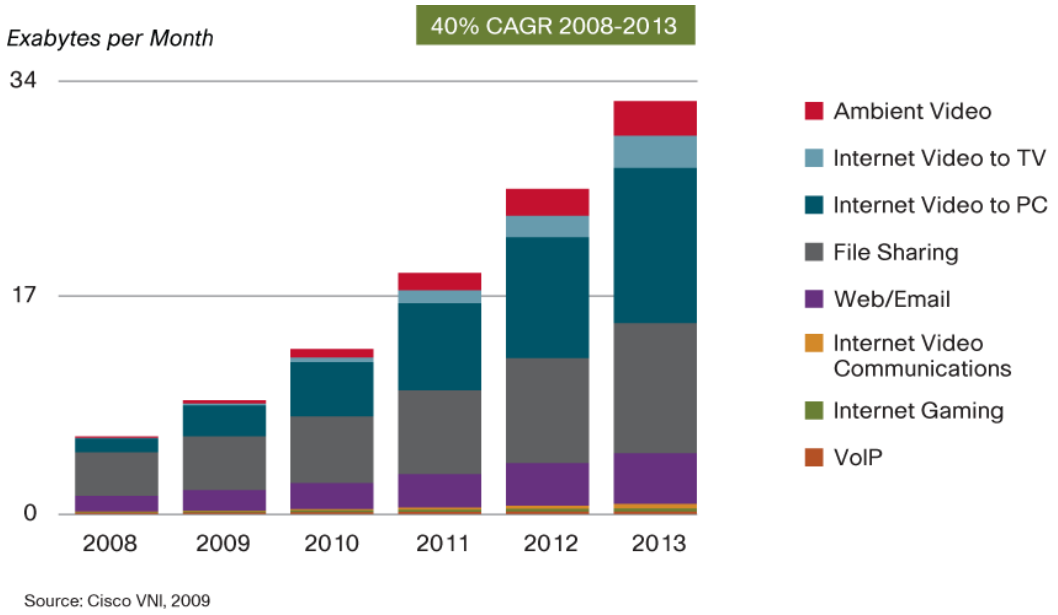


FIGURE 1: EVOLUTION OF RESIDENTIAL INTERNET TRAFFIC PER APPLICATION TYPE (SOURCE: CISCO [CIS09])

This will be particularly important with the above mentioned new generation of young people that is "always on" in Internet, is also able to create their own services, and values the connectivity as an important service, for which they could be keen to pay to have the possibility to access the wide set of Internet services. Furthermore, as stated in [CIS09], there is a clear trend to use several applications in parallel (e.g. listening to on-line music in Spotify while using a social network such as Facebook or on-line working), which is called hyperconnectivity in [CIS09]. In this scenario, the provisioning of advanced connectivity services will become a key driver for the Operators' business role in the Future Internet.

From the end users perspective, an important service is the video streaming - without considering P2P video streaming applications - which could represent the most important contributor to the traffic (the video streaming could represent the 25% of the Internet traffic in the 2013). Probably the most representative service provider is YouTube. According to [CIS09], *"YouTube traffic is both big and small: big enough to impress but not yet big enough to overwhelm service provider networks. It is nothing short of amazing that a site launched at the end of 2005 grew to take up 4 percent of all traffic by the beginning of 2007. By Cisco's estimates, YouTube accounted for 20 percent of online video traffic in North America in 2007, and online video-to-PC amounted to 19 percent of overall North American consumer Internet traffic"*.

An important aspect to consider is that well known sites such as YouTube is bring the light of social live to the video: people can comment the videos, the videos can be easily accessed and links from the social networks such as Facebook or Tuenti. According to [CIS09], *"given the varied aspects of video, it is difficult to say that "content is king." The throne appears to belong instead to the combination of communications*

and content. This combination has shown itself to be powerful enough to have enticed millions of Internet users to do something they previously showed little interest in doing: watching low-quality video on a small screen. YouTube offers more than unique content, it offers a platform for social interaction. YouTube viewers are not watching video despite the computer screen, but because of it the PC is ideal for interactivity, even if only the simple but effective ability to send a link. Traditional television may begin to seem less desirable than video that can be sent, shared, tagged, clipped, mashed up, and chatted about. Video as pure entertainment will always have its place, but even so it may turn out that to future generations, the home theatre silo and its isolated video experience will appear quaint."

Foreseen evolutions are also to more and more constrainable services, with a progressive move from massive video streaming over the Internet to more interactive videos (video chats with increasing quality between friends – in mobility or not - and between remote persons of a family like grand-parents and children), an increase of on-line game penetration, etc. In this context where the end users consume more and more traffic since they spend an important time of their social lives in Internet and where service providers are assuming in their networks traffic increases of 50% or, in the case of mobile networks, 100%, it seems that an important amount of the new incomes are going directly to service providers (such as Google or Akamai) while network operators have to assume important investments in their networks. Moreover, since social lives require some interactivity, some guarantees, the satisfaction of the end users will require the provisioning of guarantees to the end users or to the services they prefer.

But what are the investments and the problems the operators have to face and the challenges they need to face?

As a starting point, new access technologies such as FTTH and mobile broadband access connections will be deployed in order to allow end users to consume new video applications based on high definition, ultra high definition or 3D formats which are expected to produce a traffic explosion in the coming years. The deployment of new access technologies that require important investments in public sites that are not directly managed by the operators requires important investments that will require some years to be recovered. Moreover, there are some technologies such as the Mobile Broadband that require additional management to guarantee the quality of the connectivity to the end users. Probably the key element to manage the access attributes of the end users is the Access Node shown in Figure 2.

Due to the evolution in the access segment, two main scenarios should be considered:

- Due to the deployment of FTTH fixed technologies, the bandwidth generated/consumed by the end users will increase and this will have an important impact on the design and management of the metro/aggregation segments. A high increase of the upstream traffic will encourage customers to more and more actively participate to the Internet providing quality content, a "symmetrisation" of access lines, etc.
- On the other hand, due to its growing expectations and to the resource sharing among users in the access, Mobile Broadband must be studied carefully: on the one hand the operators cannot continue providing a good quality based on the over-provisioning of the network resources since the traffic is expected to continue growing at 100% (source [CIS09]) and on the other hand the effect of heavy users has resulted in the need of putting some policies to assure the fair usage of

resources, such as the limitation in the volume per month; this limitation of the volume is a policy that is included in the contract.

On the other hand the IP backbones need to support more and more traffic. Current IP backbones are often based on hierarchical architectures with high-speed point-to-point WDM links between IP routers. This architecture (Figure 2) has proven to be very useful in the provision of IP services, due to the tremendous flexibility provided by IP routers.

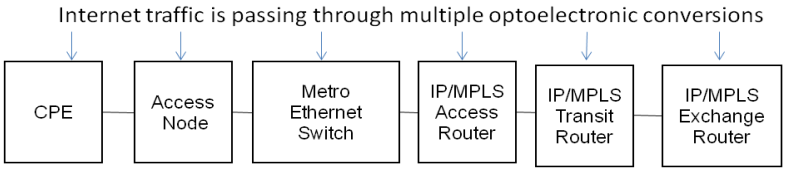


FIGURE 2: CURRENT NETWORK SCHEME

However, electronic packet switching cost is directly dependant on the transmission rate. For that reason, a traditional pure IP architecture presents economic scalability problems, mainly due to the increase of pass-through traffic (i.e. traffic whose destination is other node) which is processed at the IP layer. With the current architectures and cost models[1], network costs increase as traffic grows up, whereas with the common flat rate pricing models applied to final customers keep incomes constant. This scenario could eventually impact on the operators' margins, which would consequently lead to a lose scenario for end users that will experience an increase in their Internet access prices or a traffic rate limitation.

For these reasons, new architectural solutions will be needed in order to assure a low cost broadband Internet access with guaranteed by means on using technologies whose costs do not increase linearly with the traffic they have to carry, efficient mechanisms to provision network and the possible cooperation among the stakeholder to reduce the traffic footprint in their networks. In this context the following trends must be explored:

- The traffic engineering represents an excellent candidate to be used as the basis to build economic efficient network architecture. The final strategy (RSVP, LDP, multi-layer traffic engineering, multicast capabilities, congestion control, etc.) will depend on both technical (real capabilities of the solution, performance issues) and economical factors (the required investments and the cost of the operation of the solution). This efficient management of the network backbones and the consistency of the policies in the access networks would result in an efficient management of the whole network.
- The usage of these capabilities to easily deploy new services without major traffic impact and with an easy operation is required to really take advantage of the new solutions.

Finally, an important topic is the need of reliable networks able to satisfy the demands from the different corporations to carry out their own businesses. These demands are evolving from the basic connectivity offered to, e.g. financial companies to update the information about their customers from remote locations to the current needs to schedule tele-presence services that could reduce the travel costs.

[1] In this model the operators strongly depend on the evolution of the IP ports price, i.e. the evolution of 10G, 40G and 100G ports

It must be noticed that for the end users point of view, the quality of a service is evaluated independently of the cause of the potential problem. In fact, multiple elements can fail in an end-to-end scenario (the application server, a bottleneck in the core network, not enough bandwidth at the access or a bottleneck at the user's home), but the end users will not know about the specific source of that failure. Therefore, in order to really meet the end users' expectations, any QoS policy must be consistent in the end-to-end chain.

To sum up, according to the above text, different points should be considered:

- The first one is the evolution of the traffic. In different studies, it is stated that the traffic will continue growing at 50-60% in Internet. An important contributor to this traffic seems to be the video, so it will be important to evaluate which is the impact on the quality of the content required by the end users. Moreover, in the traffic estimations there is a growing in the traffic generated by real time applications which could result in an important challenge, especially in mobile networks where end users have to share the resources. For this traffic, if we add the dimension of the video for real time communications, the traffic demands could be important.
- We have to consider that the new speeds that could be available in access networks, due to the deployment of the FTTH, could open important new possibilities that could make the end users to consume contents with better quality (e.g. HD videos, tele-presence services, etc.). This could have a significant impact on the global traffic to be managed at the different network segments, where more cost efficient solutions should be found.
- Finally, even though it has not been considered in this section, there could be important traffic demands coming from the corporate customers. That means, the explosion of M2M, tele-presence services, or tele-medicine applications could lead to need to manage high volumes of traffic with real time characteristics[2].

2.2. DEADLOCKS AND STUMBLING BLOCKS

In the previous section, we have identified the main trends in the current Internet usage in order to infer how the future networks could look like. In particular, it has been stated that there is an important increase of all those applications that could have real time requirements (interactive applications, video communications, gaming, etc.) and/or important demands in terms of bandwidth (e.g. high definition video). It seems that there could be segments where some policies or efficient traffic management procedures should be considered. So, in order to design the ETICS solutions, it is important to understand why Assured Service Quality End-to-End (ASQ-E2E) over IP-based interconnected networks (such as Internet or private networks who address business market) has not happened yet,and which are the main challenges in the Internet scenario where multiple providers must collaborate.

In order to address this exercise, the discussion below aims to firstly discuss what could be the consequences of just offering "best-effort" services, then we analysed why the service differentiation over

[2] E.g. the tele-surgery services require real time characteristics and high quality of the video. Moreover, in order to avoid the latency caused by the compression of the images, some initial prototypes do not make any image processing.

IP-based interconnected networks did not happened and which are the main challenges in the end-to-end scenario.

Consequences of just offering simple "best-effort" services

Considering the Telco market in general we observe that today's market focus on today's services. Hence, there is no or little focus on inter-provider premium and Assured Service Quality (ASQ). The focus is on best effort content and over-the-top services, on providing Internet access at increasing speeds and a good user experience for most kinds of Internet browsing.

At the same time, we also observe some focus on intra-domain premium connections (e.g. for interconnecting corporate sites) and assured service quality within the single domain. The way to provision these services are typically based on ad-hoc solutions that are implemented locally. Furthermore, from the application/customer side or as seen from the host side, ad-hoc solutions have emerged in order to improve the user experience while just relying on a best effort network (e.g. Skype opens different connections or adapt the used codecs in order to adapt to the channel conditions).

We observe that the customers are getting used to this situation (they think that applications just work fine) and the typical attitude is to consider that the service level of the current Internet should work for all end users' applications, and that their Internet access flat fees should cover for all basic Internet services. For instance, real-time voice and videoconferencing services are working fine or well enough (based on obtained versus expected Quality of Experience (QoE) of the application service according to the application service price) over the current Internet (under low traffic conditions) for many customers (massively in the public market), and Telcos are missing customer incentives to request premium service levels.

Looking into these issues in detail from a Telco perspective, we should consider the following open questions:

- What are the demands for premium services?
- Or, what are the demands that are not met yet – or even known?
- And, what is actually the service level offered?

The answers to these questions are not obvious and are not known. What is known is: (i) that the best-effort Internet is often well dimensioned and it offers great opportunities to over-the-top players without any mechanisms in place for sharing revenues with the Telcos; and (ii) the "bill and keep" business model mainly prevails. Hence, in order to meet customer expectations, the Telcos are forced to invest in capacity without getting sufficient return on these investments to have a sustainable business.

Moreover, Telcos have limited means to respond to unexpected changes in demands. For instance, what can or should Telcos do if the demand for HD YouTube traffic increases? What could be the impact on both intra and inter-domain traffic?

In addition to the above, looking more at a intercontinental global network, the crossing of different time zones or low traffic countries (like in figure below from http://www.internettrafficreport.com/) may generate further geographic variability of the network performance that deal additional demand for premium services.

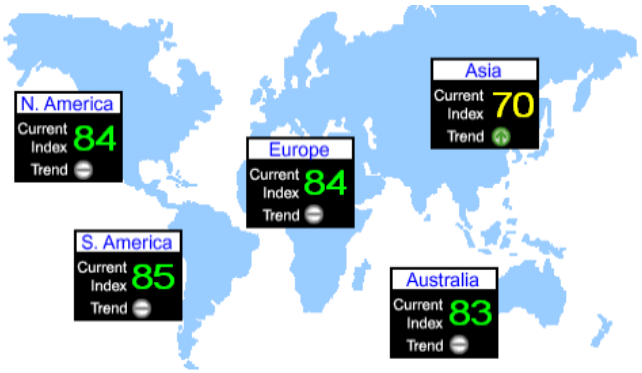


FIGURE 3: CURRENT NETWORK SCHEME (SOURCE: INTERNET TRAFFIC REPORT CET 15.50)

New reasons for premium services have also to be researched within business interconnection services for large corporation and dotcom needing to connect their business / R&D units spread all over the world. The traditional leased lines approach could be reverted in using premium services among more carriers.

Why have Services Differentiation over the Internet not happened?

One of the answers is that ASQ and differentiated services are perceived to be very costly and complex to maintain and operate. The main principle during the last years has been *"Keep it simple and cost efficient"*. Due to the complexity of their IT systems and support solutions, the initial steps to deploy and be ready for any new ASQ enabled service is very costly. Moreover, the cost-benefit analysis for premium services is by itself a challenging task, as the benefits and costs are very dependent on the operator context, and there are complex multi-actor dependencies that must be resolved.

Furthermore, a key question from the Telco perspective is – Is the willingness to pay for premium services sufficiently high? Will the customers experience a difference between premium and best effort services? Will the customers just churn over to the competitor? Is the demand sufficiently high? Many concludes by stating rhetorically – *"Show me the real need for QoS"* while summing up that,

- *We can just throw bandwidth at this problem – much less costly*
- *Content can be cached closer to the user in order to reduce latencies and minimise inter-domain traffic.*
- *The intra-office interconnect link is not the bottleneck*

As stated in section 2.1, from the economical perspective, there could be a need for QoS and there are some Telcos that have started to carefully evaluate the deployment of QoS features as a way to assure the fair usage of resources or as a way to deploy their own services in an IP convergent scenario. Typically, the driver for such service differentiation mechanisms is strongest in the mobile operator area.

Moreover, some Telcos are concerned that issues derived from the Net Neutrality debate can kick in and badly impact their reputation and public relations. Many people categorically think that any service differentiation is breaking the net neutrality principle. Hence, from a Telco point of view it is challenging to "sell" and explain the concept of and need for ASQ and premium services on the Internet. It should be noticed that the public relations of the Telco companies should highlight that the offer of premium services does not necessarily mean the filtering of the rest of application; that means, Telcos must disseminate that prioritization or assurance does not mean discrimination or filtering.

Another important part of the story, maybe partly as a result of the above, is that the Telco focus has been elsewhere, such as on operational and development cost squeeze, and on Telco support for Web 2.0, service enablers, on open service developer and delivery platforms.

The end-to-end and Inter-Telco collaboration challenges

Looking forward on how Telcos can collaborate and find approaches and solutions for offering ASQ E2E several challenges appear. These challenges and issues represent significant stumbling blocks. Here are some pertinent questions.

- What can be future-oriented and sustainable ways for sharing and pay for Internet capacity?
- What should the end-to-end resource usage fairness principles be? From the user's perspective and from the network provider's perspective?
- How should the Telcos coordinate end-to-end product and service requirements and capabilities and still stay competitive?
- What are the better revenue sharing models and mechanisms?
- How transparent can the Telcos be? Towards the end customers, OTTs and among the Telcos?
- What trust levels can be assumed?
- What services and quality levels are actually needed? Do we need a different treatment per application? Or, maybe, just two or three different levels could be enough?[3]
- How to best migrate from an environment with local solutions to an "Open Global End-to-End Service" approach?
- What are the success criterion that can enable and support Telco collaboration and coordination?
- What should be the initial bootstrapping roadmap? Are there simple but sufficient solutions to get started? How to get the snowball to role?
- How to position various future services and approaches with respect to the interconnect infrastructure that are available today? That is, the current Internet interconnect model and the current GSMA GRX/IPX interconnect models.

Technical challenges

In addition, there are several technical challenges. Technical solutions that can be used to offer ASQ E2E have not been exploited extensively. Hence, technologies are immature. For more info on Technical challenges and emerging solutions, see section 2.3.

Summary of the discussion

The situation described above is very difficult to get out of, to say the least – it can be considered as a deadlock. To get out of this deadlock, it is very important to raise awareness around these issues. Lack of sufficient industry insight is a problem by itself. While the current Internet has provided opportunities for

[3] This decision would depend, firstly on the different types of applications and requirements and the limitations imposed by the technologies.

several innovations it is also fair to say that the current Internet is blocking potential innovations that can come from offering ASQ enabled network services.

It is also recognized that there are several challenges related to market communication and public relations when it comes to explaining and motivating in a convincing way why future ASQ enabled network services are needed and how such new services can be positioned as part of a portfolio of differentiated Internet services. Furthermore, to find and define a bootstrapping case and according solutions will be a key to get beyond this deadlock. To develop and specify a way of how network service providers (NSP) can and should collaborate in order to support such a bootstrapping case and offer a premium service, including sharing revenues in a sustainable way, is fundamental.

To start with a simple service and a simple but sufficient approach and with a smaller dedicated group of NSPs may be the way to approach this – however, ensuring that the approach is extensible, scalable and sustainable in the larger or even global perspective.

In section 4, some key factors and options are identified that will impact on how NSPs can get beyond this deadlock, and accordingly, indicate what decisions different kinds of NSPs will have to make in relation to their interconnect strategies.

2.3. OVERVIEW OF QOE, QOS AND QOS ENFORCEMENT TECHNOLOGIES

Regarding QoS an important issue to consider is the meaning of the concept itself. From an end user perspective, an important concept is the QoE (Quality of Experience) since it represents the overall assessment of the overall performance of a system for the end user point of view. Therefore, QoE is a measure of the end-to-end performance at the service level from the user perspective and an indication of how well the system meets the user's needs. On the other hand, QoS is a measure of performance at the packet level from the network perspective and performance of other devices involved in the service.

So, an important decision is to select all those parameters that are considered as important in the specification of the quality of the service. It is important to characterise which parameters are specially important for the service: traditional network performance parameters such as the delay, jitter, loss ratio, latency, packet loss, delivered throughput and other parameters such the availability, reliability of the service, accessibility from different networks, nomadism, mobility, etc. At the end, all these parameters will directly impact on the experience of the end user and should be considered in the different agreements that could be done at the different levels (e.g. in order to assure the availability of a specific service, the application provider must assure the availability of the applications servers, the network operators must protect data paths, etc.).

Then, the different parameters must be moved to the end-to-end scenario. In particular, if for example the QoS path problem is presented in the inter-domain environment, different problems should be addressed:

First of all, the minimum guarantee is related to the connectivity; this implies the existence of routes for any destination. Then the choice of the route can be made on estimations based on the QoS each domain can provide. These estimations can be based on measures, eventually shared by different AS on the routes they have experimented (for example in case of alliances or use of reputation mechanisms) or on

performances announced by each AS about its own performances. Such paradigms do not need prior contract and can be supported by Best Effort routing technologies such as (enhanced) BGP protocols.

The second paradigm appears when each intermediate domain on the selected route has to guarantee some performances (bandwidth, delay, security, etc.). Thus, such routing models require *prior contracts* established with each crossed domain, this can be done following different models of interactions as it is described in Annex A (section 9.1).

As mentioned before, providing end-to-end QoS across multiple carrier networks is currently done either through overlays, or by building a chain of AS that will each participate to part of the end-to-end QoS budget. Now, we will provide an overview of the technologies allowing the second category of End-to-End QoS enforcement, as in the first one enforcement is more or less independently from the way operators managed their networks[4]. The technologies that are briefly described in this section mainly focus on the end-to-end path and the collaboration among different Autonomous Systems.

The network management architecture and provisioning processes shall also play a fundamental role into the architecture because the service contracts and SLA activations should first be pass a kind of 'collaboration' phase where each interested carrier shall process and negotiate technical and business conditions before activate computational plane. These aspects shall be further exploited in D4.1 and D4.2 in next months.

2.3.1. Border Gateway Protocol (BGP)

The basic end-to-end connectivity can be ensured by BGP the Border Gateway Protocol. BGP is massively deployed in current networks and is not performing too poorly for most of current Internet usages. However, BGP also suffers from some limitations that may infer with the global objective of the E2E QoS enforcements as summarized below:

- BGP routes are created on the tiers hierarchy with side effects such as the "valley-free" routing, while in some cases, better routes may be found with short cuts allowing having a better end-to-end QoS (for example a lower packet transmission delay).
- Without considering QoS attributes, BGP already suffers from inherent problems due the large amount of routes to be maintained imposing long convergence time in case of any modification, and instability due to diverging/competing policies of operators[5].
- Attempts to add QoS attributes either led to inaccurate solutions (due to aggregated QoS properties or to too volatile QoS data such as the delay of the available bandwidth which makes them inaccurate when propagated in the whole network) or to increase further the inherent problems of BGP (basically, the explosion of entries in the BGP routing tables).

Thus, no solution to extend BGP with QoS attributes has actually succeeded in standards, so this does not seem the way to provision end-to-end QoS paths.

[4] Note that the next subsections do not provide a complete overview of the existing technologies; they just show some technologies and the main problems they have.

[5] Moreover, due to the limitation of the number of IPv4 addresses, the addressing space is being fragmented, therefore the number of routing entries in the BGP tables is increasing.

2.3.2. BGP IP/MPLS VPNs

As explained in the annex A (section 9.2.1), the central idea of MPLS/BGP IP VPNs is to (re-)use the internal part of the BGP protocol (i-BGP) for the distribution of VPN address prefix information within individual autonomous systems (ASes). The method described in [RFC4364] is based on a "peer model": the CE (Customer Edge) routers send their routes to the PE (Provider Edge) routers; and BGP is used to exchange the routes of a particular VPN among the PE routers that are attached to the CE routers associated to that VPN. The distribution of the routes is done in such a way that ensures the routes from the different VPNs remain distinct and separate (e.g. two VPNs can have overlapping address spaces).

Each route within a VPN is assigned a Multi-Protocol Label Switching (MPLS) label. When BGP distributes a VPN route, it also distributes the MPLS label associated to that route. Before a customer data packet travels across the operator's backbone, it is encapsulated with the MPLS label that corresponds to the route that is the best match to the packet's destination address. This MPLS packet is further encapsulated (e.g., with another MPLS label or with an IP or Generic Routing Encapsulation (GRE) tunnel header) so that it gets tunneled across the backbone to the proper PE router. Therefore, the backbone core routers do not need to know the VPN routes, i.e., they may remain completely VPN-agnostic (cf. [RFC4364]).

As far as BGP/MPLS IP VPNs which cross multiple ASes are concerned, [RFC4364] envisions several different mechanisms for inter-domain VPN provisioning, with different levels of scalability and management overhead.

However, all of those mechanisms also assume a tight cooperation between the involved network operators, advancing the provisioning of multi-domain MPLS/BGP IP VPNs to become a non-trivial contractual and operational matter.

The provisioning of BGP/MPLS IP VPN traffic with the appropriate assurances completely depends upon the QoS provisioning mechanisms deployed in the underlying IP transport network. If the network operator does have the appropriate capabilities, extending the same QoS service to the provisioned VPNs represents a very straightforward, easy-to-solve task.

Regarding the support of QoS for VPN services, in a multi-domain scenario, this challenge will face the same problems that could be derived from the provisioning of QoS based solutions.

2.3.3. PRE-CONGESTION NOTIFICATION

The objective of Pre-Congestion Notification (PCN) is to protect the quality of service (QoS) of inelastic flows within a Diffserv domain in a simple, scalable, and robust fashion. The following steps achieve this. Within a PCN-domain, PCN-traffic is forwarded in a prioritised Diffserv traffic class. On every link in the domain the overall rate of PCN-traffic is metered, and PCN packets are appropriately marked when certain configured rates are exceeded. These configured rates are below the rate of the link, thus providing notification to boundary nodes about overloads before any congestion occurs (hence, "Pre-Congestion Notification"). Based on the level of marking, the boundary nodes employ two mechanisms: admission control, to decide whether to admit or block a new flow request, and (in abnormal circumstances) flow termination, to decide whether to terminate some of the existing flows. Admission control and flow termination can be used separately or in combination, according to the operator's design, to achieve QoS

for the inelastic flows. The PCN architecture is described in [RFC5559], whilst the standardised marking behaviours and encoding are in [RFC5670] and [RFC5696].

As a baseline PCN is assumed to work within a single Diffserv domain. However, a PCN-domain could cover multiple Diffserv domains that agree to cooperate (trust each other). They might count marked packets across their interconnection, as part of their SLA, but otherwise would not do any PCN-related operations at the interconnect. It may also be possible to have a PCN-domain that covers multiple non-cooperating Diffserv domains, perhaps by using the technology being developed in the IETF's CONEX WG.

2.3.4. PATH COMPUTATION ELEMENT (PCE)

[RFC4655] states that *"Path computation in large, multi-domain networks is complex and may require special computational components and cooperation between the elements in different domains. This document specifies the architecture for a Path Computation Element (PCE)-based model to address this problem space "*. One of the major problems for calculating an end-to-end constraint (e.g. QoS) path is that a single entity cannot have the full visibility on each domain topology and resource usage for obvious reasons.

The PCE architecture allows for multiple PCEs (e.g. one PCE per domain) to collaborate to compute an E2E constraint path, each PCE having the responsibility to compute an intra-domain path satisfying its part of the E2E QoS subcontract, while preserving confidentiality between operators. According to the current recommended process, the E2E path is computed using the Backward-Recursive PCE-Based Computation (BRPC) Procedure to iteratively select best QoS path in each domain and prune non compliant paths allowing the source PCE to select the best QoS path among all possible ones. While this seems to perfectly answer the E2E QoS enforcement problem, some limitations are actually remaining:

- The PCE architecture does not intend to be used at the scale of the global Internet, but **only for a limited number of domains**. It therefore does not intend to replace BGP to provide the global connectivity and be used for Best Effort paths, but instead to provide alternative routes for services having additional constraints.
- The BRPC procedure take as an important assumption that the AS path (the domain chain) is already known, which forbids to have redundant AS-path or a choice between multiple AS paths that may satisfy the E2E QoS performances.

2.3.5. SUMMARY OF SOLUTIONS

Current solutions and technologies together with emerging solutions even not fully standardised most probably shall define the set of building blocks needed to elaborate and then to propose a candidate architecture of protocols for the Inter-Carrier (IC) model and business for future networks. In particular both IC and higher levels of inter-domain communication have to be ensured. The inter-domain shall add further interconnection requirements from application layer and business layer point of view (represented in Figure 4, as "Web"). The following figure shows the concept and the interconnections at carrier level and domain level.

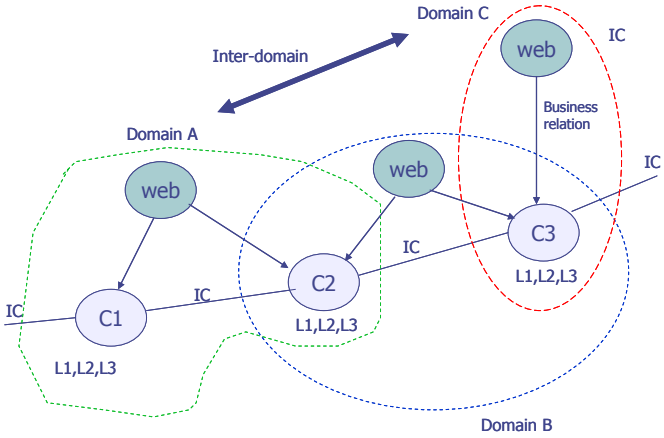


FIGURE 4: CARRIER AND DOMAIN LEVEL OF INTERCONNECTION

Therefore, in order to make a broader view of the complex problem of the interconnection, the solution and protocols to be analysed should encompass almost all relevant layers or the 3 simplified layers proposed by IPSPHERE / TM FORUM.

In order to identify various protocols and solutions that may be candidate for inter-domain and inter-carrier architectures, the following figure shows how the Inter Carrier SUITE shall be characterised and its IC Engines / Elements.

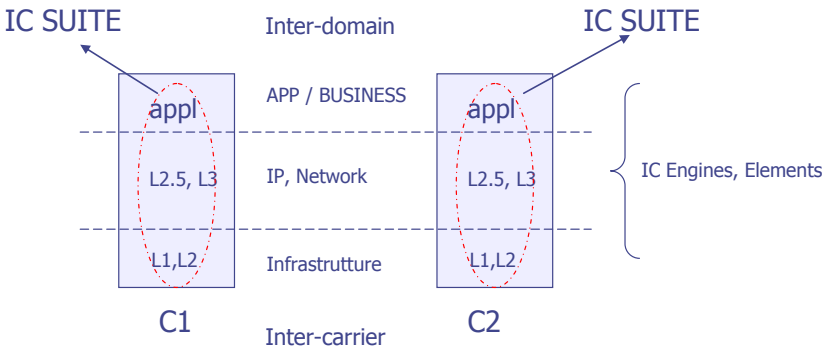


FIGURE 3: IC ENGINES

Should the project decide to limit its interests to lower layer solution, the implication from application layers and IP should be considered as well.

3. CURRENT SERVICES AND BUSINESS MODELS

This section provides an overview of the current services and business models in today's Internet but it also considers other market such as the provisioning of VPN (Virtual Private Networks). The goal of this overview is to provide the major highlights that represent the current issues that could be demanded and/or interesting for the evolution of the business models in the Internet value chain and, therefore, in the technology evolution.

3.1. VOICE SERVICES

In the context of Internet (and packet switched networks) VoIP is a relatively recent technology that has been on the rise since it first appeared. With its market penetration it was classified under the category of Disruptive Technologies. This term refers to innovative technologies that change the market in an unexpected way. And so did VoIP, due to low costs and great flexibility, it has truly changed the telephony market.

3.1.1. BUSINESS MODELS AND SERVICES

The most straightforward business model that uses the VoIP technology is traditional voice services that up to now have been using PSTN[6] as their underlying technology based on circuit switching. VoIP has been used by some telephone providers to replace PSTN as a way to provide voice services. This has given such Telcos the competitive advantage over their competitors who are still using PSTN and have to charge more (sometimes double). An important characteristic of the voice services is that in PSTN networks, the payment is based on cascading, payments, that means that all the domains are aware of all the calls that must be carried out through them and that this is a model that is being tried to replicate.

The appearance of voice over IP has allowed for other thin client voice applications to appear that work directly through a computer and allow people to make calls to other computers or even phones for free or very low rates. The most popular of these clients is possibly Skype [Skype].

In addition to traditional voice services other voice business models [Johs04] emerged such as:

Unified messaging: Unified messaging extends the basic functionality of telephony, which is streaming voice, to include other features that one typically finds on the web. The user receives all their correspondence (whether it would be voice messages, emails or fax messages) in a single inbox and has the ability to use and manipulate this data in various ways [CIS10].

Distributed call centres: a VoIP call centre is used by a company as a way to centralize telemarketing, customer servicing and ordering, and thus to optimize the marketing strategy of the company and the effectiveness of the response. It reflects the concept of a traditional call centre but with a lot of its aspects automated and optimised.

[6] Public Switched Telephone Network

Custom call handling: A lot of overhead and cost can be omitted for a business by using the flexibilities of VoIP and custom call handling. E.g. this allows the administrator to setup rules to forward calls to appropriate numbers.

IP Centrex: The IP Centrex is a set of specialized business solutions which allows the business to apply logic to the incoming our outgoing calls helping optimize and automatically handle calls at least to certain extend. IP Centrex software can understand the keys dialled from the caller and appropriately direct the call to the phone where the call should be terminated. Also such software can be developed to dial numbers and understand whether the receiver is available, busy or using call forwarding. These are just a few examples of the possibilities of IP Centrex. They can optimize the operation of a company and create new business opportunities and reduce operational costs. [IPC10]

VoIP VPN: As its name implies, VoIP VPN refers to the transfer and exchange of secure voice. It works by encrypting the voice for transmission by applying standard data-encryption algorithms. On the receiver's end the VoIP receiver decodes this stream and plays it in its original form.

Hosted PBX: PBX allows communication between VoIP and traditional PSTN. Typically PBX is used to hide a VoIP based telephony system behind existing PSTN lines. It allows all company users to use the same external lines while local calls are exchanged over the data network inside the company itself. Hosted PBX is a service often provided by VoIP providers. With this service the customer can seamlessly use their VoIP telephony without having to buy or set up PBX-specific equipment, which is all taken care of by the Hosted PBX provider.

VoIP Wholesale: This typically refers to the service of offering VoIP termination by wholesale carriers to other service providers. In this way, service providers can gain or extend their reach to the wholesale carrier's network. One example of such an interconnect backbone network is the IPX network. For more information on the IPX interconnect solution see Section 3.4.2 below.

3.1.2. ISSUES AND LIMITATIONS

Since VoIP is by definition a technology of the packet-switched network it is handicapped by the limitations and disadvantages of IP. Because of this, VoIP does not have the same reliability and reputation of traditional PSTN. In order for VoIP services and business models to become successful they need to be as reliable as the alternatives they replace. The quality of a phone call is primarily measured in distortion[7] and delay[8].

PSTN networks offer guarantees of a fixed delay and minimal distortion without a very easy provisioning of the call. As defined in the PSTN world, PSTN offers carrier class reliability and usually corresponds to 99.999% percent reliability. The single points of failure in a PSTN network are 4/5 class switches which aim to have 99.9996% availability.

In a VoIP service the system consists of the entire network or potentially sets of networks and necessary software need to provide the voice service between two users. VoIP's reliability can be measured using two

[7] Distortion is the difference between the received and original signals.
[8] Delay is the time elapsed from the origination of the speech signal until the destination user receives it.

criteria primarily: end-to-end downtime and DPM[9]. DPM can occur when three essential voice service requirements fail to take place: accessibility[10], continuity[11] and fulfilment[12]. The disadvantages of IP such as random/consecutive packet loss and excessive/variable delay are what cause increased DPM. In order to make VoIP packet loss as low as possible, networks are normally configured to give priority to such packets over others. This, still, does not guarantee carrier class reliability.

From the VoIP provider's perspective, in order to increase the trust and reliability of their service by insuring that SLAs are in place between the provider and the network providers along a VoIP call. SLAs are necessary to enforce the QoS of the end-to-end communication but they can be costly and the process of putting an SLA in place can take a long period of time.

One of the major challenges for voice services is the provisioning of VoIP in future mobile networks (as it is expected in LTE), but in this challenge, operators and the industry has considered as mandatory the provisioning of QoS guarantees for this services. In this sense, there is an important effort for the development of components as they are proposed in 3GPP and TISPAN.

As well as QoS enforcement is an important requirement, another important point is the interconnection of the voice services: what happens when several providers should interconnect in order to set up an end-to-end phone call? In PSTN all the operators were aware of the number of calls going through the different networks and the payments were based on a cascading model. It is important to highlight that, in order to translate this model to the VoIP services a need for managing the interconnection at the application level is required. That means that there should be dedicated platforms to manage the VoIP signalling messages and generate the accounting records (e.g. IPX model of the GSM Association).

3.2. INTERNET ACCESS SERVICE

Probably this is the king of current networks: the access to Internet. Internet is not just a service it is also an important revolution from the social and economical point of view. Therefore, the application of any policy to the management of this service could be due to technical and economical requirements but also to political issues.

3.2.1. BUSINESS MODELS AND SERVICES

There are multiple models in the current Internet ecosystem. But the following ones are probably the most representatives:

- End users connect to Internet and they pay flat rates to the access providers. They are not used to pay for Internet services: file sharing with P2P applications, consuming/uploading videos (from YouTube), participation in social networks (Facebook, Tuenti), web browsing to watch news, etc. In all these services the most common scenario is that the end users do not pay but they are the target of the advertising.

[9] Defects Per Million
[10] Accessibility is the ability to be able to start a voice call at any given time.
[11] Continuity is the characteristic of a phone call that is completed with no interruption.
[12] Fulfilment refers to whether the call's quality matched the expectations of the user.

- Network operators obtain money in order to provide the access to Internet. They have to optimize their network costs in order to keep the margins. They also try to provide their own services in order to obtain additional revenues. In order to provide connectivity to the whole Internet, the different network operators must interconnect between them, classically with a hierarchy between the smallest and the biggest operators (a.k.a. "Tiers" 3 to 1) as depicted in the next figure.

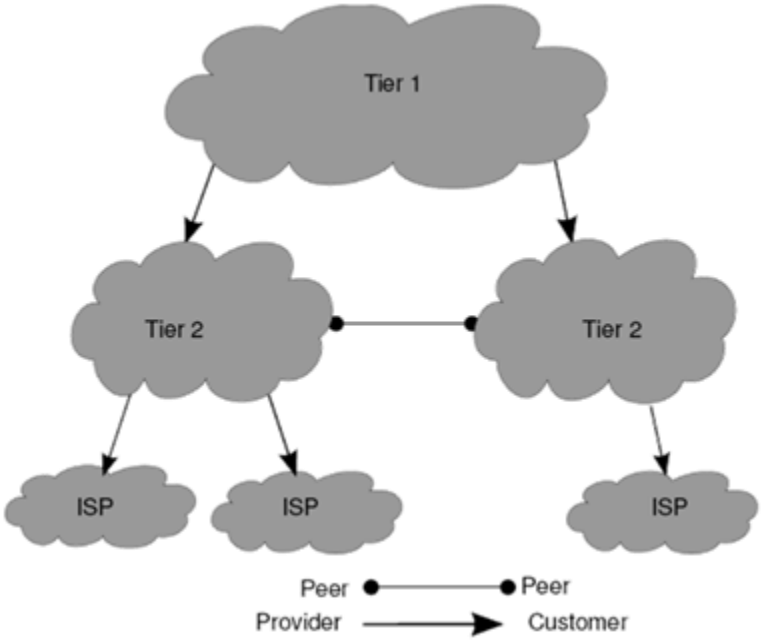


FIGURE 4: INTERCONNECTION AMONG NETWORK OPERATORS

As shown in the previous figure, there are two main types of relationships:

- Peering agreements (Peer to Peer relationship): the traffic is exchanged in a fair way. If the traffic keeps some symmetry, there is no payment. This agreement is maintained by operators of the same size.
- Transit agreements (Customer-Provider relationship): the customers pay the providers and the payment is based on the traffic volume (global traffic, dedicated ports) and there could be some discounts based on the traffic sent by the customers to the providers (95^{th} percentile rule). This type of agreements is maintained between different operators with different sizes and the normal situation is that providers send much more traffic to customers than customers do.

These business relationships have a direct impact on the way traffic engineering policies are applied: an operator will try to send direct traffic to its customers (since it is paid for that), it exchanges traffic to its peers in fair way (use peers' links to send the traffic to peers' customers) and it tries to avoid sending traffic to its providers (since these links are expensive).

- Service Providers or OTT providers maintain their services; they maintain Points of Presence in different countries for reliability. There are multiple business models:
    - CDNs providers facilitate all capabilities to content providers to distribute their contents. They receive the money from the content providers and the charging is based on the

popularity of the content. They also offer facilities to present the content (e.g. insert the advertisements for content providers), management centers, etc. A clear example is Akamai.

- The content providers get the money from subscriptions to the content and mainly from the advertisements.

- Other providers such as Google offer their services and their main source of revenue is the advertising.
- Skype offers new services that can be built over the networks.

These providers have to pay the network providers for their access to Internet but since they are the source of some important type of the traffic, any balance policy they apply in their own servers could have an important impact on the agreements between operators (e.g. if the YouTube traffic is coming from a peering link between Tiers-1, this can make that the symmetry in the exchanged traffic is lost and therefore an important cost has to be paid).

3.2.2. ISSUES AND LIMITATIONS

From the operators point of view, probably the most important limitation in the current model is that the end users are used to pay flat rates and that the last years' trends have been to try to reduce the amount for this tariff. On the other hand, as shown in section 2, the traffic is blooming and this has a direct impact on the network costs: up to now the new network technologies have made possible that the cost for bit transmission could be reduced but with the current model this is not possible any more: there is not clear commercial offer for high speed IP interfaces and the continuous needs for investments in the access technologies.

This is making that the network operators that need to maintain their infrastructures and continue the investments (considering that the Return of Investments in network investments, and especially in the access, are not obtained in the short term) in network reduce their margins and the main benefits are going to the service providers. This is creating some kind of battle such as that one between some European networks providers and Google.

On the other hand, as commented in section 2, new limitations per user are starting to be applied such as e.g. the volume cap that is currently offer for the Mobile Broadband connectivity, which tries to reduce the impact of the heavy users in the networks.

Finally, the lack of end-to-end guarantees could be a drawback to help the development of new agreements between different providers and to offer new issues, such as:

- Up to now the network operators, in order to provide the connectivity service, are only focusing on the basic connectivity dimension. They are not offering differentiated connectivity services based e.g. on a prioritization on the busy hours (that could be helpful for mobile networks), or prioritization of the video traffic in general, etc.
- The operators cannot offer worldwide service providers end-to-end capabilities since operators do not guarantee such performance between for their own services.

3.3. IPTV SERVICES

IPTV refers to the technology of transmitting digital television using the IP protocol over a packet-switched network such as the Internet. Compared to the traditional broadcasting of television, IPTV offers the opportunity to develop a more interactive interface for the television. IPTV can offer meta-information about the program the user is currently viewing. Also it offers more control to the user by letting them rewind and pause live television [Mon08].

3.3.1. BUSINESS MODELS AND SERVICES

Due to its flexibility it provides a platform which can be used to promote many business models. IPTV not only offers the same functionality as traditional television, which is the broadcasting of television programs, it also enhances the user experience with added features and extended control over what, when and how they watch something. The features you receive on your IPTV box depends on the IPTV provider and what they offer with IPTV software they use. There are many examples of IPTV features that make it different from traditional television. For example it allows the user to browse the television programs to come from a channel using their television controller and also allows them to go back in the television's schedule and watch a program they missed or want to watch again. Some providers even allow you to pause and rewind live television. Others can allow you to set reminders for your favourite programs and be reminded while you are watching television to switch to it.

Besides the enhancements to how a user experiences television, other business models appeared that took advantage of this flexibility and control over what content the user views.

VoD: VoD is a service that has really grown in the last few years with more and more people using it instead of having to go to a DVD rental store to find a film they want to watch. As its name implies it can provide users with a video whenever they choose to view it. The Video on Demand business model is very similar to that of a DVD rental store and one can see that VoD is the online version of this traditional model. As a service VoD is much more convenient for a user rather than having to go to a DVD store to rent a film, have to worry about the DVD being scratched and also have to return it when the rental expires. With VoD the user can just sit on their couch rent and watch a film with only the use of their remote control.

Targeted Advertising: Just like in almost every business, advertising is one of the greatest resources of revenue. Up to now traditional television has been broadcasted to its users without any knowledge of who was viewing a channel at any given time. Through the platform the IPTV provider has the opportunity to display advertising to the customer depending on the television program the customer is watching and information the provider has about the customer and their preferences. This way advertising become much more effective and more likely for the customer to be intrigued by the offered product or service. An example of this is offering the option of ordering a pizza while you are watching a football game.

Gaming: With IPTV increasing the interactive capabilities of television, providers can also offer games as part of their IPTV console. The IPTV provider can therefore make profit but allowing the users to purchase

the games or even renting them. Accedo broadband is an example company offering IPTV providers their gaming package for IPTV with games and puzzles for people to buy. [Acc10]

TV store: Similar to gaming, the IPTV provider can use the IPTV platform to offer products and services of third party companies through a TV store that can advertise the products and their offers while a customer is watching television. With the click of a button the customer can purchase the offered products or services using their credit card (7). Again a good example of this is being offered a pizza while you are watching a film and have it delivered to your doorstep without moving from the couch.

3.3.2. ISSUES AND LIMITATIONS

Through the business model examples mentioned above it is obvious that the bandwidth requirements are significant, and specifically its availability and reliability. Traditional television broadcasting is relatively reliable with outages usually happening because of issues with the customer's antenna. IPTV needs to be at least as reliable. One of the major issues with IPTV nowadays is that it can sometimes be lagging. This can cause the customer to become frustrated and sometimes switch back to traditional television. It is important for this issue to be overcome if IPTV is to continue to exist and appear in more households.

Just like with VoIP the issue here again is that a provider cannot always guarantee the quality of a network unless the network is fully under their control and SLAs are in place. If the QoS of the end-to-end networks was guaranteed then perhaps the consumer trust towards IPTV would increase which would mean more people would switch over to IPTV and provide revenue for the business models mentioned above.

Another important issue that should be considered is that the IPTV service is a service that has been built from the scratch: it requires a specific provisioning in the aggregation networks, there is no interconnection model that allows some roaming of the services, etc. And one important characteristic of this service that has a direct impact on the benefits is that it requires a dedicated equipment located at the end users' premises that is owned by the operator and that this is an important source of problems.

3.4. VPN SERVICES

3.4.1. BUSINESS MODELS AND SERVICES

Since the breakthrough of IP as the most important global data connectivity network in the mid 90s, companies and organizations have been intensively engaged in transitioning all (or at least most) of their data communications towards the Internet. In this sense, Virtual Private Networks (VPNs), as overlays which provide customers with transparent remote site connectivity over a shared network infrastructure, began to play an important role in this process, as remote private-domain connectivity needed to be continuously assured irrespectively of the underlying data transport technology used.

The emerging technologies for provisioning VPNs have basically been focused on two distinct paradigms:

- **Provider-provisioned VPNS:** the network operator takes full responsibility and the technical configuration burden for connecting multiple remote sites of a customer into a seamless VPN. In order to make such a form of VPN provisioning feasible, there are two fundamental prerequisites. Firstly, in case no CPE-based encryption of VPN traffic is used, the customer needs to trust the operator with the entire security of its private network traffic. Secondly, operator-managed VPNs are easily deployable

only within a single network operator domain, such that all the VPN sites of the customer should ideally be within the connectivity reach of the VPN-operating ISP.

- **Customer-provisioned.** In the case that the prerequisites mentioned in the previous paragraph are not (or only partially) fulfilled, the customer will need to assure VPN functionality on its own, normally employing VPN-capable CPE-devices implementing some specific protocol. The main drawback of this approach lies in the fact that customers in this case need to take full care of the management, provisioning and maintenance of their VPNs, which requires a substantial level of understanding of all the networking and security aspects involved. Furthermore, most of the customer operated approaches are associated to more rigorous constraints with respect to planning for highly scalable VPNs.

3.4.2. ISSUES AND LIMITATIONS

As stated in section 2.3.2, the provisioning of VPNs based mainly in IP BGP/MPLS technologies lacks from the clear provisioning of multi-domain QoS guarantees. The evolution of VPN services could be an excellent starting point for the deployment of end-to-end QoS assurance capabilities.

Due to the high number of corporate users that join their networks using customer-provisioned VPNs, a scenario where these VPNs could be provisioned in another way or that, at least, could offer more services to the end users, would be an excellent point to deploy QoS in the networks. Moreover, this approach could make that the way to provision both types of VPNs could be very similar.

3.5. INTERCONNECTION WITH IP EXCHANGE (IPX)

IPX is an interconnection model for the exchange of IP based traffic between customers of operators and service providers (e.g. ISPs). It is developed by GSM Association (GSMA) as an extension of the GRX (Global Roaming eXchange) architecture.

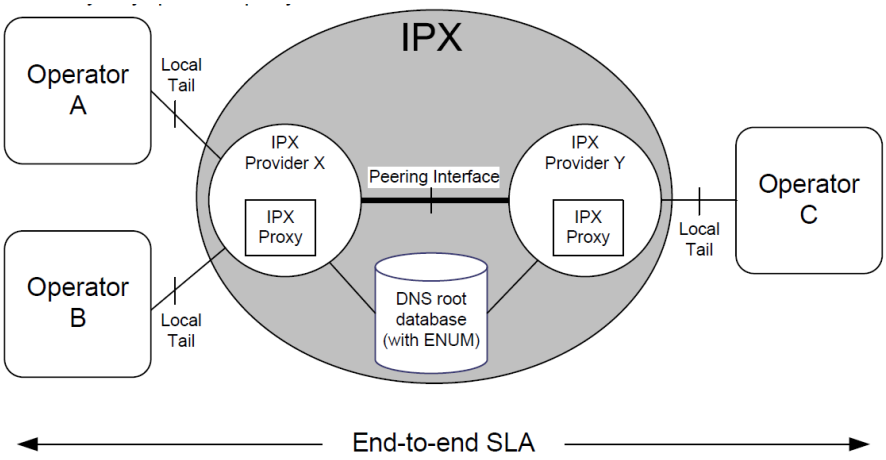


FIGURE 5: IPX-MODEL (FROM [IR-34])

The intent of IPX is to provide interoperability of IP-based services between all service provider types within a commercial framework that enables all parties in the value chain to receive a commercial return. The

commercial relationships are underpinned with service level agreements (SLA) which guarantee performance, quality and security.

The IPX is a global, private, IP network which supports end-to-end quality of service and the principle of cascading interconnect payments. In order to provide these features the IPX is service aware unlike the Internet and the GRX.

3.5.1. BUSINESS MODEL AND IPX SERVICES

The IPX environment will consist of a number of IPX carriers operating in open competition, selling interconnect services to Service Providers, mobile and fixed operators. The IPXs will be mutually interconnected where there is demand from Service Providers.

IPX supports 3 different types of interconnect models:

- IPX Transport
- IPX Service Transit
- IPX Service Hub

It will be up to the Service Provider to determine which connection model it wants to select from its IPX provider.

The use of the **IPX Transport** connectivity option implies that the two Service Providers that exchange traffic have entered into a direct contract, outlining the cost of termination for each type of traffic exchanged or service used.

In this case, service/application level charging would not be subject to an agreement between a Service Provider and the IPX Provider. This is a bilateral connection between two end Service Providers using the IPX **Transport** layer only with guaranteed QoS end-to-end. It is a bilateral agreement between the end Service Providers and any payment of termination charges is a matter for these Service Providers. Cascading of responsibilities (such as QoS) applies but not cascading of payments (Cascade billing). Each Service Provider will pay their respective IPX Provider costs for transport capacity.

The **IPX Service Transit** Connectivity Option enables a bilateral agreement and connectivity between two Service Providers utilising the IPX Transport layer and IPX Service layer provided by the IPX Provider with guaranteed QoS end-to-end and with service awareness included.

Traffic is transited though IPX Providers but termination charges are agreed bilaterally between Service Providers and settlement of termination charges can be performed either bilaterally between the Service Providers or via the IPX Providers. This is up to the Service Providers to decide.

Cascading of payments (for transport and/or service layer) may be applied depending on the service. The transit fee owed to the IPX Providers is always cascaded. Cascading of responsibilities and payments fully apply (on both IPX Transport layer and IPX Service Layer).

The **IPX Service Hub** is a multilateral connectivity using Hub functionality. Hubbing (or multilateral connectivity) means that traffic is routed from one Service Provider to many destinations/Interworking

partners through a single agreement with an IPX provider. So the IPX provider takes care of both the contract and set-up of connectivity on behalf of the operators.

This mode is operationally highly efficient for the Service Provider. The IPX Service Hub arrangement guarantees QoS end-to-end including service awareness. Cascading of responsibilities applies. This is also by default an arrangement where principles of cascading payments apply. Charging transparency is a requirement on both IPXs and Service Providers in this mode.

In the scenario of IPX Hubbing function, there should be no commercial relationship between the two involved Service Providers between which traffic is exchanged (similar to conventional international voice interconnect via carriers). Hence, service/application based charging as well as charging for the IPX Transport should be subject to the single agreement between a Service Provider and an IPX Provider. Multilateral - Hubbing will open a business opportunity for the IPX to act as an intermediary in the money stream between the service provider and the customer.

To sum up on the IPX interconnection models, IPX provides E2E QoS in all models, but only on transport layer with the Transport only option. Service Transit and Service Hub provide cascading of payments while with Transport Only termination payment takes place directly between the Service Providers. Service Hub provides a single contract and a single connection while the other models provide single contract with IPX provider but multiple contracts with the connecting Service Providers.

IPX shall support IETF Differentiated Services model (DiffServ). In [IR34] the support of DiffServ PHBs EF, AF1-4 and BE are recommended. Service providers are responsible for marking packets to correct traffic classes (PHB) with correct values in the IP packet headers (DiffServ Code Points –DSCP). They may outsource this functionality to IPX provider when suitable IPX providers may change DSCP values in their own network as long as they return values set by operators before traffic is given to another IPX/GRX provider or Service Provider and they fulfil given values for parameters per class.

An IPX provider can change the DSCP value, unless otherwise agreed bilaterally between two Service providers, to be in line with pre-agreed levels within an IPX environment. The DSCP value shall not be altered in the transport mode only in an IPX environment.

The QoS parameters shall be defined in the SLA. The QoS parameter set should be consistent and uniquely understood by all parties involved in the IP connection.

The following QoS parameters are covered: service availability, jitter, packet Loss and delay.

Service availability is a proportion of the time that IP Backbone Providers service is available for service providers (on a monthly average basis). Proposed values for availability are following:

- Availability of the IP Backbone Service Provider Core: 99,995%
- Service Providers connection to IP Backbone Service Provider core with single connection: 99.7%
- Service Providers connection to IP Backbone Service Provider core with dual connection: 99.9%

Proposed delay target values are given in IR34 for various source-destination area pairs and also jitter (variable delay) values applicable for conversational and streaming traffic classes (i.e. EF and AF4 traffic classes) are specified there.

3.5.2. ISSUES AND LIMITATIONS

Although the basic capabilities for supporting many services by the IPX interconnect infrastructure has been specified by the GSMA, the deployments of the IPX platforms are limited and many carriers are still only at the testing stage. The general opinion is that IPX will initially be used for Packet Voice interconnectivity between mobile operators, probably with the Asian market in lead. As IPX is an upgrade of today's GPRS Roaming exchange (GRX), which handle all data roaming traffic between public land mobile networks, such traffic will also typically be carried by the IPX network. However, local breakout techniques may push more of the mobile roaming traffic over to the ISP domain.

Considering future developments and deployments, we can foresee that other IP-based services will contribute to IPX traffic if the operators introduce IMS and/or launch new services. Legacy international telephony traffic may also move to IPX as transport platforms are evolving from TDM to IP. However, it remains to be seen whether the IPX platform will be applied for services beyond voice and the value added services associated with voice services. Many argue that the cost base for the IPX platform is too high for many (or even, most) future services, hence questioning the sustainability of the IPX platform.

The interesting thing here is how the two different interconnect approaches, the GSMA IPX approach and the Internet approach, will evolve in the future. Currently their business models and support of QoS are highly different, but will these models (partially) converge in the future?

3.6. INFRASTRUCTURE SHARING

Infrastructure sharing is an emerging form of business model that involves two or more mobile operators sharing specific parts of their network.

In the past, the main drivers for infrastructure sharing have been regulatory rules aimed at facilitating the entrance of new operators into the market without incurring sudden expenses for network deployment. For this reason, the most common scenarios of infrastructure sharing currently in place consist of national roaming agreements between incumbent operators and new market entrant.

Recently, network infrastructure sharing is gaining popularity among mobile operators also as a mean to reduce capital and operating expenses. Network sharing models are typically categorized in passive and active, depending whether the operators share passive infrastructure entities (e.g. sites, masts, power supply, etc.) or active network nodes (e.g. antennae, radio network controllers, etc). Passive sharing is common in many countries and usually receives good acceptance from the public due to the ecological and landscape benefits. On the contrary, active sharing agreements always face the skepticism of the regulatory bodies, which consider them a threat to competitiveness.

The interest on infrastructure sharing (in all its forms) is substantially increasing, as mobile operators are crippled by the enormous investments in network upgrades and the decrease of the revenues from the customers. As stated by Analysis Mason:

"Sharing radio access networks (RANs) can deliver major operational savings for mobile operators. But to date, the only successful RAN sharing ventures have been used for greenfield 3G roll-outs, where forming a

joint venture and sharing costs has clear, tangible benefits. In mature markets, RAN sharing is fraught with difficulty, but with mobile operators' margins under pressure, 2009 is the ideal time to try and cut costs. The key question is: do the OPEX savings gained by RAN sharing justify the CAPEX investment?

In the mature market scenario, the greatest commercial rewards are achieved by a full consolidation of two (or more) mature mobile networks to form a single, shared network using RAN-share technology. Savings are made by removing sites, reducing the volume of operational sites that need to be managed, and by removing headcount once the consolidation work is complete.

So what is the catch? The scenarios that seem to offer the greatest reward also present the greatest risk. The benefits are also time dependent, so any delays will reduce the commercial benefits, potentially threatening the entire business case."

3.6.1. Business Models and Services

There exist several business models for infrastructure sharing. The following have been used or proposed in the past:

- Unilateral service provision: This is used when the agreement involves an incumbent provider and a new entrant in the market. The incumbent provider just sells wholesale services to the new entrant, and therefore keeps the operation and management of the network.
- Mutual Service provisioning: This is typically used when two operators of comparable strength decide to cover different geographical areas and to provide mutual roaming access. The CAPEX of each operator is considerably reduced since a large percentage of population can be covered by halving the amount of deployed infrastructure per operator.
- Joint Venture: In this case the mobile operators form an equally owned joint venture (JV), which is responsible for the deployment and operation of the shared network. For example, two operators can form a JV to roll-out and operate the LTE infrastructure in specific geographical areas and manage the interconnection with both core networks.
- 3rd party network provider: In this approach a 3rd company (typically an equipment vendor) becomes a pure capacity wholesaler, which rents the spectrum from two or more operators and sells in return wholesale capacity. The provisioning of the service is regulated by service level agreements (SLA).

3.6.2. Issues and Limitations

In section 4.4, we discuss which limitations and possible cooperation mechanisms can be found in scenarios where several providers cooperate.

3.7. MOBILE APPLICATIONS AND VALUE ADDED SERVICES- TELCO INTERACTION WITH 3RD PARTY PROVIDERS

Mobile applications and value added services can be delivered only by a cellular phone and are value added, non voice services, in many cases premium. We do not include into the definition of Mobile Content the following:

- Person-to-Person Services (ex. SMS/MMS messages exchanged directly between two end-users);
- Customer Relationship Management (CRM) services;
- Mobile Payment Services;
- Advertising on cellular phones (with SMS, MMS or Video).
- Software Applications for mobiles such as Office-like applications, voice recorders, video recorders, agendas, mobile browsers (such as Opera Mini or Firefox Mobile), etc.

Most common among Mobile Content services are: customization (wallpapers, ringtones, etc.), infotainment (text news, video, etc.), communication & community (chat, blog, etc..), gaming (Java, Symbian, etc.), applications (for different devices iPhone, Android, etc.) and many others.

3.7.1. BUSINESS MODELS AND SERVICES

In the non-voice value added service and application markets, Telcos are managing the business in strict collaboration with 3rd party players who are in most cases responsible for content aggregation, content management and, in some cases, also for content planning[13]. The scheme below illustrates the value chain relationships between the Telco and 3rd party players in the Premium Mobile Content market[14]:

[13] 3rd party players are usually responsible for content planning when they are responsible for their B2c consumer brand.

[14] Premium Mobile Content market: Premium/paid Mobile Content services for use of which the final user is charged, usually with use of Telco billing via premium Sms (MO - Mobile Originated or MT- Mobile Terminated) or Wap Billing if the service is purchased through the Telco Mobile Portal. Most common premium Mobile Content services are the following: Ringtones, Ringbacktones, Images, Screensavers, Java Games, Infotainment, Video and other.

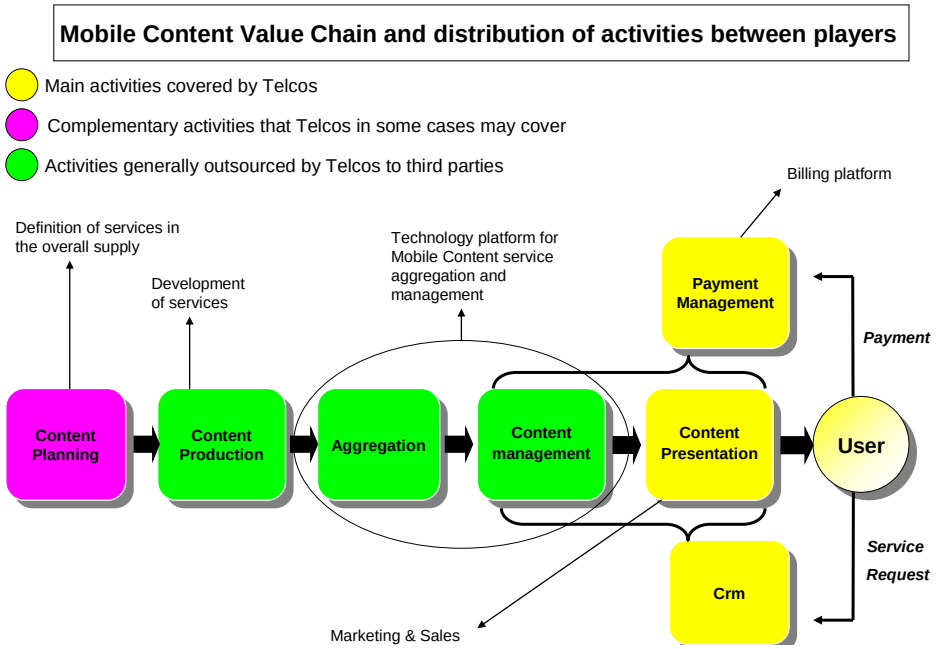


FIGURE 6: MOBILE VALUE CHAIN AND DISTRIBUTION OF ACTIVITIES BETWEEN PLAYERS (SOURCE: MOBILE CONTENT & INTERNET OBSERVATORY, POLITECNICO DI MILANO)

As illustrated in the scheme above, Telcos are managing two main activities in the premium mobile content supply: payment management (through their internal billing platform, unavailable to third parties) and customer relationship management (their own subscribers' customer base). On the other hand, content aggregation and content management are usually outsourced to 3rd party players who own special content management platforms responsible for sending the service to the Telco after each customer request. Main players responsible for this specific activity are Mobile Content Service Providers (MCSP) or Wireless Application Service Providers (WASP)[15]. MCSP can be focused on B2B or/and B2C market. B2B players deliver technological solutions to Telcos and other 3rd party players (smaller MCSP, Content Providers like Media Companies or Web Editors without own content management platform) while B2C players deliver their content to end customers using Telcos payment infrastructure and customer base[16].

Following the example of Premium Mobile Content, it is possible to draw the economic relationship existing between Telcos and 3rd party players. **The most common type of agreements between Telcos and third party players is the revenue sharing agreement**. Telco is the official "seller" of the content that "cash in" the revenues from each content sale. Based on the quantity of content sale to the final consumer, Telco pays the revenue share to MCSP.

Taking as example the Italian Premium Mobile Content Market, where standard revenue sharing between Telcos and MCSP is 50/50, and considering in this occasion the B2C player, it is possible to see the revenue flow between Telco and the 3rd party player in the scheme below[17]:

[15] The most important players in this category with global reach and multiple relationships with Telcos in different countries are m-Blocks and Sybase 365 (this players are exclusively B2b oriented).
[16] Most important B2c MCSP globally are Buongiorno and ZED.
[17] This scheme can be applied for Sms Premium Billing and Wap Billing.

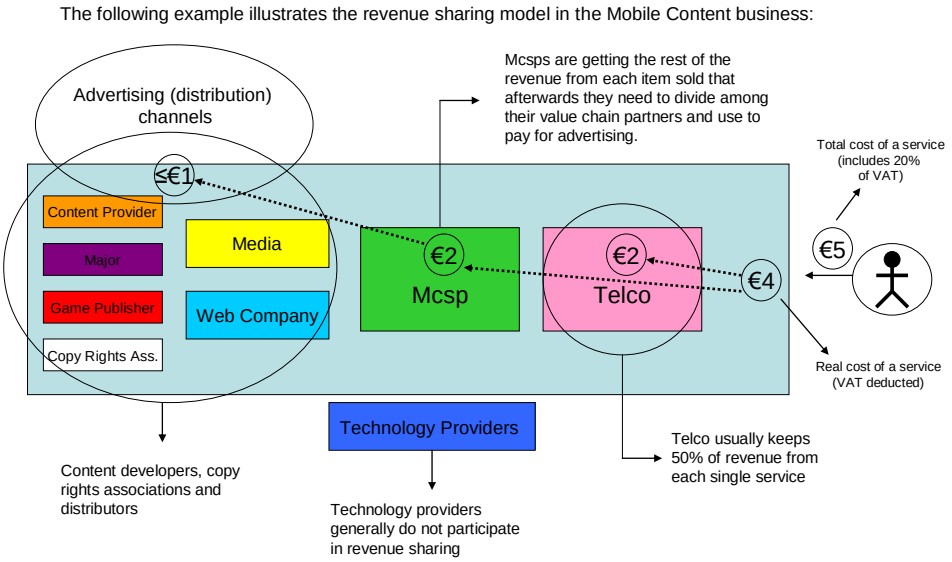


FIGURE 7: REVENUE SHARING IN THE MOBILE CONTENT BUSINESS (SOURCE: MOBILE CONTENT & INTERNET OBSERVATORY, POLITECNICO DI MILANO)

In this scheme, Telco receives payment for service from the final consumer. After deducting 20% tax Telco keeps a certain amount of money that needs to share afterwards with the 3rd party player for providing the service. MCSP after receiving its part of revenue needs to refund other 3rd party players who contributed to the service (usually with product development). This category of players can be defined as Content Providers (Media Companies, Web Editors, Game Publishers, Majors, etc.). Moreover if MCSP is a B2C player it needs to cover all the advertising costs.

3.7.2. ISSUES AND LIMITATIONS

In this actual scenario Telcos are definitely the favoured party (even if revenue sharing agreements change in different countries and MCSP can benefit more)[18]. However this situation is now changing since Apple with its AppStore[19] has somehow revolutionised the market. In front of this new market challenge and upcoming investments in Telco application stores (like Vodafone 360°) as well as in this Mobile Internet era, Telcos will probably have to re-define their relationships with 3rd party players. Moreover the introduction of new billing methods (like credit card – used massively in USA to pay for value added mobile services, and gaining more popularity in Europe thanks to Apple model), can change dramatically as the Telco billing platform is in such case not anymore indispensable for the transaction.

[18] For instance in Northern Europe the revenue share is much more favourable for Mcsps that usually keep 70% from each service sold.
[19] Apple model guarantees 60% revenue share to application developer (who in this case plays the role of content provider).

4. INITIAL ECOSYSTEM SCENARIOS AND INTERCONNECTION FUTURE OPTIONS AND CONSIDERATIONS

This section identifies and categorizes key factors and options that have significant impact on the future options regarding interconnect strategies and multi-NSP collaboration. However, the underlying solutions that can drive the feasibility of these options are not known at this stage and to-be-explored by the ETICS partners or other stakeholders.

However, before going into these subjects a preliminary reflection on some future high-level Telco options and what may characterize the future competitive landscape will be provided. These sections are followed by the identification of key options or models for multi-carrier contracts and markets. Additionally, a section addressing a scenario of cooperating access network providers that also enables cross ISP collaboration is described.

The last section will consider a bootstrapping case and provide some reflection regarding the need for a consolidated next-steps-roadmap among NSPs in order to drive the ETICS project.

4.1. Telco positioning in the ecosystem

An investigation of the Telco position in its ecosystem(s), that is, as seen from a high-level perspective is of value to ETICS in several respects. The understanding of the competitive landscape and the overall strategic priorities of the Telco sector in general or of a specific Telco, will help in understand its priorities and motivations also regarding ASQ E2E. For instance, considering the current focus by many Telcos on Web 2.0 application for mobile devices and transitioning into an agile service delivery business with support for 3rd parties partners may suggest that the ASQ E2E efforts should be positioned accordingly. Moreover, the defined role of a given Telco and how this may evolve has direct impact on the solutions and capabilities this Telco should invest in, that is, capabilities related to ASQ E2E will depend on these strategic choices. The uncertainties regarding how the value-networks will evolve will typically make it more difficult for a Telco to make decisions for the future; hence, strategic decisions may be delayed.

In the following, three main long term options for a typical Telco are presented. Note that these options represent a simplified view. There are many variants around these options. The purpose here is to identify characteristics that can have impact on the motivation for supporting and offering ASQ E2E by a Telco belonging to such a category or if such an option is characteristic for the long term future.

The following will also give some initial reflections on how the ecosystem (or systems) may define the setting for the Telco and how these ecosystems will differ in the three cases, or in scenarios built around one of such cases.

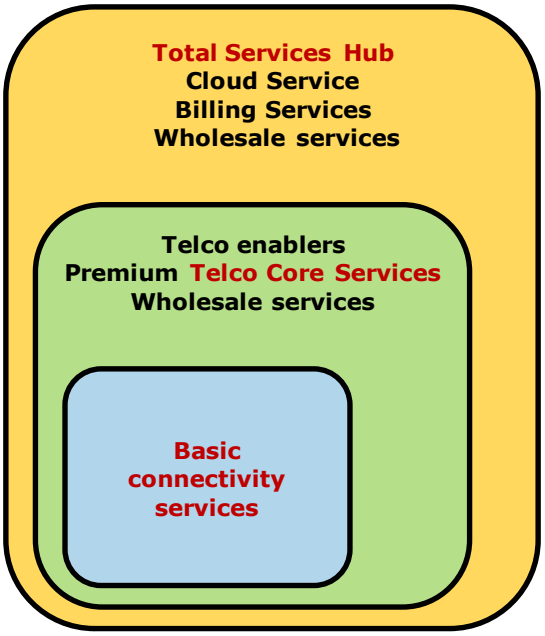


FIGURE 8: FUTURE TELCO OPTIONS

Basic Connectivity Service Provider. In this option, we should consider the following points.

- In this case the Telco is primarily a (dumb) bit pipe provider, still having the ISP role. The Over-The-Top providers (OTT) do provide what are (were) typical Telco services such as real-time voice and video conferencing services.
- Pure ISP focus, typically, the BE Internet will prevail as a preferred option while the assurance of QoS for end-to-end services would be handled by OTT players (that could ask for better connections with high bandwidths and can also evolve their services[20]). The ISP will focus on providing access pipes to the OTT services and media gateways, as well as aggregate transport between OTT servers.
- In order to get more revenues, the Telcos could explore new dimensions of the basic connectivity services to both end users and OTTs. So, e.g. some Telcos could focus on collecting transport of the big content providers.
- The Telco voice (and voice value added services) and GRX/IPX interconnect regimes become marginalized in this scenario.
- Mobile Network operators must focus their business on access services for OTT services and applications while they need to assure the fair usage of the network resources by their end users. Again, in this scenario, several connectivity profiles could be provided to the end users.
- In this scenario, where the Telcos should continue the investments in the network infrastructures, while the OTTs could get the revenues coming from the new services, will we see Telcos companies that try to enter the OTT space?
- The future communication service and application provider space could be dominated by a few big OTT players.

[20] I.e. Skype is able to adapt the codecs used for voice in order to accommodate the service to the path conditions.

- The Telco Service Provider B2B automation space is focusing primarily on automating the wholesale relationships with OTT players. These relationships could be focus on the minimization of the traffic footprint in their networks (e.g. offer of free connectivity if the PoP of the OTT is placed on their networks).

Telco Core Services Provider. In this option, the following issues could be expected.

- The Telco is focusing on premium real-time unified communication, supporting both real-time voice, video and collaboration tool services. These are considered core Telco services.
- The Telco is offering service enablers, in particular those exposing core Telco services to 3rd party Partners. Several other service enablers (SDP, Service Development Platforms) are also exposed via 3rd party APIs such as enablers related to messaging, location, presence, etc.
- A big part of the Telco sector focuses on continuing the fixed-mobile convergence with special focus on the provisioning of video services with guarantees in mobile networks.
- Premium services, beyond Telco voice service, are offered based on both the ISP and the IPX interconnect infrastructures.
- Several business-models for offering premium content delivery services exist.
- The Telco Service Provisioning B2B automation space is focusing on supporting and automating 3rd party service and application wholesale relationships.

Total Services Hub

- The Telco is positioned to be a service "hub" as seen from the end customer point of view. In addition to offering the core Telco services the Telco has expanded also into the cloud services space for consumers and/or business customers
- The Virtual Private Services concept, for business customers as well as for consumer customers, is a well established term in the market.
- Interconnect at "higher" layers is a hot concern among the Telcos as well.
- The Telco SP B2B automation space is a key to Telco success and must support many kinds of wholesale relationships, including automation of interconnection relationships at different layers.

4.2. Considering key factors impacting future interconnect options and scenarios

It is important to establish an overview of the key factors impacting future interconnection for several reasons. By identifying such key factors while avoiding most of their details ETICS should be in a better position to define and select in a clever way the more relevant and detailed scenarios for further analysis. It is important that ETICS is able to make clever selection and definition of various attractive and relevant future analysis cases in order to handle the great complexity that has impacts on the future of interconnect.

A "listing" of such key factors will also provide us an overview of the complexity that the topic of interconnect in general and ASQ enabled interconnect in particular are surrounded with. From a high level, three main key factors or drivers, considering a potential (premium) service, are:

- Willingness to pay
- Expected demand
- What are the capabilities needed and what are their costs – as compared to the as-is situation

If the NSPs are uncertain that there is sufficient willingness to pay for the given service or the demand is considered too low for the foreseeable future, they will not invest in enabling and supporting the service. This is a simple fact, however, if lack of sufficient prognosis or indicators remains one should not expect NSPs to invest. Moreover, if there are significant uncertainties regarding the capabilities needed for realizing the service in one way or another one should not expect NSPs to invest?

In addition to the above, the below key factors are structured into the following categories:

- **Key input assumptions or external factors**
    - **Service and Traffic mix.** Beyond the demand for a given service, the NSPs must consider the overall traffic mix in their networks and for their interconnections. In particular, the following main cases should be considered
        - The share of Real-time traffic is insignificant or significant. This means, if the traffic that requires strict guarantees in terms of delay and jitter (e.g. gaming, voice, videoconference, etc.) represent an important proportion of the traffic.
        - The share of video traffic (or near real time) is significant or insignificant. It should be noted that this traffic, even though has some requirements in terms of delay (that the applications can compensate due to the buffering), the mean requirement is the guaranteed minimum bandwidth that could prevent from stalling times (that has a tremendous bad impact on the perceived QoE) during the video display.
    - **Assured Service Quality End-2-End – Explicit vs. Implicit.** When ASQ E2E is considered two main cases are recognized.
        - The per-session ASQ E2E is realized using some means of explicit control end-to-end. This does not exclude treating quality assurance on an aggregate level in core and transit networks. This can be an attractive approach when dealing with a service with clear session properties, such as video conferencing. This control will be usually associated to the cascading payment model, which allows a session based control for each specific service session.
        - The ASQ (E2E) is realized using explicit (session) control only in the end customer facing edge domain. Further into the network(s) any control is at the aggregate level. Depending on how these aggregate level agreements and control mechanisms are designed some form of end-to-end service level assurance is expected as possible.

- **Internet Access Service**. If considering the traffic and services that is not assured and provisioned as ASQ, the notion of (basic) Internet Access Service is used. For this type of traffic two broad classes are foreseen
  - Simple Internet Access Service: this is the way Internet Access Service is offered today in most cases, without any access or service dependent policies than just pure access bandwidth limits.
  - Evolved Internet Access Service: in order to achieve improved fairness of network resource usage and more efficient usage of resources various mechanisms and policies can be applied. These policies may consider short time windows (e.g. minutes) in order to achieve the wanted results. A typical example is the monthly volume caps defined by mobile network operators that aim to reduce the impact of heavy users in their networks.
- **Roaming.** Today, roaming is primarily considered in mobile services and networks. However, roaming can be considered a generic service capability also applicable to the Internet or Fixed networks. Roaming capabilities in relation to specific paid services have significant impact on interconnect.
- **Application (developers) perspective.** The success of today's Internet is among several things due to simplicity from the application developers' perspective considering the current UDP and TCP/IP APIs. It must be investigated whether any new ASQ based service will require new capabilities in such APIs and if these capabilities can be universal (global) or service or NSP specific. E.g. the WAC (Wholesale Applications Community)[21] is trying to link the application developers' world with the Telco and vendor specifications.
- **Pricing model and user preferences.** The users and customers want predictability when they have to pay for a service. Moreover, they only want a single agreement per service. The flat fee model is popular in the ISP domain, while pre-paid minutes or SMSs are popular in the mobile domain. Various approaches where the customer can choose payment model must be considered. These will also impact how money can flow and revenue shared among carriers.
- **Regulation.** While today's voice interconnection is regulated to a significant degree, the ISP interconnection has little regulation, following the *ex-post* regulation scheme that usually applies to the Telco market in order to detect and solve market failures. The current and future regulation will play a key role on the interconnection, whether heavy regulation or not applies.
- **Net Neutrality** may be taken into account by regulation means. However, Net Neutrality may be a concern without or before regulation mechanisms are introduced.

- **Key end-to-end networking approaches.** In the following, key factors or approaches in the area of end-to-end or inter-provider networking are identified. The below key capabilities can in general be combined.
  - **Connectionless end-to-end connectivity or not.** One of the key principles of the Internet is the end-to-end principle and the connection-less property. On the other hand the emerging NGN networking approach does assume that NATing and/or gating are inherent properties of the

[21] http://www.wholesaleappcommunity.com/default.aspx

approach. How this may change considering a possible global migration to IPv6 should be further investigated. Hence, for the time being the following two main categories in this respect are:

- Internet end-to-end connectionless connectivity
- NGN domain, with domain gating and NATing. The IPX approach as specified by the GSMA is the industry de facto NGN approach.

- **QoS pipe capable.** The capability of offering, provisioning and managing inter-provider QoS pipes is fundamental to an ASQ capable networking solution. This notion is used generically, and the following main types are considered. The assumed dynamicity of creating and modifying such pipes will significantly impact on interconnect agreements and the management solutions for enabling and supporting such pipes.
  - Explicit pipe – some kind of path and label switching, e.g. MPLS, or GMPLS. This category can be further classified according to addressing spaces and layer(s) involved. (E.g. assuming MPLS and Internet addressing, vs. G.709 and private addressing.)
  - Implicit pipe – defined by IP aggregate flow (e.g. with respect to Internet)
    - With end-to-end significance
    - Without end-to-end significance
- **Congestion notification capable networks.** The IETF has or is defining network capabilities that enable congestion notification capable networking.
- **Application awareness.** Application awareness enables and assumes some kind of deep packet inspection (DPI) beyond just inspecting the IP header. It may allow for some value adding networking capabilities and services. Two main categories are identified here
  - Service edge application awareness
  - Interconnect application awareness
- **GMPLS capable networks.** Three main categories of GMPLS capable networks are identified here. The needed interconnect capabilities are expected to be significantly different in these three cases, although some mechanisms are general and applies to all cases.
  - GMPLS as an extension to IP/MPLS of Internet routers
  - GMPLS as an extension to IP/MPLS of private network domains
  - Sub-IP GMPLS network
- **Session control.** Considering session based services such as person-to-person (collaborative) voice and video two main categories are identified
  - Signalling based IMS is the signalling solution expected to become the solution for the future. Several variants or profiles exist. The SS7 based solutions are part of this category and expected to prevail for many years

- Management or handling based. An approach based on Web Services (and/or management) protocols. When this approach is used to manage the sessions in (near) real-time we speak of "**session handling**".

- **Naming and addressing approach.** Naming and addressing, taking into account networks, interfaces, hosts and users, is an important area that should be divided into several more specific areas. This area is not further elaborated here other than pointing to some important current approaches, such as the Domain Name System (DNS) that is an integral part of the Internet, and the Telephone number mapping (ENUM) approach for unifying POTS numbering with Internet naming. The GSMA document PRD IR.40 provides "Guidelines for IPv4 Addressing and AS Numbering for GRX/IPX Network Infrastructure and User Terminals"

- **Topological structure and types of interconnect.** There several options when it comes to interconnect topological structure and types of interconnect relationships. The below identify some of the more important. This is a non-complete list and some of the below options may be combined. The below focuses on peer-to-peer[22] services. Consideration of topologies and relationships for content delivery services is for further study.

  - Full mesh inter-carrier backbone. The GSMA IPX model has adopted this approach. There is a full mesh among all backbone carriers.

  - Bilateral with no transit. End-to-end from one customer-facing edge domain to the other customer edge domain with no transit. E.g. peering agreements in the Internet model.

  - Hub model. The interconnect carrier is a hub as seen by its customer. That is, the interconnect carrier can connect the customer network operator to any other network operator that is connected to this carrier network. Cf. e.g. the IPX network or the transit agreements in the Internet interconnection model.

To illustrate various options (none complete) of interconnect the Figure 9 is provided. As above, the following illustration focuses on peer-to-peer services.

Considering first the case of a customer-facing ISPs offering premium real-time conversational services. In such a setting and assuming per session control must be supported, then both the IPX and the ISP interconnect infrastructures appears as potential starting points for such interconnect requirements as there are pros and cons with both these interconnect infrastructures.

Considering next the mobile network operator (MNO) this kind of actor is typically served by an IPX provider. However, considering mobile broadband Internet access services in most cases today is based on interconnect with an ISP. For the future, considering evolved Internet Access (see above) or (premium) ASQ E2E over mobile broadband access we should consider both the IPX and the ISP interconnect infrastructures as potential starting points for the interconnection needs of a MNO.

In the future, if we have a situation where ASQ for a given service is supported on both the IPX and the ISP interconnect infrastructures, interconnect between actors of these two infrastructures will become an issue. However, it is observed that one actor may play the role of both the IPX and the ISP provider.

[22] Peer-to-peer is here used in a wide sense and includes e.g. VoIP.

Whether this may result in some form of convergence between the IPX and ISP can become an issue. For the foreseeable future this is not an issue.

However the analysis and development of the details of the above interconnect approaches, their requirements and solutions, are at the core of the ETICS project and will be further studied in other workpackages.

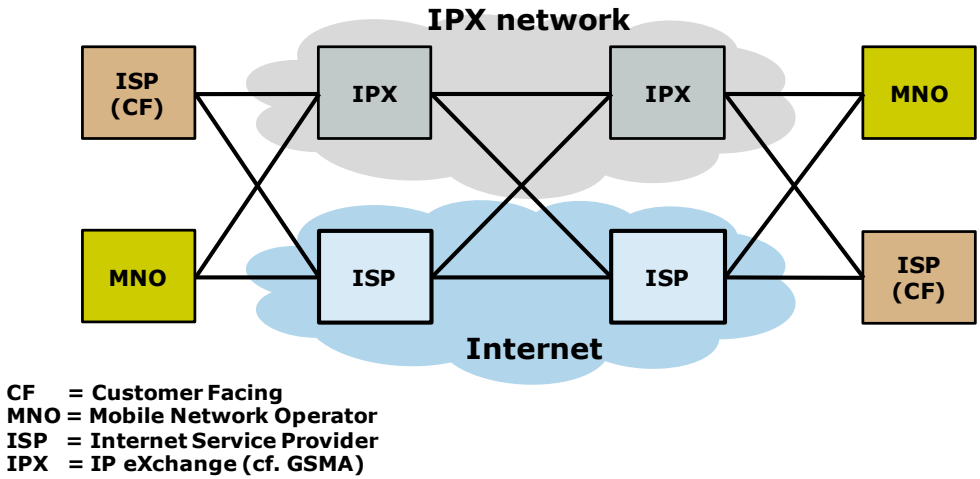


FIGURE 9: FUTURE POTENTIAL INTERCONNECT OPTIONS

- **Key revenue sharing and settlement approaches. A**s highlighted above, the revenue sharing approach is a key topic. The below identifies some main approaches or capabilities that are either in use or appear feasible. Note that some of these can be combined. More detailed approaches regarding settlement are not covered.
    - **Settlement free.** An Internet Peering agreement is settlement free. That is, the exchange of traffic is free of charge as there is a balance between the (two) partners, hence, no settlement is needed.
    - **Session and transaction based charging.** The typical approach in the POTS domain is to charge per session per minute. Variants of this approach can also be attractive for the future. Transaction based charging can be an attractive approach for instance for premium content services. These charging approaches may also allow the transit domains and terminating domains to take part in the revenue sharing as payments are typically cascaded throughout the value chain according to well-defined agreements.
    - **Traffic based charging**. IP transit service offered by an ISP to its customer ISP is typically based in traffic volume (whole traffic or 95th percentile model, where there is a discount due to the outbound traffic). Future traffic based charging may be more sophisticated taking into account volume per traffic class, quality measures, geographic region (source / destination areas). Note that Internet IP transit charging does not imply cascading of revenue sharing.

- **Revenue cascading approaches.** For further study.
- **Settlement approaches.** For further study.

4.3. A NOVEL INTERCONNECTION MARKET FUTURE ECOSYSTEM SCENARIO

In this subsection we present a novel market-based future ecosystem for the interconnection market. This ecosystem is one of the many possible one could envision and could be further studied by ETICS.

4.3.1. INTRODUCTION AND BASIC ISSUES

FIGURE 10: A TYPICAL MEDIEVAL EUROPEAN MARKET

Prior to defining the ecosystem, we briefly present some basic market terminology to facilitate the reader: A *market* is an institution where items and services, i.e. market goods, are traded. Buyers express their demand for products by means of *bids*, while sellers use *asks* to offer the market goods. In any kind of market, bids and asks are *advertised* and then *matched* by a suitable matching algorithm, thus facilitating trade. The exact definition of bids and asks and the value of their respective attributes, e.g. price, can vary from market to market and is determined by means of economics. Economics and competition also determine what kind of *market mechanism* is appropriate for the trading of a certain good.

4.3.2. MARKET AND ETICS ISSUES – POTENTIAL CHOICES

So what could a market for QoS-aware interconnection look like? An indicative yet not exhaustive list of potential market and market mechanism design options is:

- **Competitive market:** This is the case where the market is open to all, European ISPs and Over-the-top providers, intermediaries that dynamically bid in order to meet customer needs. The market matches demand and supply and clears the market, determining the price of the goods traded.
- **Commodity market:** This is the case where the market goods are simple and there is no uncertainty for their specifications. Interested parties just buy these goods from the market for a price that could be fixed by regulation but that will always depend on offer and demand.
- **Federation:** This is a version of a market where a cooperative infrastructure is built and subsequently uses to provide QoS-aware interconnection services. It is expected that in such an environment there are suitable policies and incentives for rational contribution and usage of resources, while there are also interesting priority and revenue sharing issues.

- **Hybrid:** A market that is neither completely cooperative nor competitive. One could envision a market where buyers and sellers of the interconnection goods (services) meet to negotiate over suitable contracts that match their needs.

Note that as in any market and economics environment, economics and competition determine which is the most viable and likely approach in a QoS-aware interconnection market. We intend to elaborate more on some of these issues in other ETICS deliverables, such as D3.1, while in this deliverable we focus only on the high-level future ecosystem issues that are independent of the choice of a certain market mechanism.

Having presented the main market issues, we briefly focus on the ETICS main ideas that envision QoS-aware interconnection contracts suitable to accommodate end-to-end sessions demanding QoS. One could expect an ontology of *interconnection contracts*, having the following features:

- Interconnection contracts are dynamic in the sense that these can be provided on-demand. Indeed, it suffices a new user session request in order to trigger the trading of an interconnection contract that could accommodate his needs in the market. It would be possible theoretically for the user to buy this contract directly from the market if it is offered as a simple market good. However, due to the complexity of economics institutions and markets in general, brokers usually undertake this task; there is an analogy here with the specialized stock market brokers that build their customers portfolios and perform trading according to their strategy. Thus, the traditional long-lived interconnection business agreements are complemented by **interconnection contracts** that are demanded on the fly and whose Quality of Service (QoS) attributes are suitable for supporting the features of QoS-sensitive applications. Clearly, these applications could be some of the QoS-sensitive applications presented in the various future service scenarios of this deliverable.
- The **time duration** of such interconnection contracts could vary from being large (timescale of weeks) to short as a few hours, as long as there is enough market demand for this type of contracts. That is, once there is sufficient user demand for QoS-aware applications and sessions that enforce these QoS requirements, it will be beneficial for the resource owners to provide such contracts for trading in the market. The same applies for long-lived inter-domain aggregates. Note that the former is conceptually similar to the different bearer services of the UMTS networks that are tailored to support the services of the Conversational, Streaming, Interactive and Background classes. Indeed, bearer services in UMTS are also of short time duration and characterised by a set of end-to-end characteristics with requirements on QoS that can be negotiated at service or connection establishment.
- These interconnection contracts can be traded as goods provided by means of a **market institution**, in a similar way that stocks are traded via today's stock markets, with the fundamental difference that contracts also have a time-dimension. This means that contracts are leased over a certain part of the network and for a given amount of time, as opposed to stock markets where ownership of goods is traded.

4.3.3. ECOSYSTEM DESCRIPTION

We now proceed to show how the aforementioned ETICS and market ideas blend by providing a vision of a novel market architecture that could dynamically serve the interconnection customer needs. First of all, in order to have a market, we need to define the *market good(s)*:

- **Simple:** This pertains to divisible network resources and possible sets/bundles of them. The *interconnection contract* can serve as the building block of such simple market goods.
- **Composite:** This is a composition and aggregation of multiple simple goods. This could pertain to sub-paths or end-to-end-paths or sets of paths with certain bandwidth and QoS features.

Market goods both in the ETICS and any other context are tailored to attract demand. If there is no actual demand for a good, these are not profitable and it does not make sense to have them in the market. So the ETICS market goods in this possible ecosystem are expected to be fine-tuned to accommodate services for QoS-sensitive applications; thus one can imagine an ontology of such goods whose attributes such as bandwidth and QoS attributes are targeted to certain types of applications and time-scales. It is important to stress that the goods will be determined to a large extend by the market itself. This is why it is crucial that in a market-based approach the market mechanism should be flexible and general enough to allow the market to express its needs, i.e. determine goods that are to be traded, offer and buy them. This way, buyers can just "browse and shop" and the market is expected to operate efficiently. This is to a large extend independent on whether the market will be a wholesale or retail market or mix of both where multiple business relationships and models can be envisioned on top of the ecosystem described here.

To facilitate the reader, we proceed to present a simple yet illustrative example, depicted as Figure 11. Let us assume that there are several Spanish companies, customers of TID that have branches in various French cities such as Lyon. These companies need for instance to run tele-presence applications with these branches. In order to cover this need, FT builds suitable interconnection contracts that provide the required bandwidth and QoS guarantees both in the interconnection link and inside FT's network. This market good is very attractive for TID, which purchases it in order to serve end-to-end with the demanded bandwidth and QoS the aforementioned Spanish companies.

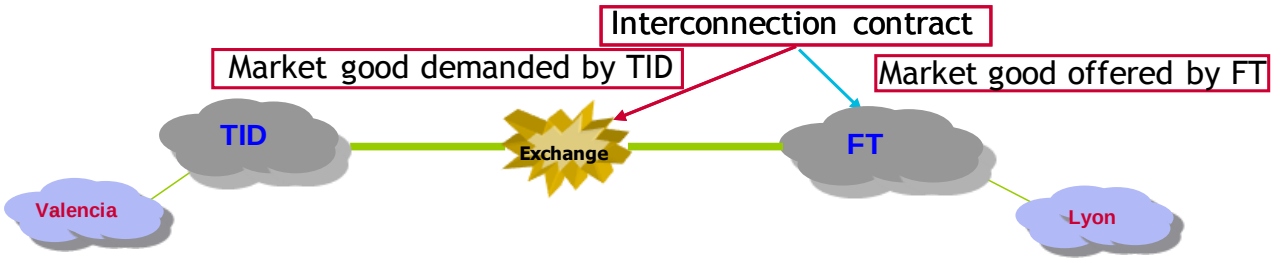


FIGURE 11: AN INTERCONNECTION CONTRACT BECOMES A SIMPLE MARKET GOOD

Note that one could envision such simple goods being combined by market participants to create composite goods for which there is demand in the market. Figure 12 depicts a case where such simple goods, similar to those of the previous example have been created by FT and DT over their interconnection links and respective networks. It is now possible for DT to buy such goods from FT and combine them with those it builds in order to build composite goods, i.e. multi-hop QoS-aware pipes that TID can utilize in order to serve users that need such contracts to interconnect e.g. Valencia and Aachen.

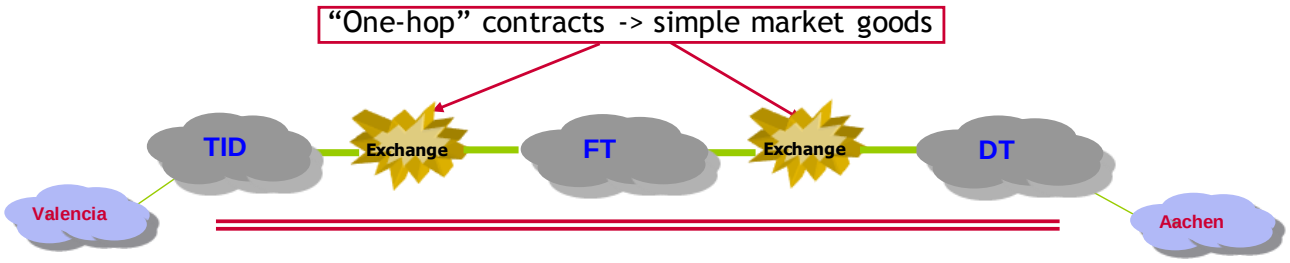


FIGURE 12: COMPLEX GOODS CAN BE OFFERED BY MEANS OF COMBINING SIMPLE MARKET GOODS

It is worth emphasizing that the aforementioned presentation and issues is so general that it is independent of the choice of a market mechanism, cooperative or not, the specific details of the market, the way market price is determined (fixed or discovered by means of e.g. an auction) or the involved TE solutions that are to be applied in the underlying networks. This comprises further evidence of the elegance, generality and strength of market-based approaches.

4.3.4. ACTORS

Next, we provide a general list of the key actors of this future ecosystem market-oriented scenario. Note that this is a general list of the possible actors one could envision for such a type of market-oriented future ecosystem scenario. The exact role and contribution of each actor can vary, according to the specific business model built on top of this ecosystem. A general discussion of such possible roles and business relationships is presented in the next subsection; a more detailed discussion is to be provided by means of the WP3 deliverables and is beyond the scope of this document.

The main key actors of this ecosystem scenario are:

- *Market Institutions - Exchanges:* The business entities hosting markets, in each of which interconnection contracts are traded. This is the meeting point of the actors where the trading of interconnection contracts by means of some market mechanism can take place. A close analogy to existing market institutions of the telecommunications sector is that of bandwidth or spectrum exchanges, such as SpecEx (www.specEx.com), where chunks of bandwidth and spectrum respectively are traded.
- *Brokers:* These are the market participants that purchase contracts through the market in order to utilize them to offer a brokering service or to provide a service to the end users who wish to run certain applications that require QoS. This is similar to the role of the stock market brokers that attempt to meet their customers' orders by means of bidding in a bid-and-ask market mechanism in the existing stock-markets. Thus, brokers apply possibly sophisticated bidding strategies (assuming a competitive market institution where e.g. a continuous double auction market mechanism is applied) in order to ensure that their clients sessions can be served.
- *Telecom providers:* These are essentially the owners of the market resources, i.e. the interconnection contracts that are built on top of the network infrastructure. Telecom providers contribute the supply of resources that are to be traded by means of the market. The existence of the telecom providers as key actor in this scenario is mandatory, since telecom providers own the network infrastructure that is required in order to provide interconnection contracts for end-to-end QoS.

- *Over-the-top providers:* These are service providers who sell to end-users certain services that require Quality of Service. For instance, this could be the case for a provider selling a voice service to end users, or high definition streaming video, tele-presence services etc. Over-the-top providers have been integrated in this scenario as key actors, due to the fact that it is possible for them to provide to their end customers the possibility of demanding sessions with QoS guarantees (e.g. high-definition TV), thus contributing to the demand supply of the market.
- *End-users:* The side of the market where demand ultimately arises. This consists of users that wish to ensure that certain services attain preferential treatment by the network, in order to meet certain QoS requirements. Eventually, networks are built and designed to serve people's needs, so end-users are mandatory key actors as well.
- *Contract advertiser/matching agent:* This is a business entity that informs interested parties on what interconnection contracts are currently offered in the market(s). Subsequently, this actor is responsible to match the demand of end users for QoS services to a set of interconnection contracts that could jointly be used to transport the user's service on an end-to-end basis and thus serve the user satisfactorily. Thus, this can be seen as a specialized repository and search engine that can efficiently match the demand provided as input with the existing supply respectively in the various market exchanges that can be utilized by the brokers in order to purchase the desired contracts of their customers. For example, this could be a central repository where all available demand and supply information for interconnection contracts is registered (i.e. 'contract advertiser') and is then searchable by means of a specialized search engine that can find the best match for a given interconnection need (i.e. 'matching agent'). The existence of a separate actor for this functional is possible but not mandatory; as also commented in the remainder of this section, this role can be absorbed by the market institution as well.

4.3.5. BUSINESS ROLES AND RELATIONSHIPS

The focal point of this future ecosystem scenario is that of the **market institution**. In practice, this role could be played by either the telecom providers or by independent entities such as Internet Exchanges. It is important that the market exchange ensures the transparency of the market, the reliability of the transactions and the policing of the market outcomes. Furthermore, it is worth emphasizing that in general multiple such market exchanges may co-exist, the same way that multiple stock market exchanges exist today. The nature of the market could also vary, depending on the amount of cooperation/competition assumed to exist among the various actors. In fact, it is possible for any telecom provider to become a market institution and in such a context a broker could combine contracts purchased in different markets.

For instance, one could assume a perfectly competitive market employing a standard market mechanism, such as the continuous double auction (with the necessary modifications) or some other customized market mechanism. An alternative would be that assuming a high amount of cooperation among the telecom operators, the market institution could diminish to being a focal point where all cooperative telecom providers make a certain amount of resources available to be used for their common interconnection needs. A hybrid situation would also include the possibility of (structured) negotiations, either bilateral or multilateral, taking place in the market among telecom providers in order to decide on

the dimensioning and the values of the respective QoS attributes of interest for various interconnection contracts.

It is also worth emphasizing that the existence of such markets does not exclude that bilateral negotiations for interconnection agreements among telecom providers also take place, bypassing the market. For instance, it could be the case that telecom providers can opt to cover their long-term needs by means of long-term interconnection agreements that are product of bilateral negotiations; they can subsequently use the market institutions in order to cover their demand spikes and needs for additional short-lived (in the time scale of hours or days) interconnection contracts that are possibly triggered by means of user session requests that require QoS in paths for which such long-term interconnection agreements do not currently exist.

Clearly, the exact specification of the interconnection contracts traded by the aforementioned alternative market institutions may vary as well.

We now focus on the role of the **brokers**. This can vary from being:

a) very simple by purchasing on demand from the market whenever there is a related request by an end customer to do so, to

b) sophisticated, where the broker purchases large wholesale aggregate contracts that will be subsequently used for the multiplexing of the expected end-user demand.

Brokers can be *independent business entities* that purchase such contracts over time from multiple telecom providers and sell them to end customers directly. Alternatively, *telecom providers* could also serve as brokers, e.g. by means of purchasing a short-term interconnection contract from a neighbouring (geographically) provider in order to build a QoS end-to-end path for the user. *Over-the-top providers* could also serve this role in order to directly purchase from the market the contracts needed to accommodate their *end-users* demand, bypassing other actors that could also play this role or make a strategic alliance with a telecom provider (see Figure 13 where such a case is depicted).

In a similar fashion, the *contract advertiser* could not necessarily be a separate entity but could be seen as part of the market institutions; in order for trading to take place, the market must advertise what is available.

The market architecture, the key actors, and the interactions among them are depicted in Figure 13:

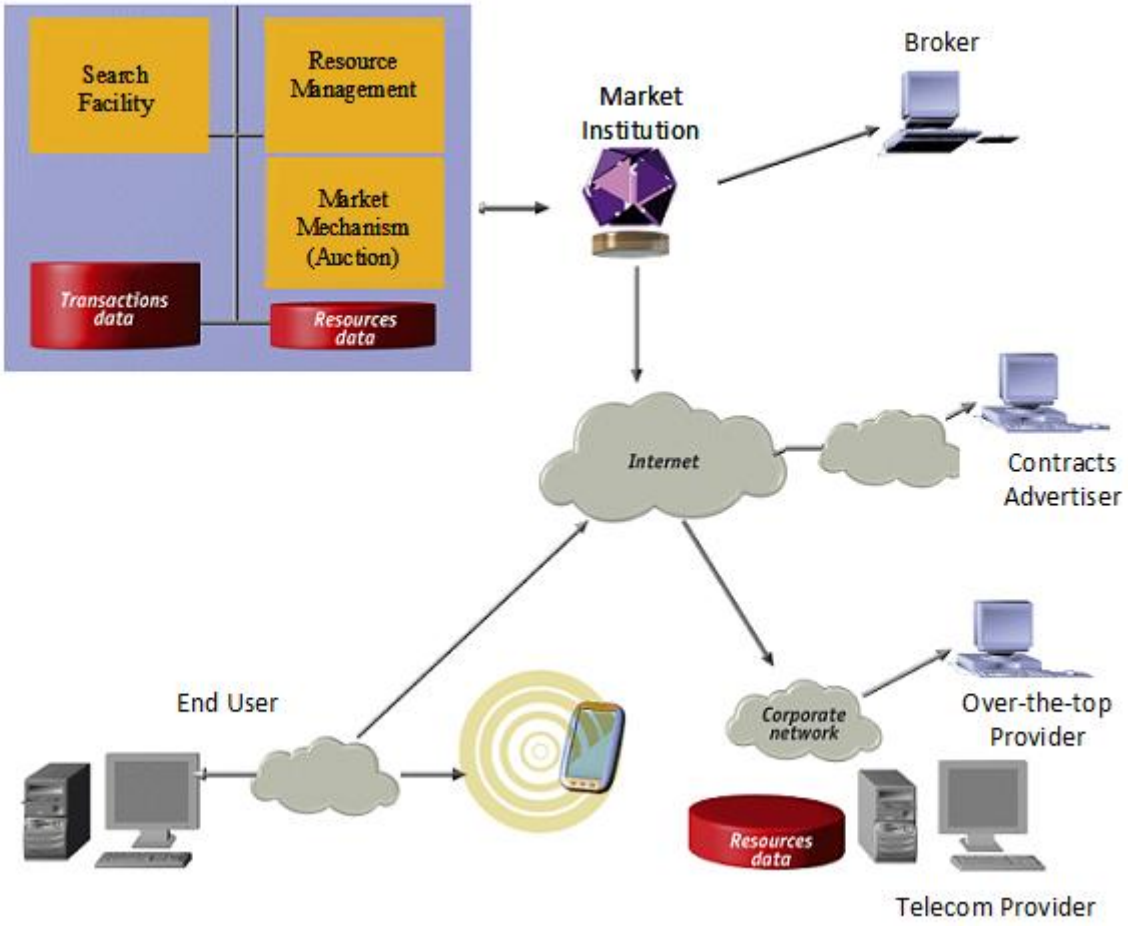


FIGURE 13: THE KEY ACTORS AND THE MARKET ARCHITECTURE

4.3.6. DISCUSSION

The use of a common market institution, i.e. a common trading platform (a possible architecture is depicted as Figure 13), brings substantial benefit to all market participants. The common contribution of resources of multiple telecom providers over the same market results in a larger market, which can cover a wide range of customer demand. Due to the fact that communication networks and markets in general have strong positive externalities, it can be reasonably expected that if a certain market reaches a critical mass threshold, then it will be viable and further grow. The multiplexing of the resources of multiple telecom providers is positive for the size of the market and is thus more likely to attract end user demand, as opposed to the case of bilateral leasing or one-provider markets.

End users benefit also from these externalities due to the fact that there is a healthy and viable market they can rely on that can indeed offer them the resources they actually need in order to meet their needs, and for the right time intervals. Also the fact that the market consists of multiple providers limits the threat of undesirable lock-in for the customers to a certain provider. Moreover, since the market is open, competition will be as high as possible, although there may be certain resources that are not widely provided. This is always to the benefit of the customers.

Furthermore, intermediaries that provide value-added services also benefit from the way interconnection contracts are traded. In particular, there is a business opportunity for intermediaries to act on behalf of the users and create by means of bidding service-aware network transport services bundles that can be sold to the end users.

Last, the telecom providers who own the various parts of the overall network infrastructure also benefit from the presence of these intermediaries since:

a) it is not required[23] for telecom providers to develop such services, which is not their core business and imposes additional costs for them,

b) market entry remains simple, since they just need to offer their interconnection contracts for sale, i.e. no entry barriers or exclusion effects,

c) due to the competition of the intermediaries, new efficient services are offered in the market which attracts more demand, thus extending the telecom sector market pie and resulting in higher revenue for themselves, and

d) spikes of demand or network failures can be accommodated on the fly by means of purchasing dynamically interconnection contracts from the market, instead of investing excessively in additional redundant network infrastructure.

The actual actions carried out by the brokers that heavily interact with the market, can vary depending on how competitive the market is assumed to be. For instance, in the case of a competitive market exchange, where an auction mechanism is used for the trading of contracts, the broker has to purchase the contracts desired by means of bidding in the auction hosted by the market institution. Alternatively, in case where contracts are decided by means of structured negotiations, the broker would be transformed to a negotiation agent.

4.3.7. SPECIFIC EXAMPLE

Having presented the key market actors and their business roles, we proceed to present an illustrative scenario that depicts the interaction of these entities in order to accommodate a high quality video real time streaming session that is demanded by a user. This specific example assumes a competitive market institution and is specified as follows:

- A decisive race of the Formula 1 championship will soon take place in Silverstone and is also broadcasted real-time over the Internet.
- Due to the fact that the race is not broadcasted by a TV channel, Jean who lives in Paris decides to subscribe to watch it over the Internet by means of an infotainment service provider, named *SportsAndNews*, which has purchased the right to broadcast it from FIA. Note that *SportsAndNews* is an Over-the-top provider that provides real-time streaming of the event on top of the Internet.
- Jean subscribes to the service of *SportsAndNews* in order to watch the race.

[23] But it does not prevent the telecom providers to do so, if they believe it is a strategic direction they should evolve towards.

- *SportsAndNews* has already purchased several interconnection contracts from a broker that assured *SportsAndNews* that it can accommodate its European customers with end-to-end QoS paths to watch the Formula 1 race. Note that in order for the broker to purchase these contracts, he has possibly interacted with multiple contract advertisers/matching agents in order to identify the desirable interconnection contracts that are available in the market.
- The broker has been purchasing these short-term interconnection contracts by multiple market exchanges from various European telecom providers, such as BT, DT, FT and Telefonica for the time interval that the racing is broadcasted in order to accommodate the demand of both the end-users of SportsAndNews and its additional customers (other over-the-top providers).
- In order to serve Jean with the demanded QoS, *SportsAndNews* demands from the broker to utilize a connection from Paris to Silverstone. Alternatively, if SportsAndNews has built a multicast VPN by means of aggregating the various interconnection contracts it has purchased from the market, he adds Jean to the multicast group of receivers and step 7 is omitted.
- The broker assembles an end-to-end QoS aware path by combining interconnection contracts it has purchased from FT and BT.
- The path is delivered by means of a suitable interface to Jean as soon as he utilizes the service session of *SportsAndNews*.

4.4. Cooperating access network providers

This scenario illustrates how two or more network access and Internet service providers cooperate in order to use the access network of the cooperating ISPs to connect their own customers to the Internet. Most ISPs today are only using their own access network infrastructure in order to connect their Internet service customers to their backbones and give them connectivity to the whole Internet. But this approach could lead to very high expenditures if the ISP wants to extend its coverage to all areas (e.g. remote small town, rural area, etc).

Instead of deploying and maintaining a new access segment (new cooper lines, new fibres, local access centres, etc.), it could be more appropriate for the ISP to use the network infrastructure of other access service providers in this area and open their own access infrastructure on the basis of a cooperation contract to the cooperating access provider.

This approach has been also imposed by regulation, which demanded to open the access network infrastructure of the incumbent players to entrant ISPs.

This cooperating access network provider scenario does not only offer some bit pipe services but should also support higher layer services like QoS guarantees. For this reason, it is assumed that the inter-provider interface acts on the IP layer[24]. Figure 14 illustrates the sketched cooperating access network providers approach and highlights that inter-provider interface.

[24] A pretty equal approach is already realized today at Layer 2 by extending the PPP(oE) sessions of e.g. DSL customers via L2TP to the BRAS of the foreign ISP. But this simpler approach is of course not feasible for instance to support inherent Multicast capabilities of the already deployed access network infrastructure

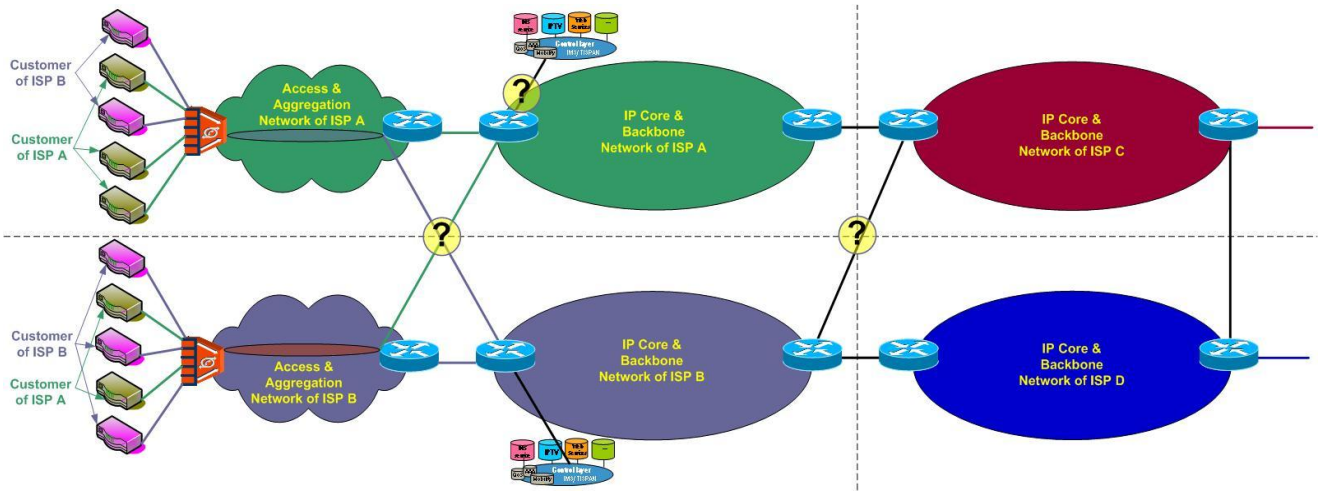


FIGURE 14: COOPERATING NETWORK ACCESS PROVIDER

The benefits for the involved stakeholders of the described cooperating access network provider service are the following:

a) Access Service Providers could get new incomes due to the high utilization of its infrastructure. The access network of the access service provider could reach a higher utilization rate and more attached customers.

b) Network providers (mainly Internet Service Providers) can reach more customers and save costs. An ISP does not need to extend their own access network infrastructure and can rely on the service functionality of the cooperating access network provider.

c) End customers could get more and value added services. Those users attached to the access network of the cooperating access service provider get the possibility to become attached to Internet service providers that provide higher value services than their original service providers.

Due to the above described architecture there exists the need to investigate and define several frameworks and interfaces in order to deliver a homogeneous end-to-end service within this network cooperation. Examples of some open issues are:

a. Which are the contractual and business relationships between cooperating providers? How they are built? E.g. if this scenarios is forced by regulation, the regulatory body has to monitor the cost models and the wholesale service price is regulated.

b. Ho to build the management interface for those customers of the ISP A (access service provider) that are connected to the access network of ISP B (network provider)?

c. The provisioning of customer configuration without the assumption of dedicated layer 2 access lines (e.g. VLANs, PPP).

d. Authentication, Authorization and Accounting. Customers of ISP B is connected via the access network of ISP A will be authorised and authenticated by ISP B, since ISP A has to be informed that the customer is allowed to be attached to the access network. Besides that it may be necessary to notify to the ISP A the IP address and prefixes assigned to the

customer in order to assure that no wrong IP address may be used by the customer (IP spoofing prevention). Possible solution approaches could be e.g: DHCP snooping or OoB signalling of the authorized address ranges of the customer. Furthermore, it might be also needed to account the traffic the customer sends and that has to be fixed as well in the agreement between ISP A and B.

e. QoS signalling and handling. QoS signalling between ISP A and B has to be unified or, at least a common understanding must be identified. This is especially relevant when, e.g. The home-gateway is not able to mark packets properly. E.g. for VoIP traffic, the access node of ISP A has to insert the right DSCP bitmap for real-time services; and the corresponding real-time service handling inside the access network of ISP A has to be "compatible" with the QoS handling for the QoS service that ISP B offers to its customer.

f. Mapping of different QoS classes between carriers. An homogeneous E2E QoS behaviour and a rule-set for mapping is required. E.g. How the DSCP code points might be mapped at the SP borders? Perhaps ISP A offers only a group of real-time transport capabilities (associated to a DSCP value) and, on the other hand, ISP B is configuring a different DSCP value for VoIP service. Therefore, it is important to have a clear mapping between the different types of services.

4.5. Considering bootstrapping and roadmap options

While section 2.2 clearly elaborates on the Internet ASQ deadlock situation and associated stumbling blocks this section briefly considers potential ways of exiting this deadlock. Section 4.2 clearly shows that a future interconnect strategy that perfectly enable and support any foreseeable service must take into account a number of factors and this is a vast and complex task. ETICS believes that for the near term it is vice to avoid striving for the perfect all-encompassing approach and as a result get into a situation where “The perfect is the enemy of the good”[25]. Rather, the question that is asked here is **– Can we find one or a few services that should be in focus as a first bootstrapping case for offering ASQ? Subsequently, later work will address the question: How should one develop a simple, feasible and sufficient interconnect approach for this service?**

Such a bootstrapping case should get significant focus by the ETICS project in an early phase. This can, considering future stages, give a good setting for extending and enhancing an interconnect strategy with additional capabilities, as parts of the solution costs are more or less covered and it may be easer to develop a positive business case given this new setting. However, care must be taken in order to analyse whether the selected bootstrapping case and its solutions have any negative potential impacts on the future option space or business opportunities.

One potential bootstrapping case may be driven by the premium video conferencing and/or tele-presence services offered to broadband and/or business customers beyond just intra-business VPN based

[25] Voltaire quote, http://www.famous-quotes.net/Quote.aspx?The_perfect_is_the_enemy_of_the_good

conferencing. It appears that there is a clear and increasing demand for such a service[26] for instance driven by climate change issues, travel cost reduction demands, and volcano ash cloud problems for the aviation sector. For many customers this service is perceived as offering high value, and hence, there should be willingness to pay for such a premium services. It is a session oriented service, which typically is best supported by some means of session and admission control end-to-end. The session is triggered by human action and can potentially take benefit from information about when it is scheduled by the user.

Moreover, beyond such a bootstrapping case the ETICS project may want to elaborate on and describe a provisional roadmap in order to indicate a rough order among different new services with respect to their timing. This can be done in according to the following categories.

- Services that have attractive properties today (should be address firstly)
- Services that are expected soon to have attractive properties (should be address next)
- Services that can wait but are expected to be attractive in not too long term (medium term)
- Services that are assumed to become significant only in a longer time horizon (long term)

As starting point, the following section identifies a set of scenarios where different services are identified considering the actors and the deployment timeframe.

[26] Note that is it is difficult to predict the demand for this service as this service as such is not in general available to the market.

5. FUTURE SERVICE ORIENTED SCENARIOS

As stated in the previous section, it is important to identify all those services and their associated requirements that are needed today and in the short term in order to define the bootstrapping scenarios. In this section, the ETICS partners have identified a set of scenarios, where different services are analyzed considering a methodology that is briefly described in section 5.1.

It should be noticed that these scenarios do not represent the scenarios that will be developed in ETICS testbeds, since they imply the usage of e.g. some application providers that are not available in the consortium. But they constitute by themselves a starting point to discuss the demonstration scenarios that will be implemented as part of WP7 work.

In order to identify all those services, we have considered different perspectives:

- First of all, it is important to identify which are the services that are expected from the end users' point of view. All these services are described in section 5.2
- Then, we identify the new services that can be provided to corporate and SME customers and users as starting point. These services are described in section 5.3.
- Finally, in section 5.4, the evolution of the wholesale services is inferred considering the new services and the new transport capabilities.

It must be noticed, that a service could be analysed from the three different perspectives, since, e.g. a tele-presence service could be initially a new service for corporate users in the short term, that requires an inter-domain deployment in the medium term and that finally could be available for residential customers.

5.1. FUTURE SCENARIOS CAPTURING METHODOLOGY

As part of the work reported in this deliverable, a methodology for capturing and analyze the future scenarios has been proposed and applied. The methodology mainly consists in a template defined for gathering all the most important aspects of each scenario in a homogeneous way and trying to duplicate information. The key aspects to be covered by each scenario are: Stakeholders, Services, Resources Mobilization and Barriers.

Following image depicts in detail the different issues to be covered by the template:

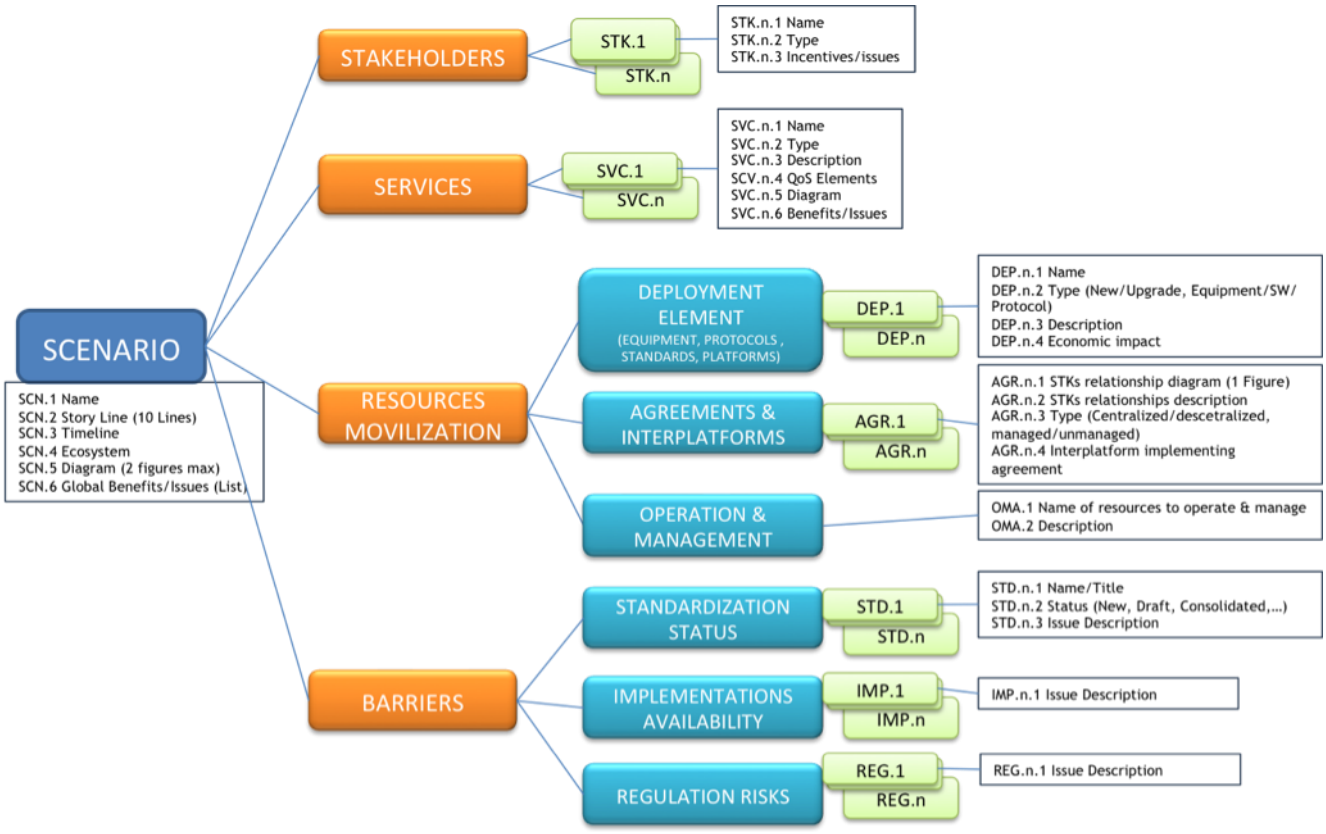


FIGURE 15: SCENARIO TEMPLATE OVERVIEW

From the scenarios information compiled in Annex C, there is proposed a scenario organization based on the target of each scenario:

a) End-user/Residential user

b) Business/Enterprise/Corporate customers and users

c) Inter-provider services and relationships

5.2. ADVANCED CONNECTIVITY AND SERVICES FOR END USERS: HOW END USERS CAN SEE THEIR FUTURE SERVICES.

From the operators' point of view, the end customers only pay for the connectivity in the Internet environment and, moreover, there are some additional services that are also offered to the end users. Considering this starting point all these scenarios aim to exploit new dimensions of the connectivity offer that could take advantage of new network capabilities as a way to obtain new sources of revenue revenue for telecom providers and as a way to get better service experience for the customers.

Among all these connectivity dimensions, several attributes should be considered, such as, e.g. to provide real time guarantees. We have made the following classification:

- Services that require real time characteristics for the end users that are presented as an evolution of current services considering different perspectives.

- Services that take advantage of Cloud Computing in order to provide advanced services to the end users.

5.2.1. Real Time Communications for End Users

This set of scenarios aim to identify the evolution of the services that are used in today's network where, considering an evolutionary approach, we have identified the set of scenarios and their associated requirements for their deployment.

5.2.1.1. Premium Real-Time Unified Communications (PRUC)

This set of services includes Triple Play Services such as VoIP (fixed and mobile), IPTV and Basic Internet Access)[27]. Residential users are demanding Triple Play services. This is not a new feature for the user, but represents a trend and a new way of delivering and offer services to the customer. Nowadays, Triple Play services are focused on residential fixed users, but a new whole mobile market demands those kinds of applications such as, **Mobile Broadband communications and Mobile IPTV services**.

Voice services

VoIP is hoped to take over the mobile calls world. Already VoIP can be used through thin mobile clients, like Skype and special Skype data packets are offered by telecoms such as 3 in the UK. Although it did not seem possible a few years ago, with the 3G coverage becoming more reliable and covering most major cities, it is hoped that mobile telephone service can be fully IP based. Mobile VoIP will possibly become more of a reality with the appearance of the 4G fully IP based mobile network with QoS capabilities. The result of having VoIP both on our mobile phone and our domestic telephone line will possibly mean that we no longer have to distinguish the two since inside our house we will be using VoIP over our wireless network and outdoors through the mobile networks, whatever that might be at the time.

IPTV services

Possibly the next greatest step for IPTV is its domination in the mobile world. With television program broadcasters already offering such applications, the mobile IPTV future is already looking bright. Although these services are available they do not offer the interactivity of IPTV. If telecoms providers can establish IPTV dominance in the household they can perhaps start moving in creating a similar platform on the mobile implementing the business models discussed above. Work is already underway to try to create an appropriate mobile platform that will be able to display IPTV via different networks, including next generation mobile networks, so that the consumer can watch non-stop while switching between networks. Again the QoS requirements for the IPTV network are as important as for VoIP and the bandwidth demand much higher. An important attribute associated to this service is the flexibility, the service should be flexible enough in order to allow its support in wide range of devices (smartphones, iLike devices, laptops, PDAs, etc.) and the adaptation of the service to the device and network conditions (e.g. if IPTV service is offered, the service should be aware of the type of access, HSPA or GPRS, and the load conditions in the cell).

[27] In the short term, tele-presence services are considered in the services to be offered to the corporate users. So this service is analysed in that section.

- Therefore, the deployment of service delivery **with assured level of quality of service is expected to become increasingly important**, making the Mobile Network Operator (MNO) go beyond pure flat rate date connectivity access services.

Therefore, even though in today's mobile networks, different levels of QoS could be provided, no strict guarantees can be applied.

Another important dimension of the evolution of the PRUC services is their availability everywhere and every time. Therefore, it is expected that roaming users would like to benefit from mobile broadband services anywhere, anytime and from any devices, again demanding information and certainty on delivered quality and pricing. Specific cases involving multiple network service providers (NSP) should be explored, addressing services that typically will benefit from various solutions and capabilities enabling assured service quality. Examples of such services are VPN access to the corporate network, digital photo (large file) upload, premium content streaming on-demand, visitor (tourist) streaming of content from his home country. All these cases involving seamless roaming onto non-3GPP networks (e.g. WiFi HotSpot); local breakout, as well as "local" access to Internet based services while roaming; and cases involving flexible content caching across MNOs can be considered in particular to accommodate roaming customers.

Once, this scenario is briefly described, which are the main challenges to implement it?

Mobile broadband is a large and increasingly important market. Typically, the network resources are limited and assured service quality end-to-end is important. It is important for the Mobile Network Operator (MNO) to offer services that can result in return on their investments. Mobile broadband-based services enable the user to accesses services anywhere, anytime and from any device. **The more mature mobile markets are focusing on improving and expanding their mobile broadband-based service offerings, while emerging markets are likely to adopt mobile broadband offerings as mobile access to Internet is typically the only option to Internet access in many areas**. In this context, in order to define the service offer for the future, it is important to clearly understand the current status of the QoS mechanisms in today's networks:

- Non-real time classes of services (Background and Interactive): define different levels of priority but do not provide strict guarantees. They are available in today's networks, where moreover, associated to the Interactive class of services, different priorities could be also specified.
- Real time classes of services (Streaming and Conversational): which are specified not only as different priority levels but they are also expected to provide strict guarantees. Therefore, there should admission control procedures associated to the provisioning of these classes of services. Even though these classes of services are defined in the standards, they are still not commercially available in the radio access segment, where the explicit QoS policies must be applied. It is expected that these real time classes of services will be available when the LTE (Long Term Evolution) mobile networks are available and voice service must be provided over IP.

On the other hand, interworking issues are particularly challenging in mobile broadband as there are many types of access network technologies as well as services session signalling and control solutions in operations today. Although 3GPP make a considerable standardization effort to minimize interworking and

compatibility issues these do represent challenges. Roaming strategies are complex as multi-provider and B2B issues are inherently complex and involve technical and business challenges at multiple levels.

This scenario should explore feasibility of and difference between an IPX based approach and an ISP based approaches or combinations of the two.

Although there is a drive to shift from access services to delivery of specific services and monetization thereof, there will still be a significant amount of non-classified traffic, traffic that can be categorized as belonging to “Basic Internet Access”. This scenario should explore an “Evolved Basic Internet Access” approach where this non-classified traffic is considered, and where fairness of resource usage (e.g. in a 5 minute interval) end-to-end (local, regional, continental, inter-continental,) is a driver while respecting non-discriminatory and transparency requirements.

Other issues are charging & billing, SLA and service assurance end-to-end, sharing and conveyance of policy and subscriber or user data, considering both the convergence of fixed-mobile as well as roaming cases.

In this kind of scenarios, the actors that are involved are Home MNO, Visiting MNO, MVNO, ISP (several roles), converged Communication Service Provider (CSP, fixed and mobile services), Content Provider, and possibly a CDN provider.

The main requirements for the implementation of the evolution of these services will be related firstly to the efficient provision of guarantees in the mobile access networks and the capability to enjoy these services in roaming scenarios.

5.2.1.2. Real Time Social Networking

In this scenario, we consider that Social Networks could become the access portal to the different Internet services as a way to enjoy your social life in Internet. Therefore, in order to really have a social life in Internet, it will be desirable to have real time communications in this environment.

Therefore, this scenario assumes that the social network platform will have access to the capabilities available in the network. The idea would be to incorporate other real time interactions (e.g. on-line gaming, Tele-presence, etc.) and services to those social networking allowing operators to get benefits too.

For the implementation of this scenario, the interaction among network providers and OTTs (such as Facebook and Tuenti) is required. The network provider can open some capabilities to third parties that could offer a service to the end users. It should be desirable that the end users could have a single agreement with the operator and a revenue sharing model between OTTs and network operators should be defined.

In order to make possible this scenario, and considering the growing usage of mobile technologies, a starting point will be the development of the PRUC services described before. Then, APIs for 3rd parties are required that, in fact, will have an important impact on the wholesale services as it will be described in section 5.4.

5.2.1.3. Remote Access Presentation

UPnP Forum[28], the ETSI TISPAN[29] and Home Gateway Initiative (HGI)[30] have started to work on the definition and specification of a Remote Access service. UPnP Forum has already published the first version of UPnP Remote Access which allows a remote user to access to a remote home through a VPN connection. Moreover, these standardization fora would develop enhance version of simple Remote Access service by allowing Home 2 Home connection, Web 2 Home, Mobile 2 Home and so on, with support and guarantee for Quality of Service (QoS). Behind the Remote Access service, two main scenario usages are already envisaged. The first one consists for customer to share their own multimedia content with their family or friends. The second concerns the Home Automation and the possibility for a customer to remotely access to its house to control heat, light, webcam, door, etc.

As ongoing work, several architectures have been proposed based mostly on VPN, IMS and HTTP web server. Even if the QoS is well controlled in the Home environment (through DLNA QoS, UPnP QoS or AVB techniques), the proposed solution suffer from a lack of QoS support in the WAN part. If in single and mono-technology some solution are envisaged (IMS RACS, DiffServ …), the **multi-domain and multi-technology are completely are left without any valuable solution**. So, Remote Access service through a variety of carrier, for people living in different country, needs a well controlled inter-carrier connection with guarantee of QoS. The key challenge for ETICS is not only to provide solution for interconnectivity but also to propose a QoS SLA service for Remote Access scenarios. Again two level of QoS guarantees are required: i) large amount of bandwidth for Multimedia Content Sharing and ii) less jitter, less delay and no loss QoS for Home Automation.

Home Automation

The service consists to remotely access to its house in a nomadic situation (from work, vacancy, business trip …) in order to access to Home Automation of the house. The actions to control the house are not bandwidth consuming (except for Live Webcam) **but required real-time connection with jitter and delay as low as possible and of course without any loss to avoid unsolicited control or action**. If the user is connected to the same network as its house, inter-carrier is not needed. But, as soon as there is more than one network operator, inter-carrier QoS negotiation and setup is necessary to ensure at high level of QoE to the user. The solution resides more in the transport of small amount of data but with a high level of priority and reliability, which is, in fact, a real challenge for mobile networks. This could be comparable to the protocol used to control the network (like routing protocol, monitoring, device control …) and as the same needs: reliability, low bandwidth requirements and real time guarantees.

Multimedia Content Sharing

The service consists of sharing multimedia content between users whatever they are located (at home, at work, from a mobile, form a cyber coffee …). Such exchanges require large amount of bandwidth to transport multimedia contents like Photo and Video and lesser for Audio stream. As the transport of data is mostly done through HTTP streaming, real-time connection is not of high priority. Data are first buffered for a few seconds before being displayed. The remote access to multimedia content allows customer to share

[28] http://www.upnp.org/
[29] http://www.etsi.org/tispan/
[30] http://www.homegatewayinitiative.org/

contents between friends or family like social tools and social networks. Thus, users are almost always connected to different network operator.

A key challenge for such service is the possibility to offer a high level of Quality of Experience to the users. Inter-carrier connection management is a key factor to provide such SLA with end-to-end QoS guarantee.

5.2.2. EXPLOITING THE CLOUD COMPUTING PARADIGM

The next set of scenarios aims to show the new services that can be provided to the end users by taking advantage of the cloud computing capabilities. Two main scenarios are considered: Gaming as a Service (GaaS) and remote virtual drive.

As it will be explained in the scenarios, the main requirements derived from these use cases are the real time capabilities and the need for guaranteed bandwidth not only from the users' side but also important throughout are expected as part of the backend services that are needed for the right maintenance of the services provided to the end users.

5.2.2.1. Gaming as a Service (GaaS)

New online games show increasing needs concerning Quality of Service. Racing and multimedia role playing games need respectively **latencies no more than 60 and 200 ms** in order to guarantee a good gaming experience. Companies (like Live Gaming) are now acting as brokers, providing to online game developer, store front, monetization and user management.

In a future possible scenario (namely Gaming as a Service), Medium and Small software developer deliver their games under SaaS (Software as a Service) approach. We can imagine that a third party actor, called E-Gaming broker, will select the best SaaS games. On the one hand, the broker provides top rate quality games and payment services to the online game users and, on the other hand, access to the market, and user management and collection services to the game software developers.

Furthermore, the E-Gaming broker, aware of specific QoS needs of the different online games, can exploit the new QoS enabled interconnection services potential in order to ensure specific end-to-end QoS to the online gamers, depending on the game they are playing, achieving economic sustainability by adapting the QoS provided to the different end-to-end services provided.

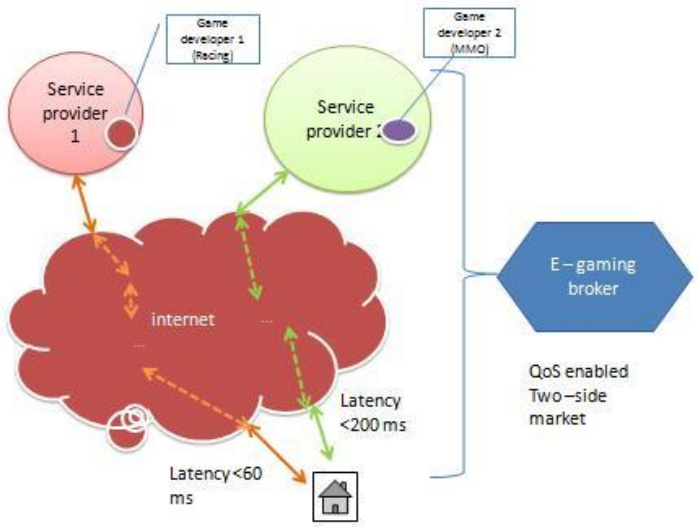


FIGURE 16: GAMING AS A SERVICE SCENARIO: ARCHITECTURE LATENCY

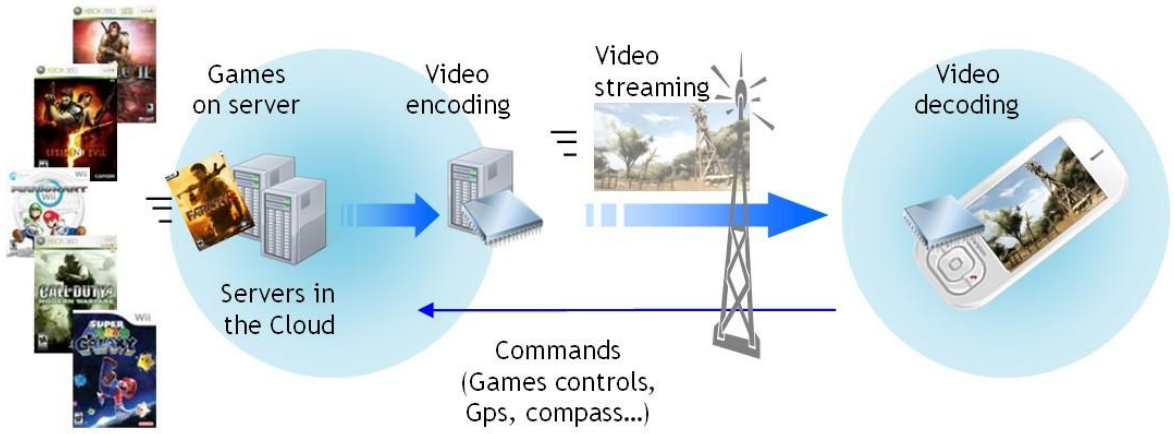


FIGURE 17: GAMING AS A SERVICE SCENARIO: PHASES OF THE SERVICE

Another approach about Games is about their more and more IT resource consumption, requiring end-users to change or update their devices (PC and Consoles) frequently. In the frame of Cloud Computing, the Software as a Service layer proposes the remote execution of software to mitigate the end-user device processing: a video stream is sent to the end-users and another flow captures the end-user interactions with the video.

Hence, gaming could be proposed on the cloud to end-users, across the Internet thus involving the following stakeholders: a Game provider, a Cloud provider, and Several Network Providers. Because of the real-time interactions, latency and jitter are critical in this scenario. As a consequence, network agreements should include requirements on these QoS parameters. As result of these agreements, game providers, network providers and cloud providers will obtain a new model of revenue.

Regarding the ecosystem integration, this service needs QoS guaranteed to offer a good user experience. Aspects such as end-to-end delay will be a key parameter in real-time applications.

Other possibilities in the GaaS scenario – considering the new interactions modes

The growing user expectations towards interaction performance and precision are strongly driven by the popularity of direct touch screen manipulations on the iPhone, natural arm gestures with the Wii, or even full body movements with the currently launched Microsoft Natal project. Especially with the trend towards distributed multiplayer and multi-device gaming, direct multimodal interaction will impose strong requirements on latency and accuracy.

These new ubiquitous interaction methods will be attractive for playful forms of learning and collaborative work, as well as for pure entertainment. The following two scenarios cover these two forms of gaming:

Playful information-sharing: People will be enabled to continuously exchange and manipulate photos and videos of their current environment via their mobile devices, such as panoramic views during tourist trips being shared live with friends and family. Devices based on gesture phone (such as the FTW's development whose details are available at [Bal09]) can be used to view this shared content projected on a nearby wall, and to use direct hand gestures to manipulate this content, which can then be communicated to the remote partner in real-time. Being based on portable devices, this form of group-based communication will not be restricted to classical home settings. At the same time, the combined use of projector-based presentation and gesture recognition will enable direct interaction with content within small collocated groups. The main requirements for the network will be bandwidth for media transmission, as well as accurate and synchronized communication of gestural behaviour.

Networked reaction games: We witness a growing popularity of action and reaction games (staring with Nintendo's Wii, but now increasingly on the iPhone), such as controlling virtual tennis or football players with finger and arm movements. These games are increasingly played in groups, and it can be well assumed that playing these games with remote friends were attractive as well. Networked reaction games take this next step. The basic idea is that everybody can come with a handheld, register to a nearby computer and a gaming platform and play games with others such as tennis or soccer, using the mobile phone as a sensor. The planned showcase application is a table tennis game, where the players' handhelds serve as rackets, and a computer screen shows the other player, a virtual table tennis table, and the moving ball. Compared to previous approaches such as [Mul07] or [Wood04], this setup would have some market potential due to the increasing number of sensor-enabled handhelds, and furthermore the setup would only require laptop, handheld, and internet connection.

5.2.2.2. Virtual Drive

The service consists of a virtual data storage space offered to users either with a web interface (e.g. Gdrive like) or with other means, more file browsing oriented (e.g. SCP, SFTP, etc.). The core of the service complexity is on the middleware that manages the actual implementation of the virtual drive space over multiple, geographically distributed physical data storages. This service will be deployed over the current Internet.

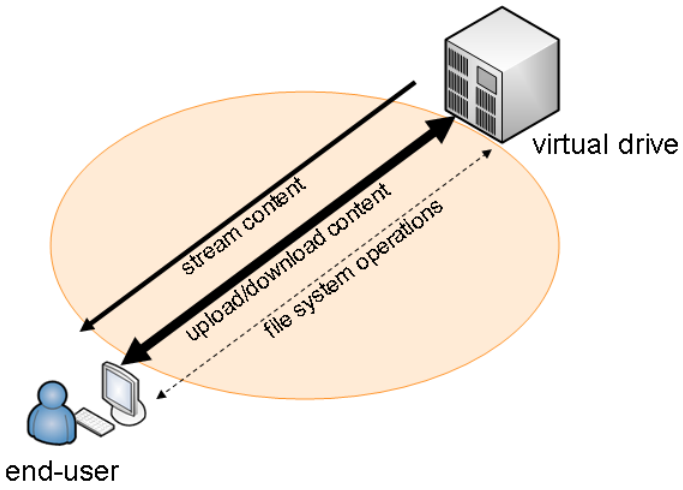


FIGURE 18: VIRTUAL DRIVE SCENARIO

However, the implications on the network services are deep and challenging. We propose to group them into two sets; the first set of requirements is to serve the end-users accessing the service with the proper QoS: high E2E (i.e. inter-carrier) bandwidth (e.g. PCI or USB speeds, ~ 400 Mb/s), very low e2e delay, limited jitter. The second set of requirements is to support the service backbone adequately: i.e. storage network management and maintenance; e.g. moving data among different data storage locations. The QoS requirements in this latter case might still affect an e2e inter-carrier chain (e.g. data storages spanning locations served by different operators), and consist of potentially huge bandwidth and reasonably limited delay.

From this general view of the scenario, the concrete services proposed are the following:

1. **Virtual disk usage by customers (file browsing operations)**. This usage constitutes the end users' dimension of the service and basically, it consists on a File browsing operation on the virtual disk.

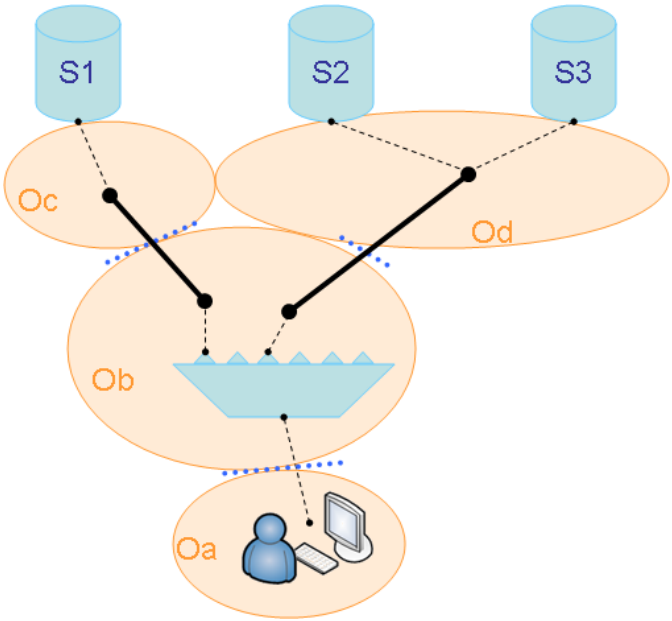


FIGURE 19: VIRTUAL DISK USAGE BY END-USERS

The virtual disk service is offered as a subset of basic file system operations: create, delete, rename, move, copy and access files. Permission management is optional.

The file browsing can be offered via web interface, SCP or SFTP tools.

File operations that require file transfers (e.g. uploading or downloading files, i.e. accessing/moving/copying files to/from the remote virtual drive) should be performed with **guaranteed bandwidth, possibly establishing the needed resources on-demand along the path** (in case this is required by an optimal and economic implementation of the network transport service).

Depending on the access type, the QoS requirements will vary:

- Moving/copying files from/to the virtual disk implies high-speed transmission rates (in order of 100s Mb/s) and the lack of losses and packets dropped. Also other aspects such as delay (up to 100/200ms) and Jitter will be considered but not critic.
- Accessing files, media files (audio/video, not images/pictures). In this case, rates are in the order of 10s Mb/s for video files or even less, depending on the codec; rates in the order of hundreds of Kbps for audio both with limited loose tolerance. E2E delay and jitter will be reduced and limited by the buffer of the media player.

The benefits of this service coincide with the benefits for the end-user of the whole scenario. Main issues for the ISPs or backbone operators are the control of BoD requests from the customers where the resources are dynamically allocated; issues are: control of fine-grain QoS, per-user/per-flow accounting of BoD and per-user/per-flow monitoring and recovery of QoS

2. **Backend services** (e.g. back-upping data, defragmenting among storages, etc.): these services are required to maintain the service and assure the availability of the end users' services. Moreover, the data can be moved to sites near the end users in order to minimize the delay.

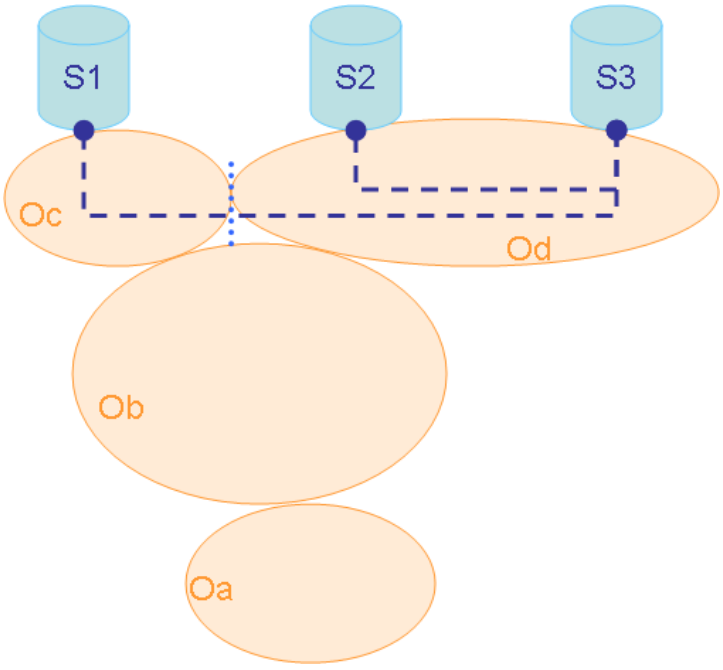


FIGURE 20: BACKEND SERVICES

As part of backend operations performed by the storage broker or the storage facility provider, **large data transfer sessions are expected**. These are mainly aimed to move users' chunks of data among storage points for various purposes: load rebalancing of physical disks, backups and restoration of lost data, etc.

These transfers can be either planned on a regular basis, or just in time based on the needs and triggers from the virtual storage service plane. The former case can be handled by scheduled or permanent connections (depending on the frequency) while the latter with on-demand connections in case permanent connections are not in place.

The quality of service requirements will mainly consist on **huge bandwidth consumption**. Actual values depend on the maximum time allowed by each storage operation in the service plane workflow (e.g. depending on the critical level of the operation).

The benefits of this service, based on when on-demand or scheduled connections are used, are the availability of big bandwidth pipes just when needed, without the need to pay long-term rentals and waste money during inactivity times.

Main issues for the backbone operators are those ones related with the control of requests for BoD and scheduling connections: control of fine-grain QoS, control of scheduled resources, per-user and per-flow accounting of BoD and per-user and per-flow monitoring and recovery of QoS

The stakeholders involved in this scenario are the following:

a) Final customers (storage end-users). Their incentives will be the availability of flexibility priced storage as much as needed. Robust Storage, since services like RAIS and backups are embedded and the lack of home storage infrastructure.

b) Storage facility providers. For them, the most important incentive will be the possibility to optimize its available storage facilities by renting out free capacity

c) Storage Brokers. Their incentives will be the possibility to run a service with basically no storage infrastructure. To obtain a multiplier effect from a very large customer base and finally to obtain revenues from a pay-per-use scheme, possibly with a free initial amount of storage (to attract customers) and a flat rate intermediate amount.

d) Network providers, both ISPs and Telcos. The key incentive is possibility to support services that need large bandwidth pipes, sharing the revenues of the service brokers and storage facility providers.

The relationships among them are the following:

- Storage broker – storage facility provider: the storage broker establishes a contract with one or more storage facility providers to buy disks or disk quotas to be used in the service.
- Storage broker – network operator/ISP: the storage broker buys connectivity to the rented storage facilities, and establishes an SLA for on-demand and scheduled transport services.
- Network operator/ISP - network operator/ISP: there will be inter-carrier contracts to support permanent or dynamic transport connections for both backend and end-user traffic.

- End-user - network operator/ISP: the end-user needs to extend its SLA with its ISP in order to be granted on-demand increased QoS for virtual drive services.

The following picture illustrates the relationship described previously:

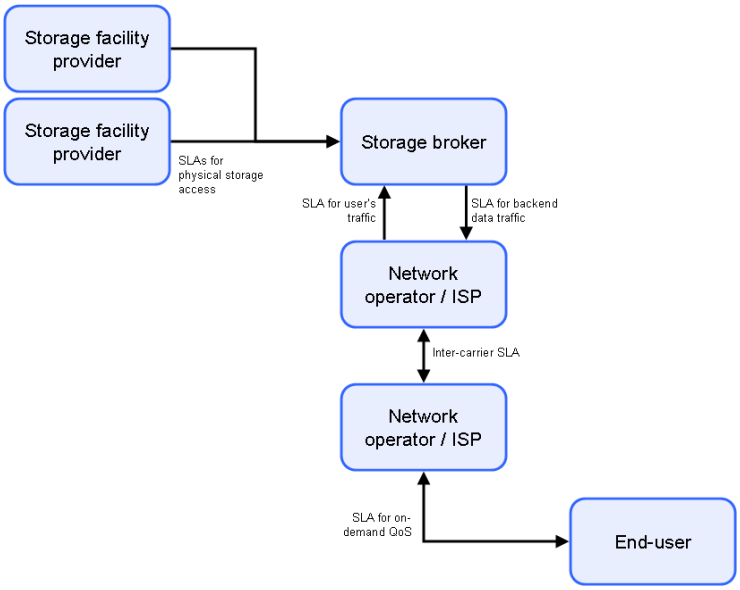


Figure 21: Stakeholders Relationship Diagram

5.2.2.3. Desktop applications as a service

This simple final scenario refers to the possibility of offering web-based applications (like Microsoft Office or Google docs) that can be offered with real time guarantees using, the Cloud Computing paradigm for the provisioning of the service itself while the network can offer real time connections for assuring the interactivity of the applications. This scenario could be implemented following a similar methodology as that one proposed in the GaaS or Virtual drive scenario.

5.3. BUSINESS ORIENTED SERVICES

The goal of this scenario is to provide to a medium or large business or corporations an integrated service that could allow them to have VPNs, Tele-presence services, secured access to main content providers, etc. One of the objectives of this new scenario is the capability to offer more services without the need of implementing ad-hoc solutions.

In addition, more dynamic services that could, e.g. take advantage of the new technologies that can, e.g. provide BoD solutions are also considered. These integrated services must include access to corporative networks using mobile technologies with guarantees. The services that will be included will be: the evolution of current VPN services, Telematic Services and Advanced Tele-presence services.

Again, it should be noticed that these services can be also extended to the residential users and that they have an impact on the wholesale services.

5.3.1. EVOLUTION OF VPN

A VPN is a logical network that creates a private and/or secure scope on underlying networks which may be public /insecure. The VPN users groups are therefore separated from other users and can communicate as though on a private and/or secure network.

There are lots of existing methods of creating VPNs, dividing into two main groups:

- Provider-provisioned, such as BGP/MPLS IP VPNs at Layer 3, and VPLS at Layer 2
- Customer-provisioned, such as IPsec at Layer 3 and SSL/TLS at Layer 4+

Traditionally provider-provisioned (PP) VPNs concentrate on providing a 'private' network with performance guarantees – for instance for a corporate spread over several sites - whilst customer-provisioned VPNs typically concentrate on providing a 'secure' network – for instance so a travelling employee can get limited access to corporate services like email, via an encrypted web-page. In some ways the evolution of VPNs is about blurring these distinctions – finding the right mix of techniques to give a simpler, more flexible VPN but with performance guarantees.

There are several potential ways that future VPNs can evolve in the short / medium term, perhaps up to 5 years. We concentrate on the following evolution scenarios that all include interconnection between service providers:

- Simpler and flexible provisioning of existing PP L2 and L3 VPN services with QoS guarantees over interconnected networks of two or more service providers. Each service provider has its own understanding and implementation of QoS. So the challenges are how to globally provision and maintain an inter-service provider VPN, jointly guaranteeing a unique end-to-end SLAs. This would require mapping the various heterogeneous QoS models and mechanisms of each service provider, and leveraging the appropriate OAM mechanisms to ensure the common SLA and reduce operational costs.
- Integration of value added services and application awareness. This means that applications /services can get appropriately differentiated QoS (more exactly, QoE) while network resources utilization is optimized based on user traffic flow classification and a clear understanding of per-flow QoS requirements. In addition application awareness allows for providing the client with advanced visibility on the VPN service usage and planning of future requirements. Among the various added value services, it is important to highlight security as a service, both for L2 and L3 PP-VPNs, applications acceleration and compression.
- Extension of Virtual Private Networks to a more general concept of Virtual Private Services (VPS) beyond networking, towards new services like IaaS (Infrastructure as a Service), PaaS (platform as a Service) and SaaS, enabling a broader scope for outsourcing of corporate Information/IT and Communications solutions. Optionally, the VPS could be based on cloud based architectures. We see this future networking, content and IT services being integrated in global service definitions within an inter-SP eco-system.
- Leveraging the global Internet connectivity to extending the reach of L2 and L3 PP-VPNs outside the footprint of an operator (e.g. using SSL VPNs) and to interconnect "non-adjacent" L2 and L3

VPNs from the same or different SPs. A key challenge here would be to keep some of the QoS benefits of existing PP-VPNs and ensure end-to-end QoE.

- Ensuring coordinated and efficient multi-service provider provisioning of client sites multi-homing in order to guarantee the required resilience level.

In all cases: the benefit will be some combination of better service for the business customers, including in some cases a single contact point for the whole coverage and services spectrum (one-stop-shopping, possibly through a broker), improved end-to-end multi-strata (network and services) QoS and QoE, more granular service guarantees and monitoring, QoS over wider footprint and greater geographical reach for VPN. For an operator, the benefits beyond cost reduction would be also moving 'up the value chain'

5.3.2. Telematic Services Using Networks Across Multiple ISPs

In this scenario we consider a telematic system distributed across several ISP domains (examples include disaster prevention and relief, traffic information, remote health services, security and safety, etc.). This scenario could be also classified into the end users' services, but we have included here since the information provided in these services will be distributed by large corporations that will have to manage several probes and access them in a reliable way.

The typical example of these services is constituted by the e-health scenario (at the end, we will have a set of sensors distributed in a home environment that must be able to connect to a corporation in a reliable way).

In this scenario we consider a telematic system distributed across several ISP domains (examples include disaster prevention and relief, traffic information, remote health services, security and safety, etc.) and some information has sense considering geographical localization. In order to provide supreme reliability for critical such services, QoS has to be assured along the entire communication chain, which may include connection-oriented as well as connectionless network domains. End-to-end QoS-enabled interconnection is guaranteed by extending the current PCE technology for bridging connectionless AS-hops.

The essential property of this scenario is that (moving or nomadic) users wish to consume geo-location related services by OTT providers with the highest possible level of QoS. As already previously mentioned, this necessitates the assurance of the path of connectivity between the relevant OTT service providers and users' network access provider, which may be quite challenging to achieve in geographically highly distributed constellations of the relevant actors. The resulting two dimensions of the geo-related coverage of the network service and the OTT (=application level) service, respectively is depicted in Figure 22.

In order to implement these services, the following points should be considered:

- **QoS Guaranteed Service**: Inter-operator QoS is the key to success in this (near) real-time scenario. E.g. in case of displaying traffic information using video streaming to end users' car devices, a minimum latency must be guaranteed. The necessary QoS agreements are shown in Figure 23.
- **OTT Service Broker**. In order to achieve transparent service usage, in which the end-users are not required to have a business relationship with multiple OTT service providers, the OTT Service Broker becomes the entity in charge of mediating the service.

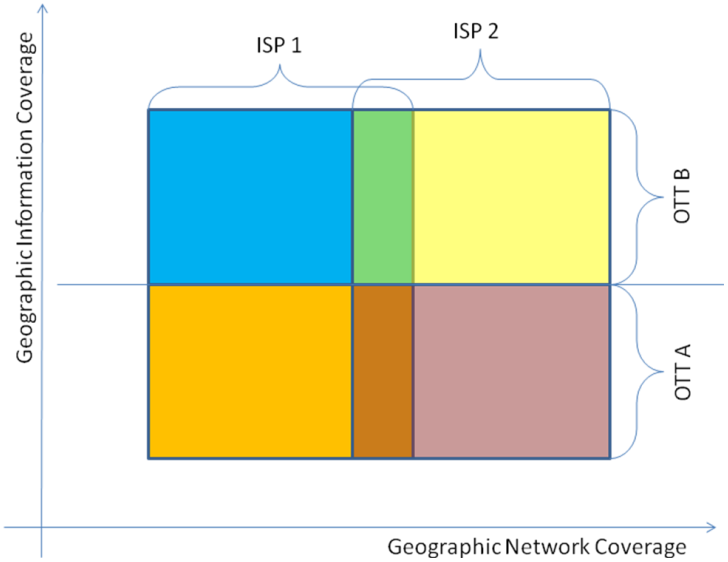


FIGURE 22: TSUNAMI GEOGRAPHIC COVERAGE PRINCIPLE

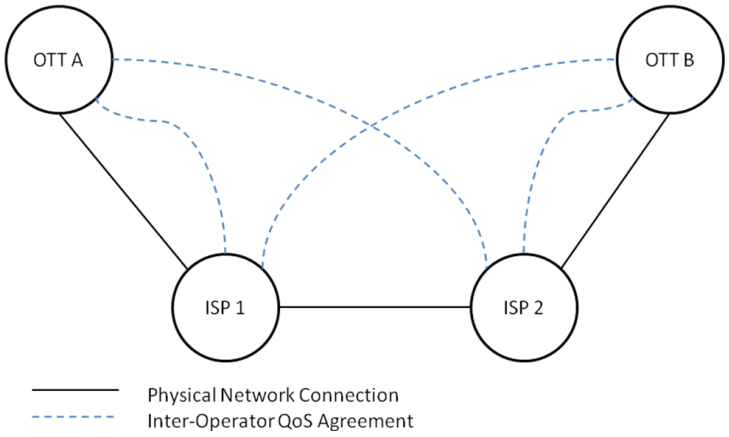


FIGURE 23: TSUNAMI QOS AGREEMENTS

Global benefits of this scenario are:

- End-user QoS empowerment: The OTT Service Broker (OSB) simplifies contractual relationships between end customer and providers down **to just one contract**. The OSB provisions QoS connectivity and mediates OTT services according to the user needs, thus ensuring end-to-end QoS between end user and OTT Service Provider.
- Identity provisioning: The OSB acts **as trusted entity** which reliably authenticates the user towards OTT Service Providers and handles charging of both provisioned QoS and content consumption.

In contrast to QoS SLAs between operators, which are usually agreed for medium to long-term periods, the OSB architecture would enable more fine-grained QoS control, both in terms of guaranteed bandwidth and session duration.

Finally, the main benefits for the stakeholders are the following:

- End-User: Seamless QoS enabled Connectivity with best OTT provider without additional management overhead (contracts).
- Network Service Provider: Increase revenue, Customer satisfaction and information gain concerning the traffic transferred for network engineering and network planning (multi-dimensional QoS traffic matrix)
- OTT Service Broker (OSB): Establish novel business model
- Authorities: Economic efficiency and legal compliance

The elements identified in this scenario are inevitable preconditions for the establishment of services which

- Are dependent from a geographical location or
- Need involvement of more than one OTT service provider.

In this sense they benefit the users because they can simply subscribe a service which involves more than one OTT service provider; the network service provider because it offers new business opportunities; and, the OTT service providers because by building alliances instead of operating on their own, they can compete with more valuable service offerings.

The crucial agreements between the stakeholders are between OTT service providers, the OTT service broker, and between the end-users, which is shown in Figure 24. In this picture, the dashed lines represent the contracts which would be necessary if an OTT service provider would not be available. We consider that it is unacceptable for a user to sign x contracts if a single service is commonly provided by x OTT service providers. In the case that an OTT service provider is involved, each user only has to establish one contract for the service (shown by bold, black lines) which could even be simplified if the OTT service broker serves multiple services and the user already has an existing contract with the OTT service provider.

Other agreements are done between the OTT service broker and the network service providers. If there is also a broker for network (connectivity) services, including QoS, which could be integrated into the OTT service broker, ideally this is a one-to-one relation.

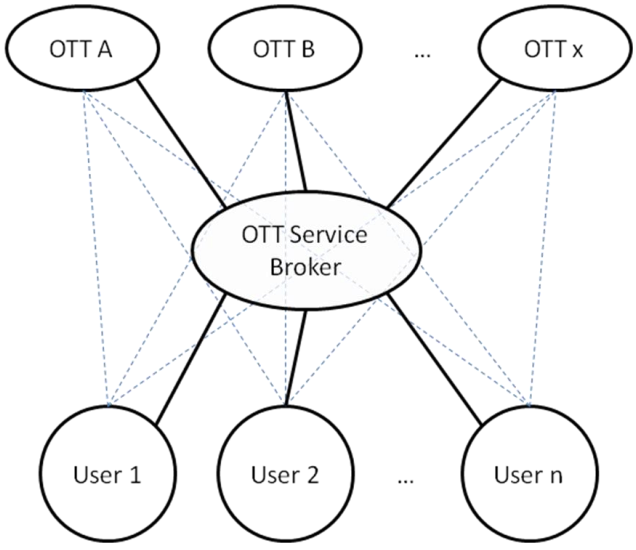


FIGURE 24: TSUNAMI MAJOR RELATIONS BETWEEN STAKEHOLDERS

As the standardization of the required technologies for realizing the described inter-domain QoS interconnections is quite mature in the area of PCE, and as the "missing" pieces are continuously being worked on within the scope of IETF, we may consider the standardization status not to represent a barrier at all.

5.3.3. ADVANCED TELE-PRESENCE SERVICES

The scenario is built around tele-presence services (TP) that span multiple sites, with access provided by different carriers over the current Internet. As starting point, we have considered that this service will be deployed in the corporate environment but that in the medium or long term there could be an offer available for the residential customers.

Advanced and full tele-presence is intended to be composed by High Definition video streams, voice streams, and additional data streams (e.g. from sensors), and its candidate application is remote collaboration on complex systems. The sites might be fixed when planning the service (e.g. corporate sites where collaboration is going to happen), or added on the fly (e.g. working sites that change dynamically). The communication paradigm can range from point-to-point or point-to-multipoint paradigms.

The QoS needs are **moderately demanding in terms of bandwidth** (up to multiple Mb/s, which is something affordable in corporate users) and **strict in terms of delay and jitter**. The service architecture, from an **inter-carrier perspective, has to face challenges like the set up and maintenance of dynamic traffic trunks supporting isolated flows to multiple endpoints** (possibly changing within the same SLA).

Additional features may include (but are not limited to): session recording mechanisms to let meeting participants to replay from storage (distributed) servers (e.g. in case of participation on different time zones), enhanced 3D visualization and virtual presence mechanisms, etc.

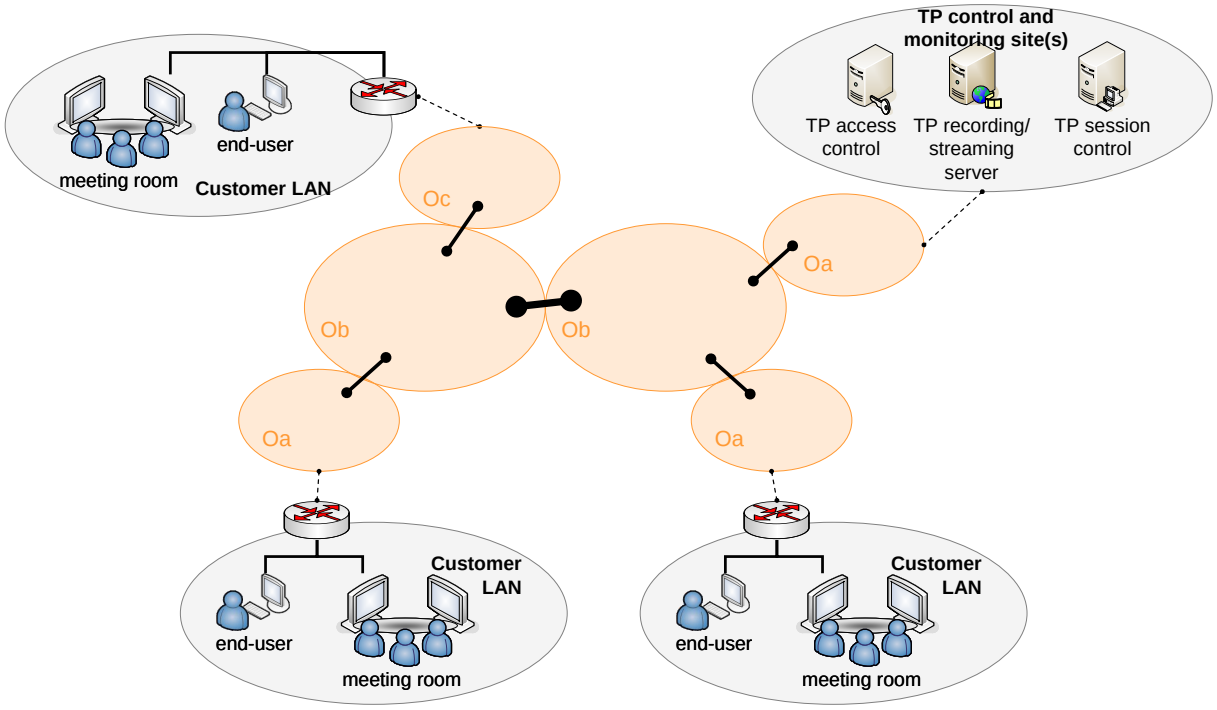


FIGURE 25: ADVANCED TELE-PRESENCE ARCHITECTURE

The main benefits of this service are the following considering the stakeholders:

- For end-users: reduced mobility for geographically dispersed teams/workers and concrete enabling of full-featured tele-working.
- For Network Operators/Service providers: increased revenues from improved VPN connectivity services.
- Issues on the network: Strict requirements on end-to-end delay and jitter for voice services and in point-to-multipoint service scenarios; need for inter-carrier authentication and accounting at different flow/connection granularities; need for on-demand connections, with increasing bandwidth (in case of aggregation in the core) from the access to the core segments.

The different services that can be deployed: TP provisioning, control and monitoring, online TP service execution (HD, 3D, virtual presence) and delayed TP playback. These services are detailed in the next subsections.

5.3.3.1. TP provisioning, control and monitoring

The TP customer asks the TP service provider to schedule a TP session among a set of participating sites and persons, with specific tools/features to be set up (or made available) at remotes sites: e.g. HD video, 3D applications, virtual presence, recording facilities, standard meeting tools like whiteboards, remote desktops, chats, etc.

The TP service provider defines the connections requirements among the sites and configures automatically the CPE equipments, the QoS-enabled network pipes, the recording facilities, etc. When the service is set up a notification to all the participants is sent.

QoS requirements basically refer to the possibility of the TP service provider to configure the end-to-end service(s) considering point-to-point or point-to-multipoint scenarios.

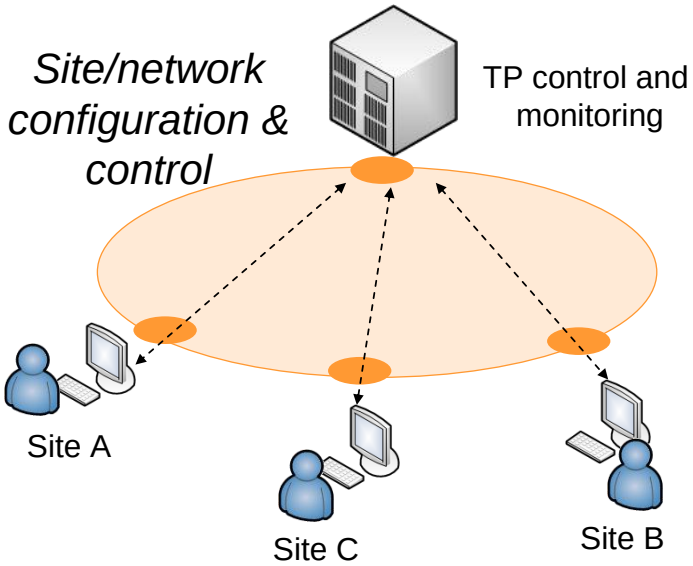


FIGURE 26: TELE-PRESENCE PROVISIONING, CONTROL AND MONITORING

Main issue for the TP service provider and/or for the Network operator is the availability of automatic control and monitoring tools, possibly integrated with network control plane.

5.3.3.2. Online TP service execution

Once the TP participants log on, the service session is executed with all the configured facilities (Voice, Video and data traffic).

The QoS requirements will mainly consist of: low E2E delay limitation (less than 200 ms), very low packet losses and jitter and scalable rates for the video traffic, ranging from a few Mbps up to tens of Mbps for HD or 3D quality. All these requirements are bidirectional and for any-to-any connection.

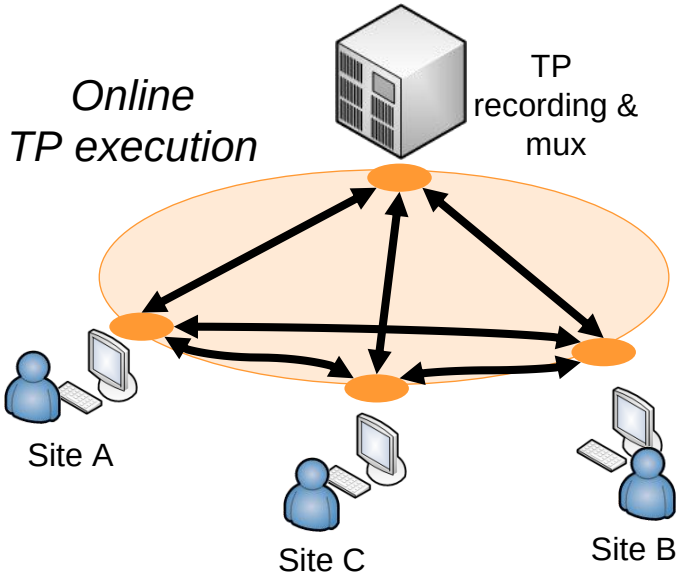


FIGURE 27: TELE-PRESENCE SERVICE EXECUTION

The main issues for the backbone operators are the following:

- Automated provisioning of point-to-multipoint connectivity trees
- Control of fine-grain QoS
- Control of scheduled resources
- Per-user/per-flow accounting of BoD
- Per-user/per-flow monitoring and recovery of QoS

5.3.3.3. Delayed TP playback

The TP participant(s) on a delayed time zone log(s) on to the system and search(es) for the specific and available session recordings. Playback is started and controlled at end-user site. QoS requirements previously described are still valid in this context, but applied just in the downstream direction, i.e. from the TP session streaming site to the specific end-user, since now there is no real interaction.

The benefits of this service are exactly the same ones that are detailed in the previous service. Moreover, main benefit is for the end-users, who reduce the collaborations issue of remote teams operation in different time zones.

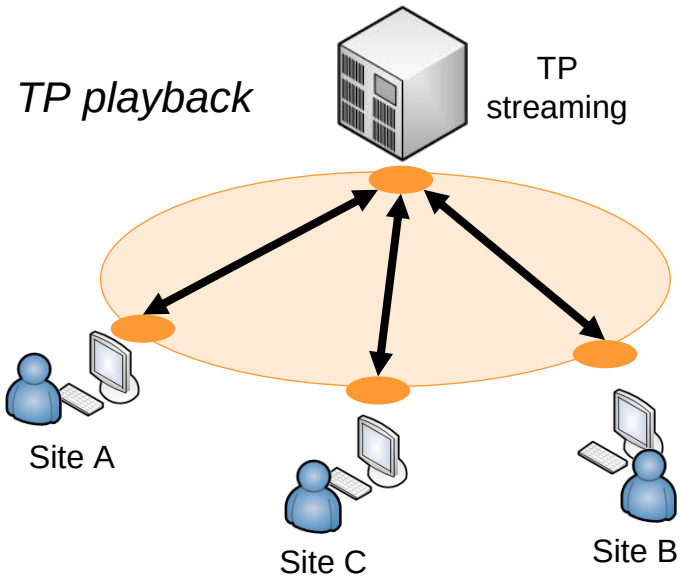


FIGURE 28: TELE-PRESENCE SERVICE PLAYBACK

Resources to be mobilized for this service include:

- TP call management system
- TP facilities at end-users' sites, including QoS-capable CPE router and voice/video/data applications and HW
- OSS systems to manage and account for connectivity and service execution

Following picture illustrates the agreements among the different stakeholders:

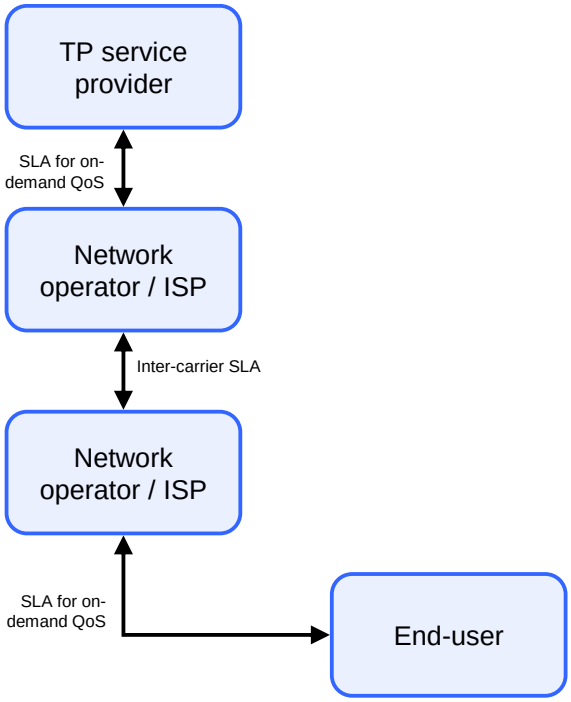


FIGURE 29: TELE-PRESENCE SERVICE STAKEHOLDERS RELATIONSHIP DIAGRAM

- TP service provider - Network operator/ISP: the TP service provider establishes a contract with one or more network operators to configure scheduled on-demand connectivity services with QoS guarantees.
- Network operator/ISP - Network operator/ISP: there will be inter-carrier contracts to support permanent or dynamic transport connections for TP traffic.
- End-user - Network operator/ISP: the end-user needs to extend its SLA with its ISP in order to be granted on-demand increased QoS for virtual drive services.

5.4. ADVANCED WHOLESALE SERVICES

This scenario will show how the operators and other parties interconnect between them considering more issues than just sending/receiving IP traffic using BGP. Therefore, all these scenarios try to infer the new capabilities that could be exposed in order to have services that could evolve from the traditional basic connectivity to Internet.

5.4.1. INTER-PROVIDER MULTICAST STREAMING (LIVE@ME)

This scenario will show how the activation of multicast features can be used to introduce High speed Real-time multimedia content in a network. Most ISPs today compete offering flat rate Internet accesses to their customers both at home premises and mobile terminals. Because of this and the decreasing number of potential new customers, added-value services research is the key to increase the business growth.

“Live@Me” is an added value service where flat rate customers willing to become live shows producers pay a monthly fee to deliver their show up to N subscribers. The maximum number of active subscribers (N) establishes the service segmentation:

- Producers expecting from 1 to 250 active subscribers at once are not expected to get revenues from their activities and thus a small fixed monthly fee should be charged.
- Producers expecting from 251 to 1000 active subscribers at once might not have direct incomes but benefit from some advertisements/marketing actions and thus expected to pay according to well-defined network usage (traffic sent) levels.
- Producers expecting more than 1001 active subscribers at once may require other specific agreements.

Besides show producers, “Live@Me” might be interesting for real-time data providers such as weather stations or authorities interested on disseminating emergency alerts where delay and congestion are key parameters. In the later case, Live@Me will not suffer congestion problems of users accessing a WEB page and thus guaranteed-delivery instructions can be delivered to the population. Network can be configured to prioritize such emergency data flows under special conditions or emergency situations.

“Live@Me” is a brand new opportunity based on the cloud computing model for long tail Internet providers to enable home or mobile customers to become live shows producers. “Live@Me” allows live content producers to set-up professional radio, TV or data-streaming stations. These producers will establish an open or pre-defined group of listeners or customers & produce real-time contents delivering them as one simple multimedia flow to the network.

With “Live@Me”, producers do not need expensive servers or broadband connections as they deliver it to the cloud (the Network) without any intermediary (such as a CDN provider), which will in turn copy & distribute the content flow to the show subscribers with low record delays and thus making live TV or radio shows a reality. Data-streaming stations or applications requesting the lowest transmission delays, such as emergency networks, might be interested as well.

Live producers might go from regular citizens creating live programs on specific hobbies to live contests, amateur singers or even famous starts answering fans requests.

“Live@Me” allows the ISP to manage the maximum number of active subscribers to a specific show, depending on the live producer profile and service agreement. This way it guarantees service quality and keeps network under control.

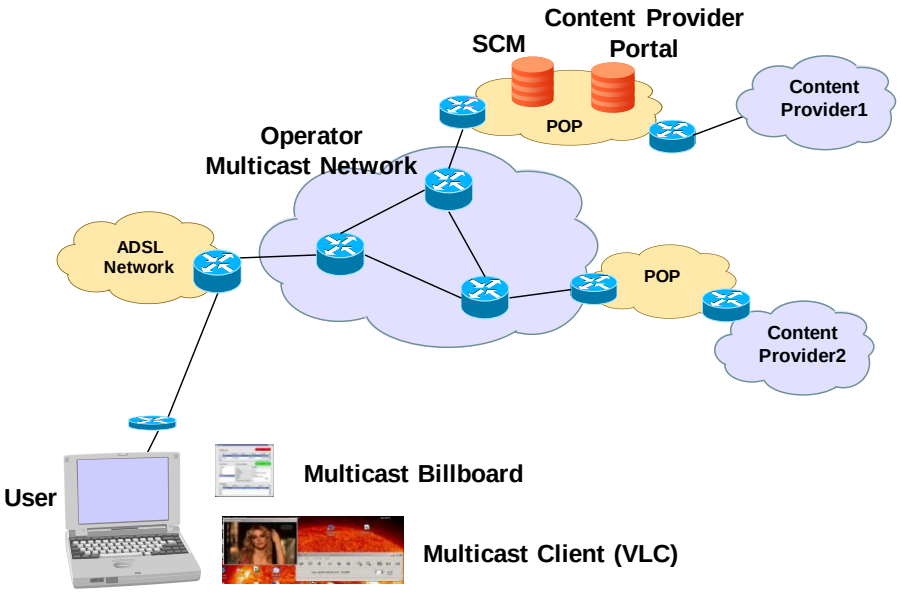


FIGURE 30: LIVE@ME SERVICE ARCHITECTURE

Regarding the technical solution and their associated requirements, in order to achieve easy content flows distribution and the lowest delays the ISP will use Multicast transport at the IP layer. Other techniques used by current solutions based on software platforms processing flows or data messages are subject to bigger delays and congestion on unexpected audience growth or emergency events. Specifically, Multicast SSM service is selected to enable a specific source to deliver content to a group of subscribed listeners.

As Multicast SSM needs a public IP to deliver the content and NATs or proxies may jeopardize the service deployment, IPv6 SSM Multicast is chosen as the transport/distribution platform for the service.

By default, standard Multicast SSM would allow any potential listener in the network to subscribe to an active source and thus preventing a proper network control and customers differentiation. To avoid this, a modified architecture is defined: a centralized platform keeps a database with the maximum number and nature of potential subscribers and manages the Multicast SSM subscriptions establishing priorities and keeping the number of active subscribers within the agreed broadcasting limit.

Potential subscribers may receive the contents with any of the numerous existing IPv6 multicast SSM enabled clients and a WEB portal will keep information on active open and accessible groups.

Due to the architecture of Multicast SSM and in order to provide a universal service across the Internet a method and platform to establish and implement inter-provider agreements is needed. The definition of such method and platform enabling other ISPs to join "Live@Me" services is beyond the scope of this document.

The stakeholders involved in this service will be the following:

d) Prosumers. Those users (fixed or mobile) that can create and distribute multimedia contents into the network.

e) Final customers. Those end users (fixed or mobile) that can access and view the multimedia contents offered in the network.

f) Network providers, both Internet Service Providers and Telcos; will enable and provide the necessary multicast infrastructure to provide this service to their end-users (fixed and mobile) and the agreements with other Telcos & ISPs users that want to access the service.

5.4.2. CARRIER-DRIVEN CDN

Due to the tremendous impact that the content distribution is having in operators' traffic, operators could provide CDN based solutions that can be used by the content providers in order to distribute their content. These CDNs could be directly provided by the ISPs or the ISPs could have bilateral agreements with CDNs providers.

An important change in these carrier-driven CDNs will be that the content distribution will be done considering the capabilities available in the network.

5.4.3. APIS FOR 3RD PARTIES – WESTBOUND INTERFACES

This scenario assumes that the end users have a connectivity profile, where they have bandwidth for Internet access and they also have a set of services provided by the ISP. This scenario, offers the possibility of adding a new set of services that are provisioned for the end users by OTTs.

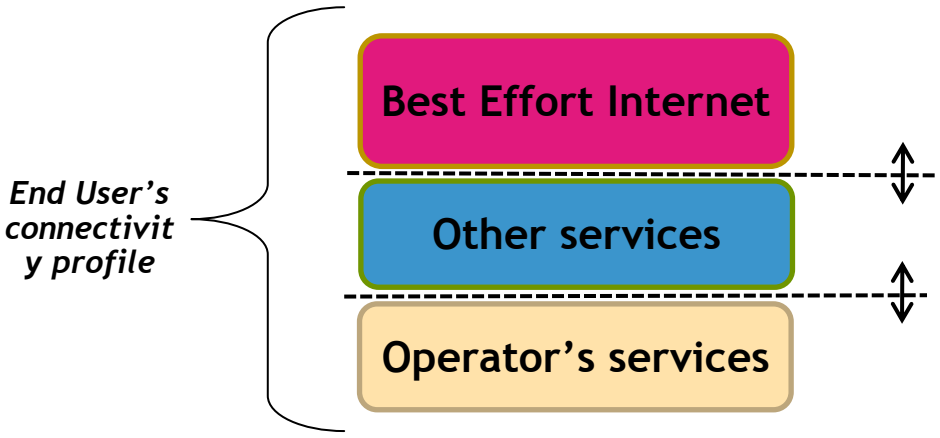


FIGURE 31: APIS FOR 3RD PARTIES AND END USERS' CONNECTIVITY PROFILE

In this scenario, we consider the open capabilities that can be offered by network operators to:

- The social networks to e.g. implement the real-time social networking scenarios that we have explained before.
- Gaming servers if multi-party gaming applications are deployed.
- Or to OTTs, such as Google, in order to enjoy their interactivity services with real time guarantees and minimum bandwidth (e.g. Google docs with low latency or better visualization of the figures) .

In this scenario, it is important to establish the incentives for the main stakeholders:

- For network operators, they can improve their users' loyalty since they can experience better quality in the services they use. Moreover, for the main carriers, there could be another possibility to offer services following a model similar to the hub model or to the transit agreements. Effectively, small operators will not have enough negotiation power to sign agreements with important OTTs (such as Google or Facebook) in order to offer them their network capabilities. Therefore, in this scenario, we also consider the possibility of having the agreements with the main carriers and making these agreements consistent in the end-to-end path in order to allow the small Telcos to offer their capabilities to major Internet OTTs. This scenario is depicted in the next figure.

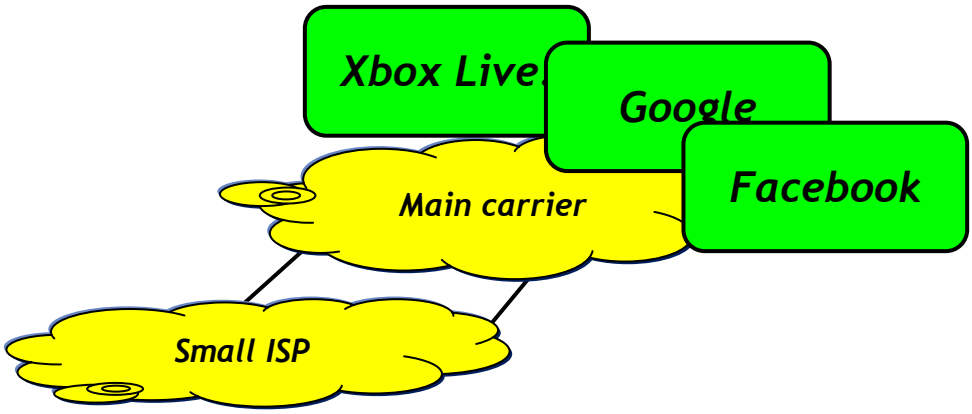


FIGURE 32: HIERARCHICAL MODEL IN THE API FOR 3RD PARTIES

- Finally, for the OTTs, these capabilities offer them the possibility of controlling some properties of the last mile.

In order to implement this model, there are important issues that should be considered:

- First of all, the API must be reliable. There should be no problems associated to the openness of the model that could make the network unstable.
- There should be effective ways to manage the last mile and enforce policies and changes in the user's profiles.
- Finally, there should be a clear business model where the revenue sharing schemes are detailed and then implemented in the technical solution. In this sense, it is important to highlight that multiple stakeholders claim that the end users only want a single contract for their payments.

5.4.4. OPTIMIZATION OF DATA TRANSPORT PROVISIONING

This scenario collects a set of technologies that could make the data transport more efficient and, therefore, could make possible the implementation of the above scenarios. All these scenarios should not considered scenarios by themselves but these new network capabilities should be considered in the specification of wholesale services.

5.4.4.1. BoD and dynamic contracts

The BoD technology could be an excellent candidate to optimize the setup of connections between different domains and stakeholders. More details can be seen in annex B.

5.4.4.2. Green Communications:

Optical switching seems more efficient in terms of CO2 emissions, the possibility of introducing BoD could allow the reduction of the number of provisioned links, since the PCE architecture allows the usage of your own (operators' ones) algorithms, these algorithms can be updated to considering green issues, etc.

Network of future as any other electrical system in the world has to cope with energy efficiency issues by implementing effective power consumption monitoring and controls in order to meet both performance requirements and CO2 emissions. The scenario consists of 3play service by providing on northbound interfaces monitoring of power consumption information and to cope with incoming and out coming signalling carrying 'green communication info'. Also, the services provide and cope with Inter-carrier 3Play connectivity having a label like 'low, medium, reduced, traditional' energy efficiency (network service, 3play), providing billing and tracing report with Co2 emission info.

Regarding the QoS elements of this scenario, best effort, premium / gold connectivity and leased lines classes based on 'low, medium, reduced, traditional' energy efficiency parameters. Costs based on power consumption effectiveness.

Current Internet is the Ecosystem used for this service. IPv6 is presented as candidate protocol with additional 'virtual layer' dedicated to the green networking.

Main benefits of this scenario are:

- To provide efficient communication from the energy point of view. Connectivity provided on low consumption network areas will cost less to the providers that will have reduced OPEX (energy) for the equipment and web farm involved. The low energy traffic shall have thresholds because intrinsically more traffic generates more energy consumption. Special connectivity tariffs can be applied for these routes. Benefits are addressable to providers and end user and event to the countries involved.
- New business model / billing mechanism based on power consumption
- To route communication sessions through low power consumption network areas.
- To speed up industry dealing with low consumption equipment

The main stakeholders of this service could be network and service providers, electronic providers, etc.

Direct incentives are foreseen for the network / service providers that will have the possibility to reduce energy consumption costs by delivering standard green networking connectivity / services at established SLAs. End user or large communities of end users (corporate, academies, NRENs) may receive special

condition on case of ‘green networking’ services. Government together with standardisation will have the possibility to regulate globally these services and prepare incentives for providers applying.

6. HIGH-LEVEL REQUIREMENTS AND NEXT STEPS

The analysis of current services and business models, the identification of potential factors that could influence on the future positioning of the different stakeholders and the identifications of future service oriented scenarios have been established above as the setting to take into account in the further analysis by ETICS. In this section, we aim to provide a first initial set of high level technical and business requirements that constitutes the basis for further work in WP2 and other ETICS WPs as it is described at the end of the section.

6.1. HIGH LEVEL BUSINESS REQUIREMENTS

Considering the different types of scenarios, an important common point is that multiple types of agreements have been considered among different parties and that they all have an important impact on the network capabilities that must be provisioned.

Therefore, **the ETICS framework must be designed and developed in such a way that multiple types of interconnections can be managed.**

This framework must provide support to the automation of the process among the different stakeholders (network providers, content providers, service providers, main corporations, service and application developers, etc).

The framework must, therefore consider how to publish, fulfil and monitor the different agreements and services to be managed. Moreover, an important requirement is to assure the **reliability** of the proposed solution and model, since network operators will not offer information and/or services if a sustainable and stable model is not defined.

The ETICS model must assure a sustainable ecosystem, where carriers have incentives to continue the deployment of advanced network infrastructure and where the development of new innovative services will assure enough Return on Investments for all actors.

In this sense, the **following main points must be considered when analyzing the different business models**:

- For the different services and their associated models, the positioning of the different stakeholders in the model itself and their position in the future value chains and the market in general.
- Revenue sharing models considering that end users will not like to sign multiple contracts but do require predictable service fees.
- The cost to exploit the services provided by the own NSPs in comparison with the costs and revenues associated in a collaborative environment.

- The possible creation of new roles (e.g. brokers, intermediaries) that will be in charge of creating some kind of market place where the different parties can exchange their contracts. The benefits of these new roles and their impact in terms of regulations should be carefully analysed.

According to the identified scenarios, at least the following agreements are considered:

- Agreements between NSPs and OTTs through some kind of API to third parties where the NSP exposes its capabilities.
    - The NSP exposes network capabilities they can offer to the OTT the management of the end users' connectivity profile.
    - In order to support different business models, this API must implement monitoring information visualization/notification in order to verify the fulfilment of the agreement and must also implement an environment where the OTTs can be authenticated, authorized and charged in a reliable way.
    - These agreements should also have an associated fair usage of the resources among the different stakeholders (e.g. use of codecs that do not require too much bandwidth for mobile users if they are using a smartphone in order to avoid an unnecessary traffic in a loaded cell).
- Since not all the NSPs have the same negotiation skills, a **hub model** where main carriers have agreements with major OTTs and these agreements are implemented in the end-to-end chain should be supported. This will have benefit for the main carriers (since they will provide more advanced interconnection models) and the small ISPs since they will have the opportunity of composing their network capabilities with service providers in order to provide carrier-class services.
- The agreements in scenarios where services are provided to companies for the provisioning of advanced services must be considered in the framework. These services will not be pure Internet services but could guide the evolution of the future networks.

The exposition of capabilities among the different stakeholders must be technology agnostic. Therefore, the business logic should be able to adapt seamlessly and agnostically to the network and transport conditions.

Therefore, in order to assure the success of the solutions, the business and policy logic must be able to translate from this network agnostic description to each specific network infrastructure.

The **business relationships and agreements must be described in a SLA (Service Level Agreement)**. The ETICS project must provide a clear data modelling of this SLA.

Considering the different scenarios that have been considered in this deliverable, at least, the following elements should be included:

- Description of the service in terms of bandwidth, delay, jitter and availability.
- Description of the duration of the service.
- Description of the involved parties, their functionalities and the accounting/charging procedures.

- Description of the way to monitor the contract fulfilment.
- Description of the actions to be taken in case of violation of the contract.

Since multiple stakeholders could be involved in the provisioning of a single service, **this SLA must be composable** in order to have the way to calculate the final characteristics of the services that are provided.

This SLA must be defined with clear semantics; it must be extensible and flexible to support the different lifecycles of the services. In order to support high dynamicity in the relationships among stakeholders, the business logic must support and enable **negotiating the contracts in a dynamic way**.

This point seems fundamental considering the potential market impact of cloud services where software, IT infrastructures, etc. may be accessed dynamically for a given period of time and charged at the usage. The network(s) providing the connectivity between the different extremities in the cloud should therefore be at least as dynamically (and openly?) composable as these cloud services.

6.2. HIGH LEVEL TECHNICAL REQUIREMENTS

Firstly, the ETICS technical solutions should consider the **requirements at the data plane** for the different proposed services. According to the studied and identified scenarios, the following main requirements are considered:

- Delay and jitter are quite important for the provisioning of real time communications.
- Reduced packet loss paths are required in order to support remote access presentation services.
- Assured bandwidth for the end users is needed and large bandwidth guarantees should be available for the backend services that are specified for some cloud computing based services.
- Mobility is considered a key issue in the scenarios where end users are involved. In all these scenarios, roaming and service ubiquity must be also provided.
- The availability of the service must be assured.

The way to provide all these features in the different technologies is different, therefore the ETICS framework must:

- support the **dynamic and automated inter-carrier SLA establishment** and their mapping to the specific technology options.
- support E2E connectivity services across carriers **using different transport and QoS technologies**.

The provisioning of the service quality assurance capabilities must consider:

- Different routing and traffic engineering options (complex, pre-computed, policies, etc.)
- Policies to assure the fair usage of the network services (e.g. traffic shaping)
- Control plane capabilities, such as PCRF (Policy and Charging Rules Functions).

- Dynamic management of the end users' profile.
- Interaction with cloud computing and other network service platforms.
- Solutions that cover connection oriented or connectionless solutions.
- Integration or not of the application signalling.
- Scalable set of rules to be managed at the data plane.

ETICS should also consider services that enable the avoidance of specific equipment at the end users' premises. An important benefit of scenarios such as the GaaS or remote virtual drive is that end users do not need specialized devices at home (a simple equipment able to reproduce a streaming and send the required actions is enough); therefore, if the service is provided as a complete solution, the stakeholder that makes the final offer do not really need to provide and maintain specialized devices (any device simply works, since the intelligence of the service is in the network side) at the end users' locations, which , as commented for the IPTV services could have an important operational costs .

ETICS services could be realized on top of existing services (recursive service setup), to build more complex ones as a way to provide an efficient support for the composition of SLAs.

All the layers and components in ETICS should have open and standardized(/standardizable) management interfaces in order to ease its integration in the existing networks.

The network services must have a set of **automated and dynamic phases**, which include:

- The SETUP of the service where the agreement is negotiated and enforced in the networks. This setup should also consider protection techniques to guarantee the reliability of the services.
- The VERIFICATION/TESTING mechanisms must be available in order to guarantee the effectiveness of the services.
- If failures are detected, RECOVERY mechanisms should be considered. It should be noticed that this phase could represent the major load to be supported at some equipments, e.g. if a failure is detected, the PCE will have to support multiple requests.
- Finally, general Operation, Administration and Maintenance phases should be considered.

The ETICS framework must be designed in order to support multiple charging schemes: per volume, per session, cascading payments, etc.

6.3. NEXT STEPS

After the identification of the scenarios and high level requirements, ETICS will proceed in the following way:

- WP2 will take this initial set of high level requirements and will develop the set of technical and business requirements that will define the needed capabilities and the boundaries of the ETICS framework. This will be reported in deliverable D2.2. The most urgent, short and medium term services

and capabilities will be in focus for D2.2 while more long term capabilities will also be considered in the final deliverable on requirements, that is, in D2.3.

- WP3 will take the proposed scenarios and requirements as the input to elaborate the business models.
- WP4 will take all these requirements (the high level ones and those ones that will be further elaborated) as the initial input to elaborate the high level architecture of the ETICS framework.

Finally, the proposed scenarios will be carefully analysed in order to pick a selected group to demonstrate the ETICS solutions in WP7.

7. REFERENCES

[ABI10] A BI Research's "Mobile Subscriber Database": http://www.abiresearch.com/press/1064

[Acc10] Leading App Store for IPTV and Connected TV. [Online] Accedo broadband. [Cited: 03 23, 2010.] http://www.accedobroadband.com/index.php?page=products.

[ALLB10] "Research and Markets: A Detailed Analysis of Mobile VoIP Evolution & VoIPo3G Business Models", link available at: http://www.allbusiness.com/media-telecommunications/11405747-1.html

[ASP10] IPTV ASPIS. [Online] [Cited: 03 23, 2010.] http://iptv-aspis.com/.

[Bal09] Baldauf, M., Fröhlich, P., Reichl, P. (2009). Gestural Interfaces for Micro Projectorbased Mobile Phone Applications. Adj. Proc. Ubicomp 2009.

[Bar05] D. Barth, L. Echabbi, C. Hamlaoui, and S. Vial. An economic and algorithmic model for QoS provisioning BGP interdomain network. In EuroNGI Workshop on QoS and Traffic Control. Deliverable reference number.D.SEA.6.4.2, 2005.

[CIS09] Cisco, "Hyperconnectivity and the Approaching Zettabyte Era", June 2009

[CIS10] Unified Messaging. [Online] Cisco. [Cited: 03 19, 2010.] http://www.cisco.com/en/US/docs/ios/solutions_docs/voip_solutions/UM_ISD.html.

[Corr04] J.R. Correa, A.S. Schulz, and N.E. Stier-Moses. Selfish routing in capacitated networks, November 2004.

[Czu02] A. Czumaj, P. Krysta, and B. Vocking. Selfish traffic allocation for server farms. In 34th ACM Symp. on Theory of Computing (STOC '02), 2002.

[Fall10] Falluca, Valerio. VoIP The future is bright. [Online] 12 2009. [Cited: 03 22, 2010.] http://www.slideshare.net/ValuePartners/voip-the-future-is-brightvalue-partners.

[GENZ06] The generation Z connection: teaching information literacy to the newest net generation. Teacher Librarian (February, 2006)

[How05] M. P. Howarth. et al. Provisioning for Interdomain Quality of Service: the MESCAL Approach. IEEE Communications Magazine, vol. 43, no. 6, pp. 129–137, June 2005.

[INST10] "Number of Mobile VoIP Users Will Approach 300 Million by 2013", link available at: http://www.instat.com/press.asp?ID=2747&sku=IN0904428MCM

[IPC10] What is IP Centrex? [Online] IP-Centrex.org. [Cited: 03 19, 2010.] http://www.cisco.com/en/US/docs/ios/solutions_docs/voip_solutions/UM_ISD.html.

[IR34] GSMA PRD IR.34 "Inter-Service Provider IP Backbone Guidelines", version 4.8, September 2009

[Johs04] R.Johnson, Carolyn, et al. VoIP Reliability: A Service Provider's Perspective. IEEE Communications Magazine. 07 2004.

[Kam04] CA Kamienski, D Sadok. The case for interdomain dynamic QoS-based service negotiation in the internet. Computer Communications, 2004 - Elsevier

[Mon08] Monaco, Raymond. The Future of IPTV. [Online] 2008. [Cited: 03 23, 2010.] http://www.slideshare.net/raymondmonaco/the-future-of-iptv-presentation.

[MOT10] MOTSWAN. [Online] [Cited: 03 23, 2010.] http://www.celtic-initiative.org/projects/motswan/.

[Mul07] Müller, F., et al. "Sports over a distance, personal and ubiquitous computing", Volume 11, Number 8 /Decmber 2007.

[RFC4364] E. Rosen, Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, February 2006.

[RFC4655] A. Farrel, J. Vasseur, and J. Ash. A path computation element (PCE) based architecture, IETF RFC 4655, August 2006.

[RFC5559] P. Eardley, Pre-Congestion Notification (PCN) Architecture, IETF RFC 5559, June 2009.

[RFC5670] P. Eardley, Metering and Marking Behaviour of PCN-Node, IETF RFC 5670, November 2009.

[RFC5696] T. Moncaster, B. Briscoe, M. Menth, Baseline Encoding and Transport of Pre-Congestion Information, IETF RFC 5696, November 2009.

[Skype] Skype. [Online] [Cited: 03 22, 2010.] http://www.skype.com.

[Ver07] FL Verdi, MF Magalhães, E Madeira. A Welin Using Virtualization to Provide Interdomain QoS-enabled Routing. Journal of Networks, 2007.

[Yah06] A.D. Yahaya, T. Suda. iREX: Inter-domain QoS Automation using Economics. IEEE CCNC 2006 proceedings.

8. ACRONYMS

API	Application Provider Interface
ASQ	Assured Service Quality
ASQ-E2E	Assured Service Quality End-to-End
B2B	Business to Business
B2C	Business to Customer
BE	Best Effort
BGP	Border Gateway Protocol
BoD	Bandwidth on Demand
BRPC	Backward-Recursive PCE-Based Computation
CAPEX	CAPital EXpenditures
CE	Customer Edge
CRM	Customer Relationship Management
DHCP	Dynamic Host Configuration Protocol
DiffServ	Differentiated Services
DSCP	DiffServ Code Points
E2E	End-to-End
ETICS	Economics and Technologies for Inter-Carrier Services
GaaS	Gaming as a Service
GPRS	General Packet Radio Service
GRX	Global Roaming eXchange
GSMA	GSM Association
HD	High Definition
HGI	Home Gateway Initiative
HSPA	High Speed Packet Access
iBGP	internal BGP
IC	Inter-Carrier
IMS	IP Multimedia Subsystems
IPX	Interconnection with IP eXchange

MCSP	Mobile Content Service Providers
MMS	Multimedia Messaging Services
MNO	Mobile Network Operator
MPLS	Multi-Protocol Label Switching
NSP	Network Service Provider
OoB	Out of Band
OPEX	OPerational EXpenditures
OTT	Over The Top (Providers)
PCE	Path Computation Element
PE	Provider Edge
PoP	Point of Presence
POTS	Plain Old Telephone Service
PRUC	Premium Real-time Unified Communication
QoE	Quality of Experience
QoS	Quality of Service
SLA	Service Level Agreement
SMS	Short Message Service
SSM	Source Specific Multicast
TDM	Time Division Multiplexing
TISPAN	Telecom and Internet converged Service Protocols for Advanced Networks
TP	Tele-Presence
UPnP	Universal Plug and Play
VoD	Video on Demand
VoIP	Voice over IP
VPN	Virtual Private Networks
VPS	Virtual Private Services
WASP	Wireless Application Service Providers

9. ANNEX A: SOTA – TECHNICAL DETAILS

9.1. OVERVIEW OF INTERACTION MODELS AMONG DOMAINS

This establishment can be done in different ways we detail now:

1. In a ***cascade model*** (see Figure 33.1 and Figure 33.2), a QoS demand is sent to each AS on the route selected by the original AS. In case of positive answers, a SLA is then fixed between each pair of consecutive AS on the route. Such a cascade model can be managed in a distributed way, i.e., each AS on the route negotiates with the next one (see Figure 33.1) or in a centralised way, i.e., the origin AS negotiates with each one on the selected route (see Figure 33.2).[31]
2. In a ***reverse cascade model*** (see Figure 33.3), each AS can first buy a route with QoS guarantees to some of its neighbours and then can sell parts of it to other neighbours. This model follows the classical routing paradigms in which routes are chosen from the destination to any origin AS. Here, QoS guarantees consist in a stock management strategy, based on learning and yield management.
3. **In a *centralized model*** (see Figure 33.4), each AS delegates the route and QoS management to a same centralised broker. Such a model can only centralizes a cascade or reverse cascade process, in which the broker respects the competition between AS and in which each AS has its own commercial strategies and benefits. It can also implement an alliance approach in which the broker both manages the combination of resources of each AS in the alliance and the benefits sharing. Note that a distributed management of the alliance can also be considered.

[31] Note that this model is also well adapted to hierarchical routing: a physical route with QoS guaranties, considered as an "enhanced pipe" is established through this model from origin to destination; then, the origin AS can provide logical routes to its own final customers without too long signalling time processes.

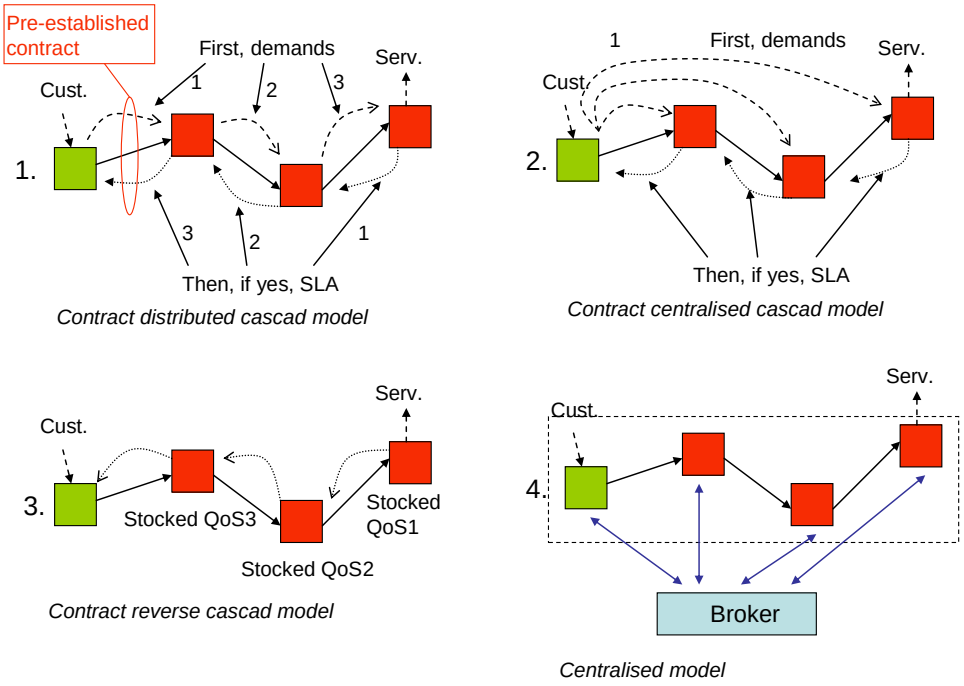


FIGURE 33: PARADIGMS OF CONTRACT ESTABLISHMENTS IN A PRIOR QOS CONTRACT MODEL

The final model (or models) will be selected according to the trust level between operators, on the business relationships, etc. Now the question is related to the technical solution to implement such scenarios, in this sense, one of the most promising technical approaches could be based on the PCE (Path Computation Element) [Farr06].

The example about the negotiation of the QoS parameters is just presented as an example of the business and technical solutions that could be deployed in the Future Internet for the inter-connection among ASs with QoS requirements.

9.2. OVERVIEW OF THE CURRENT STATE OF THE ART OF VPN (VIRTUAL PRIVATE NETWORKS)

Since the breakthrough of IP as the most important global data connectivity network in the mid 1990ies, companies and organizations have been intensively engaged in transitioning all (or at least most) of their data communications towards the Internet. In this sense, Virtual Private Networks (VPNs), as overlays which provide customers with transparent remote site connectivity over a shared network infrastructure, began to play an important role in this process, as remote private-domain connectivity needed to be continuously assured irrespectively of the underlying data transport technology used.

The emerging technologies for provisioning VPNs have basically been focused on two distinct paradigms:

- **Network provider operated VPNs.** In this concept, the network operator takes full responsibility and the technical configuration burden for connecting multiple remote sites of a customer into a seamless VPN. In order to make such a form of VPN provisioning feasible, there are two fundamental

prerequisites. Firstly, in case no CPE-based encryption of VPN traffic is used, the customer needs to trust the operator with the entire security of its private network traffic. Secondly, operator-managed VPNs are easily deployable only within a single network operator domain, such that all the VPN sites of the customer should ideally be within the connectivity reach of the VPN-operating ISP.

- **Customer operated VPNs.** In the case that the prerequisites mentioned in the previous paragraph are not (or only partially) fulfilled, the customer will need to assure VPN functionality on its own, normally employing VPN-capable CPE-devices implementing some specific protocol. The main drawback of this approach lies in the fact that customers in this case need to take full care of the management, provisioning and maintenance of their VPNs, which requires a substantial level of understanding of all the networking and security aspects involved. Furthermore, most of the customer operated approaches are associated to more rigorous constraints with respect to planning for highly scalable VPNs.

Following this introduction, we will take a closer look at several concrete VPN technologies belonging to both of the mentioned paradigms. As far as the network operator managed solutions are concerned, we will focus on BGP/MPLS IP VPNs (Border Gateway Protocol Multiprotocol Label Switching IP Virtual Private Networks), as they represent the most widely deployed and the most popular solution of this kind. Concerning customer operated VPNs, we will address IPsec- and SLL-based VPN solutions, elaborating on the technical specifics of the respective technologies, the feasibility of different deployment configurations based on the currently available solutions, and we will also discuss the manageability of such networks with respect to assuring optimality of the network traffic routed between the different CPE hubs and end-points.

VPNs basically connect two (or more) physically separated networks (in the following called “associated networks”) of the same type together via a transport network which is not necessarily compatible to the associated networks. An associated network in this respect can either be a whole subnet or a single host.

Besides the before mentioned categorization of VPNs, and despite the fact that this is the main categorization according to IETF documents, other categorizations are also often used, depending on the type of associated networks (or nodes) connected together, like e.g.:

- **Site-to-End.** In this case, devices are (separately) connected to the assigned network via a special “gateway” and thereby they become part of the assigned network. Since this creates the illusion of the device directly being part of the assigned network, the assigned networks as a whole are named “virtual network”.
- **Site-to-Site.** In this case, instead of the single device, a whole subnet is connected to the assigned network, by which it also becomes part of the assigned network.
- **End-to-End.** In the case that only single endpoints are connected together by means of VPN, without a real (physical) network involved, the VPN is called End-to-End.

This categorization aids in determining which technology is best suited for the respective type of VPN in question.

9.2.1. BGP/MPLS IP VPN – AN OPERATOR MANAGED VPN SOLUTION

9.2.1.1. BGP/MPLS IP VPN Functionality

The central idea of MPLS/BGP IP VPNs is to (re-)use the internal part of the BGP protocol, i.e. I-BGP, for the distribution of VPN address prefix information within individual autonomous systems (ASes). As described in [RFC4364], this method is based on a "peer model", in which the customer's edge routers (CE routers) send their routes to the operator's edge routers (PE routers), and BGP is then used by the operator in order to exchange the routes of a particular VPN among the PE routers that are attached to that VPN.

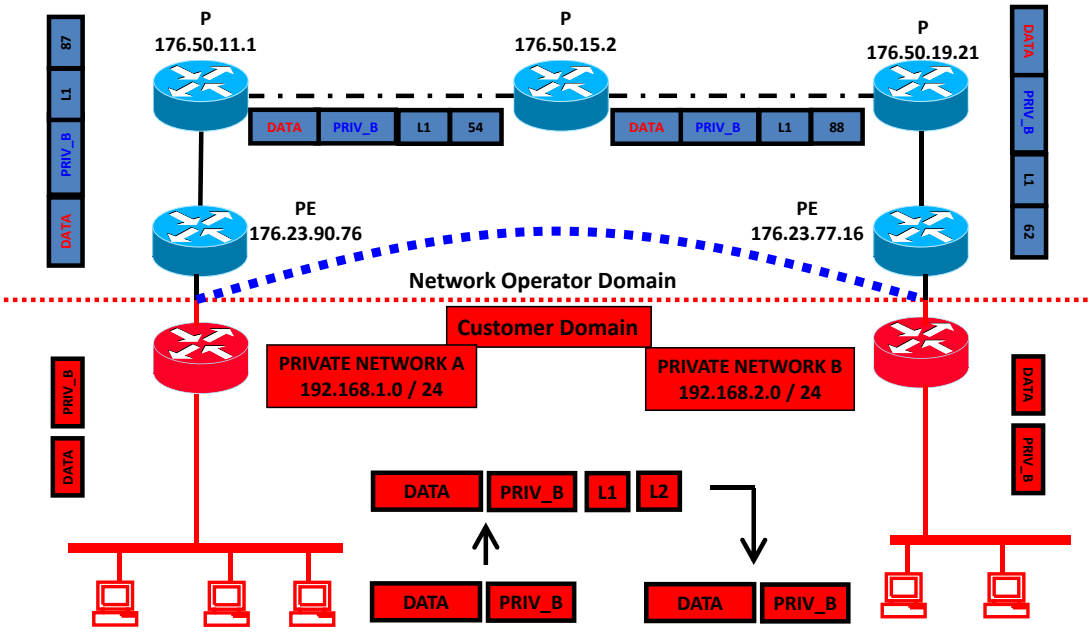


FIGURE 34: A SCHEMATIC REPRESENTATION OF BASIC BGP/MPLS IP VPN OPERATION

This is done in a way which ensures that the routes from the VPNs remain distinct and separate, even if two VPNs have overlapping address spaces. The term "IP" in "IP VPN" is used to indicate that the PE receives IP datagrams from the CE, examines their IP headers, and routes them accordingly.

Furthermore, each route within a VPN is assigned a Multiprotocol Label Switching (MPLS) label, and when BGP distributes a VPN route, it also distributes an MPLS label for that route. Before a customer data packet travels across the operator's backbone, it is encapsulated with the MPLS label that corresponds (in the customer's VPN) to the route that is the best match to the packet's destination address. This MPLS packet is further encapsulated (e.g., with another MPLS label or with an IP or Generic Routing Encapsulation (GRE) tunnel header) so that it gets tunnelled across the backbone to the proper PE router. Therefore, the backbone core routers do not need to know the VPN routes, i.e., they may remain completely VPN-agnostic (cf. [RFC4364]).

As far as BGP/MPLS IP VPNs which cross multiple ASes are concerned, [RFC4364] envisions several different mechanisms for inter-domain VPN provisioning, with different levels of scalability and management

overhead. However, all of those mechanisms also assume a tight cooperation between the involved network operators, advancing the provisioning of multi-domain MPLS/BGP IP VPNs to become a non-trivial contractual and operational matter.

Regarding QoS, the provisioning of BGP/MPLS IP VPN traffic with the appropriate assurances completely depends upon the QoS provisioning mechanisms deployed in the underlying IP transport network. If the network operator offering BGP/MPLS IP VPN does have the appropriate capabilities, extending the same QoS service to the provisioned VPNs represents a very straightforward, easy-to-solve task.

9.2.1.2. BGP/MPLS IP VPN Applications

The primary use case of BGP/MPLS IP VPNs is to enable VPN connectivity to customers who obtain IP backbone connectivity from a network operator with whom they maintain a long-term contractual relationship for all of their connected sites. Typically, customers interested in this type of VPNs are governments and large public sector entities, as well as large corporations the physical sites of which are within the reach of a number of ISP willing to offer VPN connectivity to their customers.

One of the main benefits of BGP/MPLS IP VPNs lies in its great flexibility in terms of supporting a variety of possible topologies, facilitating the creation of different company-internal network configurations. Additionally, as VPNs using this technology are managed by the operator, most part of VPN provisioning complexity is being outsourced from the customer's domain, which might represent an attractive option for many customers interested in optimizing the cost structure of their IT operations.

However, BGP/MPLS IP VPN convenience is also associated to a number of technical and operational limitations. Firstly, whereas this technology does enable isolated, private address space connectivity of geographically distributed sites (which comprises only one part of the VPN concept), it does not in any way provide for the cryptographic protection of the carried traffic, therefore requiring full trust in the network operators involved in the provisioning of the VPN, and also making private traffic potentially subject to legal interception by law enforcement agencies. In other words, if customers do have a strong need for keeping their VPN communications strictly private, they will need to implement an inter-site authentication and encryption scheme themselves, which however substantially reduces the attractiveness of BGP/MPLS IP VPNs as an outsourcing solution. Secondly, the feasibility of organization-wide BGP/MPLS IP VPN deployment crucially depends on the question of whether it is possible to connect all individual sites via a single network operator offering such VPN functionality, or at least a set of operators who wish to offer such a service in cooperation. This second aspect is particularly limiting to organizations which operate globally, as it is often hardly feasible to find a set of operators who can seamlessly provision BGP/MPLS IP VPNs across a highly geographically distributed set of sites.

9.2.2. CUSTOMER MANAGED VPN SOLUTIONS

9.2.2.1. IPsec VPNs

IPsec VPNs are nowadays among the most powerful and most widely deployed VPN technologies, the functionality and applications of which we discuss in more detail in this chapter.

9.2.2.1.1. IPsec Functionality

IPsec functionality builds on 2 fundamental parts:

1. Key exchange, and

2. Data transport.

Key Exchange:

For exchanging keys, the Internet Key Exchange (IKE) protocol, version 1 and 2, is used. IKEv2 was specified in December 2005 in RFC-4306 and contains the following enhancements to IKEv1:

- Protection from DoS attacks which use spoofed packets, and
- Detailed specification of NAT traversal mechanisms.

Data Transport:

IPsec has two basic modes of operation, which are denoted by their headers:

- Authentication Header (AH), and
- Encapsulating Security Payload (ESP).

In AH mode, only the integrity and the authenticity of the IP packets are guaranteed, but the payload is NOT encrypted. AH is an extension of the IP header.

IN ESP mode, integrity, authentication, and confidentiality of the packets are guaranteed. ESP operates on top of IP and thus has an IP protocol number of its own (50).

Each of the basic modes of operation can either be operated in transport- or in tunnel mode. In tunnel mode, the original IP header is copied and thus preserved, such that a tunnel can be built and the original IP header can be reconstructed ad the receiver's side. This is not possible in transport mode. The different modes are depicted in Figure 35 and Figure 36.

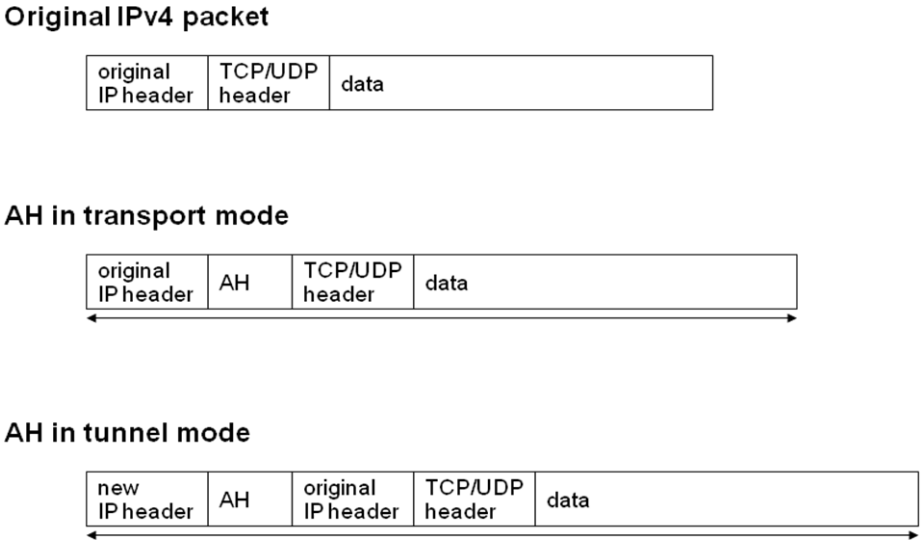


FIGURE 35: AUTHENTICATION HEADER (AH).

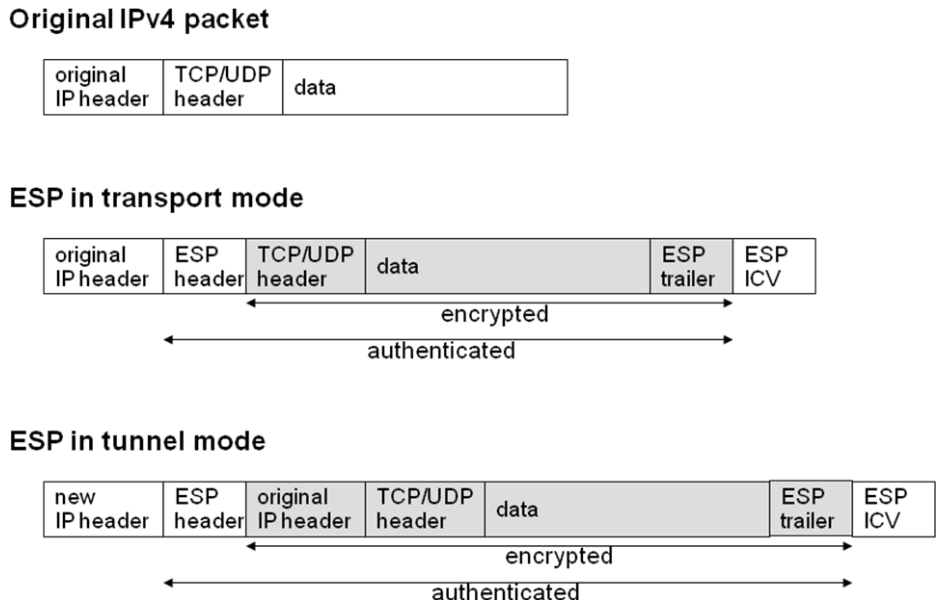


FIGURE 36: ENCAPSULATED SECURITY PAYLOAD (ESP).

In the original IPsec specification [RFC1825] NAPT[32] traversal was not considered. Consequently, due to the ubiquity of NAPT devices, the original IPsec specification is often not applicable nowadays.

This shortcoming was fixed in later IPsec specifications, namely RFCs 3715, 3947, 3948, and 4306. According to these standards, in the ESP mode of operation, a UDP header is inserted right in front of the ESP header field, which ensures reliable VPN operation across NAPT boxes. This is exemplarily shown on the ESP in tunnel mode in Figure 37.

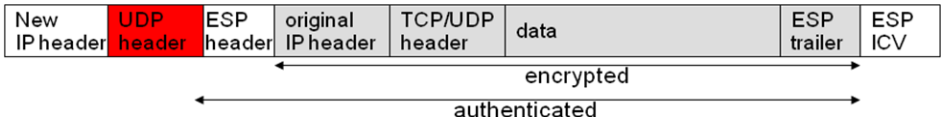


FIGURE 37: ESP IN TUNNEL MODE WITH NAPT EXTENSION.

9.2.2.1.2. IPsec Applications

As shown in the previous chapter, IPsec is a very powerful specification and as such has its typical applicability in the area of Site-to-Site VPNs, i.e., where private networks, mostly for companies geographically split over a number of locations, need to be seamlessly connected together. Indeed it is possible to implement also Site-to-End and End-to-End solutions by means of IPsec, especially by the mechanism shown in Figure 37.

Another factor for IPsec being the predominant solution for Site-to-Site VPNs is the fact that it is very powerful with respect to security related features, as it provides *per definitionem* for the for authenticity, integrity, and confidentiality of the carried traffic.

Furthermore, IPsec is a very efficient technology. Due to its two different headers, providing different levels of security, which can be operated in two different modes, either for simple data transfer of for tunnelling entire sub-networks, overhead is kept very small. Additionally, a single IP packet of the associated network

[32] *Network Address and Port Translation*, often untruly referred to as „NAT“ which, according to IETF, is wrong since the vast majority of NA(P)T boxes also translate the port and consequently have to be referred to as NAPT.

is always mapped to exactly one IP packet on the transport network which avoids additional delay and is specifically advantageous for (near) real-time applications like VoIP.

On the other hand, IPsec introduces a significant administrative overhead and requires a substantial planning effort, especially for large, i.e., physically strongly distributed deployments. As shown in Figure 38, IPsec needs to be configured on each edge-router for every pair of IPsec associations[33], which represents a considerable effort.

Also, special care must be taken with respect to the planning of the VPN topology. Besides managing the expected volume of traffic of the virtual network, VPN designers must also considered how to map this traffic to the (capacity of the) underlying physical network topology.

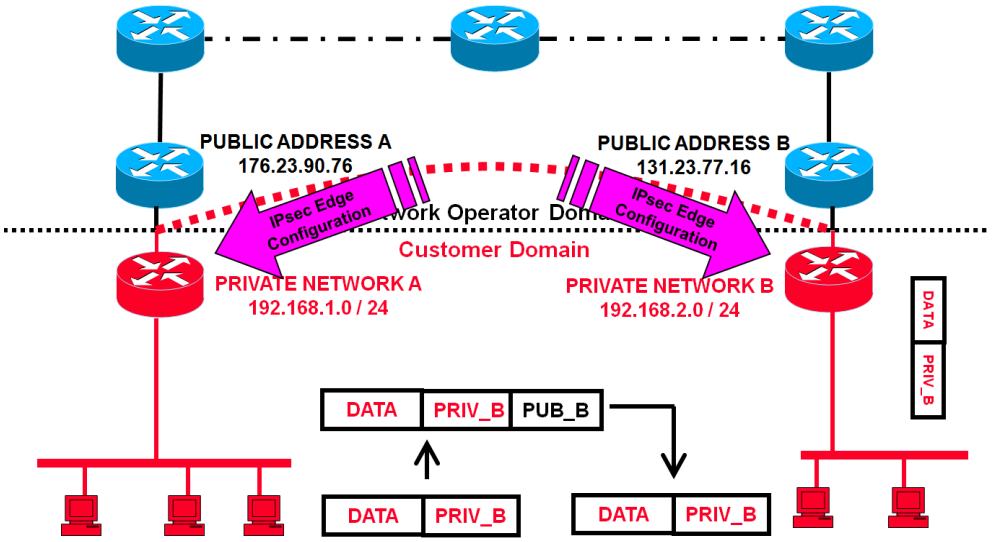


FIGURE 38: IPSEC CONFIGURATION.

Due to the fact that a full-meshed topology is usually not possible due to scalability issues, other topologies have to be considered, which could be:

- Star- (Hub-),
- Ring-,
- Tree-

topologies, or a combination of those.

Concerning QoS, IPsec has a significant drawback compared to an operator managed solution. While in the operator-managed scenario, the provider-edge (PE) router can analyze the packet and assign the corresponding QoS according to some rules, this is not so easily achievable in the IPsec scenario where the packet is encapsulated by the customer's border router. However, since there is a 1:1 mapping between IP packets of the virtual network and IP packets of the transport network, it seems feasible for the customer's border router to examine the IP packet of the virtual network and assign the appropriate QoS marker, e.g. by means of *DiffServ* codepoints, to the IP header of the transport network. Of course this needs an agreement between the customer and all ISPs *en route* to the destination domain.

[33] IPsec is always configured between exactly two communication endpoints

9.2.2.2. Other technologies

9.2.2.2.1. SSL-VPNs

The term "SSL VPN" is not clearly defined and is used by different parties to denote very different things. The only common understanding of this term is that TLS[34] is used for data transport.

We identify two main types of "SSL-VPN" systems:

1) The VPN mechanism works much like described in chapter 9.2.2.1.1 (IPsec functionality) with the difference that TLS, instead of IPsec, is used as transport protocol, i.e. whole IP packets are encapsulated and transmitted with TLS as the Layer-3 tunnelling mechanism instead of IPsec.

2) The data stream of the application is conveyed by TLS, either directly by the application or indirectly by tunneling the application stream over TLS in which case there is a 1:1 relation between the application's data stream and the TLS connection.

Sometimes also systems which offer only HTTP access via TLS (=HTTPS), where the user can access resources on the server in a groupware style, are referred to as "TLS-VPN"; however such systems are out of scope of this deliverable.

System (1):

The vast majority of all implementations use port 443 for the transport connections, which is the port used for HTTPS. Since HTTPS is usually allowed even in very restrictive firewall-configurations, this solution offers the important advantage that it can pass through most firewalls and that is also compatible with NAPT boxes in the path, which makes it perfectly suitable for Site-to-End applications.

The large drawback of this solution is that QoS is very hard – if not impossible – to enforce, due to the fact that a packet oriented protocol (IP) is encapsulated into a single stream of data (TLS, or TCP, respectively) and that therefore the 1:1 mapping of packets from the assigned (virtual) network to the transport network is lost. For the same reason, the timing advantage of UDP is lost as well, effectively rendering the use of UDP useless, which is an enormous drawback for (near) real-time applications like voice, video, and gaming.

System (2):

This system has a very charming property that it needs no installation at the client side at all. The basic concept is as follows: If the application in question supports a secure channel (e.g., TLS), usage of the same is enforced by the VPN server and no further action is taken. An example for this is a Remote Desktop Connection (RDP) which has got built-in mechanisms for authentication and encryption. If the application in question does NOT support a secure channel, a TLS tunnel is set up automatically and the application is forced to use this tunnel. In this case, a local port simply becomes the local tunnel endpoint (TCP or UDP, i.e., on OSI Layer 4) and the application connects to this port on the local-host rather than to the remote system directly; the remote end of the tunnel itself is connected to the remote service where the application originally would have connected to. The tunnelling software could either be pre-installed on the local system or it could be a piece of Java-code which is automatically transferred by the web-browser such that this system would work on any client which has a web-browser with Java enabled.

Concerning QoS, for UDP the same is valid as previously stated for System (1). However, regarding QoS enforcement otherwise, this solution has the advantage that each application stream is mapped to

[34] Transport Layer Security (RFC-2246), formerly known as „Secure Sockets Layer (SSL)"

exactly one stream of data on the transport network. Consequently it seems easily achievable for the tunnelling software, e.g. OpenSSH, to map the QoS from the associated network to the transport network.

9.2.2.2.2. Layer-2

Layer-2 VPNs connect (usually a single computer) to a site such as it were connected to a local LAN which means that this mechanism is usually used for Side-to-End and End-to-End scenarios.

The advantage of this scenario lies in the fact that broadcasts and/or multicasts are often also supported which increases the number of supported services. A typical example for this is browsing the network neighbourhood for finding devices on the network like printers and media servers. Most games which support multiplayer LAN games also need this.

On the other hand, this mechanism does not scale very well because broadcasts, which contribute a considerable amount of traffic to today's LANs, also need to be transmitted to all active hosts. In a large network, only a limited number of users should be allowed to connect to such a virtual subnet.

PPTP:

Since a PPTP implementation ships with the Microsoft Windows product family, it is most likely the most common VPN mechanism used by end systems. Originally it was developed by a consortium which consisted of Microsoft, Ascend Communications (today part of Alcatel-Lucent), 3Com, and others. In 1999 it was published as [RFC2637]. Since 2005, a "Microsoft compatible" PPTP implementation is also available for Linux-type operating systems.

PPTP uses a TCP control connection on port 1793 and the payload is encapsulated by General Routing Encapsulation (GRE) packets. In the PPTP implementation, GRE is not fully compliant to the GRE specification as used, e.g., in MPLS.

GRE is a very slim protocol; in its minimal version the header consumes only 4 bytes. Even though an optional "key" field is defined to distinguish between sessions, this field is usually not used, assumingly because it would increase the overhead. Consequently, if GRE is used directly on top of IP, which is often the case, its functionality in conjunction with NAPT is very limited in the sense that the NAPT device can only handle a single GRE (and thus also PPTP) tunnel, as the session identifier (i.e., the UDP/TCP port) is missing. This fact makes it a bad choice for companies which want to offer their mobile employees VPN access to the corporate network.

Concerning QoS, the 1:1 mapping between packets of the assigned network to packets of the transport network is advantageous in the sense that potential QoS algorithms could easily map the QoS to the transport network.

L2TP:

In simple terms, L2TP could be described as "advanced PPTP". It is used on top of UDP (port 1701) and control- and data packets are transported within the same connection. Consequently the protocol can pass NAPT devices which support L2TP. It is common to use PPP inside the L2TP tunnel. Since L2TP does not support encryption, IPsec is also often used to secure the connections.

The "improvement" over PPTP is that a number of tunnels, as well as a number of sessions within each tunnel, are supported. This makes it well suitable for cable- or ADSL network providers.

Concerning possible QoS mechanisms, the same is valid as noted for PPTP in the previous section.

9.2.3. SUMMARY

Summarizing our outline on the technologies for building network provider operated and customer operated VPNs, we conclude that the currently available solutions like BGP/MPLS IP VPNs, IPsec-based VPN, SSL-VPNs, etc., offer an exhaustive plethora of very thought-out and mature options, which in most cases cover the customer requirements very well.

However, as an exception to this, we also note that QoS enablement has not yet become a fundamental part of wide-area VPN service specifications, especially in cases where VPNs span multiple administrative network domains. Therefore, we identify *inter-operator QoS enablement for VPNs* as a current technological stumbling block, which will be overcome by the architectural inter-domain solutions for QoS currently under development in the EU FP7 ETICS project.

9.3. TECHNOLOGIES FOR TRAFFIC ENGINEERING, RESILIENCY AND QOS PROVISIONING IN SUB-IP NETWORKS

Core sub-IP (backbone) networks have been evolving from many years driven by three main triggers:

- a continuous increase of raw bandwidth at the Transport Plane for new emerging application needs (up to 10Gbps for end-users and 40/100Gbps for the core backbones)
- the availability of highly dynamic and automatic procedures at the Control/Management Plane for the provisioning, survivability and management of network services
- the progressive simplification of the technologies overlays at the transport plane (packets, TDM, optical, fibres), and the continuous trend towards Transport Ethernet and Wavelength Switching.

Traffic Engineering (TE), network resiliency and QoS provisioning are the main objectives of Control Plane architectures elaborated in the last years. In fact, they conjugate both network operators' needs for a more efficient utilization of the heterogeneous networks resources (i.e. Traffic Engineering) and the users' needs for tailored network services (i.e. based on QoS specifications), with differentiable survivability levels (i.e. with automatic protection and/or restoration of the carried traffic). All with possibly more and more dynamism/automatism in the network, controlled (triggered) at standardized network interfaces.

A common and well-founded desideratum for Network Control Plane is for a generalized architecture that can ease the operation of these complex systems by network operators and allow the on-demand access by end-users. Moreover, such a generalized architecture must seamlessly adapt to the large geographical networks, and interface the emerging Wavelength Switching Optical Networks (WSON) deployed on the backbones to other legacy access transport technologies like Ethernet, SDH/SONET or even IP packet networks.

Among the Control Plane architectures developed for telecommunications systems in the past decades, IETF GMPLS is considered the standard de-facto for managing the physical core tunnelling technologies of both Internet and L1/L2 service providers. The GMPLS architecture as defined within the IETF CCAMP WG is designed to provide automatic provisioning of connections with traffic engineering, traffic survivability (i.e. protections, restorations), automatic resource discovery and management. The core GMPLS specifications are fully agnostic of specific deployment models and transport environments: they are built upon the MPLS procedures and broaden the applicability of those mechanisms beyond the single data plane envisioned by the original MPLS specifications. However, due to the features of the underlying Transport Plane technologies, some specific procedures and protocol extensions have been defined in GMPLS to control transport networks as diverse as SDH/SONET, DWDM-based OTNs, OTNs incorporating G.709 encapsulation, and Ethernet. This process of enhancement of the GMPLS protocols foundations is still active in IETF, because it needs to cope with new emerging transport technologies, such as the Carrier-grade Ethernet (i.e. PBB-TE) and the dynamically reconfigurable optical devices (ROADMs) which are key elements in WSONs.

The IP-based GMPLS Control Plane architecture is built around three main functionalities:

- the transport resource discovery with a link-scope via link-management (LMP) and with network scope via intra-domain and inter-domain routing protocols (e.g. OSFP-TE with ASON extensions for hierarchical routing, E-NNI-OSPF);
- the transport resource reservation (e.g. RSVP-TE at the different network interfaces – UNI, I-NNI, E-NNI);
- the constraint-based TE path computation (e.g. based on a centralized or distributed Path Computation Element architecture – PCE).

All the GMPLS protocols and engines deal with topology, reachability, and addressing details about the controlled transport networks and run on a data network (Signalling Control Network – SCN) that can be either separate from the transport network (e.g. out-of-fibre), or shared with it (i.e. in-fibre/in-band or in-fibre/out-of-band, i.e. on a separate wavelength).

The IETF Path Computation Engine architecture (PCE, RFC4655, RFC4657) elaborated by the PCE WG complements the GMPLS architecture by defining a basic toolbox for decoupling the computation function (possibly based on highly complex algorithms, and therefore CPU consuming) from the route usage functions by the raw GMPLS protocols. The PCE architecture is built around two major entities:

- the Path Computation Client (PCC), which is the entity requesting path computation services in the form of explicit routes matching the n-tuple <request type, ingress/egress, constraint>, and may be embedded in Network elements, NMS, Diagnostic tools or other PCE (in a inter-PCE cooperative model)
- the Path Computation Element (PCE), which is the entity performing path computations on behalf of PCC clients, and to this purpose accesses network topology data (e.g. via TEDB), performs graph-theoretic computations on that data, coordinates the route assembling process with other adjacent PCEs in case the full network topology is not locally known.

The mechanisms and tools made available by these Control Plane architectures simplify some critical issues of the sub-IP network operation and maintenance, in the perspective of carriers providing their users with network connection services installed and maintained in a more and more dynamic, automatic way. Some of these issues are briefly discussed in the following subsections.

9.3.1. ISSUE #1: SUPPORT OF MULTIPLE SWITCHING TECHNOLOGIES UNDER THE SAME NETWORK CONTROL PLANE

Backbone networks are composed by transport nodes offering multiple data plane layers of either the same or different switching technologies. These networks are referred to as multi-region and multi-layer networks (MRN/MLN). In standard GMPLS, a MRN is a Traffic Engineering (TE) domain supporting at least two different switching technologies (e.g., lambdas and Ethernet, or lambdas and SDH/SONET). A GMPLS MRN/MLN is controlled by a single Control Plane instance to limit the complexity and processing at protocols level and overcome the inefficient overlay of different instances and partitions.

IETF CCAMP and MPLS WG produced several solutions to cast multi-region and multi-layer networking in the GMPLS Control Plane. Two main solutions have been identified:

- *LSP nesting*, which consists of aggregating LSPs to create a nested hierarchy [RFC4206, RFC5212, RFC5339]. Three different kinds of LSP nesting are available:
    - *Hierarchical LSPs (H-LSP),* which are created in one layer and appear as a TE-links in higher layers. One or more LSPs in a higher layer can traverse this H-LSP as a single hop.
    - *Forwarding Adjacency LSPs (FA-LSP),* which are H-LSP that are advertised as a TE links in the same TE domain. This kind of approach could be used in case of GMPLS integrated model. An FA-LSP may also be advertised in other TE domains, obtaining a sensible improvement in terms of scalability of inter-domain routing and signalling.
    - *Virtual TE-Links,* which are TE links between two upper layer nodes that are not actually associated with a fully provisioned FA-LSP in a lower layer. A virtual TE link represents the potentiality to setup a FA-LSP in the lower layer, and as soon as an upper-layer LSP tries to use it through signalling, the underlying FA-LSP is immediately signalled and provisioned (provided there are available resources in the lower layer).
- *LSP stitching,* which consists in building an LSP from a set of different "LSP segments" (S-LSPs) that are connected together in the data plane, in case of switching technologies with equal granularity. While LSP nesting allows more than one LSP to be mapped to an H-LSP (or FA-LSP), with stitching at most one LSP may be associated with an S-LSP.

The PCE architecture is applicable also to the MRN/MLN scenario and in this context it implements the inter-layer route computation task by processing topology and resource information from the different layers. Due to the potential complexity of the multi layer routing algorithm, the PCC-PCE signalling protocols (PCECP) is under extension by PCE-WG to specific a set of MRN/MLN route request specification and the related procedures (ref. draft-ietf-pce-inter-layer-ext-03.txt)

9.3.2. ISSUE #2: FAST RECOVERY OF NETWORK SERVICES

Concerning the network service recovery, the GMPLS protocols can make use of a set of procedures to provide protection or restoration of the data traffic, and the PCE is requested to compute node and/or link and/or risk disjoint paths to re-route the traffic appropriately. There are many recovery schemes mostly

deployed in the intra-domain scenario. They require different levels of resource/route pre-provisioning and could be categorized by time and resource consumption, setup vulnerability, quality of protection, packet loss, failover coverage [RFC 3469]. A list of some common recovery strategies is provided in the following table:

Recovery type	**Description**	**Recovery time [0, 10] scale**	**Resource consumption [0, 10] scale**
Full LSP Re-routing	This is a recovery scheme without path pre-computation, which is the most flexible solution.	10	2
Pre-planed LSP Re-routing without Extra-Traffic	This is a recovery scheme resource pre-selection and without resource pre-allocation. It can be referred also as “shared mesh” recovery (there is no resource pre-selection).	6	4
LSP Protection with Extra-Traffic	This is a recovery scheme with resource pre-allocation. Protection LSP can be used for transporting additional low priority traffic. Resource consuming optimization is available by N:M path mapping type.	4	6
Dedicated LSP Protection (1+1)	This is a recovery scheme with resource pre-allocation which doesn't allow sharing of the recovery resources. It is the best time performance and the less flexible option.	2	10

TABLE 1: COMMON GMPLS RECOVERY SCHEMES.

Every recovery model has advantages and disadvantages but from a user point of view protection schemes achieve the best performance. However, using protection schemes, especially in dedicated 1+1 scheme, implies a higher consumption of network resources in entire operator's network.

Recovery types can be also divided by spatial factor. A label-switched path may be subject to local (span), segment and/or end-to-end recovery [RFC 4426]:

- span protection is protection of link between two neighbouring switches,
- segment protection refers to recovery of an LSP segment (Sub-Network Connections in the ITU-T terminology),
- end-to-end protection refers to protection of entire LSP from the ingress to the egress port.

The main issue with recovery schemes is their effectiveness and applicability in inter-domain scenarios (both in terms of resource consumption and of time to recover from failure). In this context, the IETF CCAMP WG limited its activity to a problem analysis (ref. RFC5298).

9.3.3. ISSUE #3: INTER-DOMAIN NETWORK SERVICES

Network operators are used to partition their infrastructures into domains to protect business practices or to comply with managerial and/or policy issues or to represent the transport network heterogeneity in terms of technologies or control and management models/mechanisms.

The GMPLS architecture by IETF and the ASON by ITU-T provide two different models to cope with the multi-domain scenario in terms of routing and signalling.

In IETF GMPLS/PCE, the routing model is based on the "flat" Autonomous System peering at given border nodes, augmented with inter-PCE communications in case of distributed inter-domain route computations (e.g. ref. to Backward-Recursive PCE-Based Computation – BRPC – defined in RFC 5441). The procedures of Interior and Exterior Gateway Protocols guarantee the flooding of reachability information among AS-es and routing areas (including TE information within a routing area). Concerning signalling, the GMPLS signalling (RSVP-TE) is intrinsically end-to-end, i.e. specific inter-AS/domain actions and procedures are implemented in border nodes on the reception of the single-layer signalling flow.

On the contrary, ITU-T ASON is built on top of sharp network reference points (the domain boundary ones in particular, UNI and E-NNI), which reflect the administrative partitioning and break the signalling and routing flows accordingly. The ASON routing model is further completed with the concept of hierarchically nested Routing Areas (RAs), in which I-NNI routing control domains represent the base layer and each subsequent E-NNI layering above constitutes an ancestor level (ref. Figure 39). At each hierarchical level, one OSPF instance and area is defined (as per G.7715).

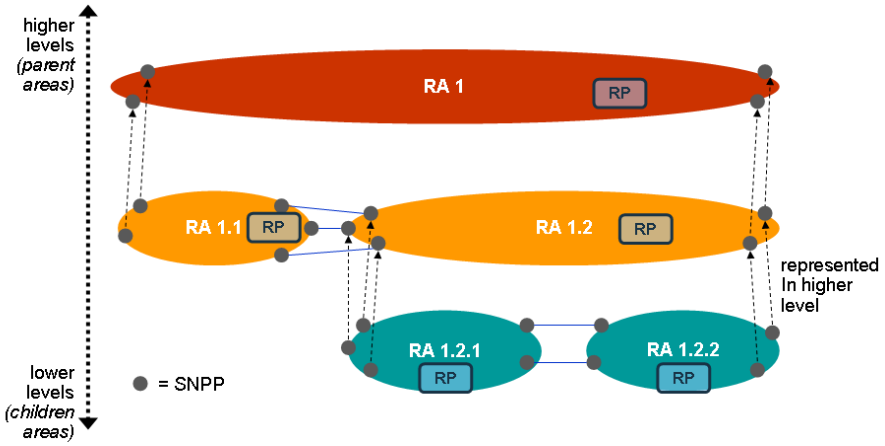
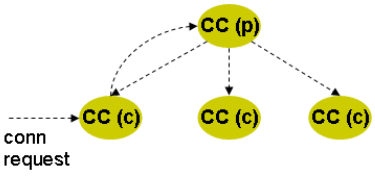


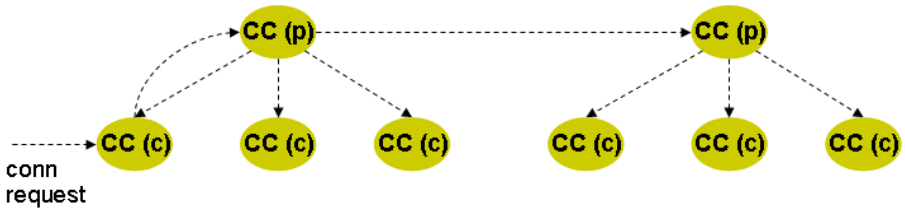
FIGURE 39: ASON ROUTING HIERARCHIES.

Inter-domain operations in the ASON framework are based on the concept of "federation", i.e. the community of domains that co-operate for the purposes of connection management. Three types of federation are identified in ASON (ref. Figure 40):



a)

b)



c)

FIGURE 40: ASON FEDERATION MODELS: A) JOINT FEDERATION, B) COOPERATIVE, C) COMBINED FEDERATION.

- **joint federation model**: one CC ("parent") coordinates the operation of the other CCs involved ("children")
- **cooperative model**: no single point coordination; each CC takes care of its part of connection
- **combined federation model**: a mix of the two above, i.e. the CCs are partitioned into groups where the joint federation model is applied; inter-group operations are coordinated according to the cooperative model.

9.3.4. ISSUE #4: SECURITY ISSUES IN NETWORK CONTROL PLANE

The Control Plane controllers are an attractive target for intruders who want to disrupt or spoof or gain free access to telecommunications facilities, just like the network elements in the Transport Plane. A brief summary of the main attacks that can be launched to (G)MPLS-based networks is provided in Table 2 .

	Main attacks
[GMPLS/PCE] Control Plane	• Traffic Analysis • Routing topology spoofing • Theft of Service • Unauthorized LSP creation • Unauthorized PCE requests • Interception of Service • Control Plane messages interception • Denial of Service • Storms of messages with critical CPU utilization and/or rapid memory exhaustion (both up to system crash), e.g. • Storms of LSP creation (RSVP, LDP), • Storms of Hello messages (LDP, OSPF, LMP), • Storms of PCEP requests • Storms of Graceful restarts (RSVP, OSPF) • Storms of diagnostic procedures (e.g. OAM MPLS-ping, LMP link verification)
Data Plane	• Traffic Analysis

	• Unauthorized observation of data traffic (Man in the Middle, M-i-M) • Unauthorized Traffic Pattern Analysis • Interception of Service • Unauthorized Deletion • Degradation of a provider's service quality • Spoofing and replay of a provider's or user's data • Theft of Service • Data traffic modification • Denial of Service • Injection of inauthentic data into a provider's or user's traffic stream

TABLE 2: MAIN ATTACKS FOR GMPLS-BASED NETWORKS.

Threat	Major Impact	Current mitigation mechanisms
Traffic Analysis	Service disclosure	IPSec ESP with encryption (RFC2406)
Denial of Service	Service disruption	IPSec ESP with authentication (RFC2406) RSVP integrity (RFC2747) OSPF Crypto Auth (RFC2328)
Interception of Service	Service deception	IPSec ESP with encryption (RFC2406) IPSec ESP with authentication (RFC2406) RSVP integrity (RFC2747) OSPF Crypto Auth (RFC2328)
Theft of service	Service usurpation	IPSec ESP with authentication (RFC2406) RSVP integrity (RFC2747) OSPF Crypto Auth (RFC2328)

TABLE 3: MAJOR NETWORK CONTROL PLANE SECURITY THREATS AND AVAILABLE MITIGATIONS.

Therefore, core (backbone) networks can be exposed to security threats similarly to the common packet networks like the Internet, but with a bigger service impact because of the higher granularity of the disrupted entities (lambda/fiber connections, full node controllers, network management systems, etc.).

The mechanisms available in the state of the art to protect the GMPLS and PCE mostly rely on the coordination of secrets, keys, or passwords between sender(s) and receiver(s) of protocol messages. The GMPLS protocols have specific extensions to implement these mitigation mechanisms (ref. Table 3), which proved to work properly in intra-domain environments, i.e. when the administrative boundaries are not crossed and the establishment of trust relationships and exchange of keys can be simple. Another stronger mitigation strategy consist of establishing IPsec bidirectional adjacencies between protocol peers with the

desired level of protection (ESP), like in OIF UNI/E-NNI. This approach separates the protection mechanism from the Control Plane, thus avoiding per-protocol extensions, but is not flexible (security adjacencies are pre-established) and can clash with protocol specific behaviours encoded in the protocol headers (e.g. use of Router-Alert in classic RSVP).

Moreover, when the sender and receiver lie in different administrative domains, the security coordination between network administrators is not automatic or provided through the use of a specific protocol. In this context, most of the security protection is delegated to operator-defined policy rules/filters, like:

- avoid the distribution of sensitive (internal) topology information (TE routing data),
- hide the explicit route objects related to internal TE resources (e.g. Path Key sub-object, RFC5520 and RFC5553),
- rate-limit LSP setup requests or error notifications from a particular domain
- disallow recording of hops within the domain (RRO) or drop the domain-internal parts out of the RRO
- limit the end-to-end connectivity verification procedures (where available) or to bypass them at domain level

10. ANNEX B: CURRENT SERVICES AND BUSINESS MODELS - DETAILS

10.1. BANDWIDTH ON DEMAND (BOD) IN SUB-IP NETWORKS

From a long time network operators have been consolidating means for dynamic and efficient Bandwidth-on-Demand services (BoD) both towards their "power" (business) end-users and with peering operators (inter-carrier problem, core ETICS focus). Despite of the numerous BoD solutions developed in specific research networking contexts[35], ASON/GMPLS and PCE promises to be jointly the more effective BoD solution for commercial inter-domain backbone networks.

Until today, BoD implemented through ASON/GMPLS and PCE have been just demonstrated by many carriers and vendors around the world, above all in OIF interoperability events like the latest 2009 OIF Worldwide Interoperability Demonstration "Enabling Broadband On-Demand Services" (http://www.oiforum.com/public/OIF_Networking_Demo_2009.html).

Nevertheless, BoD technologies have not been widely deployed in operational networks up to the end users, above all in commercial infrastructures. They have been rather used in an intra-carrier mode by the Network Operators to

- minimize their CAPEX/OPEX on the infrastructure (also supported by the use of traffic restoration techniques instead of hw-embedded protection schemes)
- drastically reduce time-to-market for connectivity services
- homogenize the co-existence of equipments and technologies by different vendors deployed under the same administrative ownership (i.e. more interest by carrier on the inter-vendor issues than on the inter-carrier).

Despite the current GMPLS User-Network Interfaces (UNI) can support dynamic service requests from users (both as specified by IETF and by ASON/OIF), network operators still tend to keep full control and supervision of the transport service provisioning via NMS. One of the most advanced use-cases of commercial deployment of BoD is the US operator Verizon Business, which commercializes (just in New York area and for business customers only) SONET and GbE BoD tunnels with one working day time-to-market. This service is obtained through a GMPLS Control Plane mediated by a NMS engine [http://www22.verizon.com/wholesale/solutions/solution/bod.html].

Main stumbling blocks for the effective and large deployment of GMPLS and PCE-controlled inter-carrier network services basically derive from:

- the immaturity of these architectures to operate in real inter-carrier business cycles, with consolidated procedures and mechanism for

[35] Example of BoD in research networking that have been rolled out in the last years are: AutoBAHN by GÉANT2, OSCARS by ESnet, DCN by Internet2, D-RAC by Nortel/SurfNet, UCLP by Canarie.

  - a seamless end-to-end network service negotiation and installation (across heterogeneous transport plane technologies)
  - an efficient end-to-end network service monitoring
  - proper mechanisms for Authentication, Authorization and Accounting (AAA) integrated with the dynamics of the GMPLS-controlled network service
- the lack of huge users' demands and long-term incentives for carriers to offer dynamic and automated network connectivity services
- the inappropriateness of the peer Internet AS-model and the deriving routing mechanisms (which are the foundations of the GMPLS architecture by IETF) to the specific inter-carrier context, made of distinct and often competing administrative domains with strict requirements on internal topological data privacy.

This framework might be definitely evolved by the emerging Cloud computing services for the Future Internet. Many commercial frameworks are emerging in the Cloud arena to cope with a large set of potential beneficiaries, both as end-users and as technology developers. They are produced and operated by some of the major IT players[36] (Over-The-Top, OTT), who offer and implement different types of resource abstractions, ranging from the virtualized hardware platforms (Infrastructure as a Service – IaaS) up to the distributed development platforms (Platform as a Service – PaaS) and application layer (Software as a Service – SaaS). Clouds highly rely on the on-demand and pay-by-use paradigms of networked IT resources over the Internet, with a wide cross-impact on applications, service platforms, computing, storage and network hardware resources. The network connectivity service is a basic enabler for Clouds, but it is often treated as an "always-on" and over-dimensioned dumb connectivity pipe, completely transparent to the service orchestration flow both in terms of service and, subsequently, of carriers' revenues.

The increasing scale of the Clouds, both in geographic dimensions and in size (i.e. users), and the need to guarantee high throughputs, QoS differentiation and service resiliency to the Cloud services across all the involved components, including the network, are strong motivations for the deployment of Control Plane technologies among carriers.

[36] Amazon, Sun, Google, Yahoo, Microsoft just to cite some of them.

11. ANNEX C: Scenarios Description Methodology

This annex describes the template used to collect the ideas from the different partners to specify the scenarios that are described in the main document.

11.1. TEMPLATE MODEL

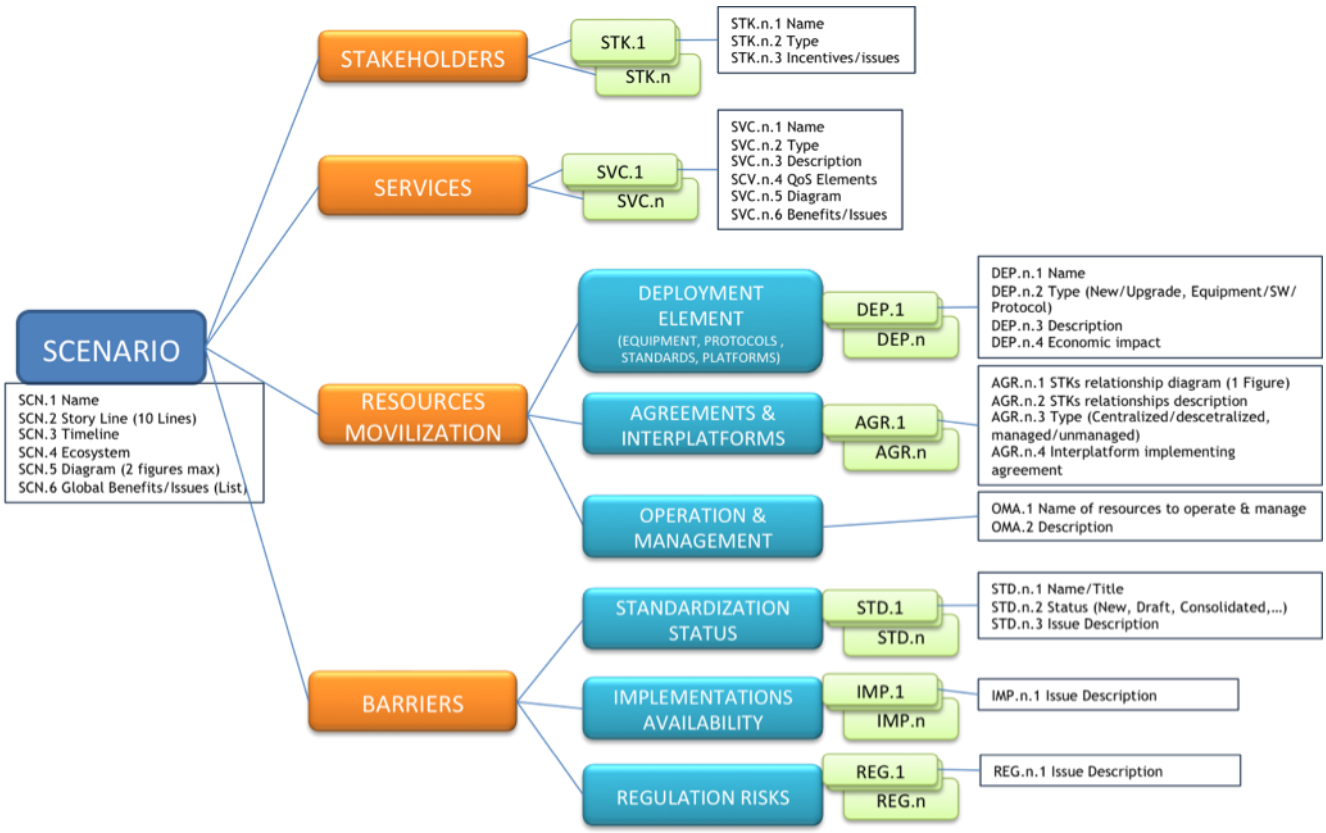


Figure 1. Scenario Template Overview

11.1.1. SCENARIO DESCRIPTION

- SCN.1: Name:

[Descriptive name of the scenario]

- SCN.2: Sort description: (10 lines max)

[Describe as clear as possible the proposed scenario, dealing with its main benefits, involved services and stakeholders as well as the technical solution or agreements needed]

- SCN.3: Time-line:

[Possible values: – Sort term: from 0-2,5 years – Medium term: 2,5-10 years – Long term: 10 years or more] [Starting from the end of the ETICS project]

- SCN.4: Ecosystem:

[Possible values are items in section 4 of this document. Examples: Current Internet, Transition from IPv4-IPv6, Internet IPv6, others… Add others and explain them if its needed]

[Ecosystems are such groups of scenarios or those complex environments where there will appear a new set of scenarios]

- SCN.5: Diagram: (2 figures max)

[Representative diagrams that describing the current scenario. Example: could depict the interconnection relationship and/or the technical solution]

- SCN.6: Global Benefits/Issues:

[Provide a bullet-list with the main benefits and Issues of the scenario from a general point of view with impact in the overall Internet (5 items per bullet-list max). A scenario can be composed by several services but those services should not be covered in this point (use SVC.n.6 for this)]

11.1.1.1. STAKEHOLDERS: (from 1 to N)

[List of the stakeholders involved in the proposed scenario]

[Note: relationship between stakeholders must be described in the Agreement & Inter-platform section (AGR.n) not here]

- STK.n.1: Name:

[Name of the stakeholder]

- STK.n.2: Type:

[Select from: CDN, OTT Provider, Telco Operator, Vendor, Research Entity, Virtual Operator, Final Customer, other (please specify)]

- STK.n.3: Incentive: (max. 3 lines per item)

[Provide a bullet-list with the main incentives for each stakeholder explaining the incentives of the service described for the stakeholders]

11.1.1.2. SERVICES: (From 1 to N)

[Enumerate and describe the different services that compose the main Scenario]

- SVC.n.1: Name:

[Service name]

- SVC.n.2: Type:

[Service Type: High Quality Teleconference, Private Cloud Services, Cloud rendering, Online gaming, Remote Tele-operations, Tele-presence, IMS… (Add any other if needed)]

- SVC.n.3: Description: (30 lines)

[Provide a detailed description about the service including its main functionality]

- SVC.n.4: QoS Elements:

[Provide a detailed description about the elements provided in this service that guaranties the quality (any quality value) in the service]

- SVC.n.5: Diagram: (1 Figure if needed)
- SVC.n.6: Benefits/Issues:

[Provide a list of the benefits and issues that a service provides, from a technical/functional perspective and, more important, from an economical point of view]

11.1.1.3. RESOURCES MOVILIZATION:

[Description of all necessary resources involved in the Scenario deployment]

11.1.1.3.1. DEPLOYMENT ELEMENTS (EQUIPMENT, PROTOCOLS, STANDARDS & PLATFORMS) (From 1 to N)

- DEP.n.1: Name:

[Name of the equipment, protocol, standard or SW-platform used to deploy a service/scenario]

- DEP.n.2: Type:

[Select from: New/Upgrade; Equipment/Software/Standard/Protocol]

- DEP.n.3: Description: (5-10 lines)

[Describe the element deployed or used in the scenario and its functions from a technical point of view]

- DEP.n.4: Economic Impact:

[Value the costs of deploying an element and why: Development/CAPEX, OPEX]

11.1.1.3.2. AGREEMENTS & INTERPLATFORMS (From 1 to N)

- AGR.n.1: Stakeholders Relationship Diagram: (1 figure)

[Provide a diagram of the relationship among the stakeholders]

- AGR.n.2: Stakeholders Relationship description: (10 Lines)

[Explain the inter-relationship agreements among the stakeholders from both technical (platforms/services/mechanisms that support those agreements) and market point of view (template of the agreement…)]

- AGR.n.3: Type:

[Centralized/decentralized; managed/unmanaged]

- AGR.n.4: Inter-platform implementing agreement:

[Explain the agreements among the different parts of the Scenario]

11.1.1.3.3. OPERATION & MANAGEMENT

[Concerns with the impact over the main scenario. It's NOT necessary to explain in detail all the O&M issues of the services]

- OMA.1: Name of resources to operate & manage:

[Descriptive name of the resources/elements/services to operate and manage]

- OMA.2: Description: (5 lines)

[Describe the main O&M impact of deploying this new scenario]

11.1.1.4. BARRIERS:

11.1.1.4.1. STANDARDIZATION STATUS (From 1 to N)

- STD.n.1: Name/Title:

[Name of the standard]

- STD.n.2: Status:

[Status values: New (clarify: ETICS contribution or not), Draft, Consolidated, Implementation Available, others…]

- STD: Issue Description: (10 lines)

[Describe the barriers and current deadlocks of the Standard referred]

11.1.1.4.2. IMPLEMENTATION AVAILABILITY (From 1 to N)

- IMP.n.1 Issue Description: (10 lines)

[Describe the available/current implementation and its main issues]

11.1.1.4.3. REGULATION RISKS (From 1 to N)

- REG.n.1: Issue Description: (10 lines)

[Describe the available/current risks from a regulatory point of view]