Final ETICS Architecture and Functional Entities High Level Design

Deliverable/Milestone: D4.4

Date: 05/2/2013

Version: 1.0

Editor:	P. Zwickl, H. Weisgrab, FTW Telecommunications Research Center Vienna
Deliverable nature:	R
Dissemination level: (Confidentiality)	PU
Contractual Delivery Date:	31/12/2012
Actual Delivery Date	05/02/2013
Suggested Readers:	Public
Total number of pages:	155 (with annexes 212)
Keywords:	Interconnection, QoS, Assured Service Quality (ASQ), architecture, ETICS overlay model

ABSTRACT

The provisioning of Assured Quality (AQ) services in interconnected networks is accompanied with many technological, architectural, and economic challenges. A feasible architecture targeting this usage field needs to flexibly compensate domain-specific requirements by utilising AS-specific optimisations and enabling the formation of sustainable business cases around it. Beyond that, upcoming architectures need to accommodate the diversity of use cases being realised on top of network services, as well as the varying interests of different types of buyers of such services. Compiling these aspects in a single ETICS architecture thus results in platform yielding inter-carrier harmonisation on various levels, e.g. regarding network resource trading or cooperation aspects. As such the central concept of Assured Service Quality paths are introduced as powerful and generic Service Level Agreement-based mechanism for the interconnection of network services on aggregate resource level. This is subsequently complemented by explicitly introducing concepts for enabling session-services and application-side utilisations on top of aggregate resources. The focus is set on Network Service Provider to Network Service Provider services, which corresponds to the core topic of the ETICS project. However, in order to better utilise such services, also the realisation of enterprise services is discussed. Through the capability of enabling route diversification, low latency services (e.g. video conferences) can also be specifically addressed. Aligning business interests with technical capabilities, a series of monitoring solutions are introduced for verifying the conformance of the provisioned service with signed agreements. The present deliverable as such can be considered to be the compendium of architectural work done in ETICS integrating main achievements or previous deliverables to formulate a final architecture definition. As such this document summarises the final ETICS architecture as well as a technical roadmap for its realisation in practice by also incorporating valuable feedback from other work packages.

DISCLAIMER

This document contains material, which is the copyright of certain ETICS consortium parties, and may not be reproduced or copied without permission. All ETICS consortium parties have agreed to full publication of this document. The commercial use of any information contained in this document may require a license from the proprietor of that information.

Neither the ETICS consortium as a whole, nor a certain party of the ETICS consortium warrant that the information contained in this document is capable of use, or that use of the information is free from risk, and accept liability for loss or damage suffered by any person using this information.

This document does not represent the opinion of the European Community, and the European Community is not responsible for any use that might be made of its content.

IMPRINT

Full project title: Economics and Technologies for Inter Carrier Services
Inter-carrier high level technical architecture for end-to-end network services
Document title:
Editor: P. Zwickl and H. Weisgrab, FTW Telecommunications Research Center Vienna
Workpackage Leader: A. Cimmino[1] and G. Parladori[2], Alcatel-Lucent Italy
Project Co-ordinator: Nicolas Le Sauze, Alcatel-Lucent Bell Labs France
Technical Project Leader: Richard Douville, Alcatel-Lucent Bell Labs France
This project is co-funded by the European Union through the ICT programme under FP7.

[1] Until Dec 16, 2012

[2] From Dec 17, 2012

EXECUTIVE SUMMARY

The provisioning of Assured Quality (AQ) services especially in the context of interconnected networks entails a series of challenges resulting from the distinctive level of heterogeneity whether on the axis of technologies used, business interests or socio-economic determinants in using the services. Such determinants are reflected in realistic use cases spanning over multimedia services required by end customers (retail) to Virtual Private Network (VPN) services provided by enterprise/business customers, or the support for intra-Network Service Provider (NSP) trading of network resources (wholesale). As a consequence, architectural mechanisms are required allowing NSPs to find network resources in interconnected networks required for satisfying given customer demands. A feasible architecture targeting this usage field needs to *flexibly compensate domain-specific requirements* (resulting from currently deployed and used technologies) by avoiding centralised technological or economic regulations or prescriptions, but rather *facilitating cooperating* among "unequal" partners.

The present deliverable formulates the final ETICS architecture deliberately focusing on the ETICS service management component with strong links to the Interconnected Networks, Service Enhancement Functional Area (SEFA), Virtual Private Network (VPN), and charging of session & application services as illustrated in the big picture of FIGURE 9. Driven by the challenge of required cooperation among heterogeneous ASes, ETICS has on the one hand proposed to use a network overlay model abstracting underlying network infrastructure in order to ease the trading of network resources, i.e. generic and flexible *Assured Service Quality (ASQ) paths*. Such ASQ paths accommodate for a great diversity of different use cases and further extensions, which are negotiated on the basis of Service Level Agreements (SLAs). On the other hand, the operationalization is targeted by *ETICS Community* (see [ETICS-D3.5]) types assisting the cooperation of NSPs. The resulting architecture will thus be presented as compendium of architectural works in ETICS rendering previous deliverables of the series obsolete, i.e. [ETICS-D4.2] and [ETICS4.3], by also incorporating lessons learned e.g. from [ETICS-D5.6], [ETICS-D3.3], and [ETICS-3.5].

In detail, besides the recognition of interesting use cases for premium session-aware services in interconnected networks (see Section 3.3), an intentional focus has been laid on "big pipes" (see Section 3.4 and 4.1), i.e. *ASQ paths* aggregating a series of micro-flow demands, serving as fundamental basis for more fine-grained utilisations. In particular, the present architecture has the capability to support Point-of-Interconnect to Point-of-Interconnect (*PoI-to-PoI*) as well as *PoI-to-Region* services on the basis of generic ASQ paths. By the assistance of SEFA (see Section 6.1), more light-weight session-aware services may be released on top of PoI-to-Region services.

SEFA has specifically been presented as a flexible mechanism for realising *advanced functionalities* on top of generic ASQ paths. Exemplarily, Graceful Denial of Service (GDoS), session-based connectivity service, and Congestion Exposure-based (ConEx; cf. Section 6.2) capacity sharing have been illustrated.

SLA-based offers (based on network capabilities) for an ASQ path may be exchanged between customer and supplier on the basis of a "ready to wear" (push) or "made to measure" (pull) paradigm. In both cases, a series of network resources spanning one or more NSPs are traded, while the computation of ASQ paths

may be organised in various degrees of centralisation. Inferred from a series of related architectural options a *multitude of deployment scenarios* have been reported in [ETICS-D4.3], which have been *consolidated* in the present work. In particular, through preferring the automation of some deployment scenarios over others, by better integrating the usage of SEFA and other complementary technologies, and especially through the introduction of a roadmap concept, rendering technologies subject to rollout phases, a clearer picture is drawn for the readers. A bootstrapping concept is constructed around the investigation of straightforward manual, only bilateral, and also PoI-to-Region based configurations. This bootstrapping scenario shall represent the first step for the introduction of ETICS concepts by NSPs at a rather short term with limited technological gaps. Future steps will be deployed according to the level of maturity of the demands (type of QoS services required by customers) and of the market (increasing size of communities with more and more involved NSPs, a possible increasing need for more automation and extended relations beyond bilateral links, etc.). Such evolutions being uncertain at the present time, a flexible and adaptable ETICS architecture configurable in different collaboration modes was therefore a necessity for a possible usage by NSPs at medium and longer term.

Special emphasis has also been lead on developing a monitoring architecture with clear interfaces to the main ETICS architecture. Monitoring solutions are required in order to be able to check the conformity of provisioned network services to SLAs. For this purpose, an *extension of the known intra-domain monitoring mechanism OAM as well as a newly designed active/passive centralised monitoring architecture* has been given. For the *bootstrapping phase, monitoring solutions may be omitted* to keep the required investments low, whenever sufficient trust between cooperating partners exists.

Originating from two main service types, i.e. *ASQ* path-based services - ASQ Tunnel services between Points of Interconnect (PoI) and ASQ Traffic termination to end point regions - and *End-user ASQ connectivity* (microflow-based ASQ traffic for end users realised on top of ASQ paths), the final architecture also takes various buyer/supplier modalities into account (see Section 4.4; also see [ETICS-D3.5] and [ETICS-D2.3]). While the present architecture essentially builds on directional charging, preferably *Sending Party's Network Pays* (SPNP, also see [ETICS-D2.3]), more advanced variants *with Initiating Party's Network Pays* (IPNP) mechanisms may be realised on top of SPNP. Tunnel services like Virtual Private Networks (VPNs) may in parallel be directly constructed with an IPNP paradigm due to their bilateral nature. The investigation of VPNs as technically been followed up by elaborating their realisation on top of an ASQ paths infrastructure.

In the analysis being detailed in Section 7 promising scalability results for the ETICS architecture have been observed: on the Network Service and Business Plane acceptable runtimes have been shown. On the Control & Data Plane the on-demand (pull) deployment scenarios have yielded good convergence times (under 500ms) and easily accommodated for 250 ASes in a community. Also the required monitoring solutions have been shown to scale up to the bandwidth required for current networks.

The deliverable concludes by enumerating a series of further work aspects mainly beyond the scope of ETICS (cf. Section 9): Amongst others, intra-domain route determination/diversification, PoEI integration, or hardware deployment issues remain untouched.

TABLE OF CONTENTS

1. INTRODUCTION

1.1. MOTIVATION & PROBLEM

The provisioning of Assured Quality (AQ) services especially in the network interconnection context entails a series of challenges resulting from the distinctive level of heterogeneity whether on the axis of technologies used, e.g. connection-less versus connection-oriented technologies, various business interests, e.g. geographical or market differences, and socio-economic determinants in using the services. Such determinants are reflected in realistic use cases spanning over multimedia services required by end customers (retail) to Virtual Private Network (VPN) services provided by enterprise/business customers, or the support for intra-Network Service Provider (NSP) trading of network resources (wholesale). As a consequence, architectural mechanisms are required allowing NSPs to find network resources in interconnected networks required for satisfying given customer demands.

A feasible architecture targeting this usage field needs to *flexibly compensate domain-specific requirements* (resulting from currently deployed and used technologies) by avoiding centralised technological or economic regulations or prescriptions. As such the central concept of Assured Service Quality (ASQ) paths will be introduced as powerful and generic concept for realising and trading AQ services on the basis of interconnected heterogeneous ASes. The cooperation of individual networks is further operationalized in ETICS by the ETICS community (see [ETICS-D3.5]) types as mechanisms for enabling AQ services across a set of ASes.

Emphasising the requirements of providing a coherent economic, business, and technological view on network interconnection, the ETICS architecture aims at not restricting business model decisions (in terms of offered QoS, pricing, target customer and other combinations) through avoidable architectural boundaries. In turn, by utilising the concept of generic Service Level Agreements (SLAs) ASQ paths are established as tradable and flexibly composable goods for aggregate-level network resources, i.e. ASQ goods. The important special case of low latency services requires the architecture to accommodate for route diversity (especially in respect to Best-Effort services) and determination being orthogonally assisted by sophisticated monitoring solutions fitting to individual network solutions.

Despite setting an architectural focus on NSP-to-NSP services (with reference to the Peering, and Transit markets), the present work accommodates for more flexible actor interactions and use cases, as inferred from the actor analysis of [ETICS-D3.2] and the ETICS reference model depicted in [ETICS-D2.2].

For detailed requirements to be taken into account we refer to [ETICS-D2.2], but also to its on-going revision in [ETICS-D2.3] focusing on more specific aspects.

1.2. CONTRIBUTION

Completing the series of deliverables on the ETICS architecture, the present deliverable D4.4 will give the final architecture as compendium of ETICS architectural work resulting in a *self-contained* reference for the

ETICS architecture. As such D4.4 will not only essentially build on the results of the preceding deliverable [ETICS-D4.3], but also essentially integrate lessons learned from the detailed specification [ETICS-D5.6] and economic implications on the architecture, cf. [ETICS-D3.3][ETICS-D3.5].

In addition to what has been presented in [ETICS-D4.3], the present deliverable fundamentally aims at drawing a final big picture illustrating the environment of ETICS architecture and associated immanent interrelations. Further specifications of the components of this big picture will be the goal of the final ETICS technical deliverable (D5.8).

Availing ourselves of the main ETICS principles (reiterated and further elaborated in Section 4.1) a systematic consolidation of views resulting from requirements (WP2), economics (WP3), and specification/simulation work (WP5) is aspired by aligning definitions, integrating feedback loops, and intentionally positioning the ETICS architecture in the picture of the overall ETICS system. Along with that, a clean separation between ETICS architecture and its core-system architecture, i.e. focusing on key ASQ concepts traded between Network Service Providers (NSPs), is established.

By explicitly addressing challenges in operating, configuring and rolling out the ETICS architecture in practice, an orthogonal view to the architectural definitions is formed capturing the variety of potential (partial) ETICS architecture usages – thus, telling a more complete ETICS story from the beginning bootstrapping phases via its timely and system-wise evolution to full deployments.

Recognising essential use case scenarios in network IC, the present deliverable also dedicates specific resources to business and enterprise service, as well as to the integration of session services complementing the aggregate resource view applied by ETICS.

1.3. STRUCTURE

The remainder of this deliverable is structured as follows:

Starting by revisiting the state-of-the-art in inter-carrier technologies, concepts, and services potentially being supportive in realising Quality of Service (QoS) differentiation or even AQ services in Section 2, Section 3 heralds the start of the refocusing on the well-chosen concepts of the ETICS approach. In particular, prerequisites such as required definitions, the use cases in the ecosystem and required inter-carrier concepts are discussed.

This is followed up in Section 4 by giving an updated overview of the overall ETICS solution being constructed around a set of fundamental visions and goals. This lays the foundation for discussing various interactions and realisations of the ETICS system, e.g. automated versus manual basic mode or the set of associated deployment scenarios (cf. roadmap in Section 4.5). Please take note that the revised presentations of these variations over [ETICS-D4.3] incorporate feedback from [ETICS-D3.3] and [ETICS-D3.5]], and from the ETICS advisory panel, as well as some detailed advancements.

Diving deeper in the core-system architecture, an essential part of the ETICS architecture, corresponding details are discussed in Section 5 in a top-down approach. In addition, monitoring concepts and their interrelation with the rest of the architecture are specifically highlighted.

For the sake of supporting advanced features such as session support or particular business/enterprise services, Section 6 aims at providing a series of functional extensions to the ETICS architecture. While these elements are intentionally placed as extensions, we would like to highlight their high interrelation with the basic concepts previously given in Section 5.

Orthogonally, scalability challenges corresponding to individual parts of the ETICS architecture are reviewed in Section 7. Finally, Section 8 and 9 provide a brief overview of questions remaining for further work, as well as concluding comments respectively.

The main corpus of present document is complemented by an extensive set of further materials in the Appendix (Section 11), e.g. further related works, scalability details, and a glossary.

For a list of **changes over the preceding deliverable** [ETICS-D4.3] we kindly refer the reader to the annex Section 11.1.

1.4. TARGETED AUDIENCE

For readers familiar with [ETICS-D4.3] we recommend focusing the reading on modified sections as previously indicated. For new readers, the present deliverable (as refined and extended compilation of most essential blocks discussed in [ETICS-D4.2][ETICS-D4.3]) represents a well-suited entry point for the ETICS architecture.

As audience we target experts in the network core and interconnection field focusing on higher network layers (routing and above). Amongst others, this may encompasses NSPs looking for solutions for technical inter-carrier collaboration or replacements for existing solutions.

This deliverables does not set an explicit focus on access technologies, while establishing links to session-enabled end-to-end services.

For business models, market, and economic investigations we especially kindly refer to [ETICS-D3.5] and subsidiary deliverables [ETICS-D3.3] and [ETICS-D3.4].

While the present deliverable also targets rollout phases for the ETICS architecture, further aspects are also captured in [ETICS-D3.4] (market quantifications) and especially [ETICS-D2.3] (bootstrapping, rollout phases).

A detailed specification of architectural building blocks and algorithms can be found in [ETICS-D5.6] and [ETICS-D5.8].

1.5. GLOSSARY

Subsequently, a few fundamental notions will be introduced, which are used throughout the document. A detailed summary and definition of used abbreviations is given in Section 11.9.

- **ETICS Solution**: The overall solution proposed by ETICS considering the inter-networking solution and network services as well as the ETICS core-system solution providing, managing and supporting inter-carrier ASQ paths and connectivity.

- **ETICS Core-system solution**: Part of the overall solution focusing on ETICS Services and SLA management as well as providing links to underlying interconnected networks (and their technologies) and applications realised with the help of ETICS, e.g. session services.
- **ETICS Architecture**: Architectural concepts enabling the realisation of the ETICS solution and its deployment in practice. Resulting deployments may be referred to as *ETICS System*. The interaction between the ETICS System and customers, e.g. NSPs, are handled by the help of the *ETICS Portal*.
- **ETICS Community**: The set of ETICS NSPs as suppliers and buyers of ETICS network services are called ETICS Community. The degree of cooperation and alignment between members of a community is defined by community types, i.e. open alliance, federation, alliance. For a detailed analysis and definition of community types we kindly refer to [ETICS-D3.5] as refinement and extension of the concepts given in [ETICS-D4.3].
- **Assured Service Quality (ASQ):** Assured Service Quality (ASQ) is used as adjective referring to quality aspects - i.e. QoS guarantees, availability, etc. - for provisioned network services. Please, refer to more specific notions for more detailed usages.
- **Assured Quality (AQ) Service:** An ETICS network service is more precisely termed Assured Quality (AQ) network Service, or for short “AQ Service”. AQ services refer to network services with an ASQ nature (see ASQ).

2. RELATED WORK

In this section, we briefly summarise the existing individual solutions, which only partially address the Network Service Providers' (NSPs') inter-operator QoS requirements, whereby we highlight the underlying problems that need to be resolved. This section will intentionally focus on missing technical functionalities in the current inter-carrier connectivity approaches around stumbling blocks identified in the ETICS project.

Beyond the overview given in previous deliverables, the present related work compilation deeper investigates the (potential) interplay of ETICS-external solution approaches, as well as their relationship to ETICS. In addition, the present section puts emphasis on the great variety of inter-carrier QoS requirements as seen in the context of ETICS, as well as the QoS configuration in practice. For brevity reasons we kindly refer the reader for a more general introduction of technologies to [ETICS-D4.3] and the subsequently used references.

The remainder of this section will be organized as follows: starting with an investigation of basic layer-3 inter-carrier connectivity in Section 2.1, we will continue with an investigation of data transport in roaming scenarios (cf. Section 2.2). Shifting the focus towards end-to-end session service support, access network QoS technologies are revisited by the well-known example of the IP Multimedia Subsystem in Section 2.3. On this basis, the interplay of data transport in roaming scenarios with access network QoS and ETICS will be analysed in Section 2.4. Thereafter our attention will shift more towards connection-oriented technologies being useful for the provisioning of low latency services (route determination). Amongst others the relationship between the Multi-Protocol Label Switching (MPLS) protocol suite and ETICS will be discussed. Section 2.6 will deeper investigate the example of Virtual Private Networks (VPNs).

2.1. BORDER GATEWAY PROTOCOL (BGP)

There are important stumbling blocks that prevent the integration of inter-carrier QoS connectivity provisioning with the Border Gateway Protocol (BGP) being cardinal for inter-carrier layer-3 connectivity – see [RFC4271]. The lack of choice of multiple paths between two remote autonomous systems (ASes) represents one of the most important obstacles. Consequently, each AS chooses and advertises only a single neighbour AS for forwarding traffic towards a specific destination address range, which potentially inhibits the setup of the desired level of inter-carrier QoS.

Beyond the lack of flexibility concerning the choice of inter-carrier network topology, the integration of QoS capabilities with BGP would introduce further scalability issues to this protocol. Reinforced by the tremendous growth of globally advertised prefixes, e.g. stemming from multi-homing, a further explosion of the BGP forwarding table size in backbone routers (cf. [BGPMEA]) would be inferred. Paired with the global advertisement of individual prefix reachability with the current protocol, also significant communication overhead would result (network connection and computation resources in BGP routers; cf. [RFC4271]), as any information on instabilities of the numerous connections belonging to multi-homed end-systems is automatically propagated in numerous update messages

Nevertheless, more recently two particular functions for the exchange of essential inter-carrier information between domains have gained some favour at the IETF leaving the standard BGP protocol itself untouched: Multi-path enhancement for BGP [IETF-DR-1, IETF-DR-2], which allows the announcement of alternative routes, and conveyance of *link state information and traffic engineering information* [IETF-DR-3].

2.2. IP EXCHANGE (IPX)

As far as the business capabilities of GSMA's IP eXchange (IPX) [IR.34] are concerned, the specified three interconnection models do envision different options for the realization of service agreements, financial flows, and technical responsibilities. Building on these capabilities, a variety of inter-carrier business models can be implemented, such that we may conclude that part of the ETICS business requirements could be addressed within the IPX business architecture. However, some aspects of QoS provisioning and the automation thereof seem to be not sufficiently addressed to fulfil ETICS requirements, i.e. NSPs' requirements.

It shall be noted that IPX itself does not mandate specific mechanisms for the provisioning of QoS for end-user sessions and flows, but rather leaves it up to network and IPX platform operators to make their individual choices (cf. [IR.34]). This may accommodate for the high heterogeneity in today's Internet landscape, which correlates with ETICS endeavours of providing technology-agnostic inter-carrier services.

Concerning its current status, IPX is increasingly becoming the technology of choice for exchanging high-value IP traffic between mobile network operators, i.e. IP data and VoIP traffic Even though IPX foresees *DiffServ* as the standard inter-operator QoS mechanism [IR.34], this QoS solution has not yet been adopted by operators, probably due to the difficulties in aligning all operators to a single scheme of QoS classes.

However, as soon as a suitable QoS mechanism is used within individual IPX systems, ETICS solutions could be used to transport the traffic between remote IPX providers with QoS guarantees. In other words, , the ETICS system does not necessarily impose to replace, but could also complement IPX; Whenever IPX is already deployed and used, ETICS could rather augment it with an inter-IPX provider, QoS-enabled connectivity service.

2.3. IP MULTIMEDIA SUBSYSTEM (IMS)

The perceived potential in terms of value added services is mostly related to the capability of the IP Multimedia Subsystem (IMS) to assure QoS at the granularity of end-user sessions, opening the way to per application quality assurances (cf. [IMS][IMS-2][3GPP-3]), e.g. carrier-grade voice service (cf. [IR.92]). Apart from mobile networks, this situation is also quite similar for the fixed network flavour of the IMS, i.e. TISPAN[3] (cf. [ETSI-1] and [ETSI-2]), where the IMS could assure the quality in the last mile for premium content traversing the IP data connections of the customers. Support for fixed line networks has been added to IMS in 3GPP release 7 (finalized beginning 2008) which translates not only to opening the network for other (i.e. non-mobile) network operators, but actually also to actively supporting those types of

[3] Telecoms & Internet converged Services & Protocols for Advanced Networks, http://portal.etsi.org/tispan/TISPAN_ToR.asp, last accessed: Dec 13 ,2012

networks in order to improve the range of service coverage and consequently to improve economic attractiveness of the IMS system.

Whereas the IMS does foresee mechanisms for assuring QoS in the access network part, at the same time it lacks similar provisions for inter-carrier sections of the end-to-end path (cf. [3GPP-1], [3GPP-2]). While [3GPP-2] in Annex A.2 does identify a number of scenarios for the realization of end-to-end QoS for IMS-based services, the proposed solutions for the inter-carrier part merely sketch the possibilities for IP QoS differentiation on inter-operator links, and they do not propose concrete mechanisms which would logically associate end-user sessions to the individual classes of traffic at the network edge.

By potentially using the ETICS solution, complexity were significantly reduced by merely focusing on aggregate interconnection traffic on the transit-ISP level, which were subsequently assisted by mechanisms attaching sessions to assured-quality resources by means of Service Enhancement Functional Area (SEFA; cf. Section 6.1). In turn, interplay opportunities between ETICS and IMS, but also with IPX are enabled, which are subsequently described.

2.4. INTERPLAY OF IPX, IMS, AND ETICS

In combination with the previously described IPX, the IMS could indeed offer an end-to-end solution for inter-carrier QoS. Whereas a comprehensive summary of the benefits and drawbacks of such an approach is hardly possible to come up with due to a lack of practical experience by the operators even on a pilot project scale, we dare to suggest that while IPX and IMS do represent extensively specified and commercially available solutions, the large overhead they introduce in terms of initial investments, deployment and operations renders this combination a relatively monolithic and quite resource-intensive scheme.

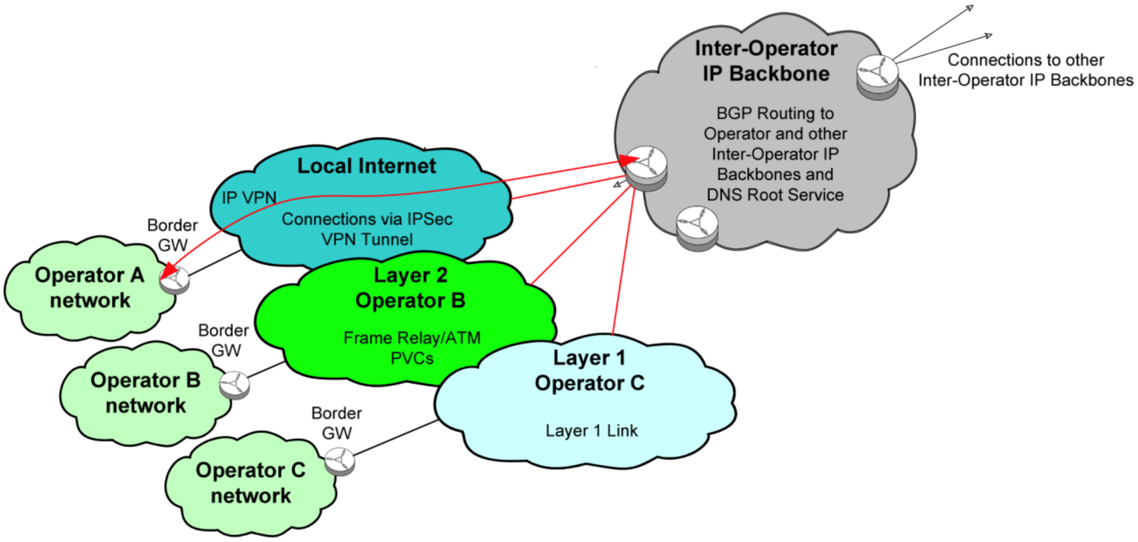

Inter-Operator IP Backbone
BGP Routing to Operator and other Inter-Operator IP Backbones and DNS Root Service
Connections to other Inter-Operator IP Backbones
Local Internet
IP VPN
Connections via IPSec VPN Tunnel
Border GW
Operator A network
Layer 2 Operator B
Frame Relay/ATM PVCs
Border GW
Operator B network
Layer 1 Operator C
Layer 1 Link
Border GW
Operator C network

FIGURE 1: IPX OVERVIEW [IR.34]

However, there are indications that due to the ever increasing traffic volume in the access networks, bandwidth problems are emerging also in the core part of the access network, especially on the backhauling trunks. In order to solve this problem, often IMS with its Policy Charging and Rules Function (PCRF) is the proposed solution to this problem by identifying the traffic and shaping it according to its needs. The Policy Charging Enforcement Function (PCEF), in particular, is executing the required access control within IMS in a manner comparable to an advanced firewall. Consequently, the number of IMS enabled networks with PCRF (which is necessary for QoS enforcement) is increasing.

IMS + ETICS: As already explained, IPX could be used to connect IMS enabled access NSPs, but QoS between IPX networks cannot yet be assured. Consequently, instead of IPX, ETICS mechanisms could be used to interconnect IMS access NSPs via ETICS enabled transit NSPs. In this case, the traffic on the edge of the IMS network (whereby the network operator knows exactly about the type of traffic and consequently its QoS requirements) would be put into an ETICS ASQ, and the ETICS system would ensure transport of the traffic with QoS assurances to the remote IMS access network. Within the remote IMS network, IMS-native QoS mechanisms would again take over (provided that the remote IMS network is QoS-enabled), effectively providing for a truly end-to-end QoS-enabled service. While IPX interconnection may in practice be stronger linked with the peering-point paradigm, where traffic is exchanged at defined points, route determination mechanisms (as described in Section 2.5 and envisioned by ETICS) could provide a more fine-granular control over traffic flows.

While this may provide a very interesting technological combination, ETICS will aim at providing QoS mechanisms mainly on an aggregate flow level. On this basis, more light-weight approaches enabling quality assured session services may be placed (see Section 6.1). This may in turn strongly reduce the involved complexity and thus mitigate scalability problems.

IPX + ETICS: There may be several mutually interesting scenarios of IPX and ETICS interplay:

- **QoS parameterisation**: IPX provides mechanisms for harmonizing QoS classes beyond the reach of a single NSP. Established practices, agreements, or techniques aligning quality classes in ETICS may be inspired by IPX solutions. While the ETICS approach may highly advocate a flexible quality class arrangement per request for aggregate quality-assured network resources (over all involved ASes), on an operational level an alignment of QoS classes may be realistic (cf. Section 11.3).
- **Transport technology for the IPX cloud**: ETICS may provide mechanisms for interconnecting several IPX inter-operator IP backbone clouds (denoted by a grey cloud in Figure 1).
- **Reachability of inter-operator IP Backbone**: Especially smaller NSPs, which are not IPX providers themselves, may look for mechanisms connecting themselves with the nearest IP Backbone provider. The arising interconnection traffic may then be handled by using an ETICS solution (denoted by the red connections in Figure 1).

Beyond that, ETICS may also be used as alternative mechanism for quality assured Transit or Peering services (see Section 3.3) going beyond the roaming use case targeted by IPX and the access network perspective applied by IMS. Especially a more flexible and dynamic arrangement of assured quality interconnection agreements for such services may be beneficial.

2.5. CONNECTION-ORIENTED OPERATOR INTERCONNECTION

Thanks to traffic engineering extensions, Multi-Protocol Label Switching (MPLS) provides network operators with tunnel placement capabilities, thus allowing to select the network QoS parameters per (aggregate) traffic flow. Traffic engineering parameters are not flooded beyond the boundaries of each Intra domain routing protocol (IGP) area. To address this issue, the Path Computation Element (PCE) function can be used. The PCE architecture provides a communication protocol (PCEP) enabling to request a path calculation from a remote entity, i.e. a PCE.

However, when it comes to multi-AS and multi-carrier routing, it is unlikely to find an entity in the network which is aware of the necessary information from every sub-network. In this case, a sequence of PCEP requests between all the network domains is to be spanned by the user traffic. The “Backward-Recursive PCE-based Computation Procedure” (BRPC) is a standard specification, which uses PCEP to address that problem in case of limited numbers of interconnected domains and in case of chains of domains known a priori. However, there is still a lack in standards on the automation of the calculation of this domain sequence when domains interconnected together are numerous.

In terms of fault detection, usual MPLS mechanisms can be used without change in a multi-AS context, e.g. for loss of signal on equipment interfaces, IGP hellos, Bidirectional Forwarding Detection (BFD), etc., while no current inter-carrier deployment is to the best of our knowledge present.

Service monitoring usually happens on a per carrier basis, with information relevant to the corresponding AS. The IETF has recently standardized the spatial composition of metrics: [RFC6049] allows combining the traffic metrics from a set of sub-networks or ASes in order to build a complete path metric from the sub-path metrics. What remains to be defined is the way this information is collected among the different ASes and how it is put together; this relationship could be handled by peering or as an agreement part of a consortium of carriers.

Despite all the capabilities provided by MPLS or Generalized MPLS - (G)MPLS -, such mechanisms still require a broad technological inter- and intra-AS convergence of NSPs, while in contrast ETICS envisions a more technology-agnostic solutions approach. By utilizing AS-optimal solutions, we envision a higher global efficiency of the ETICS system with broader global reach and hence mitigated scalability concerns. Nonetheless, the route determination capabilities of (G)MPLS associated with inter-carrier assistance of PCE, as well as reducing the size of forwarding tables utilizing the MPLS labels are key technologies in the ETICS context.

2.6. VIRTUAL PRIVATE NETWORKS (VPN) SERVICES

In the following we provide a reference model for the provision of business VPN connectivity services along with an overview of alternative, currently widely established *provider provisioned* VPN service types i.e., BGP MPLS based VPNs[4].

In the considered VPN service model, a business customer maintains a set of sites (e.g., headquarters and branches) that must be interconnected (see Figure 2). Though the same network infrastructure may be employed for several VPN service instances, each VPN can have its own routing address space, while different VPNs may have the same routing address space. On the data plane, traffic remains strictly within the boundaries of the VPN, without traffic ever reaching sites not participating the VPN. In most cases, the interconnection of sites is further characterized by well specified QoS metrics.

[4] It is noted that the provided VPN types overview targets at providing the technical context for the subsequent ETICS related material, rather than a complete description of the corresponding protocols and standards.

The interconnection of the sites is established through the Provider Edge (PE) routers which are operated by the NSP(s). On the Business Customer (BC) side, each site interconnects with one or more PEs through a Customer Edge (CE) device. A PE typically serves multiple CEs for different (and isolated) VPNs. A CE delivers traffic to a PE with the purpose of reaching an end host located in some other BC site[5]. For this purpose, the packets/frames emitted by the CE contain the location identifier (address) of the targeted end host. For the data to be routable, the source PE i.e., the PE receiving the packets from the CE, maintains a forwarding table with mappings between location identifiers of destination hosts and location identifiers of the corresponding (destination) PEs. The delivery of the data then relies on the existence of a routing and forwarding fabric among PEs which is used to tunnel the data between source and destination PE pairs. Once the data reaches the destination PE, a VPN identifier is used to select the correct VPN interface and the data is delivered to the CE.

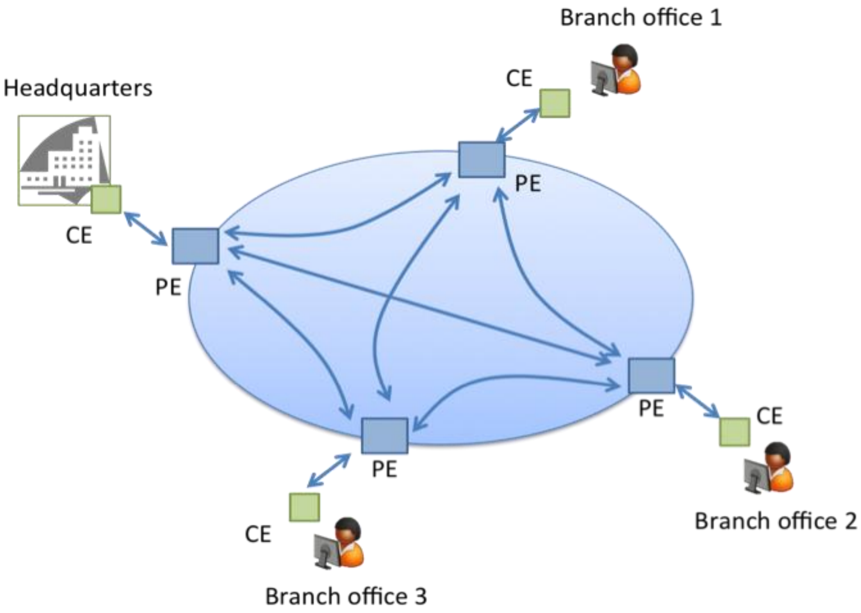


FIGURE 2: VPN REFERENCE MODEL

The establishment of the VPN requires populating the routing and forwarding tables at each participating PE. This procedure varies with the exact type of the VPN. In the following we elaborate on two common VPN types both based on BGP and MPLS but differentiating on the layer of connectivity they establish between VPN sites.

2.6.1. Layer 3 BGP/MPLS VPNs

In this type of VPNs (RFC 4364), an NSP provides Layer 3 connectivity to the BC sites based on the use of BGP and MPLS for the exchange of routing information between the PEs and the routing/forwarding among the PEs. A PE maintains the aforementioned routing and forwarding information in a Virtual Routing and Forwarding (VRF) table for each VPN supported. A VRF maintains the following information:

[*IP Destination prefix*, *Tunnel label*, *VPN Route Label*]

[5] Other modes of communication are also possible, such as a multicast and broadcast, but these are not addressed here for simplicity reasons.

Given a destination IP, a PE uses the Tunnel label (an MPLS label) to identify the path that must be used to reach the destination PE for this packet, and the VPN Route Label to identify the targeted VPN at the destination PE.

The establishment of the VPN routing information is based on the BGP protocol. In the general case, all PEs of a NSP are connected in a full mesh of BGP sessions. In order for a PE A to announce reachability to some destination IP within the VPN site it supports, PE A sends a BGP update message to all PEs indicating itself as the Next Hop towards the destination IP. At the same time it populates the VPN Route Label to be used for the populated IP. In order to ensure the isolation of the addressing spaces of the multiple VPNs (i.e., two or more VPNs may use the same address space), the populated destination IP is augmented with a unique *Route Distinguisher* (RD) prefix. Since not all PEs necessarily participate every supported VPN, a VPN provider (NSP) further uses *Route Target* (RT) identifiers to control the establishment of the VPN routing state. An RT value is configured by the NSP at each VRF. Each BGP Update message also carries an RT value in the extended community attributes. The routing information of a BGP message is installed only at the VRFs that have the same RT value configured with the RT value of the BGP message.

The use of a full-mesh of BGP sessions among all PEs introduces scalabilities concerns as the number of PEs rises. A proposed solution is the use of Route Reflectors (RR) that can introduce a hierarchical structure to the route distribution mechanism. In this case, PEs establish BGP sessions with RRs, instead of each other, which are then responsible of redistributing the routing information.

2.6.1.1. Inter-NSP deployments

The establishment of a Layer 3 BGP/MPLS VPN as described above is first of all based on the BGP sessions between the participating PEs (or the PEs and the RRs) for the exchange of routing information. In cases however, where the different VPN sites reside in different Autonomous Systems (or in the context of ETICS, in different NSPs), these sessions cannot be assumed to exist. [RFC 4364] proposes three different options for the establishment of VPN connectivity in this context.

(a) In the first option the Autonomous System Border Routers (ASBRs) mutually as PE-CE couples in each direction i.e., each ASBR receives VRF information from the PEs in its AS and provides it to its neighbour ASBR which acts as a CE. This method necessitates the maintenance of VPN routing information at the ASBRs resulting in scalability concerns.

(b) In the second option ASBRs use EBGP to distribute the routes they receive from their ASes' PEs to neighbour ASBRs. In this case the ASBRs need not maintain VRFs, but they need to inspect the distributed VPN information.

(c) In the third option the ASBRs neither maintain nor do they distribute VPN routing information. Instead direct (multi-hop) EBGP sessions are established between the participating PEs. RRs may also be used to convey VPN routing information through EBGP for scalability reasons.

2.6.2. Layer 2 VPNs

Although Layer 3 VPNs are widely deployed, there is also an increasing interest in L2VPNs so that customers can manage their own IP routing. This interest is especially growing in the context of cloud business connectivity services, where Layer 2 connectivity facilitates WAN VM mobility e.g., in the case of *cloudbursting* and *follow-the-sun* service models.

RFC 4664 describes a framework for Layer 2 VPNs with the CE and PE devices terminology already introduced in the layer 3 VPNs framework. RFC4664 defines two main kinds of layer 2 VPN services: Virtual Private Wire Service (VPWS) and Virtual Private LAN Service (VPLS) [RFC4761].

A VPWS is a layer 2 VPN service that provides a layer 2 point to point service between two CE devices. It can be an Ethernet point to point (E-Line) service for instance. On the contrary, a VPLS service appears as a LAN service across a metro or wide area network (thus providing a multipoint connectivity between more than two end-points).

An alternate solution to L2VPN services (either VPWS or VPLS) based on BGP is L2VPN services using the Label Distribution Protocol (LDP) signalling. As explained in the previous chapter, PEs communicate via Pseudowires (PWs). These pseudowires can be either Single Segment Pseudowires (SS-PWs) or Multi-segment Pseudowire (MS-PW). LDP can be used for establishing SS-PWs and MS-PWs as per RFC 6073. These PWs themselves rely on Label Switched Path (LSP) tunnels that can be set up with the RSVP-TE protocol in conjunction with the OSPF-TE routing protocol.

In ETICS context, the difficulty is to determine dynamically the S-PE devices to establish the ASQ path. Two options can be considered: either extend the LDP protocol or take benefit from the H-TE Path Computation Element (PCE).

3. ETICS NETWORK SERVICES

Assured quality paths (ASQ paths) are quantifiable network connectivity services that guarantee quality data carriage (e.g. bandwidth, delay, etc.). Today, such services can only be sold, if they are provided by a unique network service provider with full control of its network. Examples of such services are those offered by some ISPs to their business customers in order to interconnect their distant sites and data centres spread in different locations.

It is the goal of ETICS to establish assured quality paths that go beyond the boundaries of single NSP domains. In this section, we present use cases for such inter-carrier services: In particular, we focus on the following issues: Which actors can benefit from *inter-carrier ASQ paths*, for which purposes do they use them, and how can they use them? The various actors' needs of inter-carrier services have an impact on the definition of the ETICS inter-carrier services themselves. Based on these needs, we take a first step towards defining the set of ETICS inter-carrier ASQ services. The rest of this section is organized as follows: Section 3.1 starts with some useful definitions. Section 3.2 presents the various actors of the ETICS ecosystem, which are followed by a summary of inter-carrier use cases in Section 3.3 (for an actor-based analysis we kindly refer to Section 11.4). Based on the analysis of Section 3.3, Section 3.4 finally enumerates the basic inter-carrier ASQ paths that the ETICS architecture needs to provide. Finally, an inter-carrier ASQ path results from the composition of different *single-NSP ASQ paths*. *A single-NSP ASQ* path is an ASQ path provided by a single NSP.

3.1. DEFINITIONS

An *ASQ path* is an assured quality traffic delivery service from one (or multiple) point(s) to another point (or multiple points).

An Inter-carrier ASQ path is an ASQ path that crosses multiple NSP domains. It results from the concatenation of two or more *single-NSP ASQ path*s.

A single-NSP ASQ path is an ASQ path provided by a single NSP. Thus, per definition, an Inter-carrier ASQ path results from the concatenation of single-NSP ASQ paths.

3.2. THE ETICS ECOSYSTEM

In Figure 3, we briefly detail the different actors that are part of the ETICS ecosystem. A central actor in the ETICS ecosystem is the network service provider (NSP) who is responsible for providing inter-carrier ASQ paths or connectivity to non-NSP (final) customers. Different NSPs need to collaborate in order to provide inter-carrier ASQ paths. In terms of roles, NSPs can further be divided into *edge or access NSPs* which directly connect end users, and *transit NSPs* which do not connect end users and are mainly responsible for carrying traffic between the different edge NSPs. Note that in practice, some NSPs play these two roles simultaneously (they are edge NSPs and transit NSPs at the same time) although, for specific services instances you can identify the NSP role either as *edge* NSP or as *transit* NSP.

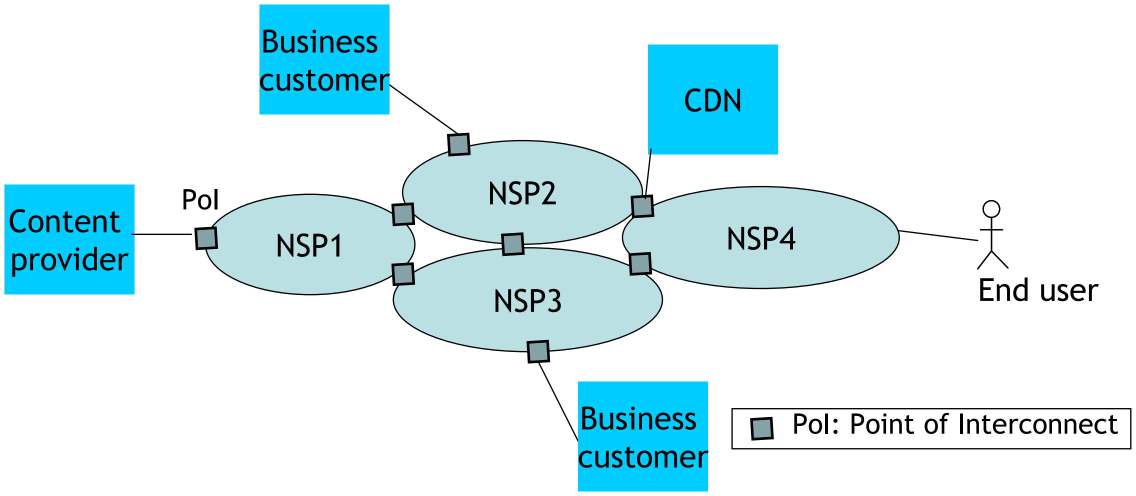


FIGURE 3 THE ETICS ECOSYSTEM AND THE ETICS ACTORS

The other actors of the ETICS ecosystem are potential final customers of the NSPs that might be interested in having inter-carrier ASQ paths. These actors are (1) the content or application providers whose role is to provide content and/or application services to end customers, (2) the residential or consumer end users, (3) business customers which are enterprises that can have sites spread in different locations and, finally, (4) content delivery networks (CDNs).

The different actors of the ETICS ecosystem interconnect through what we call *Points of Interconnect (PoI)*, as shown in Figure 3. A PoI can be seen as a delimiter of the administrative and network topological boundaries between two different actors.

The communication between the different ETICS actors is performed by so called interfaces. These interfaces are specified by the ETICS actor role model, see Section 4.4.2.

3.3. ETICS USE CASE EXAMPLES

The focus of ETICS is on assured service quality (ASQ) network services traded between multiple NSPs in order to establish inter-carrier ASQ paths across the different NSPs. Use cases which could benefit from an ASQ path provided by the ETICS architecture are manifold. However, the use cases supported by the ETICS framework can be divided into two major use case areas, such as:

- Aggregate level ASQ path use cases

  Use cases which are focusing on the NSP-to-NSP ASQ path level can be understood as aggregate level ASQ path (generic ASQ path) use cases targeting to establish an assured service quality (ASQ) infrastructure.

- “On-top of” aggregate level ASQ path use cases

  Use cases which use the aggregate level ASQ paths can be understood as “on-top of” aggregate level ASQ path use cases targeting to provide or support specialised transport services towards their final customers, as for instance managed quality connectivity services.

In the following both use case areas are presented in more detail.

3.3.1. AGGREGATE LEVEL ASQ PATH USE CASES

Aggregate level ASQ path use cases focuses on the establishment of an assured service quality (ASQ) infrastructure interconnecting multiple ETICS actors in order to transport the application data in required quality. The actor's content provider, business customer, content delivery network (CDN) and network service provider (NSP) shown in Figure 3 can benefit from the generic ASQ path inter-carrier service.

One example of an aggregate level ASQ path use case is the "assured-quality destination-based Internet transit". Today, when content providers do not use the services of a CDN, they buy bundled connectivity services to reach the entire Internet. Such a service is called the transit service, its pricing is volume-based. An Internet transit service is a-best-effort reachability service to all potential destinations that are connected to the Internet.

In the context of inter-carrier ASQ services, content providers might want to buy a *destination-based assured quality transit service*, that is an ASQ path to a given set of destinations within one (or multiple) access ISP(s). This set of destinations within one (or multiple) access ISP(s) is what we call a *region*.

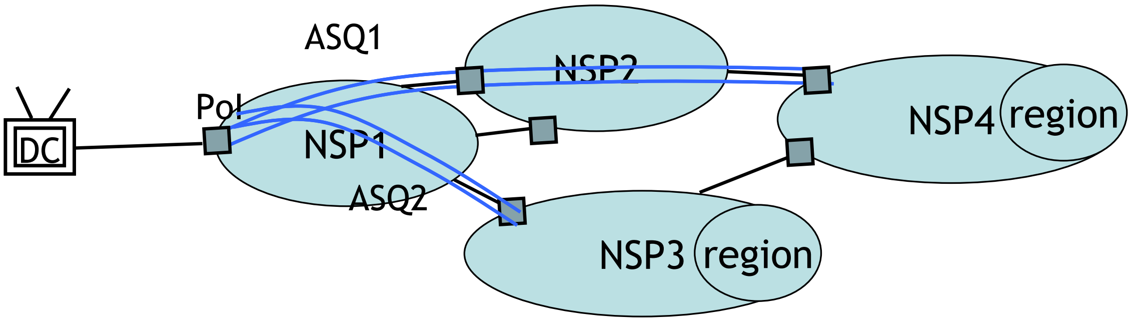


FIGURE 4: ENHANCED INTERNET SERVICE FOR CONTENT PROVIDERS

The transit provider has only control on the traffic that transits within its own network and has no guarantees therefore on the end-to-end paths. To illustrate this, in Figure 4, a destination-based transit provided by NSP1 to the Data Centre on the left cannot have guarantees on how the traffic will flow in NSP2, NSP3 and NSP4. However, an inter-carrier ASQ path to the different regions in NSP4 and NSP3 would give such end-to-end guarantees. (Note: Since the network parts of an ASQ path are operated by the NSPs themselves a real guarantee can only be given by an NSP for the own network part. For the cascading scenario all the "following" NSPs operate their own networks and the "sending" NSP can only trust them. Real end-to-end guarantees can only be given by the ETICS consortia as a whole.)

It is important to note that the traffic in such ASQ paths goes through two phases. (1) A first phase in which it is aggregated (e.g. through NSP1 and NSP2 for ASQ1, and through NSP1 for ASQ2). (2) A second phase, at the ingress point of the last access NSP, in which it is "disaggregated" to reach the different destination hosts within a region. Although it is relatively easy to give traffic delivery guarantees to the aggregated level (1) (e.g. a given aggregate bandwidth and a given delay), it is harder to quantify such guarantees for the "disaggregated" traffic (2) as the traffic will split to reach the different destinations.

Use cases which are related to "disaggregated" traffic and not on the aggregated ASQ path could be covered by the "on-top of" aggregate level ASQ path use cases.

3.3.2. "ON-TOP OF" AGGREGATE LEVEL ASQ PATH USE CASES

"On-top of" aggregate level ASQ path use cases are built "on-top-of" the ASQ path. However, the ASQ path may be transparent for transported applications or network services. Based on such an ASQ path infrastructure the NSPs can offer specialised transport services towards their final customers (e.g. content provider, business customer, neighbouring NSPs etc.), as for instance managed quality connectivity services. In order to realise such "managed connectivity" use cases that are built "on-top-of" the generic ASQ infrastructure additional functionalities have to be implemented and deployed, for instance to enable guarantees for "disaggregated" traffic. These additional functionalities may also have more or less impact at the NSP-to-NSP interactions in order to enable and support/enhance such end-to-end ASQ inter-carrier connectivity as well as other services on-top of the ASQ path.

Content or application providers can demand an ASQ path to a large set of *potential* destinations, called *region*, with only a limited subset of users within the region that can use it simultaneously, as shown in Figure 5. One example of an "on-top of" aggregate level ASQ path use case is the "managed connectivity service for consumer end-users or SMEs". Residential end users may need inter-carrier services for one of the following reasons:

- Have Assured Quality connectivity services to certain content providers' content (that are not directly connected to their access NSP) either in a permanent way or on-demand for a limited amount of time (e.g. watch an HD/3D movie).
- Have Assured Quality connectivity services to certain destination end-hosts in order to have assured quality communication services (e.g. HD Videoconference, telepresence, HD audio).

For both cases the ETICS customers (Edge NSP, Content provider) will request an inter-carrier ASQ path between the involved ingress and egress edge NSPs. Figure 5 presents the use case for managed connectivity service with session handling. In order to support a finer granular service level as for instance to enable session-based connectivity service, additional mechanisms and functionalities have to be implemented. Such additional functionalities could be for instance mechanisms to investigate the available performance (e.g. bandwidth) on the last mile to the end customer and / or to ensure a dedicated performance on the last mile to the customer.

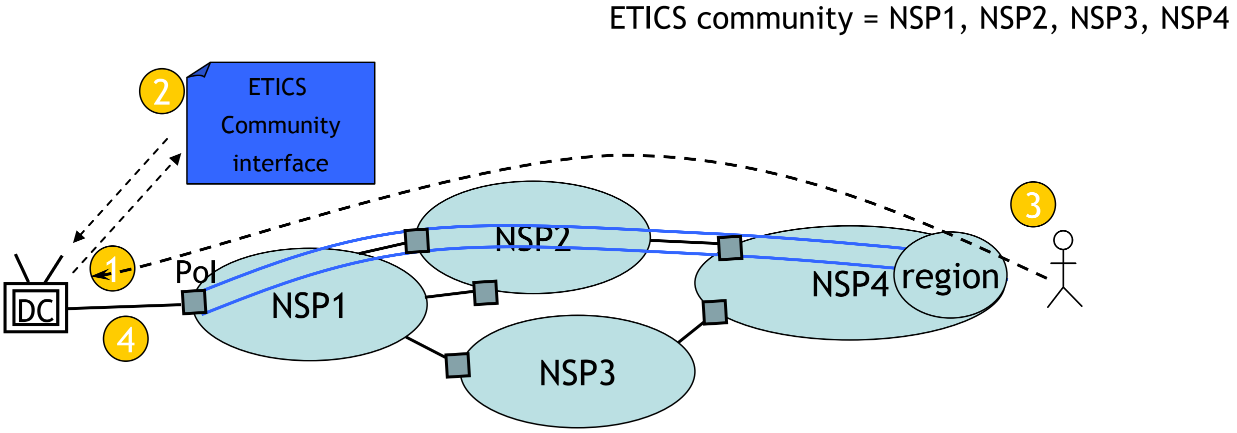

ETICS community = NSP1, NSP2, NSP3, NSP4
2
ETICS Community interface
3
1
PoI
DC
4
NSP1
NSP2
NSP3
NSP4
region

FIGURE 5: ASQ PATH TO A REGION WITH SESSIONS (MICRO FLOWS) HANDLING

These additional functionalities for such a (more fine-granular) individual session-related connectivity based on the generic ASQ path service can be provided by the extension to the core ETICS architecture framework – e.g. with the Service Enhancement Functional Area (SEFA). The SEFA approach is illustrated and described in more detail in Section 6.1, as a generic extension of the ETICS architecture [ETICS-D4.3]. The intention of SEFA is to enrich the basic ETICS architecture and to provide the base for individual as well as specific value added functions and services "on-top of" the ETICS ASQ path (e.g. for establishing and handling of managed (end-user) connectivity services that are carried "on-top-of" the infrastructure level ASQ paths..

3.4. REQUIRED INTER-CARRIER ASQ PATHS

The present section will start with refined service type definitions on the basis of individual ASQ path service requirements and will subsequently aim at giving an inferred schematic ASQ path solution approach for the ETICS architecture. Please, note that technical details will be discussed in subsequent sections.

Leveraging on the overview of ETICS services [D1.6]) and in line with the (on-going) detailed service description provided in [D2.3], the two main types of AQ service requirements are the following:

- **PoI – to – PoI (PoI2PoI)**: Point of Interconnect (Po) to PoI services mainly targeting transit services at the reference points E2, 3, 4 (see actor model in Section 4.4.2), and the related sub-case
    - **PoI-to-PoEI (PoI2PoEI)** where the PoI is replaced by a Point of Enterprise Interconnect (PoEI) service along the E6, 7 reference points (see actor model in Section 4.4.2)

  Please take note that PoI – to – PoI services may are in the general case session-unaware due to scalability reasons. Supporting extensions are presented in Section 6.

- **PoI – to – (Destination) Region (PoI2Region)** with set of IP prefixes as destination identifiers. On the contrary to PoI2PoI services, PoI2Region services may be easily individual session aware or unaware according to NSPs technical constraints or customer needs.

Orthogonally, especially for enterprise services the distinction of inter-carrier Layer 2 or Layer 3 solutions may be of relevance. Thus, the ETICS core architecture will focus on the inter-carrier service needs expressed by all above-mentioned services.

3.4.1. COMPOSED INTER-CARRIER ASQ PATHS TO BE IMPLEMENTED

Through abstracting on the discussed service requirements, the present section aims at schematically describing the ASQ path being capable of addressing each required service type as listed above.

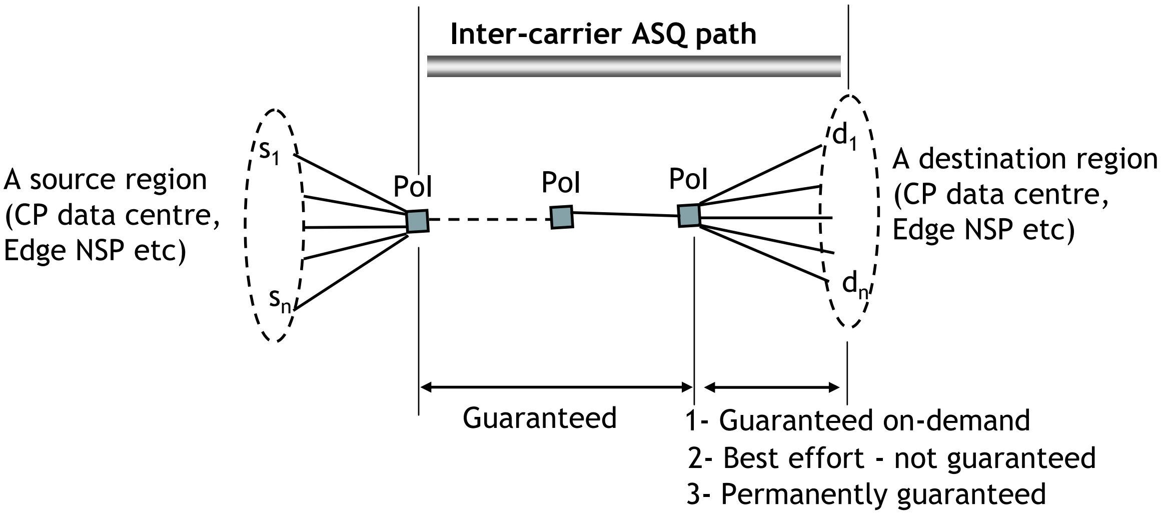


FIGURE 6: COMPOSED INTER-CARRIER ASQ PATHS TO BE IMPLEMENTED

The role of an inter-carrier ASQ path provided by ETICS is to finally deliver traffic from a source region (a set of source hosts) to a destination region (a set of destination hosts) with the special case where the number of destination hosts is one (PoI-to-PoI services). Figure 6 summarizes the inter-carrier ASQ paths that need to be implemented and thus provided by the ETICS solutions. The figure shows a source region, which **needs** an inter-carrier ASQ path to reach a destination region. Please take not that in the figure we assume that the requester of the inter-carrier ASQ path is the source region, it is not therefore included in the ASQ path. These services may be further assisted by the so called Service Enhancement Functional Area (SEFA) being further detailed in Section 6.1 enabling **Host-to-Host communication** (end-user-to[-end-user]/[content-server]) on top of them.

Note that both regions, the source and the destination, **can belong to any of the ETICS actors**. A region in this figure can be a content provider's internal network (e.g. data centre), an edge NSP, or business customer site. Therefore, as represented in the figure, regardless of which actor it serves, an inter-carrier ASQ path, that needs to be implemented by ETICS, is composed of two parts:

- **Part 1: A PoI-to-PoI ASQ path:** This is an "aggregated traffic" part that consists in handling traffic starting from a PoI and transporting it to another PoI (or PoEI respectively). This part always provides guarantees on the traffic delivery (e.g. guaranteed aggregated bandwidth, delay, loss etc).
- **Part 2: A PoI-to-region ASQ path:** This is a "disaggregated traffic" part in which traffic is handled starting from a PoI and is delivered to a region, that is, multiple different destination hosts. This part of the ASQ path can be:
  - **Best effort or not guaranteed:** In this case, the traffic delivery from the PoI to the region does not obtain any traffic delivery guarantees. Note that an inter-carrier ASQ path terminating on this type of best effort region can be seen as only a PoI-to-PoI ASQ path that connects to this best effort destination region.
  - **Permanently guaranteed**: In this case, the resulting inter-carrier ASQ path (Part1 + Part2) gives traffic delivery guarantees to each single destination host in the region (This lasts the entire life of the ASQ path).

- **Guaranteed on-demand:** In this case, a given host in the destination region will be assigned traffic delivery guarantees only upon an explicit demand of the host, i.e. by specification through a Service Level Agreement. Therefore, one of the main tasks of this service is the ability to dynamically associate given traffic flows (individual sessions) on-demand to the ASQ path, e.g. realised with SEFA (cf. Section 6.1).

3.4.2. Single-NSP ASQ Paths For Composition

An inter-carrier ASQ path results from the concatenation of single-SP ASQ paths, whether PoI-to-PoI or PoI-to-region. We remind that a single-NSP ASQ path is a network connectivity service provided by a single NSP. We enumerate the following single-NSP ASQ paths that each NSP, participating in the establishment of an inter-carrier ASQ path, needs to implement.

- PoI-to-PoI ASQ paths (provided mainly by transit NSPs; please take note that PoI-to-PoEI may represent a special case not further detailed in this section)

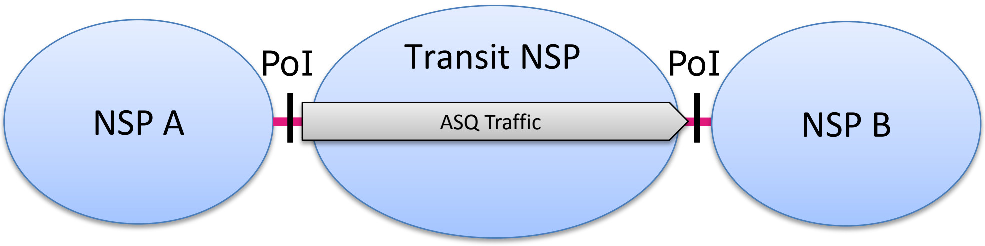


FIGURE 7: SINGLE NSP POI-TO-POI ASQ PATH

- Guaranteed on-demand PoI-to-region path (provided by an edge/access NSPs).

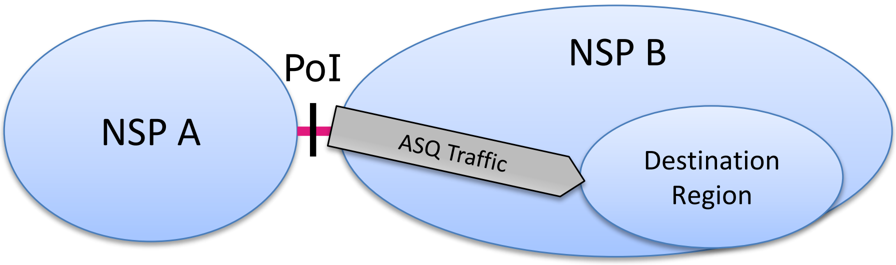


FIGURE 8: POI-TO-DESTINATION REGION FOR A SINGLE NSP

- Best effort (non guaranteed) PoI-to-region path (this service does not need to be implemented but may be rather advertised to make a binding between A PoI and the regions that it serves)
- Permanently guaranteed PoI-to-region ASQ paths (Provided by edge/access NSPs)

4. OVERVIEW OF THE ETICS SOLUTION

This section presents the overall ETICS solution and its main principles, solution elements and deployment modes. The ETICS solution represents the holistic view where the vision for the interconnected networks and the network services are pointed out. It also establishes the context for the ETICS core-system architecture for managing the ETICS assured quality (AQ) network services and SLAs. The ETICS solution enables the participating NSPs to trade network services among each other and as a result provide inter-carrier AQ network services to the end-customers. The NSP-to-NSP network services and charging principles are an integral part of the ETICS solution. Taking the current regulatory, business and technological contexts for NSPs into account, such a holistic approach is not straightforward. Therefore, the suggested roadmap helps to understand how we recommend progressively introducing such AQ Services and its different technical components, from a realistic bootstrapping phase to hopefully a mature solution facilitating more advanced elements and improved operational and ecosystem efficiencies.

The overview of the ETICS solution is structured into the following subsections. In Section 4.1, we provide the “big picture” overview of the solution together with the vision and goals that have directed the solution design. Section 4.2 introduces the ETICS core-system architecture and accordingly the main deployment modes (scenarios). Section 4.3 identifies buyer-supplier actor roles and relationships in the context of the main collaboration modes, and briefly reminds of main charging principles, while Section 4.4 concludes the ETICS solution overview by suggesting a roadmap for the various deployment scenarios and collaboration modes considering each of the main solution building blocks.

4.1. VISION AND GOALS

Subsequently we will present an illustrative view of the overall “ETICS solution” enabling inter-carrier AQ network services. For this purpose, we capture our vision, assumptions and high-level goals. The AQ network services offered by an ETICS-enabled NSP either toward other ETICS NSPs or external actors are managed and controlled by a set of ETICS core-system elements, which enables the automated order, activation and assurance of the ETICS network services and SLAs.

The overall “ETICS solution” takes the interconnected networks and their services to be managed (blue element in FIGURE 9), as well as the ETICS core-system elements for managing these AQ inter-carrier network services (green elements in FIGURE 9) into consideration.

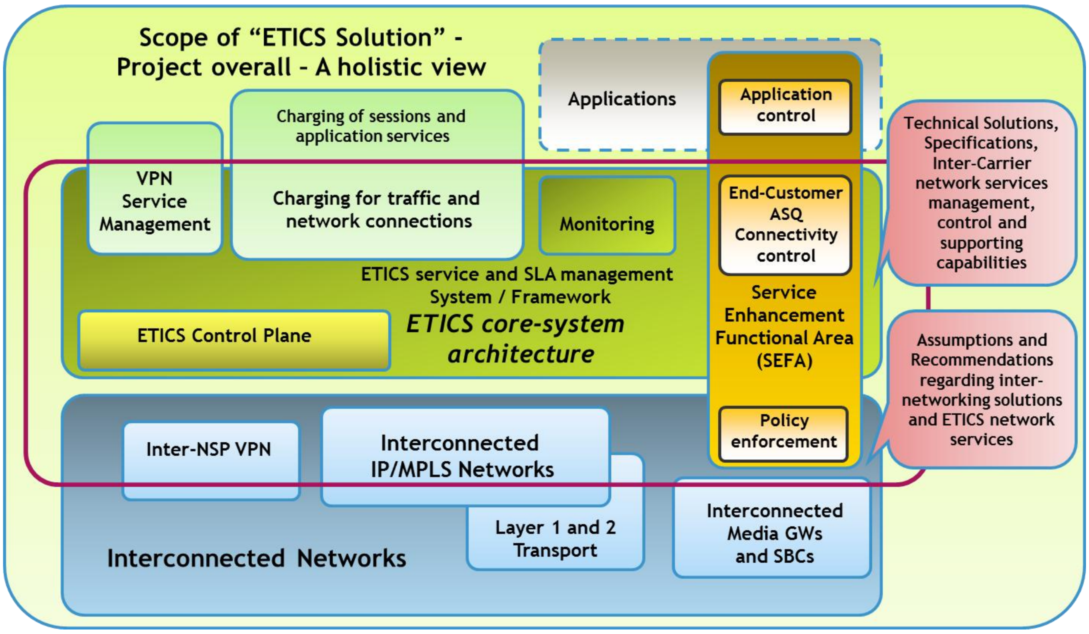

Scope of "ETICS Solution" - Project overall - A holistic view
Applications
Application control
Charging of sessions and application services
VPN Service Management
Charging for traffic and network connections
Monitoring
End-Customer ASQ Connectivity control
Technical Solutions, Specifications, Inter-Carrier network services management, control and supporting capabilities
ETICS service and SLA management System / Framework
ETICS Control Plane
ETICS core-system architecture
Service Enhancement Functional Area (SEFA)
Assumptions and Recommendations regarding inter-networking solutions and ETICS network services
Policy enforcement
Inter-NSP VPN
Interconnected IP/MPLS Networks
Layer 1 and 2 Transport
Interconnected Media GWs and SBCs
Interconnected Networks

FIGURE 9: "ETICS SOLUTION" OVERVIEW

The interconnected networks in focus by ETICS are based on IP/MPLS and to some extent Layer 2 networks. This includes VPN networking solutions. On the other hand interconnected media gateways and session border controllers or not within direct scope. The ETICS core-system for managing AQ services and resources at the inter-NSP level enables and support proper Service and SLA life-cycle management including SLA assurance. Monitoring are considered in a separate building block that can be deploy and integrated in several ways. Charging at the ASQ traffic and connection level is an integral solution element, whereas session and application level charging is out of scope. Inter-NSP VPN management is positioned as a potential solution extension. For the purpose of handling and supporting the end-customer or end-user ASQ connectivity services that are carried "on-top-of" the ETICS ASQ paths (see Section 3) the Service Enhancement Functional Area has been defined as an additional extension building block. For more details see Section 6.

The following statements reflect the projects iterative efforts and attempts to consider both business and economic aspects on one side and technology considerations on the other. More recommendations from the business and economic side are provided by [ETICS-D3.5] (see for instance Section 2). While taking these recommendations into account, this section presents the more technical oriented **directions, assumptions and vision**.

Note that while the solution vision is provided in this section it has not been possible in the time-frame of the project to implement all aspects of the ETICS solution, nor all the solution elements. Moreover, the vision or direction statements below must be balanced according to the recommended gradual approach and step-by-step roadmap proposed in Section 4.5.

(1): A common platform for the future Internet and Extranets

The future networking solutions are evolving into a more unified overall networking platform still allowing a heterogeneous set of networking technologies. Each of the networking technologies has their merits and a context where they are more optimal. The future networking platform will include a *network service and business plane* that can enable and support the inter-NSP trading of network services and the automated management and control of services and SLAs according to industry drivers and evolution. The overall solution will support the addressing space of the Internet, as well as the private address spaces for private and public extranets such as enterprise VPNs and NSP backbone extranets. The solution will accordingly ensure connectivity with assured quality between compatible end-points.

(2): Evolution from the current BE Internet

An Internet revolution is not foreseen, rather a step-by-step evolution. The project is suggesting a bootstrapping phase as well as an ecosystem evolution or roadmap in this respect. In particular we expect that the Internet BGP will mostly stay unchanged[6]. The introduction of AQ interconnection for Internet traffic delivery should be introduced as a complement to the existing agreements, e.g. by proposing suitable services and collaboration models for AQ interconnection.

(3): A generic ASQ internetworking solution supporting a future rich set of Internet Application Services

The future Assured Quality Internet will offer a limited set of AQ connectivity services for the end-customers. Likewise, the basic future inter-carrier AQ service business models will serve a feasible and straightforward design (e.g. bilaterally cascading), i.e. primarily based on the SPNP principle evolving side-by-side to the current business models. This will constitute an internetworking solution and business model that can be the generic connectivity platform for the enablement of the anticipated rich set of future Internet application services.

(4): Connection oriented vs. connection less

In order to maximize the impact and capture the largest number of NSPs, NSPs have the freedom to choose whatever technology they find suitable for implementing ASQ paths. It can be either connection-less (e.g. pure IP, Diffserv, MPLS with LDP) or connection-oriented architectures (e.g. MPLS with RSVP), as they are currently doing.

(5): Scalability and operational efficiency

The design of the ETICS core-system has been driven by attention to operational efficiency, as well as for scalability both at the ETICS system level as well as at the internetworking level.

(6): Big ASQ pipes (paths) vs. end-customer ASQ connectivity services (e.g. micro-flows)

In addition to scalability concerns, it is not realistic to set-up or compute inter-carrier ASQ paths each time there is a demand for an individual user (session) connectivity service or even for managed ASQ connectivity services for small and medium enterprises (SME). ASQ paths are set only to provision the “big

[6] Note that this may still not exclude updated best current practice usage of the current Internet BGP system, for instance for the purpose of disseminating some limited ASQ related capability information of an IP prefix. This is a topic for further study beyond the project.

AQ pipes" at the infrastructure level enabling ASQ paths for significant bandwidth demands of enterprise customers (such as content and application providers).

(7): Modular ETICS core-system architecture

The core-system architecture must be flexible and modularly prepared for many deployment scenarios and different preferences by different NSPs. The ETICS core-system can be considered as a solution framework that can be utilized and configured in several ways by each NSP, by selecting the needed solution elements or modules and adapting them by further configuration to its particular business role and strategy for trading ASQ goods. Moreover, one may also speak of the set of such interoperable ETICS core-system instances constituting an ETICS community-wide system. This modularity plays a special role in combination with the alignment to various market phases, i.e. an architectural roadmap (see Section 4.5), and appears to us as fundamental, knowing the current uncertainties on the future of such AQ service market.

(8): Ready to wear versus made to measure

In order to allow for flexibility, ETICS NSPs have the choice between (1) offering "ready-to-wear" single-NSP ASQ paths with predefined quality and (2) providing "made to measure" single-NSP ASQ paths upon a request. In the latter case, the NSPs would need to divulgate loose network capabilities (with more or less details) in order to maximize their chances of getting selected to provide a single-NSP ASQ path. Providing "made-to-measure" (aka. ***pull***) ASQ path does NOT mean that the NSPs will necessarily build the single-NSP offer only upon the reception of a request for it. NSPs are free to pre-build or pre-compute offers if they find it suitable. In the same way, providing "ready-to-wear" (aka. ***push***) single-NSP ASQ paths does NOT force the NSPs to pre-provision in their network explicit paths satisfying these offers. We expect that in a mature market, push and pull modes may be used simultaneously to propose lower cost standard/classical pushed offers for well-known needs, and higher costs pull offers for specific needs.

(9): Governance assumptions

Again, for the sake of flexibility, the ETICS project has not taken a decision so as to which actor will take critical decisions in terms of service composition or traffic routing. Instead, we provide a rich set of deployment scenarios that are different ways to utilize the ETICS core-system architecture, in order to trade, realize and manage the different types of network services as recommended by ETICS. The idea is to let the market decide so as to which scenario is better in the different contexts that will exist side-by-side.

4.2. FULLY AUTOMATED SOLUTION: A USER GUIDE

We first present a high level picture of how the ETICS solution would work in a fully automated setting. An ETICS customer willing to obtain an inter-carrier ASQ path from the ETICS NSPs will go through two steps.

4.2.1. STEP 1: DISCOVERY OF "ASQ-ENABLED" REGIONS

Having no idea about which destination regions are served by the ETICS community, the ETICS customer will first connect to the ETICS portal. The *ETICS portal* is the interface between the ETICS community and the ETICS customers. In this portal the customer can find information about all the regions that are served by the ETICS community. A possible description of regions is further specified in Sec. 5.3.1.2. Depending on the type of service through which the region is reachable, it can be either:

1) A bare description of the region that is served. This is the case when the access NSP which serves the region does not provide guarantees to the individual hosts within the region. An example of such description might look like: The region of Berlin, IP prefixes covering the region, through which PoI it is reachable etc.

2) A more complete description of the capabilities associated to the region. This is the case when the access NSP can provide guarantees (permanent, or on-demand) to the individual hosts within the region. An example of such description would contain, in addition to the previous example, the type of service that the region can serve (guaranteed, guaranteed on-demand), together with the necessary quality (e.g. Number of individual sessions, and some of the (non-additive) QoS parameters that can be associated with each session).

4.2.2. Step 2: Find an inter-carrier ASQ path to reach the desired region(s)

Now that the customer has an idea about what regions it can reach through the ETCS community, the second step is to look for an inter-carrier ASQ path to reach one of these regions. As multiple paths might exist between the customer PoI and a given region, multiple combinations of different single-NSP ASQ paths can satisfy the customer's inter-carrier ASQ path need. We refer to the task of finding the right combination of single-NSP ASQ paths to form an inter-carrier ASQ path as *service composition*. This task can be either done by the community of NSPs or by the customer itself.

4.2.2.1. Service composition is done by the community of NSPs

In this second step, and as illustrated by Figure 10, the customer delegates the service composition task to the ETICS community. The customer will use the information obtained from step 1) to formulate an inter-carrier ASQ path request from the ETICS community. The ETICS community will propose a set of inter-carrier paths to satisfy the customer request. The customer then chooses one of these offers and orders it. We will precise later which actor within or outside the ETICS community will play the role of this interface, and make the composition. Depending on this decision, there will result multiple deployment scenarios of the ETICS architecture, each of them having a different consequence both on the business and on the technical solution.

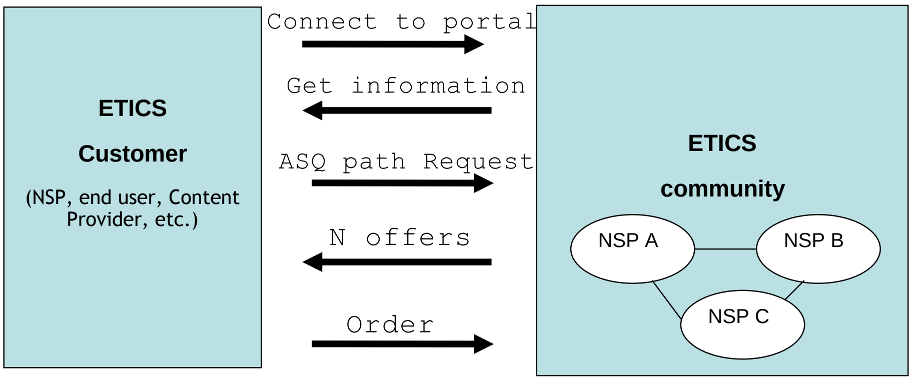


FIGURE 10: REQUESTING INTER-CARRIER ASQ PATHS FROM ETICS (WHEN THE ETICS CUSTOMER DOES NOT DO THE SERVICE COMPOSITION)

4.2.2.2. Service composition is done by the customer itself

Note that another alternative to the scheme described in Figure 10 is the following. The ETICS community is not responsible for composing the service and providing an offer. Instead, the ETICS customer is able to access a sort of catalogue where it can find information about all the single-NSP ASQ paths (from the different NSPs). In this case, the ETICS customer is responsible for **composing its own inter-carrier ASQ path**, i.e. deciding which single-NSP ASQ paths to buy and to concatenate in order to have an inter-carrier ASQ path. Although from a technical perspective, such variant is similar to the previous one, the economic consequences might differ a lot.

4.2.3. SLA LIFECYCLE (PROCESS WORKFLOW)

The ETICS architecture is governed by different processes that must follow a given work flow. The goal of this work flow is to implement the inter-carrier ASQ paths described in the previous section. Figure 11 below shows the different steps that govern the life cycle of an inter-carrier ASQ path.

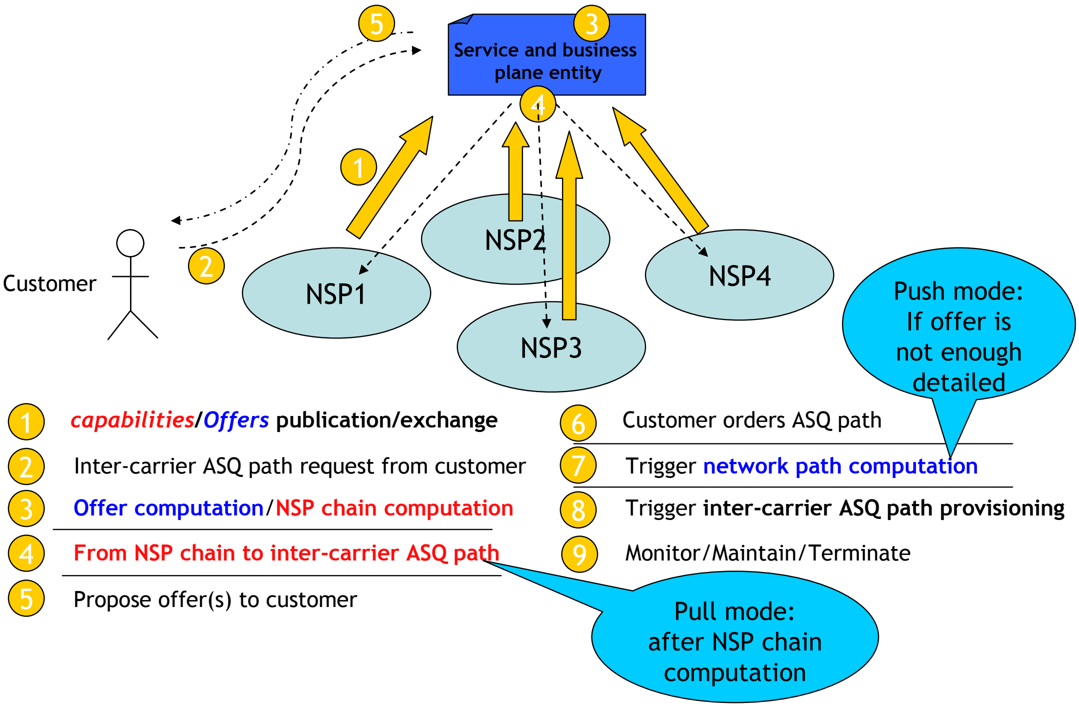

5
3
Service and business plane entity
4
1
Customer
2
NSP1
NSP2
NSP3
NSP4
Push mode: If offer is not enough detailed
1 capabilities/Offers publication/exchange
2 Inter-carrier ASQ path request from customer
3 Offer computation/NSP chain computation
4 From NSP chain to inter-carrier ASQ path
5 Propose offer(s) to customer
6 Customer orders ASQ path
7 Trigger network path computation
8 Trigger inter-carrier ASQ path provisioning
9 Monitor/Maintain/Terminate
Pull mode: after NSP chain computation

FIGURE 11: ETICS WORK FLOW

These steps could be grouped by means of phases in the ETICS work flow (these phases were first summarized in Deliverable D1.6 [ETICS-D1.6]:

- Step 1 corresponds to an initialization phase in which the NSPs of a community get ready to receive inter-carrier ASQ path from ETICS external customers.
- Steps 2 to 5 that correspond to the Invocation phase where an ETICS customer requests an inter-carrier ASQ path,
- Steps 6 to 8 design the phase where NSPs are enforcing the inter-carrier ASQ path in their respective networks,

- Step 9 simply designs the monitoring and management phase of the ASQ path including its termination.

Almost each one of these steps has a corresponding subsection in Section 5 that details it.

In a bit more details, the first step (1) is to get the service and business plane aware of the different network capabilities or ASQ path offers of the different NSPs of the ETICS community. This is covered by the **capabilities/offers publication/exchange** functionality. This functionality is detailed in Sections 5.3.1. Then, the ETICS system is triggered by an ETICS customer inter-carrier ASQ path request (step2). This will trigger first an Offer computation in the push mode and an NSP chain computation in the pull mode (step3). The difference between pull and push will be explained in the next section. This step is covered by the **Offer computation/NSP chain computation** functionality and is detailed in section 5.4 .

Next, in the pull mode, the next step (4) is to go **from the NSP chain to a precise inter-carrier ASQ path**. This step is only needed in the pull mode, because in the push mode, the Offer computation has already determined an inter-carrier ASQ path. This functionality is covered in section 5.6.3.

Once we have one or multiple inter-carrier ASQ path(s), we offer it to the customer (5). If the customer accepts the offer, it orders it (6). If the inter-carrier ASQ path is not enough detailed, an extra step is needed (7). This step is detailed in section 5.6 (and Section 5.6.2 in particular).

Once the network path is computed, the next step (8) is to **provision the inter-carrier ASQ path.** This functionality is covered in section 5.7.

During the life of the inter-carrier ASQ path, the ETICS architecture needs to **monitor and maintain** the ASQ path (step 9). The monitoring step is covered in section 5.8.

Finally, the last step is then to terminate the service once the contract takes end.

4.3. DIFFERENT DEPLOYMENT SCENARIOS TO SATISFY INTER-CARRIER ASQ PATH CUSTOMER REQUESTS

In its automated flavour, finding an inter-carrier ASQ path that satisfies a customer request is the goal of the ETICS architecture. The service composition refers to the task of finding the right set of single-NSP ASQ paths in order to satisfy a given request. We have previously seen in this section that there exist at least two possibilities depending on who does the service composition: first, the customer is responsible for the service composition. Second, the ETICS community of NSPs is responsible for such task. Although from a technical perspective, the difference between these two cases is limited, the business and economic implications are of higher importance.

In order to satisfy a variety of differing business demands partly associated with different phases of the rollout of the ETICS system, multiple deployment scenarios have been defined. With this, **we aim at providing *the set of tools or capabilities to implement all these different scenarios, and let the market decide on which approach to deploy in which rollout phase*.** *Further details on the phased rollout process are introduced in Section 4.5 and complemented by driving business considerations in [ETICS-D3.5].*

Depending on which actor will make the service composition and which information is disclosed by NSPs, we define in ETICS different deployment scenarios that we briefly present later in this section. The reader should refer to deliverables D4.3 [ETICS-D4.3] and D5.2 [ETICS-D5.2] for further details about these deployment scenarios. The deployment scenarios essentially follow two dimensions, a governance dimension, and an information disclosure dimension:

i) The governance dimension defines **who** takes the critical decisions in terms of service composition:

- **Customer-decision** (in reference to Section 4.2.2.2)
- **Community-decision** (in reference to Section 4.2.2.1) in which case it can be:
  - **Fully-Centralized mode**, a centralized entity, called *facilitator*, takes the decisions. The facilitator is regarded to be a neutral technical entity being not controlled by a(n) (single) NSP.
  - **Per-NSP Centralized mode**, an NSP (not necessarily always the same one) centrally takes the decisions and offers its composition service to other (potentially smaller) NSPs which do not own the capabilities or state of information to complete a request on their own.
  - **Distributed mode**, decisions are taken following processes distributed along the NSPs.

ii) The information disclosure (publication) dimension which defines whether the service is tailored or provided for a demand or built or provided prior to any demand. This dimension affects the granularity of the disclosed information at a service plane level, but also how the service composition is done. We distinguish between two modes:

- **Push mode**: service **(single-NSP ASQ path) offers** are **published** prior to any explicit request for them. Such services are published – but not necessarily reserved - *prior to any demand*. Offers contain detailed properties of single-NSP ASQ paths and are *ready to be ordered* by ETICS actors. In such mode, the role of the service composition entity consists mainly in **computing the best combination of single-NSP ASQ path offers that satisfy a given inter-carrier ASQ path request**. Entailed by the analysis of [ETICS-D3.3] and [ETICS-D3.5] it may be stated that such an approach may be trimmed towards more mature markets with more predictable demands. This results from the binding character of offers, as well as the flooding of potentially unnecessary information in the network, which represent two internalized risks.
- **Pull mode:** In this mode, service **(single-NSP ASQ path) offers** are only provided upon an explicit demand for them. In this mode, in order to help the entity which does the service composition, service **capabilities** are **published** (instead of offers). Service capabilities can have the same format as offers except that they can be less detailed, and if they are as much detailed, they **cannot** be considered to be legally binding offers. They can vary from simple reachability information (Simply entry point, exit point and interconnection) to more detailed capabilities that can include delay for example. Regardless of their level of detail, the role of service capabilities is **only** to **help the service composition entity to select the NSPs from which it can request single-NSP ASQ paths**, and use them to form an inter-carrier ASQ path. As such, the pull mode may complement the push mode in phases where the demand may not be sufficiently predictable for NSPs. However, due to the complexity of budget splitting (see [ETICS-D3.5]), an automated configuration of instances of the pull scenarios need to be avoided. We tag them as non-automatable in this deliverable.

The above-mentioned dimensions lead to different combinations or deployment scenarios. Most of these options or deployment scenarios have been clarified in the architectural deliverable [ETICS-D4.3], and analysed within the framework of the WP3 "business" feedback on the architecture [ETICS-D3.3, section 5][ETICS-D3.5]. Based on this feedback non-automatable solutions are extracted from the initial set of deployment scenarios. These modes essentially affect the critical decisions associated to the publication of capabilities or offers, and the selection of NSPs that will participate in providing an inter-carrier ASQ path. Hence, all the network functionalities relative to the provisioning of the inter-carrier ASQ (path signalling, traffic identification and monitoring etc.) are either agnostic or barely affected by these modes (for instance, the governance dimension might impact the data collection of monitoring).

We next go through the different deployment scenarios of the ETICS architecture consisting of a combination of decisions on the governance and information disclosure dimension. We only present the community–decision governance model, but we remind the reader that the customer-decision governance model is also possible (introduced in Section 4.2.2.2). For the sake of simplicity, we assume that the final customer has connected to the ETICS portal (the step in Section 4.2.1) and is able to formulate an inter-carrier ASQ path request. We illustrate all the scenarios through a toy example representing a community of four NSPs and a user which requests an inter-carrier ASQ path.

4.3.1. Fully-centralized push deployment scenario

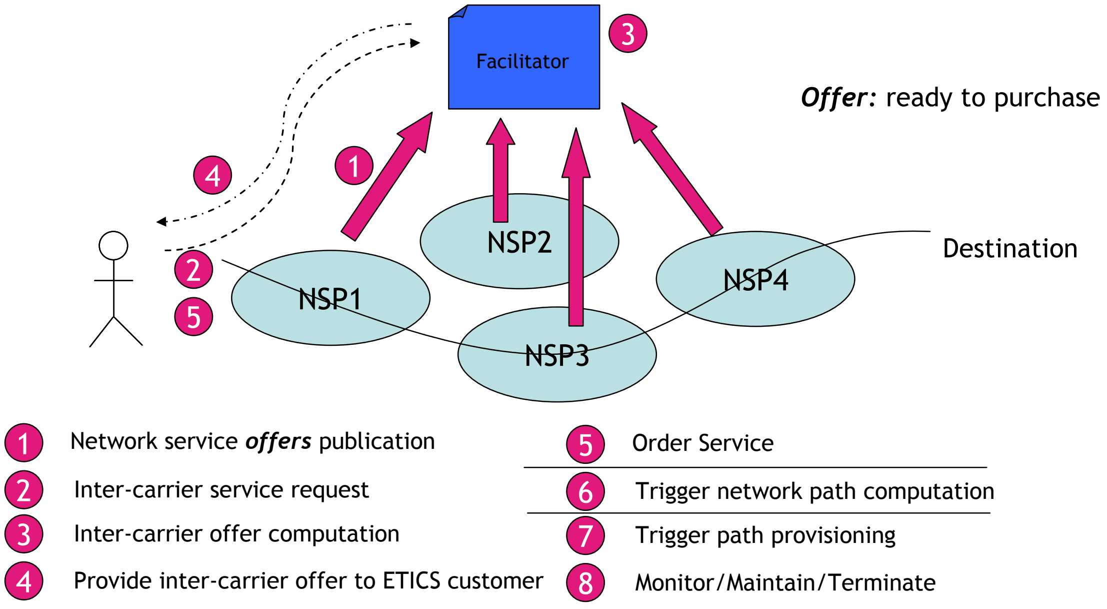

Facilitator
3
Offer: ready to purchase
4
1
NSP2
Destination
NSP4
2
NSP1
5
NSP3
1 Network service offers publication
2 Inter-carrier service request
3 Inter-carrier offer computation
4 Provide inter-carrier offer to ETICS customer
5 Order Service
6 Trigger network path computation
7 Trigger path provisioning
8 Monitor/Maintain/Terminate

FIGURE 12: STEPS OF THE FULLY-CENTRALIZED PUSH DEPLOYMENT SCENARIO

In this scenario, the service composition is done by a central facilitator, an entity which is not an NSP.

This scenario is one of the two scenarios that have been chosen by ETICS for the detailed specification and the prototypes at an earlier stage of the project. Figure 12 shows how this scenario works by showing the different steps involved in the service composition. In this scenario, NSPs publish their single-NSP ASQ path offers (PoI-to-PoI or PoI-to-region) to the centralized facilitator (1). This allows the facilitator to perform the service composition (an offer computation) (3) once it receives an inter-carrier service request from a customer (2). The facilitator proposes the offer(s) to the customer who might order them, in which case the inter-carrier path is provisioned.

4.3.2. Per-NSP centralized push deployment scenario

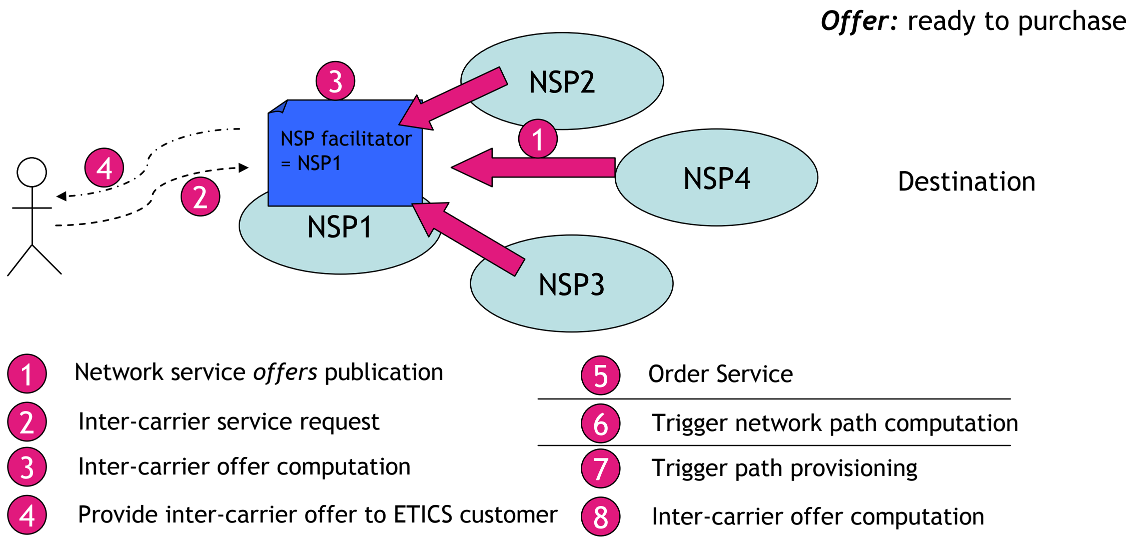

Offer: ready to purchase
3
NSP2
NSP facilitator = NSP1
1
NSP4
Destination
4
2
NSP1
NSP3
1 Network service *offers* publication
2 Inter-carrier service request
3 Inter-carrier offer computation
4 Provide inter-carrier offer to ETICS customer
5 Order Service
6 Trigger network path computation
7 Trigger path provisioning
8 Inter-carrier offer computation

FIGURE 13: REQUESTING INTER-CARRIER ASQ PATHS FROM ETICS (WHEN THE ETICS CUSTOMER DOES NOT DO THE SERVICE COMPOSITION)

This scenario is very similar to the scenario in Sec.4.3.1, except that the central facilitator role can be played by one (or any) of the NSPs in the community. Figure 13 shows the different steps of this deployment scenario from the offer publication till the offer fulfilment. In terms of prototyping, components are exactly the same than for the fully centralized push scenario, so that this scenario is also implemented when having at least two facilitators deployed in two different domains.

4.3.3. Distributed push deployment scenario

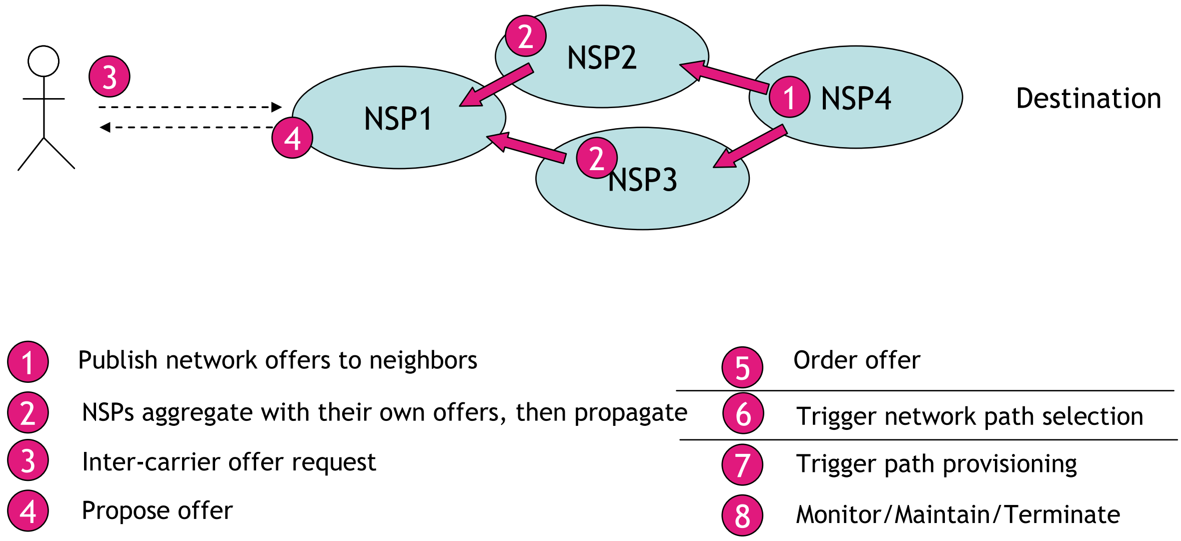

2
NSP2
3
NSP1
1 NSP4
Destination
4
2
NSP3
1 Publish network offers to neighbors
2 NSPs aggregate with their own offers, then propagate
3 Inter-carrier offer request
4 Propose offer
5 Order offer
6 Trigger network path selection
7 Trigger path provisioning
8 Monitor/Maintain/Terminate

FIGURE 14: DISTRIBUTED PUSH DEPLOYMENT SCENARIO (STEPS OF THE SERVICE COMPOSITION)

Figure 14 shows the different steps involved in this scenario. This scenario is push, meaning that NSPs publish ASQ path offer prior to any explicit request for them. This deployment scenario mimics the philosophy of BGP but at the service plane level. The idea of this scenario is that each NSP will publish its ASQ path offers to its neighbours (step1). An NSP which receives an ASQ path offer aggregates it with its own offers then send it as well to its neighbours (step2). After a “convergence” step, all the NSPs in the community will have in their catalogues ASQ paths to reach all the destinations reachable by the community. Once an NSP receives an inter-carrier ASQ path (step3), it will immediately know by looking into its offers catalogue or database if such offer is available through the community. It can then offer it to the customer (step4). If the customer is willing to take the offer, it can order it (step 5) and the remaining steps 6 to 9 are similar to the ones already described in the previous sections.

4.3.4. Distributed Pull deployment scenario

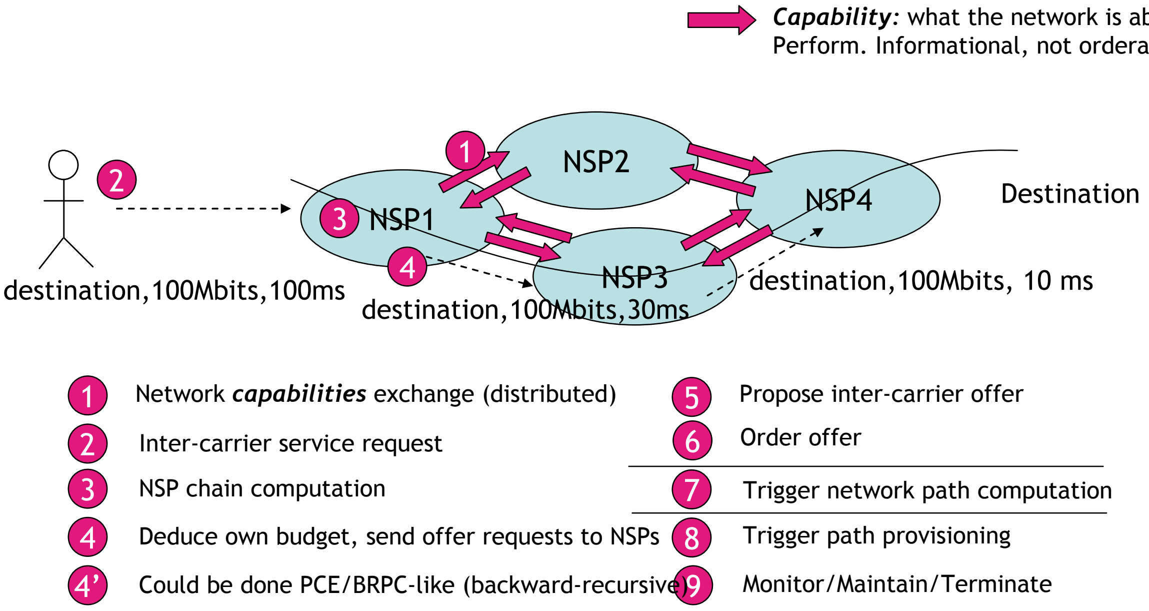

Capability: what the network is able to Perform. Informational, not orderable
NSP2
1
2
3
NSP1
NSP4
Destination
4
NSP3
destination,100Mbits,100ms
destination,100Mbits,30ms
destination,100Mbits, 10 ms
1 Network *capabilities* exchange (distributed)
2 Inter-carrier service request
3 NSP chain computation
4 Deduce own budget, send offer requests to NSPs
4’ Could be done PCE/BRPC-like (backward-recursive)
5 Propose inter-carrier offer
6 Order offer
7 Trigger network path computation
8 Trigger path provisioning
9 Monitor/Maintain/Terminate

FIGURE 15: STEPS OF THE DISTRIBUTED PULL DEPLOYMENT SCENARIO

In this scenario, the service composition is done in a distributed way and all the NSPs participate in the decisions. In this scenario as well, NSPs do not publish single-NSP ASQ path offers but rather loose capabilities. The goal of these capabilities is only to help the other NSPs make the service composition. The NSPs obtain single-NSP ASQ paths from each other by explicitly requesting them.

The stepwise process of this scenario is given in Figure 15. In a first step, service capabilities are flooded between the different NSPs (1). If a service request is received by an NSP(2), it uses these capabilities to compute an NSP chain (3) that is more likely to satisfy the customer request. The second part of the service composition procedure is done then in a distributed way(step 4 or 4’ or step 4 refined later by step7).

One way of doing this distributed procedure is the following. The first NSP (NSP1) will deduct its own QoS budget (example how much delay it will “consume” in its own network) from the end-to-end budget, and then send a new inter-carrier ASQ path request with the remaining budget to the next NSP(s) in the NSP chain(s). Once the request reaches the last NSP, this NSP will answer by giving an ASQ path offer to the upstream NSP from which it received the request. The upstream NSP will use this offer, concatenate it with its own offer, and send it back to its upstream NSP. This procedure is repeated in all the NSPs in the chain, until the first NSP obtains the inter-carrier ASQ path. This solves the problem of the budget splitting since each NSP is responsible of deducing its own budget from the entire end-to-end QoS budget.

In another variant of step 4, a similar procedure can be done "backward recursive" in a very much similar way to the backward recursive path computation procedure of the PCE architecture [RFC5441].

Finally, once the distributed procedure is completed, the first NSP can propose the computed offer to the customer which can order it.

ETICS considered two variants of this deployment scenario. In the first, ASQ paths returned by the different NSP during step 4' are enough detailed to allow for direct provisioning. This is the case of the H-TE solution described in Section 7.3.1. In the second (service plane offer negotiation approach), the ASQ paths returned by the different NSPs are not precise in which case the additional step 7 (network path computation) is needed.

The first variant of this scenario (H-TE) is one of the two scenarios that have been chosen by ETICS for the detailed specification and the prototypes. [ETICS-D3.5] analysed this deployment scenario and **recommended that the price setting (putting a price on an offer) should not be automated**, or price fluctuations may be fierce and ruin the system ("price and niche wars"). As such, human intervention in deciding the prices should be maintained.

The variant of this scenario that was chosen for prototyping (H-TE) does not automate the price setting. The prices are rather configured by the administrator of the network. Configured prices are conveyed within the capabilities messages during the capabilities exchange step. The final price is obtained during the BRPC composition procedure.

4.3.5. Non-automatable Deployment Scenarios

The following three on-demand (pull) deployment scenarios were not recommend for automation by the ETICS workpackage 3 (WP3) for business considerations [ETICS-D3.5] (also see [ETICS-D4.3],[ETICS-D5.2]). This reasons have been already been briefly recaptured in the introduction of Section 4.3.

4.3.5.1. On-demand (pull) Fully Centralized entity deployment Scenario

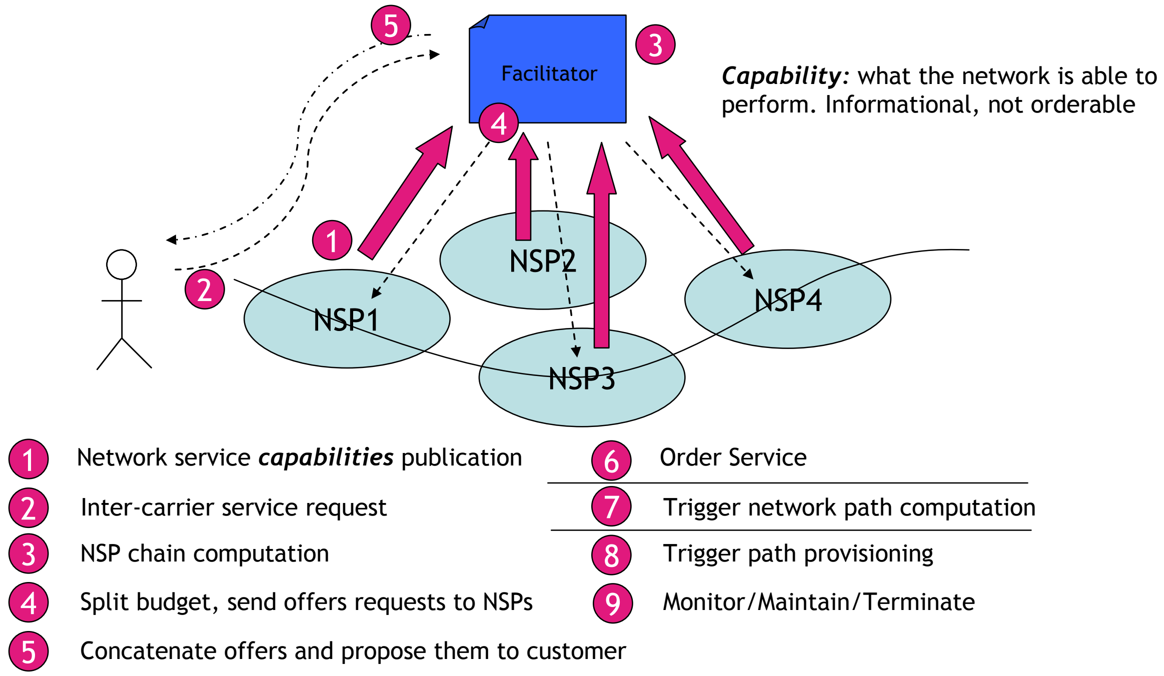

5
Facilitator
3
Capability: what the network is able to perform. Informational, not orderable
4
1
2
NSP1
NSP2
NSP3
NSP4
1 Network service capabilities publication
2 Inter-carrier service request
3 NSP chain computation
4 Split budget, send offers requests to NSPs
5 Concatenate offers and propose them to customer
6 Order Service
7 Trigger network path computation
8 Trigger path provisioning
9 Monitor/Maintain/Terminate

FIGURE 16: STEPS OF THE PULL CENTRALIZED ENTITY DEPLOYMENT SCENARIO

In this scenario, the service composition is done by a central facilitator – a technical entity whose functioning shall not be dominated by an NSP (see [ETICS-D3.3][ETICS-D3.5]). In addition, NSPs do not publish single-NSP ASQ path offers but rather loose network capabilities whose role is only to help the facilitator to make the service composition. The central facilitator obtains single-NSP ASQ paths by requesting them from the different NSPs.

The stepwise process of this scenario is given in Figure 16. In a first step(1), NSPs publish their network service capabilities to the centralized facilitator. These capabilities can range from only topological (connectivity) information to more detailed description of network capabilities. The more detailed are the capabilities, the easier would be the service composition. The service composition is triggered by an inter-carrier ASQ path request that an end user sends to the facilitator (step2). The service composition is two-fold. First, the central entity computes the NSP chain(s) that is (are) more likely to satisfy the user demand (3). Second, it splits the end-to-end QoS budged between the NSPs in a chosen chain (e.g. how much delay to request from each NSP) then send a separate request to each of the NSPs in order to obtain single-NSP ASQ path offers.

Note that if the capabilities are too loose (e.g. only interconnection topology), too many NSP chains can possibly satisfy the desired customer request (in step 3). Some studies in ETICS [Djarallah11] *tried to tackle this problem by proposing solutions to explore multiple NSP chains in parallel (in step 4).*

Finally, the facilitator will concatenate these single-NSP ASQ paths to form one (or more) inter-carrier ASQ paths. It then sends the most suitable one(s) to the customer, which can order an offer if it satisfies its demand.

Due to the budget splitting difficulties discovered in [ETICS-D3.5], this deployment scenario was not recommended by WP3 (business aspects).

4.3.5.2. Per-NSP centralized Pull deployment scenario

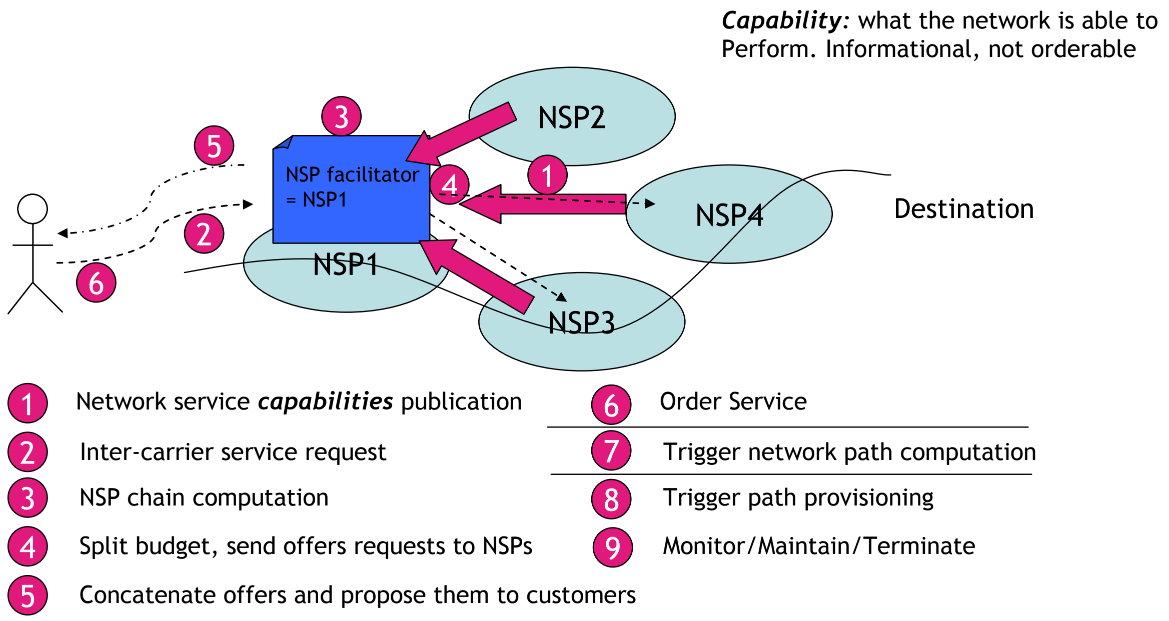

Capability: what the network is able to Perform. Informational, not orderable
NSP2
NSP facilitator = NSP1
NSP4
Destination
NSP1
NSP3
1 2 3 4 5 6
1 Network service *capabilities* publication
2 Inter-carrier service request
3 NSP chain computation
4 Split budget, send offers requests to NSPs
5 Concatenate offers and propose them to customers
6 Order Service
7 Trigger network path computation
8 Trigger path provisioning
9 Monitor/Maintain/Terminate

FIGURE 17 STEPS OF THE PER-NSP CENTRALIZED PULL DEPLOYMENT SCENARIO

This scenario is similar to the previous scenario except that the central facilitator role is played by any of the NSPs, typically, the one NSP who will receive an inter-carrier ASQ path request. The stepwise process of this scenario is given in Figure 17. Similarly to the previous scenario, a phase of service capabilities publication would allow the central entity (here an NSP) to perform the service composition. The service composition is done in two phases: First compute the NSP chain(s). Second, split the budget and request inter-carrier ASQ path offers from the NSPs. Finally, form the offer and propose it to the customer.

Note that if the offers are not enough detailed to allow for path provisioning, an extra step is needed to compute the exact network path of the inter-carrier ASQ path offer (step7).

Due to the budget splitting difficulties discovered in [ETICS-D3.5], this deployment scenario was not recommended as well by WP3 (business aspects).

4.4. BUYER AND SUPPLIER SCENARIOS

This section presents and elaborates on buyer-supplier scenarios and actor roles that can be played by different types of NSPs, as well as the main types of NSP-to-NSP AQ network services that are traded between NSPs.

The different buyer-supplier scenarios introduced below stand in close relationship to the different main charging principles:

- Sending Party Network Pays (SPNP)
    - Pure cascading
    - Hybrid
- Initiating Party Network Pays (IPNP) on top of SPNP

- Tunnel-oriented charging, supporting both SPNP and stand-alone IPNP

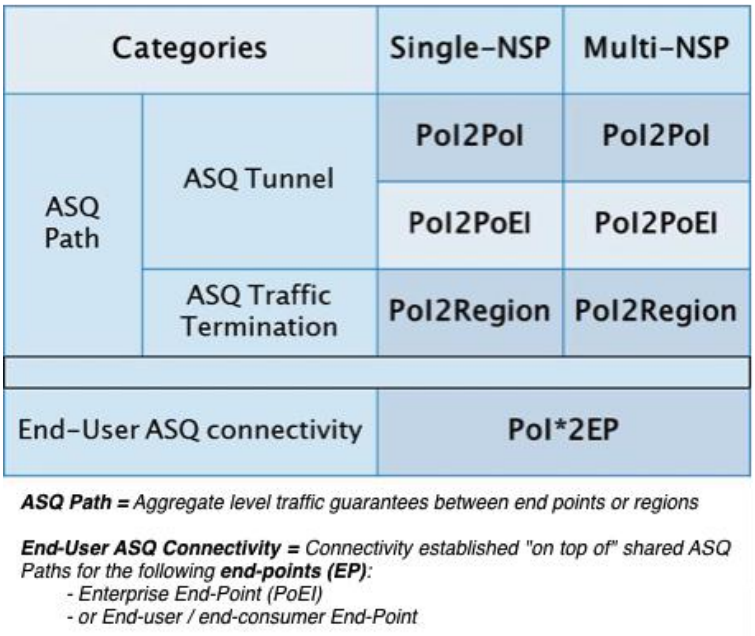
Please note that the alternative charging principle of “Receiving Party Network Pays”[7] could be realised in correspondence to SPNP, but is taken as out of scope due to recommendations from [ETICS-D3.2][ETICS-D3.3][ETICS-D3.5].

An introduction of corresponding charging mechanisms is out of scope of this deliverable. A detailed definition and investigation will be provided in [ETICS-D2.3].

4.4.1. NSP-TO-NSP NETWORK SERVICES

While section 3.4 introduced the “Inter-carrier ASQ path” that the ETICS core-system will realize this section presents the main types of ETICS network services that supplier NSPs can offer[8] to other (buyer) NSPs. These network services are called NSP-to-NSP services, and they are the basis for providing a structured presentation of what services are traded between the ETICS NSPs in the different buyer-supplier scenarios and collaboration models.

The main NSP-to-NSP network services are presented in the following figure (table).



Categories		Single-NSP	Multi-NSP
ASQ Path	ASQ Tunnel	PoI2PoI	PoI2PoI
		PoI2PoEI	PoI2PoEI
	ASQ Traffic Termination	PoI2Region	PoI2Region
End-User ASQ connectivity		PoI*2EP	

ASQ Path = *Aggregate level traffic guarantees between end points or regions*

End-User ASQ Connectivity = *Connectivity established "on top of" shared ASQ Paths for the following* ***end-points (EP)****:*
- Enterprise End-Point (PoEI)
- or End-user / end-consumer End-Point

FIGURE 18: MAIN TYPES OF NSP-TO-NSP NETWORK SERVICES

In accordance with the inter-carrier ASQ paths defined in Section 3.4 we have the following definitions.

An NSP-to-NSP ASQ tunnel is an AQ network service from a specific PoI to either another PoI (i.e. a **PoI2PoI service**) or to a point of enterprise interconnection (PoEI) (i.e. a **PoI2PoEI service**). It can be a single-NSP service where the other end-point (PoI or PoEI) belongs to the supplier NSP, or it can be a multi-NSP service where from the buyer’s NSP point of view there is at least one transit NSP (T-NSP) involved. While the

[7] In ETICS [ETICS-D2.2] and in [ETICS-D4.3] this was referred to as ASQ Traffic Origination

[8] Note that here, an offer can be a “Product Offer” with a price tag on it, or a Service Offer where the price is pre-established or know by other means.

intermediate PoIs may or may not be known to the buyer NSP the supplier NSP and any other contributing NSPs do have awareness of this particular service instance at all intermediate PoIs. An appropriate traffic identification scheme for this ASQ tunnel awareness at each ASBRs involved in the path must thus be supported. These types of network services are ASQ paths, and their traffic directionality can be any direction as well as bi-directional.

The ASQ traffic termination services are IP packet delivery services, for the AQ Internet, or for IP packet delivery based on a private address space (e.g. for IPX IP interconnect). So far, the ASQ traffic termination service as defined by ETICS is the PoI2Region service which is an infrastructure level ASQ path. (More specific ASQ traffic termination services can be foreseen but this is so far out of scope.) The sending NSP, which may not necessarily be the buyer, delivers the traffic at an agreed PoI and the destinations are any end-points as defined by the Region. Typically, a (destination) Region is defined as a set of IP pre-fixes. While the end-points of a single-NSP PoI2Region services are within the supplier NSP, some or all the end-points of the multi-NSP PoI2Region service on the other hand is located in one or more E-NSPs beyond the supplier NSP. This implies that as perceived from the buyer NSP at least one T-NSP may be involved before reaching a specific end-point.

While the above AQ network services are ASQ paths the next AQ network service is specifically related to an end-customer ASQ connectivity, whether human end-users or data hosts. In general, this type of network service is called **End-user ASQ connectivity** and is realized “on-top-of” already existing infrastructure level ASQ paths. From an NSP-to-NSP point of view this service is the so-called **PoI*2End-Point** network service. Hence, this type of service is used to enable and support end-user ASQ connectivity, where the ASQ connectivity services can be either business related (e.g. End-Point is a PoEI) or end-consumer customer or end-user related (i.e. End-Point is a consumer end-user end-point). This service is used where the end-customer specific demand for bandwidth is not feasible (not big enough) for establishing an end-customer dedicated ASQ path. The handling (management and control) and support of such “on-top-services” are enabled by the so-called service enhancement functions (cf. Section 6.1).

End-user ASQ connectivity service enforcement policies are typically only relevant at the customer facing service provider edge node and will not result in state-awareness at the NSP-to-NSP border nodes (e.g. the ASBRs). The exact PoI might be unknown by the supplier at a given point in time, i.e. being reflected in the the PoI* notation. The choice of PoI is up to the dynamic ASQ traffic steering decision of the upstream (e.g. buyer) NSP. The corresponding inter-NSP SEFA level interactions to enable and support these services can be directly E-NSP-to-E-NSP even where one or more T-NSPs are involved at the data plane. (Cf. E1’ and other reference point described further below). Hence, the distinction between single-NSP vs. multi-NSP does not have the same direct relevance as for the ASQ paths.

In general, note also the following:

- The ASQ traffic termination services can and should typically be supported by the SPNP charging principle.
- The ASQ tunnel can be based on pure IP, IP/MPLS, or any Layer 2 technology
- An NSP-to-NSP ASQ path can be considered as an infrastructure ASQ path

- The PoI2Region service is based on a "destination oriented" service model where at the PoI (point of service delivery) only the destination IP address is of concern, in addition to any policy based routing decisions.
- Strictly speaking, if the SPNP is assumed for the PoI2PoEI ASQ tunnel for ASQ IP traffic termination this type of network service can also be considered as a special type of ASQ traffic termination service.

More information on the different NSP-to-NSP network services is provided in ETICS Del.2.3.

4.4.2. ETICS Actor Role Model

This section presents the ETICS actor role model and reference points, and explains how these relates to ETICS network services and deployment scenarios (modes). The notion of collaboration mode (model) is introduced where the deployment scenario and community type concepts (see [ETICS-D3.5)] are coupled together in order to capture the two main inter-NSP collaboration (coordination) modes; that is, i) the **bilateral cascading** collaboration mode (distributed push model with open association), and ii) the **coordinated ASQ composition** mode. These high-level modes will be explained just below. This introduces a simplified view of the various options that the ETICS architecture enables as the NSPs overall can still utilize all deployment modes in a variety of combinations according to the specific business strategy and roles of each NSP. These two main collaboration modes will be considered further in relation to the proposed roadmap presented below.

The (updated) ETICS actor role model is shown in the figure below.

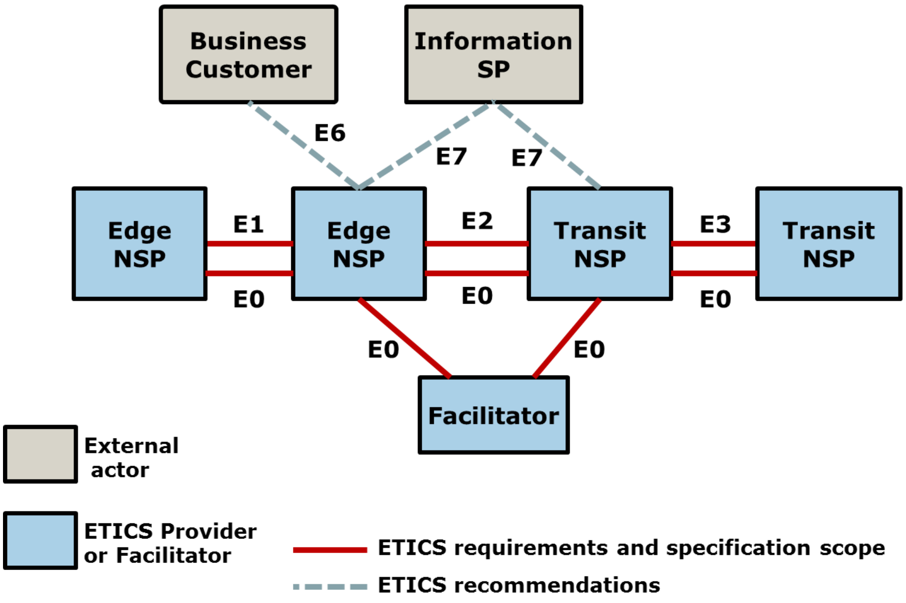


FIGURE 19 ETICS ACTOR-ROLE MODEL (UPDATED[9])

The Edge-NSP is offering AQ network services to end-customers (i.e. consumer, business, and Information Service Providers (InfSP, aka. Content and Application Providers, and OTTs). The E6 reference point is

[9] See Annex 4.4.2 for more information, including suggestions for E4 and E5 reference points.

between the Business Customer buying network services from an (edge) NSP, while the InfSP can, depending on his role, buy network services from NSPs in any roles (E7). The actor role model does not consider any reference point with the consumer end-customer in any direct manner.

The reference points E1, E2, and E3 are used in the context of the bilateral cascading collaboration mode capturing the association of the distributed push deployment scenario (see Section 4.3.3) with the open association community type (a loose cooperation community type being defined in [ETICS-D3.5]). This mode relies on a myopic coordination approach. Note also that in this section and for the roadmap (Section 4.5) we also consider as applicable the option of manual operations in addition to the automated approaches. For instance, a manual business agreement negotiation phase (similar as today) preparing for automated ordering and activation seems attractive.

In this mode the NSP-to-NSP network services these reference points refer to are the PoI2Region and PoI2PoEI services (excluding PoI2PoI). Here, these services are offered and delivered with respect to a given PoI and where the two NSPs meeting at the two sides of the PoI both play the buyer and the supplier roles in a reciprocal fashion. Hence, in this mode we can also speak of interconnection services[10]. The E1 reference point supports only single-NSP services, while the E3 supports only multi-NSP services. Considering the E2 reference point the E-NSP is selling single-NSP services and buying multi-NSP services. The bilateral cascading mode has similarities to the “distance/path vector” routing approach.

As a complement to the PoI2Region service the PoI*2EP services are facilitated by the SEFA solution for the support and management of specific inter-NSP end-user ASQ connectivity services and charging solutions. The corresponding inter-NSP SEFA level interactions to enable and support these services can be directly E-NSP-to-E-NSP (E1’) even where one or more T-NSPs are involved at the data plane. However, this can create scalability issues and a unmanageable high numbers of E1’ relationships. Hence, it is foreseen that a SEFA hub role can become relevant (in a similar fashion as an IPX provider), and correspondingly E2’ (E-NSP -to- SEFA hub) and E3’ (SEFA hub -to- SEFA hub) reference points.

The E0 reference point is introduced to enable and support the various deployment scenarios that enables a coordinated ASQ composition and path computation, whether assuming per-NSP path computation or fully centralized path computation utilizing a central facilitator. The distributed pull (HTE) deployment scenario also belongs to this high-level mode as this scenario also enables a non-myopic ASQ composition. The business and economics analysis documented in Del3.5 suggests that these deployment scenarios are coupled with the federation community type for the per-NSP path computation mode and the alliance community type for the fully centralized mode. In general, we speak of these as the **coordinated ASQ composition** collaboration mode.

What characterizes the coordinated ASQ composition and E0 reference point as compared with the E1, E2, E3 is that it enables also the PoI-to-PoI service, and services that are referring to PoIs that are not a PoI belonging to the buyer NSP. Hence, this enables to build up and maintain an inter-NSP level multi-domain topology view that can support ASQ path computation at the inter-NSP level by composing single-NSP

[10] Note that the concept of “Interconnection” can be used as a general high-level concept or more specifically when speaking of “interconnection service” as in this context. When using the term “interconnection” it will be clear from the context which of these two levels we have in mind. (The notion of “interconnection” as used with “Point of Enterprise Interconnection” will be explained in more detail below.)

services as elements of the composed ASQ path. This approach has similarities to the "link state" routing approach. The NSP that is delivering the composite ASQ path is called the **primary NSP** while the NSPs that are contributing the elements[11] of the ASQ path are called **subordinate NSPs**. In addition, the E0 is also applicable toward the facilitator where (in this particular scenario) the single-NSP services are published to the central facilitator that performs the ASQ path computation. For more information of the coordinated ASQ path computation oriented deployment scenarios see Section 4.3. (Note the exception of Section 4.3.3, which belongs with the bilateral cascading mode).

While the main focus of the core-system architecture has been on the coordinated ASQ composition mode and on composition of element services in a direct 1:N containment relationship for establishing the inter-carrier ASQ paths, the bilateral cascading mode relies on the notion of indirect composition where an *interconnection service* may be an element of multiple composite network services which again themselves can be interconnection services. The automation of the coordinated ASQ composition mode is more straight forward, while the indirect composition will typically involve a more complex business process that requires manual considerations.

Note that when considering just one specific E1 relationship, that is, just between two specific Edge NSPs the difference between these two main collaboration models disappears.

It is also important to note that these two main collaboration models may be combined when looking at ASQ communities of NSPs. One interesting case is where edge NSPs are selling single-NSP network interconnection services as well as buying interconnection services (cf. E1, and E2), while the T-NSPs offering E2 interconnection services to the edge NSPs have themselves a closer community relationship (federation, or alliance) and deploy the coordinated ASQ composition mode and the E0 reference point.

The various main charging options and principles are presented in [ETICS-D2.3] (to be released), along with explanation of how these relates to the main collaboration modes.

4.5. ROADMAP

The ETICS project has spent significant efforts on analysing the current Internet situation and trying to understand the roadblocks to introducing services that go beyond the best effort service. This has motivated efforts on considering and pointing out the first "bootstrapping" and coordinated steps NSPs should take to establish AQ connectivity services over the Internet. This section provides a brief introduction to the so-called ETICS bootstrapping proposal[12] as well as an introduction to a suggested roadmap beyond the pure bootstrapping phase, including proposals for longer term solutions and mechanisms including fully automated ETICS deployment scenarios.

Note that the roadmap proposal is just an indication, as different NSPs will have different preferences and will prioritize advancements differently according to their business profile and strategy. For instance, an

[11] The elements are single-NSP services. However, future solutions should also consider multi-NSP services that can be elements of a composite ASQ path. This can extend the ETICS architecture to support hierarchies of network resources in a more efficient manner.

[12] This is work in progress and is likely to continue beyond the ETICS project ending by for instance ETICS NSPs.

NSP focusing on consumer and business end-customers may prioritize differently from an NSP that is focused on international backbone network services.

The following diagram illustrates the roadmap and indicates the anticipated timing for the various AQ network services and also suggesting when to introduce the various collaboration modes (introduced in Section 4.4.2) and their service selection and composition scenarios.

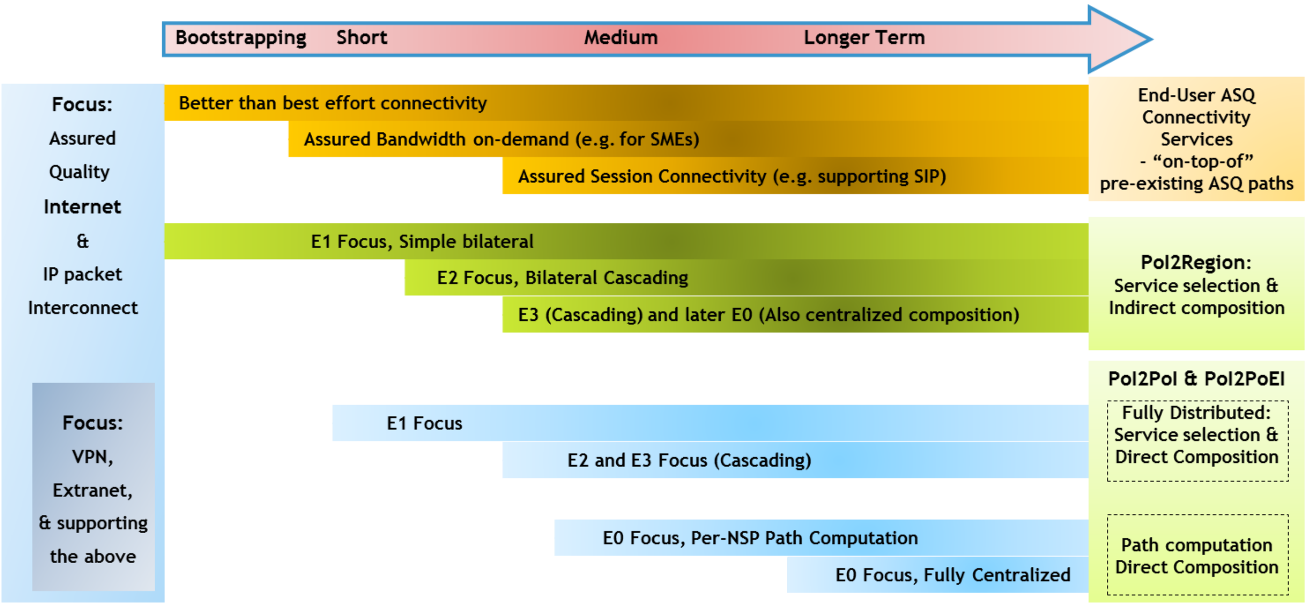

Bootstrapping
Short
Medium
Longer Term
Focus:
Assured
Quality
Internet
&
IP packet
Interconnect
Focus:
VPN,
Extranet,
& supporting
the above
Better than best effort connectivity
Assured Bandwidth on-demand (e.g. for SMEs)
Assured Session Connectivity (e.g. supporting SIP)
End-User ASQ Connectivity Services - "on-top-of" pre-existing ASQ paths
E1 Focus, Simple bilateral
E2 Focus, Bilateral Cascading
E3 (Cascading) and later E0 (Also centralized composition)
PoI2Region: Service selection & Indirect composition
E1 Focus
E2 and E3 Focus (Cascading)
E0 Focus, Per-NSP Path Computation
E0 Focus, Fully Centralized
PoI2PoI & PoI2PoEI
Fully Distributed: Service selection & Direct Composition
Path computation Direct Composition

FIGURE 20: ROADMAP OVERVIEW - NETWORK SERVICES & COLLABORATION MODES

Guided by the recommendations from the ETICS ecosystem and business analysis conducted in [ETICS-D3.3] and concluded in [ETICS-D3.5] the bootstrapping and short term solutions should be driven by a limited set of quality parameters for commercial ASQ traffic exchange based on bilateral agreements and existing route announcements between NSPs with a relatively high level of trust due to common interests. A limited community of bootstrapping NSP (e.g. in Europe) can establish simple E1 PoI2Region services to end-points inside their own domain. Having established such simple interconnection agreements on the aggregate level ETICS suggests that the SME market is addressed in the bootstrapping phase. The very first and simplest step is to offer a better-than-best-effort service with bartering on the ASQ traffic exchanged. The end-customers or their applications can help checking individual end-point ASQ capability. However, this approach will in general soon result in questioning the real added value for the customers and the NSPs should quickly prepare to offer assured bandwidth on-demand (also see Section 3.3.2 and Section 6.1) with additional quality parameters, in particular for the delay. Hence, SEFA based capabilities should be introduced to facilitate such end-user ASQ connectivity services (cf. the PoI*2EP service) for the SME market, which is expected introduced by the NSPs in the short term phase.

While starting with a limited set of "bootstrapping NSPs" the intention is to prepare the simple bilateral PoI2Region interconnection approach (basic mode E1) and in later phases also with cascading properties (i.e. multi-NSP PoI2Region services) as part of the open association community type (ETICS-D3.5). In this context it is expected that the E2 relationship will become attractive already later in the short term phase. The bilateral cascading approach with the PoI2Region services should be based on the sending party network pays (SPNP) as the foundation layer charging, recognizing that not all ASQ traffic exchange will allow easy identification of an initiator side willing to pay for both traffic directions. This supports a design for independent ASQ traffic routing in the two directions between two endpoints.

Looking beyond the bootstrapping and short term phase with simple SPNP mechanisms, more sophisticated SPNP with destination oriented pricing models are expected. This approach is considered quite powerful and will enable the customer NSP to select the more attractive PoI (traffic exit point) to reach a destination while considering the price vs. quality selection strategy in their traffic routing decisions. Furthermore, the initiating party network pays (IPNP) charging mechanism in relation to end-customer ASQ connectivity is expected already from the short to medium term. This will allow charging in relation to specific end-customer or end-user end-points and as an attractive complement to SPNP when needed. Hence, the SPNP approach is still the “base layer” charging at the aggregate level and for transit NSP PoI2Region services (E2 and E3).

On the other hand when looking at demands for business (enterprise) VPNs we expect that E1 solution for ASQ tunnels for business end-customers will also be supported in the short term. This will imply E1 interconnection agreement for offering PoI2PoEI tunnel service. In the medium term and beyond these are expected to evolve and E2 and E3 relationships in the bilateral cascading collaboration mode are expected.

While the above have primarily been oriented toward the bilateral cascading collaboration mode interesting opportunities are expected from a federation of T-NSPs for offering E2 services to E-NSPs (as well as buying E1 services). It is expected that the coordinated ASQ composition mode will enable improved routing decisions and network resource utilization efficiencies among the T-NSPs as this supports a non-myopic approach. Hence, the T-NSP will in this case need the E0 capabilities among themselves.

The federation community and the coordinated ASQ composition mode as suggested above can be extended to also include E-NSP (cf. E0) in particular those NSPs that are targeting business VPN customers. Again, this operational mode is expected to enable improved routing decisions and network resource utilization efficiency, in addition to automation benefits. For instance the PoEI2multi-PoIE (cf. E6) is expected to be more efficiently supported.

The fully centralized path computation collaboration mode is in particular difficult to place in the roadmap. Del3.5 anticipates that this collaboration mode will require an (closed) alliance community type. This allows conducting joint cross-NSP business by including efficient revenue sharing and routing decisions as well as penalty regulations for non-compliance. This mode can be an option for a (limited) set of highly trusted NSPs with a common goal and the alliance can evolve into a Virtual Network Operator. If this collaboration mode becomes successful it is foreseen that several alliances will be competing as the market matures. However, the more advanced forms of this mode are anticipated for the longer term.

The voice market segment is of high importance for the Telcos. An interesting topic is how the interconnection solutions and regimes for voice and future rich multimedia communication services will evolve. The ETICS solutions with ASQ PoI2Region services at the general ASQ traffic interconnection level using SPNP, and the generic SEFA solution elements that allows for SIP and IMS solutions to trigger and control end-user session services ASQ connectivity is anticipated as a future-oriented approach. In this way the IPNP approach is facilitated by SIP/IMS (SEFA layer) while SPNP is the foundation layer for charging at the aggregate level. Early trials can be anticipated later in the short term phase evolving into more mature solutions for the medium and longer term.

It is important that the ETICS specifications, first for the so-called basic mode (those associated with simple bilateral E1 relationships and services) are evolved into standardized solutions for NSP-to-NSP B2B interaction. The very first step is to have a “bootstrapping community” agreeing on the PoI2Region service type and the associated SLA and quality parameters. As bootstrapping trials provide more experience the automation proposals by ETICS, at the SEFA layer as well as at the ASQ path layer should be further progressed into standardized solutions. These efforts will be important for the evolution of the industry and for helping NSPs prepare and transition their businesses for assured quality and differentiated end-to-end network services.

More information about charging principles and mechanisms and more details around the roadmap are provided in [ETICS-D2.3] (to be released). This includes a closer consideration about inter-NSP automation (publishing, discovery, ordering, activation, assurance), SEFA capabilities, monitoring and dynamic pricing. A closer look at the mechanisms and the solution options that can come together with the federation and the alliance community types are also considered.

5. ETICS CORE-SYSTEM ARCHITECTURE BUILDING BLOCKS

The present section deeper investigates the core-system architecture as a sub-element of the overall ETICS architecture. For this purpose, the essential building blocks are reviewed by utilising information being available from [ETICS-D4.3] and incorporating feedback from [ETICS-D5.6]. Originating from the access to ETICS services, a top-down approach is applied for describing the interwork of fundamental building blocks. In consequence, a picture from the representation and exchange of ETICS offers/capabilities to the computation of offers and network paths is drawn. This subsequently complemented by the actual provisioning of inter-carrier ASQ paths, as well as by monitoring solutions and their interfaces to the rest of the architecture. We start this section by showing the big core-system architecture picture and where the different building blocks fit in it.

5.1. ARCHITECTURE & DEPLOYMENT

As seen before in section 4 and in previous iteration of deliverables (ETICS-D4.2 and ETICS-D4.3), the SLA management within the ETICS system corresponds to a dedicated life-cycle that follows a given workflow.

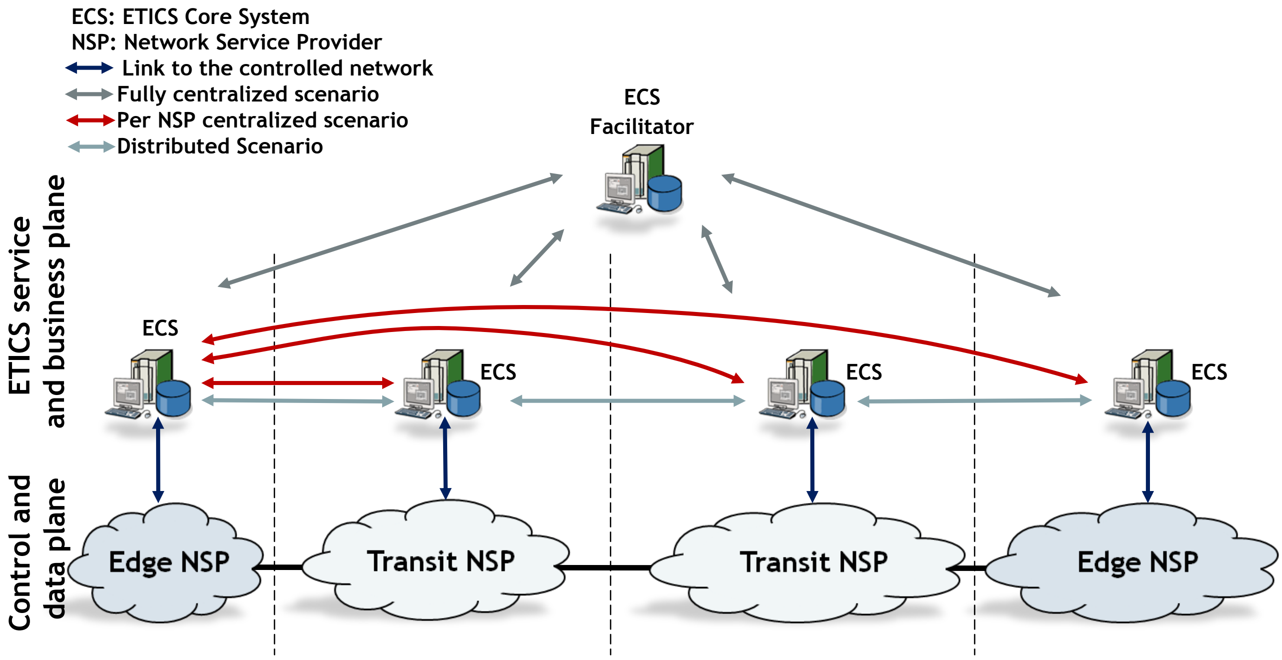

ECS: ETICS Core System
NSP: Network Service Provider
Link to the controlled network
Fully centralized scenario
Per NSP centralized scenario
Distributed Scenario
ECS
Facilitator
ETICS service and business plane
Control and data plane
ECS
ECS
ECS
ECS
Edge NSP
Transit NSP
Transit NSP
Edge NSP

FIGURE 21 ETICS ARCHITECTURE AND SCENARIOS

The different steps are divided in 3 specific groups that we presented previously in Figure 11:

- Learning phase that corresponds to the creation and publication of SLA offers or Network Capabilities,
- Composition phase that groups the Negotiation, Validation and Enforcement of the ASQ connectivity raised when the ETICS core system receives a user request,

- And the Monitoring and Termination that play once the ASQ connectivity is in place.

We have seen also that different deployment scenarios (that reflect different collaboration modes between the actors) can be envisaged for the ETICS solution deployment. Figure 21 below summarizes the different scenarios exposed previously and shows where the different functionalities need to be deployed in each scenario. The ETICS Core System "boxes" correspond to the ETICS software that will perform the ETICS service and business plane functionalities. We next "zoom" on a typical ETICS core system "box".

5.2. ETICS CORE-SYSTEM OVERVIEW

The architecture depicted in Figure 22 below illustrates the ETICS Core system. It consists of six main building blocks:

- **SLA Manager:** groups all the functions involved in the management of the SLA (Negotiation, Validation, Termination, and Service Assurance steps of the SLA life cycle). This block has a relation with almost all the others: it is the central part of the core system as it orchestrates and triggers most of the steps described in the ETICS workflow Figure 11. It is also the first external interface (E1/E2 - S) to the neighbouring ETICS Core System.

- **SLA Offers:** groups all the functions relative to the creation, certification and publication of SLA offers (only PUSH model). It uses the Policy Rules to determine which SLA offers will be communicated to the SLA Manager for the Service Composition step. It also communicates with neighbours ETICS Core System to **exchange** SLA offers. This is the second external interface (E1/E2 – B) of the system. This block implements **step 1 (offers publication/exchange)** of the workflow of Figure 11.

- **Network Capabilities** groups all functions devoted to the Inter-carrier Routing behaviour. In particular it defines the way Network Capabilities are built against Network Topology as well as how this information is exchanged with neighbouring NSP(s) through the third external interface (E1/E2 – B)[13]. This block is in charge of the creation and publication actions in the PULL model only, which corresponds to **step 1 (capabilities publication/exchange)** of the workflow of Figure 11. Network capabilities mainly correspond to the H-TE solution being detailed later on.

- **Business & Policy** groups all functions involved in the business and policy rules that govern how the SLA offers are built, but also how the service composition shall be done by the SLA Management. "Billing and Accounting" as well as the standard Authentication Authorization & Accounting complete the block.

- **PCE** This entity corresponds to the Path Computation Element function standardized by the IETF. The SLA Manager triggers the PCE to refine the **network path computation (step 7 in** Figure 11) when the offers are not enough detailed (Push model). It also trigs the PCE without AS path as parameter to request a dedicated ASQ (Pull model). In this latter case, the PCE uses Traffic

[13] Note that the E1/E2 – B interface is currently different between the Pull and the Push model due to implementation choice but could be the same.

Engineering information learnt from the underlying network it controls and from the Network Capabilities block.

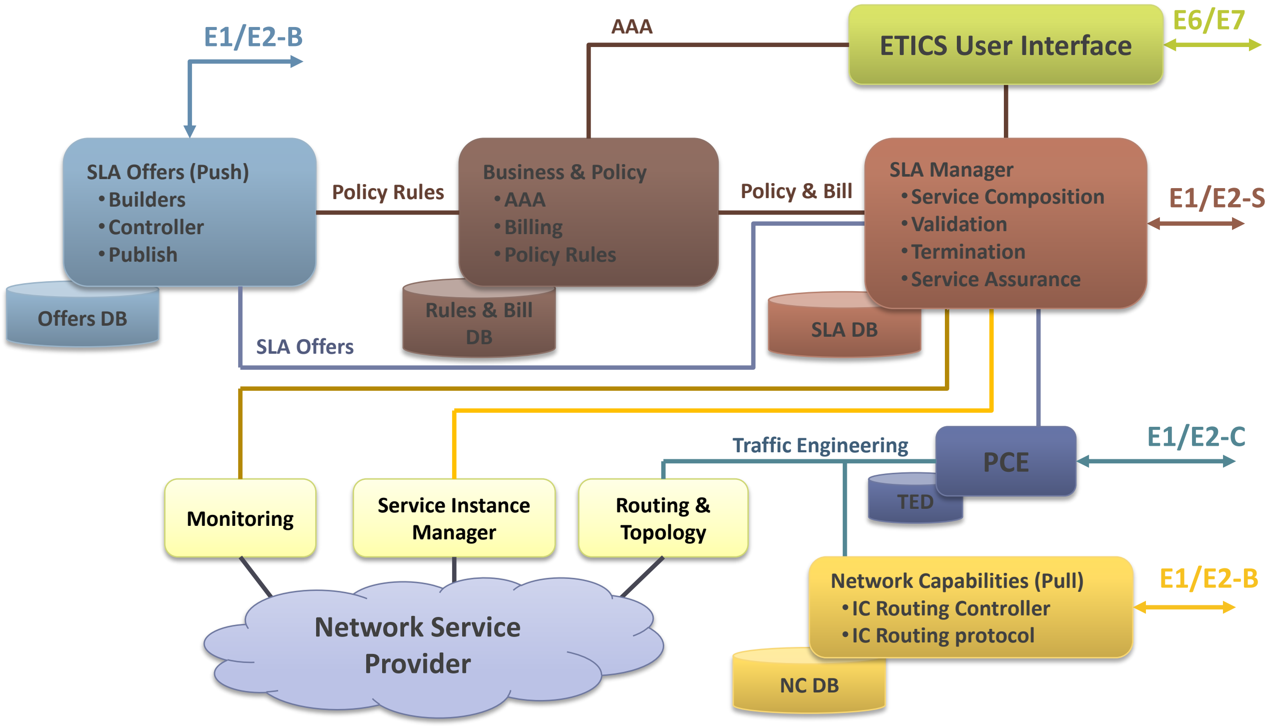


FIGURE 22 ETICS CORE SYSTEM BUILDING-BLOCKS

Peripheral blocks complete the ETICS Core System:

- **A first group** provides interface to the network. These are the Monitoring and Service Instance Manager (including the network configuration) that interface with the SLA Manager.
- **ETICS UI** or **Service access point**: groups all features related to the User Interface that the system exposes to its customer. It models the interface *E6* and *E7* disregarding the protocol, interface, or API exposed by this block. The purpose is to show that SLA composition could be triggered by both, another SLA Manager instance (from a neighbouring ETICS system through *E1*, *E2*, or *E3* interfaces) or by an ETICS users through the *E6* or *E7* interfaces.

5.3. LEARNING PHASE (CAPABILITIES/OFFERS EXCHANGE)

The first step in the ETICS overall workflow corresponds to the learning phase. Indeed, the ETICS Core System must not only controls its underlying network but mainly acquire information from its neighbouring NSPs. Two models have been designed by the ETICS consortium: the PUSH model based on SLA offers and the PULL model based on Network Capabilities (a.k.a. Traffic Engineering information) exchange. The following sub-section describes in details the learning phase for these two models capitalising on deliverable D5.6 [ETICS-D5.6]: it provides a summary of the latter while at the same time presenting some incremental updates and clarifications.

5.3.1. Offers representation and exchange (Push Model)

In the push model, the different NSPs exchange offers between each other. Each offer is described by a set of parameters that can be split into:

- Technical parameters: describe the guarantees of the ASQ path, as well as how it can be provisioned.

- Business parameters: any business related information like the price or the validity of the offer.

- Administrative parameters: like the boundaries of the ASQ paths

As we saw in Section 3.4, ETICS ASQ paths can be of two main types: PoI-to-PoI and PoI-to-region.

5.3.1.1. PoI-to-PoI ASQ path offers description

The PoI-to-PoI ASQ path offers contain the following information that is summarized in this table:

Ingress PoI	Egress PoI	Traffic id	Bandwidth	Delay	Jitter	Loss	Availability
Establishment delay	Price	Duration	Validity	NSP	ASN		

We subsequently breifly describe these parameters. The technical guarantees are expressed with a confidence interval.

- Bandwidth in Mbits/s: It defines the guaranteed rate of the data traffic handled by the ASQ service.

- Delay in µs: Represents the maximum time that data takes to be transmitted between the ingress PoI and the egress PoI of the ASQ path.

- “Jitter” in µs (optional): Represents the maximum delay variation in ms between two successive packets.

- Loss: Expressed as a percentage of lost packets.

- Availability: expressed as a percentage of availability time of the connectivity service. More complex representation of the availability (for example taking into account minimum time between two failures, maximum time to repair…) could be proposed but the composition of a richer parameter between different operators is considered too complex for a rapid adoption of the ETICS solution by different actors (operators, customers…).

- Establishment delay in ms: represents the maximum delay to establish the connectivity service from the reception by the provider of an order of the service sent by the customer.

Even though the proposed ETICS monitoring solutions aim at being as cost-efficient as possible, in certain situations it may be suitable to avoid monitoring tools in the core where – due to high bandwidths – monitoring is by nature somewhat costly. In such situations, we take the following assumption that traffic policing mechanisms are deployed at the entries of the ASQ paths (at the “edge”) to ensure that the transmitted data (aggregated traffic flows) is compliant with the SLA. By default, packets violating the contract will be dropped to not introduce jitter, or impact other ASQ services managed by the community of NSPs. Using this strategy it is not mandatory for the ETICS community to deploy such a mechanism at each PoI.

- Boundaries of the ASQ path: PoI naming:

First, the ASQ path takes place between an entry point and an exit point. These points need to be defined in the ASQ path offer. As such the double **(ingress PoI, egress PoI)** needs to be part of the offer. Now, as far as the **PoI naming** is concerned, there are mainly three directions.

The first is to base the naming on the IP addresses of AS border routers (ASBRs) and how they connect together. As such, a PoI would be an ASBR of one NSP with the following convention: the ingress PoI is the ingress ASBR of one NSP and the egress PoI is the ingress ASBR of the next NSP. With this convention, each NSP is responsible for its own network plus the interconnection link to the next network.

The second is to create a new namespace to name PoIs. A PoI can be for instance the name of the Point of presence (PoP) where NSPs interconnect (e.g. PariX IXP). This approach has the drawback of assuming that all the NSPs that are present in the PoP are fully meshed, co-located and therefore are willing to do "ETICS business" with each other. This approach has, on the other side, the advantage of being more scalable for the offers computation (reduces the number of PoIs for the offer computation algorithm).

As a first step, we opt for a restricted version of the second approach relying on a new namespace to name ETICS PoIs. In our solution, we limit the creation of PoIs between NSPs only if they are effectively fully meshed: each NSP has at least one interconnection link (between ASBRs) with each of all other NSPs in the PoI. In this case, interconnection links' properties between NSPs of a given PoI have to be homogeneous (similar length, traffic identification capabilities…).

For this reason, when an NSP is not physically located by its ASBRs in a POP but only interconnected to other NSPs by mean of a longer links coming from another POP, then it is mandatory to create a new set of PoIs, each of them dedicated to a bilateral interconnection between this remote NSP and each of locally present NSPs that is interconnected to it.

When NSPs agree on the creation of a PoI, they have to respect the following rules for the Traffic Delivery Points (TDPs) of all their network connectivity services terminating or starting at this PoI:

- Network services of an NSP terminating at a PoI terminate at the ingress ASBR of the next neighbour NSP in this PoI.
- Consequently, network services of an NSP starting at a given PoI starts at its ingress ASBR.

This choice has to be respected by all the NSPs at all PoIs. A PoI naming convention, preserving name collision, has to be used. For example, it can be the aggregation of the name of the PoP and an integer identifier value incremented for each of the new PoIs created at the PoP.

This approach has several advantages. First, the higher is the number of NSPs joining a PoI, the "smaller" is the (network services) graph of an ETICS community (and consequently, the smaller is the number of network service offer publications). If a PoI matches an interconnection PoP, then the size of the graph can be drastically reduced (cf. analysis of scalability studies in Section 7).

Second, in a push model, an NSP can easily add or upgrade an ASBR at a given PoI without impacting the network service plane and the network service offers of NSPs concerned by this PoI. Identically, when a NSP would like to join an already existing PoI, the other NSPs do not have to modify their already published network service offers.

- Traffic identification: This part was first introduced in [ETICS-D5.6] in order to leave a placeholder to express on which traffic fields the ASQ path is able to identify traffic in order to give it its proper treatment.

The traffic that goes within the ASQ path from a PoI to a PoI needs to be identified at the ingress ASQ path. There are several identifiers (belonging to different technologies) that allow a router to identify the traffic (IP addresses, DSCP field, MPLS header, VLAN id etc). In a similar fashion to what has been described in Table 1, each NSP express its capabilities in terms of fields on which it is able to perform traffic identification. The goal is to answer the following question: on which traffic identifiers (which fields in the header of data packets) is the ASQ path able to handle traffic? Is it only MPLS labels, or only IP addresses? or more complicated filtering capabilities?

- Price: The first pricing model alternative for PoI-to-PoI ASQ path is flat rate pricing (not reflecting the real consumption of the service, i.e. not reflecting the real volume of data traffic), depending on the duration of the service. Flat rate pricing is most straightforward, as it does not imposing volume measurements at PoIs, while it is also not capable of compensating issues resulting from traffic asymmetries or traffic peaks.

A variation of this model could in addition take into account the volume of traffic that flows between two PoIs (e.g. using 95th percentile measure).

The first pricing model could be therefore: $P = A \times T$

Where P is the final price that will be charged to the customer, A is a price coefficient depending of the required bandwidth for the connectivity service, T is the total duration of the service.

The second pricing model could be: $P = A' \times T + B \times V \times T$

Where: P is the final price that will be charged to the customer. A' is a price coefficient that is relative to the provided bandwidth. T is the total duration of the service. B is a price coefficient that is relative to the volume of traffic that flows inside the ASQ path. V is a fixed price coefficient that depends on the provided bandwidth.

- Duration: this parameter represents the couple of minimum and maximum duration of the connectivity service in seconds, once the service is purchased. Note that the maximum component of the duration parameter can be optional. Such a maximum value could be the possibility to the offering NSP to renegotiate the PoI-to-PoI offer (e.g. to revise the price etc.)

- Validity of the offer: when an offer is released, and while it is not order by one of the ETICS actors, it contains an expiration date. Any actor who is willing to order the offer must do it before the expiration date. Any provider who is not able to provide an offer that did not expire exposes itself to penalties. According to the business policy rules of alliances (a highly integrated form of an ETICS community, cf. [ETICS-D3.5]), penalties may be shared between the different actors. In push model, emitted offers have to provide minimum validity duration in order to avoid potential scalability issues that might arise due to too many frequent publications.

- NSP name: Who owns the ASQ path offer.

- AS number: The Autonomous System number of the NSP that is issuing the ASQ path.

5.3.1.2. PoI-to-Region ASQ path offers

We identified three types of PoI-to-region offers that NSPs can present in the ETICS service access point (**Sec. 5.3.2)**: best effort, permanent guarantees and on-demand guarantees. A first detailed description of these offers was first provided in [ETICS-D5.6]. This section builds on that description and partially updates it. For clarity, we divide the permanent guarantees into permanent soft (better than best effort) guarantees and permanent hard guarantees per host, which gives us four types of PoI-to-region offers that will be described in this section. Similarly to the PoI-to-PoI ASQ paths, the parameters describing each of the three types of PoI-to-region offers will be divided into (1) technical, (2) business and (3) administrative parameters.

It is important to note that such offers are intended to be available for customers at the ETICS service access point described in Sec. 5.3.2.

5.3.1.2.1. Best effort PoI-to-Region description

The PoI-to-region Best effort offer can be described by one of the following information:

Ingress PoI	IP prefix	N_user	NSP	ASN	Pricing	Duration	Validity	Additional info

or

Ingress PoI	IP list	N_user	NSP	ASN	Pricing	Duration	Validity	Additional info

Or in the hybrid case:

Ingress PoI	IP prefix(es)	IP list(s)	N_user	NSP	ASN	Pricing	Duration	Validity	Additional info

Note that this offer does not contain guarantees. We next detail the different parameters that describe this offer.

- Ingress PoI: ingress Point of Interconnect of the region

- Region naming: This defines the set of hosts that are "reachable" through the PoI. This can be an:

<IP prefix>: A contiguous IP addresses range: A range of IP addresses that cover the hosts of the region. It can be encoded as an IP prefix expressed in the CIDR notation.

<List IP> An explicit list of IP host addresses: Unlike IP routing where the CIDR was introduced to enhance scalability by allowing finer-grained aggregation, in our case, we don't need to necessarily use the CIDR notation. In fact, the region naming is not necessarily intended to be used for routing. In practice, the list of IP hosts can be made available at the ETICS portal. ETICS customers can simply download it.

<Hybrid list>: A list of both IP prefixes and lists of IP addresses: This can be seen as a generalisation of the above two region addressing schemes. A region can be therefore defined as a list of zero or multiple IP prefixes and zero or multiple IP lists.

- Traffic identification: since the region is offered on a best effort basis, the traffic identification is not an issue. Once delivered to the PoI connecting the region, the traffic can flow using the classical destination-based forwarding in the region.

- Price: The pricing of such a best effort service depends on the business model adopted for such service. In a model like the best effort Internet of today, the current PoI-to-region service would correspond to a peering relationship between the ETICS customer that will use the PoI-to-region service and the (edge) NSP that owns the region. This “peering” can be either (1) for free (meaning that both parties see a similar benefit from this interconnection) or (2) paid such that one of the actors compensates the other (a rather seldom case).In the case the pricing decision (free peering or paid peering) depends on the customer that will buy the region ASQ path, the pricing should be set on an on-demand fashion when requesting the service.

- N_user: number of potential active users: This number is a good indicator for ETICS customers to estimate the market they can target (simultaneously active but also inactive users). Since the IP addresses space defined earlier does not allow inferring the number of customers, N_user represents the number of active destination hosts within the region (i.e. the number of Internet Access Service contracts).

- Duration

- Validity of the offer

- Additional information: The above mentioned information may be augmented with additional profiling information on the users in the region.

- NSP: NSP name

- ASN: AS number

- Administrative boundaries: from the ingress PoI to each host in the region (or to the NSP “box” that connects the user network).

5.3.1.2.2. Permanent hard guarantees PoI-to-Region ASQ paths

A PoI-to-region permanent guarantee ASQ path defines for each sub-region the following information:

Ingress PoI	Sub-region identifier	N_user	NSP	ASN	Traffic id	Price	duration	validity	Guarantees

- Guarantees: the guarantees for such a service are “hard” for each of the hosts in the region. The type of provided guarantees will be typically the same as the ones offered by an access ISP for its managed services (e.g. guarantees for its IPTV sub-system in a triple-play access service), but now offered to third party services.

> **- Per-host guarantees**: the same parameters that were used for the PoI-to-PoI service can be used: **Bandwidth in Mbits/s:** Dedicated bandwidth allocated to the host (home network) **Delay in ms:** Maximum delay from the PoI till the host. As well as **“Jitter” in ms, Loss**, **Availability** and **Establishment delay in ms.**

Since such a PoI-to-region offer can be stitched with other PoI-to-PoI offers from other NSPs, **it is up to the ETICS customer to estimate** (depending on the demand for its services) **the aggregated “capacity” that it needs to purchase for the PoI-to-PoI part**, before reach the final PoI that connects the region. This aggregated capacity can be communicated to the edge (access) NSP.

- **Ingress PoI:** ingress Point of Interconnect where the region starts (same naming as PoI-to-PoI ASQ path above)

- **Region naming**: This defines the set of hosts that can be served through the PoI. Since permanent guarantees are defined for the region, we divide the region into subsets (**sub-regions**) *that have an equal treatment in terms of guarantees*. A region is therefore defined as the union of these sub-regions. **Sub-regions** can be addressed/represented by:

<IP prefix>: A contiguous IP range: As described in Sec.5.3.1.2.1.

<List IP> An explicit list of IP host addresses: As described in Sec.5.3.1.2.1..

<Hybrid list>: A list of both IP prefixes and lists of IP addresses: As described in Sec.5.3.1.2.1..

- **Traffic identification**: When providing the region service, the access NSP needs to precise, in a similar way to the one described in Table 1, its capabilities in terms of traffic identification. This shall specify up to what granularity the access ISP is able to identify the traffic in order to give it the right treatment till the host. This can be done by "ticking" the field of Table 1 on which the NSP is able to identify traffic.

- **Pricing:** Since the access NSP offers all its users while only a small subset of them simultaneously use the network at a given period of time, the best pricing model for such service is usage-based. This can be on the basis of the number of connections or on the volume.

- **Number of potential active users N_user**

- **Duration**

- **Validity of the offer**

- **NSP name**

- **ASN: AS number**

- **Administrative boundaries:** From the ingress PoI to each host in the region.

5.3.1.2.3. Permanent soft guarantees PoI-to-Region ASQ paths(better than best effort)

A PoI-to-region permanent soft guarantee ASQ path defines for each sub-region the following information:

Ingress PoI	Sub-region identifier	N_user	NSP	ASN	Traffic id	Price	duration	validity	Guarantees

- **Guarantees:** The guarantees for this type of offers are soft. These guarantees can either relate to the entire region or only to a subset of a region. An example of such soft guarantees is the use of Diffserv to serve the region. In such case, the ASQ traffic goes with the best effort traffic except that it gets a higher priority. These soft guarantees can be augmented with approximate information about delay, loss, access rates of hosts inside the region.

- **Ingress PoI:** ingress Point of Interconnect where the region starts (same naming as PoI-to-PoI ASQ path above)

- Region naming: This defines the set of hosts that can be served through the PoI. Since permanent guarantees are defined for the region, we divide the region into subsets (**sub-regions**) *that have an equal treatment in terms of guarantees*. A region is therefore defined as the union of these sub-regions.

- Sub-regions can be addressed/represented by:

<IP prefix>: A contiguous IP range: As described in Sec.5.3.1.2.1..

<List IP> An explicit list of IP host addresses: As described in Sec.5.3.1.2.1..

<Hybrid list>: A list of both IP prefixes and lists of IP addresses: As described in Sec.5.3.1.2.1..

- Traffic identification: When providing the region service, the access NSP needs to precise, in a similar way to the one described in Table 1, its capabilities in terms of traffic identification. This shall specify up to what granularity the access ISP is able to identify the traffic in order to give it the right treatment till the host.

- Number of potential active users

- Pricing: The pricing for such a service can be either flat rate depending on the aggregate capacity that the ETICS customer has purchased, or volume based depending on the volume of traffic that flows within the region.

- Duration

- **Validity of the offer**

- NSP name

- ASN: AS number

- Administrative boundaries: From the ingress PoI to each host in the region.

5.3.1.2.4. On-demand guarantees PoI-to-Region description

The following parameters identify

Ingress PoI	Sub-region identifier	N_user	NSP	ASN	Traffic id	Price	Duration	validity	Guarantees

- Guaranties: In such region offers, the NSP offers the guarantees on a per-host basis. The type of provided guarantees will be typically the similar to the ones applying to the per-host permanent guarantees, however the bandwidth is not fixed per host, but the NSP is instead responsible for sharing the allocated capacity between concurrent premium connections..

As such the following parameters are needed to describe the guarantees of the on-demand PoI-to-region ASQ paths:

- The offer has to provide the information of the **maximum number of parallel on demand individual connectivity sessions** that can be provided at the same time through the offer,
- The **maximum and the minimum allowed bandwidths** (peak) for each on demand individual connectivity session,

- The **authorized bandwidth step** between the precedent minimum and maximum bandwidths,
- The **maximum delay in ms to establish** an on demand individual connectivity session from the reception of its demand,
- The **maximum delay in ms to release** the individual connectivity session from the reception of this operation request sent by the customer. This parameter is important when the number of current established sessions in parallel is near to its maximum authorized value.

The other parameters that were described earlier in this section can be also added:

- **"Jitter" in ms**
- **Loss**
- **Availability**

- **Ingress PoI:** ingress Point of Interconnect where the region starts (same naming as PoI-to-PoI ASQ path above)

- Region naming: The region should define the set of hosts that can be served through the PoI. Since on-demand guarantees are mainly defined on a per-host basis, and since different host in an edge NSP can have different guarantees, we divide an NSP region into subsets (**sub-regions**) *that have an equal treatment in terms of guarantees*. An entire region is therefore defined as the union of these sub-regions. Sub-regions can be addressed/represented by:

<IP prefix>: A contiguous IP range: As described in Sec.5.3.1.2.1..

<List IP> An explicit list of IP host addresses: As described in Sec.5.3.1.2.1..

<Hybrid list>: A list of both IP prefixes and lists of IP addresses: As described in Sec.5.3.1.2.1..

- Traffic identification: When providing the region service, the access NSP needs to precise, in a similar way to the one described in Table 1, its capabilities in terms of traffic identification. This shall specify up to what granularity the access ISP is able to identify the traffic in order to give it the right treatment till the host.

- Business parameters: Compared to the per-host permanent guarantees, on-demand services would require at each connection costly additional network configuration and checking operated by the NSP. To provide incentives to NSPs to provide such kind of enhanced services it is reasonable to reflect these operations in the charging. As such, each connection can be charged a fixed amount of price.

Then, the duration of these individual connectivity sessions may be also considered in the pricing model. The different requested bandwidths for each of the individual connectivity sessions may be also taken into account. In this case, either flat rate per bandwidth/capacity pricing, either volume based pricing or a mixture of both could be applied on these individual connectivity sessions.

To provide a comprehensive illustration, we describe in the following an example of pricing model for an on demand PoI-to-region service:

$$P = N \times C + \sum_{i=1}^{N} \left(A_i \times T_i \right)$$

Where P is the final price that will be charged to the customer, N is the total number of on demand established individual connectivity sessions during all the duration T, C is the operational price for each establishment of an individual connectivity session, A_i is the price coefficient depending of the ith individual connectivity session required bandwidth, T_i is the duration of the ith individual connectivity session. Finally, in addition to the pricing, a number of other parameters can be described:

- Number of potential active users: Number of users that can profit from the on demand individual connectivity session part of the offer.

- Duration (global): maximum duration of the overall contract to the region.

- Duration (per-connection): Such parameter can also be required to specify the minimum and the maximum time allowed for each individual connectivity session.

- Validity of the offer

- An **individual connectivity session duration**

- Administrative boundaries: The administrative parameters are the same as for the previous types of region services.

- NSP name

- ASN: AS number

5.3.1.2.5. Composition considerations for the PoI-to-Region ASQ path

As we saw, PoI-to-region offers do not necessarily contain guarantees in terms of aggregated bandwidth.

With the convention that was used for the PoI naming and setting, it is up to the upstream NSP (the one before the edge PoI-to-region NSP) to specify the aggregated capacity. The aggregated capacity is either: (1) The bandwidth of the PoI-to-PoI ASQ path that will be stitched to the PoI-to-region ASQ path. Or (2) In case of a direct purchase; the dimensioning of the interface from which the traffic will get inside the region. In any case, after the composition, this parameter needs to be passed to the edge NSP that is offering the region.

5.3.1.3. Offers Exchange and storage

5.3.1.4. Offers Exchange and storage Details about the offers exchange and storage will be provided in the next final detailed specification deliverable D5.8. This section, which is based on the D5.6 deliverable [ETICS-D5.6] to complete the high level ETICS architecture presentation, only gives a high level overview on how offers are exchanged between the different NSPs as well as between NSPs and the centralized server. This includes the technology used (web services) as well as the main types of messages/actions.

A single design has been created to support all ETICS centralized push model scenario, i.e. per-NSP centralized or fully centralized (see for instance section 4.4 and 4.5 of ETICS deliverable D5.2 [ETICS-D5.2]). The provided sequence diagrams precise at a higher level the different states and main information exchanged between entities and their principal block functions.

Our service plane relies on a Service Oriented Architecture (SOA) based on *Rest* (mainly using *post*, and *get* of HTML protocol in our case). Two main software applications are used:

- The “NSP service server” which *creates* offers (as specified in Sec.5.3.1), *sends* them for publication to one (fully centralised) or several (per-NSP centralised) “Facilitator servers”, *requests* connectivity service offer propositions (under QoS constraints) to one of the potential “Facilitator servers”, and finally *orders* offers to one of the potential “Facilitator servers”.
- The “Facilitator server” which *receives* offers from one or different “NSP service servers”, *sends* end-to-end offer based on a composition process (using the algorithms described in Sec.5.4.1) to “NSP service server” who requested an offer, *requests* connectivity service offer propositions (under QoS constraints) to one of the potential other “Facilitator servers”.

The different deployment scenarios of these two software applications allow the representation of the different push scenarios.

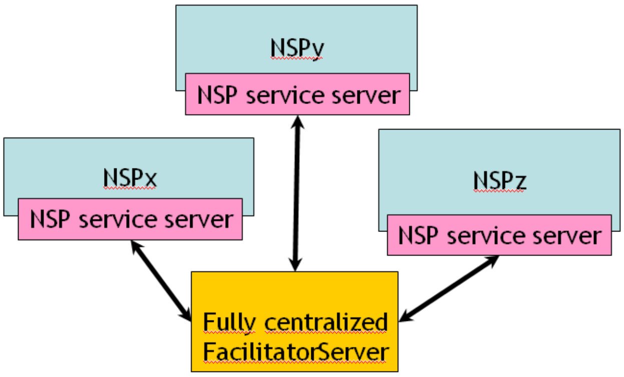


FIGURE 23: FULLY CENTRALIZED FACILITATOR SCENARIO

In the above example of the “fully centralized scenario”, each NSP of the ETICS community has its own instance of the software application “NSP service server” that interacts with a unique instance of the software application “Facilitator server” that runs for the entire ETICS community. Note that this “Facilitator server” can be either hosted on an independent device managed by the community or hosted and possibly managed by one of the NSPs of the community.

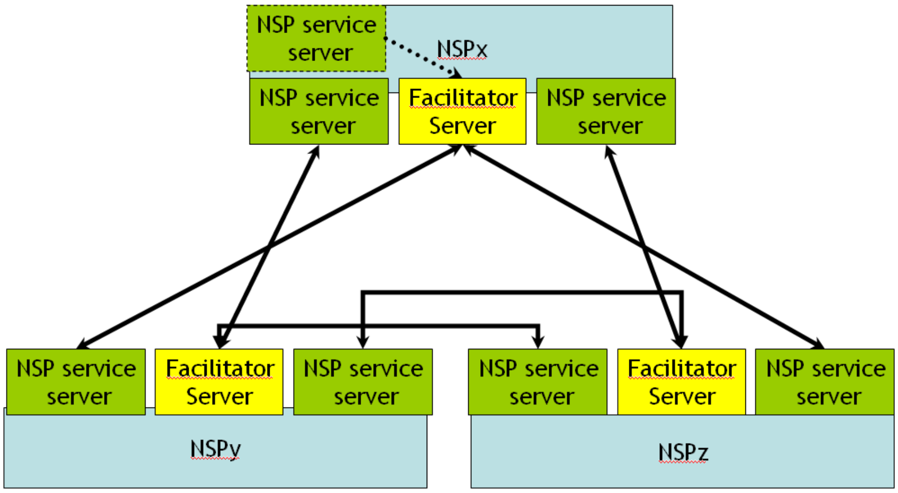


FIGURE 24: PER-NSP FACILITATOR SCENARIO

In the above “Per-NSP facilitator scenario”, a different instance of the “facilitator server” is hosted and managed by each NSP of the ETICS community, which is willing to play this role. In this case, each NSP can have different instances of the software application “NSP service server” that allow the usage of different policies for the publication process according to the targeted NSPs. In addition each NSP should have (as illustrated in dashed lines for the NSPx) another instance of the “NSP service server” to publish its own local offers to its facilitator server, which can then composed end-to-end taking into account its own possibilities.

5.3.2. CAPABILITIES DESCRIPTION AND EXCHANGE (PULL MODEL)

Compared to the PUSH MODEL, the PULL MODEL is based on the exchange of network capabilities between the different NSP. Then, from these network capabilities, ASQ are computed on demand based on what the different networks have previously announced and the requested parameters.

As already described in previous deliverables (D4.3, D5.2 and D5.6), the PULL MODEL uses a standard Interior Gateway Protocol (IGP) with Traffic Engineering capabilities to convey the network capabilities. Of course, mostly for scalability reasons but also for confidentiality and security issues, the network capabilities are not exchanged at the standard level and the IGP-TE is not running between ASBR. Indeed, the IGP-TE runs at a higher hierarchy level i.e. on an abstract view of the different networks. Named Hierarchical Traffic Engineering (H-TE), the IGP-TE used by the PULL MODEL runs between dedicated servers in an overlay mode (See figure below). These specific machines are named Autonomous System Virtual Router (ASVR). Finally, you could consider the interconnection of the different networks (at the AS level) that formed the ETICS community, as a simple network formed by the interconnection of ASVR that represents the ETICS Network Service and Business Plane. Thus, we could simply apply the same TE rules and Tool Box on the set of ASVRs as the ones used to manage a standard network. The challenge is that instead of a single network administrator, we must handle one network administrator per ASVR as well as their coordination to ensure a proper configuration of the ASVR and thus a coherent configuration.

The next sub-sections describe the different parameters exchanged by ASVRs and which kind of offers is proposed when ASVRs are solicited.

5.3.2.1. Network Capabilities Parameters

First of all, by using a standard IGP-TE protocol, we could convey between ASVRs the standard TE parameters:

- TE metric: A TE administrative weight,
- Bandwidth in bytes/s: Multiple bandwidths could be announced. The Maximum, Maximum Reservable, the Unreserved (one for each of the 8 classes supported by standard TE).

In addition to these standard values, the support of draft-ietf-ospf-te-metric-extension allows us to advertise these complementary parameters:

- Available and Residual bandwidth in bytes/s,
- Delay in µs with the possibility to advertise anomalous value (particular useful in case of problems),
- Jitter in µs,

- Link loss in % with, like delay, the possibility to advertise the anomalous case too.

Complementary to the technical parameters, we support also the price information. Two options are possible:

- Through the addition of a new TLV parameter to convey the Link price in cents/Mbit/s,
- By using the TE Metric field. In this case, all operators within the ETICS community must agree to use the TE Metric to describe the Link price.

5.3.2.2. PoI-to-PoI ASQ path offers

There are no proper ASQ path offers in the PULL MODEL. ASQs are built on demand from the computation of the shortest path on top of the network capabilities exchange between ASVRs. This means that not only TE parameters need to be exchange between ASVRs, but also path, network prefix in the same manner of a standard routing protocol.

As already said, mostly for scalability and confidentiality purposes, ASVRs do not exchange all the details of the network topology. In fact, only a subset i.e. an abstract view of the real network topology is announced between the ASVRs. The aggregation method (i.e. the algorithm that computes the abstraction model) is up to the NSP. For example, we could refer to [WH07]. But the network schema announced by each NSP remains more or less the same: ASVRs announce through the H-TE area two kinds of information:

- Pseudo-Links for which attached TE parameters describe the network capabilities,
- And Pseudo-Nodes that advertise the node of the abstract network topology.

In this way (see figure below), the network could be seen as a black box where only entry/exit points (i.e. the PoI) are announced, together with the TE parameters that represent the performance for the traffic crossing the network from one PoI to another one. More precisely, the pseudo nodes that correspond to:

- The ASVR itself that corresponds to the nucleus (N1 and N2 in our example below) in the abstract view of the network. It could be considered as the barycentre of the network,
- The different ASBRs that connect the AS to its neighbour ASes ($ASBR_{11}$ and $ASBR_{21}$ in the figure below).

The nodes are interconnected by the pseudo-links (in red in the figure below) which are of two types:

- Inter-AS TE links: By using the RFC 5392, the IGP-TE could collect precisely the TE parameters of the Inter-AS link that connect it to the remote AS i.e. the link between the ASBR that run BGP (the link between $ASBR_{11}$ and $ASBR_{21}$ in the figure below). These information are redistributed between the ASVRs in the H-TE area,
- Intra-AS TE Link: 2 new pseudo-links ($PL_{ASBR11N1}$ and $PL_{ASBR21N2}$ in the figure below) are advertised to describe the TE parameters to go to/from the nucleus from/to an ASBR.

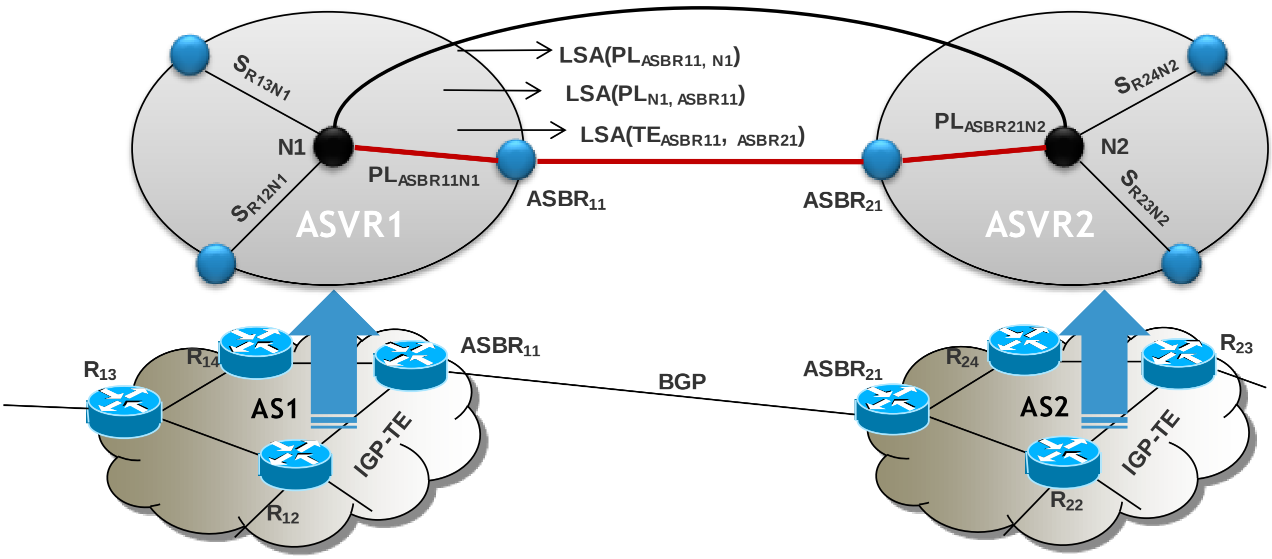


FIGURE 25 TOPOLOGY ABSTRACTION FOR H-TE

In addition, network prefixes are also announced in order for each ASVR to localise the region (i.e. to determine which AS owns a given prefix). This is optional as the information could be extracted from the BGP table. In our case, only own network prefix are announced by the ASVR to complete the H-TE topology. We see in the next subsection how regions can be taken into account.

The PoI is described by the knowledge of the pair of ASBRs and the inter-AS link that connected these two ASBRs. Of course, a PoI could contain more than two ASBRs. If this has some advantages (i.e. a precise knowledge of TE parameters of the Inter-Carrier links inside the PoI) this has also a drawback: it is impossible to determine or name precisely a PoI. Only inter-AS link and connected ASBR are known in the Pull Model. However, during the path computation (i.e. service composition, see section 5.4 below) the Pull Model will use the precise topology information of the PoI (i.e. the inter-AS link and connected ASBR capacities) instead of the PoI itself. Then, the PoI is automatically selected as during the second phase of the service composition (when the PCE uses the BRPC algorithm), it is possible to impose which ASBR has to be used, and so not only the PoI, but also the inter-AS link located in this PoI.

When requesting an inter-carrier ASQ path, the client could precise the scope (i.e. where the ASQ starts and stops) in two ways:

- Directly by providing the IP address of the Ingress and Egress nodes. These nodes could be located inside the PoI, e.g. by providing the loopback IP address of the ASBR, or outside if the ASQ start (respectively end) from (respectively to) the middle of the network,
- Indirectly by giving the source and destination IP addresses in the form of decimal dot notation followed by a slash and the size of the mask e.g. 10.1.2.3/32. This way, it is also possible to ask for network prefix instead of IP address. Then, the service composition will retrieve from its topology database where the source and destination are connected, and so the Ingress and Egress node of the ASQ.

5.3.2.3. Taking into account regions

As mentioned previously for the PoI-to-PoI ASQ path offers, the region are taken into account during the service composition following the Inter-carrier ASQ path request. For that purpose, it is mandatory to request one or more network prefixes instead of a simple IP address.

To determine which network prefix are directly attached, or served by the different ASes, the ASVR needs to know the own network prefix of the ASes as well as the networks that are directly connected to them (in case the prefix is not assigned by the AS). Such information can be obtained in two ways:

- By learning the BGP routing tables and extract from the whole Internet routing table which ones are suitable for the ASVR,
- By advertising also the network prefix into the H-TE area.

If the first solution does not require an additional protocol, it suffers from a higher degree of complexity and dependency. Indeed, the service composition module must maintain a mapping table between the network prefixes and the ASBR. This not only require to search in the BGP database the network prefixes that are attached directly to a given AS, but also to maintain up to date the mapping table each time BGP announce a modification. It is not trivial and generates a risk not to learn all networks prefixes if a BGP router has decided to aggregate several network prefixes. In the second solution, we simply use the native IGP features to convey the own network prefix like the IGP does in the standard.

5.3.2.4. Capabilities exchange

In the PULL MODEL, we have chosen to use a standard IGP-TE protocol to convey the network capabilities. Two choices are possible: OSPF-TE or IS-IS-TE. Both are adaptable and the IETF provided several RFCs to exchange Link State Advertisements (LSA), which can be used to represent the abstract view of the Hierarchical Topology. For example, the network capabilities necessary for the PULL MODEL are conveyed through the Opaque LSA type 10 (Area flooding) and Type 11 (AS flooding).

5.4. ALGORITHMS FOR OFFERS COMPUTATION/NSP CHAIN COMPUTATION

5.4.1. ALGORITHMS FOR OFFERS COMPUTATION/NEGOTIATION

Algorithms for offer computation and negotiations have been described in detail in Deliverable D5.6 [ETICS-D5.6], which is internal to the ETICS project. These algorithms will be available in the upcoming and final detailed specification deliverable D5.8 [ETICS-D5.8], which will be publicly available. This section presents only an overview of these algorithms. The reader should refer to the upcoming Deliverable D5.8 for more details about them.

5.4.1.1. Inter-carrier ASQ Path Computation/NSPs selection

Some of the algorithms presented in this section have been previously described in [ETICS-D3.3]. [ETICS-D5.6] or its final public release as [ETICS-D5.8] provide a specification of these algorithms for an implementation purpose. Hence, some models have been refined and some choices have been made on the basis of the simulation results provided in [ETICS-D5.4].

5.4.1.1.1. Negotiation under different operational modes

In a centralized architecture the problem of negotiating occurs especially under the push option. In the pull option the central entity does not chooses the offers but the NSPs does. The problem is thus repeated at the NSP level. Hence, we rather focus on the centralized push and consider how a central entity can both optimize the NSPs revenues while satisfying customers.

5.4.1.1.1.1. Full or Per-NSP Centralized push operational mode

The Service Facilitator retrieves offers and associated prices from one or several repositories it can access to and which are fed by the NSPs. The Service Facilitator is the entry point for customer requests and treats them according to their arrival time. The whole negotiation process can be summarized as follows:

1. A customer makes a request at time t which consists in the selection of a normalized QoS profile and an (entry, exit) points couple.
2. The Service Facilitator identifies the feasible offer chains and selects one following a criterion (e.g. the chain of maximal benefit).
3. The Service Facilitator negotiates each offer with each corresponding NSP: each NSP checks that its network is able to support the selected offer.
4. All NSPs confirm or infirm their offer availability.
5. An inter-carrier ASQ path is made to the customer who decides whether to accept it or not. If yes, the offer is instantiated.

The problem faced by the Service Facilitator is to maximize the ETICS community revenues while making a proposal that fits the customer expectations.

5.4.1.1.1.2. Distributed (Push & Pull) operational mode

The problem for an NSP in a distributed architecture is very different to the centralized counterpart where only one entity (for a given request) centralizes the choices with partial information (especially on NSP capacities) but also more complete ones (on the possible NSPs paths). In a distributed architecture a NSP is facing two choices:

- To which neighbour NSP - allowing reaching the target - should the remaining request be propagated?
- Which offer to select for the request so that the remaining request succeeds and gives the highest revenue?

For the first decision, in the pull operational mode, the NSP knows, through the capability disclosure, its neighbours and which NSPs they are a gateway to. In a distributed push operational mode, the NSP paths are built by the aggregation of offers. The aggregation of offers allows an NSP to know its neighbouring NSPs and the destinations (PoI or Region) they allow reaching.

FIGURE 26 describes the distributed process involving a learning algorithm in a **distributed pull mode**. This process is composed of two phases:

1. a choice phase, highlighted in yellow,
2. an observation phase, highlighted in red, where the NSP observe the results of its choices.

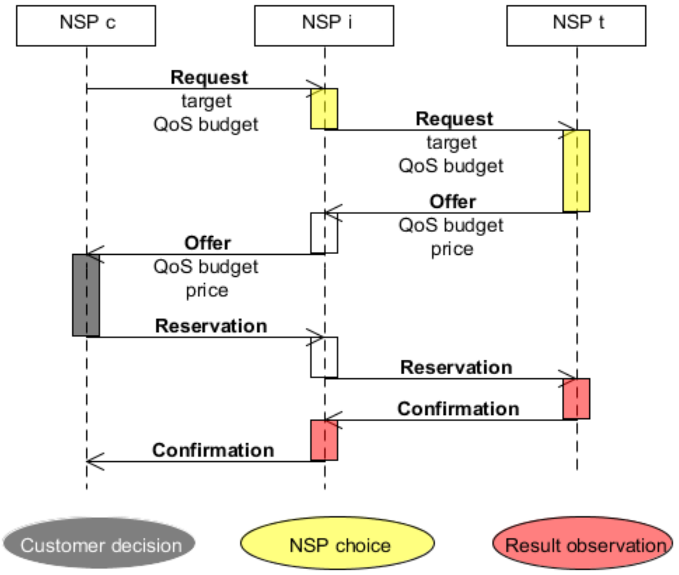


Figure 26: Distributed framework for ETICS pull option

FIGURE 27 illustrates the process of a **distributed** negotiation in a **push** architecture: offers are disclosed as well as some indicators about the availability of offers. The NSP is thus facing two slightly different choices:

- Which published offers to aggregate with its own offers?
- Which request to satisfy and with which aggregated offer (assuming that an aggregation of QoS could actually lead to several aggregated offers)?

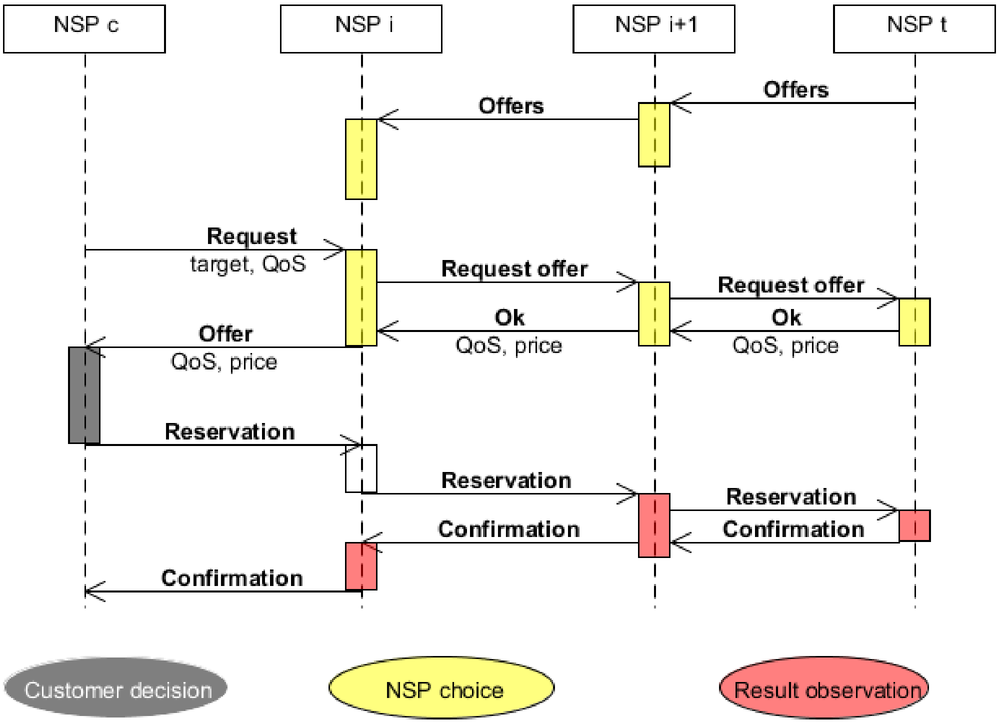


FIGURE 27: DISTRIBUTED FRAMEWORK FOR ETICS PUSH OPTION

5.4.1.1.2. Optimization Algorithms for ASQ path computation

Previous work has studied the problem of combining several QoS offers as an optimization problem for a given customer demand.

5.4.1.1.2.1. Model specification in optimization approaches

In a centralized model, the problem can be formulated at a Multi-Choice Multi-Dimensional Knapsack Problem (MMKP) for each NSP chain. Such chains are determined by the graph of NSPs. The MMKP can be built as follows: the NSPs being the classes of items and each item being a multi-dimensional offer.

In a distributed pull mode, the exploration is quite particular and described in Sec. 5.4.1.2.

In a distributed push mode, the optimization is quite limited and the function to optimize must endeavour to track both the economic and the technical interests of offers. For these reasons, we have not considered such an application.

5.4.1.1.2.2. Optimization algorithms

The problem being NP-Complete due to the multi-dimensionality of QoS [WC96] and to the fact that in the worst case all NSP chains should be explored, different approaches have been considered by various authors:

- Deterministic algorithms: enumeration, branch-and-bound, dynamic programming [DP09] which benefits from some space reduction if the published offers are similar, graph-based approaches [XB05].
- Meta-heuristic algorithms: genetic programming [HG06], ant colony optimization [DP09].

We have implemented such algorithms in the platform for the centralized push mode. The whole process is consisting in 1) transforming the published offers into the solution space of an MMKP and 2) Applying one of these algorithms.

5.4.1.1.2.3. Implementation specification

The implementation specifics and details about the mapping of the ASQ path offers into the solution space of MMKP were described in the private deliverable D5.6 and will be made publicly available under the upcoming Deliverable D5.8 [ETICS-D5.8].

5.4.1.1.3. Learning Algorithms

Optimization approaches are adapted to treat one request. But extending them to treat several requests simultaneously would lead to complex and maladjusted solutions. In fact, this would impose queuing requests until they are treated, which might be adequate for specific contexts, but may be not for taking the business and real-time contexts that we consider. We explore therefore a different approach: Reinforcement Learning (RL) techniques.

5.4.1.1.3.1. Model Specification for reinforcement learning algorithms

Reinforcement Learning techniques are based on Markov Decision Processes whose description for the specific problem of NSP/offer selection have been described in [ETICS-D3.3]. In this section, we specify the application of these models to the different modes described in the previous section.

A Markov Decision Process (MDP) is a tuple *(S,A,R,P)* modelling the environment of a learning agent where *S* is the set of states of the environment, *A* is the set of actions the agent can choose, *R* (s,a) is the reward function defining the gain the agent can have using an action *a* at a state *s* and *P (s,a,s')* is the transition function giving the probability to be in a state *s'* when applying an action *a* at state *s*. The function *P* verifies the Markov property.

In the **centralized push mode**, the learning agent is the Service Facilitator. The corresponding MDP should be defined as:

- Each state is couple (requested QoS of an inter-carrier ASQ path, success/failure) where the requested QoS is the one of the requests at decision epoch *t* and success/failure tracks the success of the proposed end-to-end offer at epoch *t-1*,
- Each action is a combination of NSPs' offers,
- The reward function is equivalent to the price of the combined offers if both the customer accepts and by the NSPs acknowledge the availability of their offers,
- The transition function is a combination of: the customer likeliness of acceptance, the customers' request law of arrival and the NSPs acknowledgement probability.

In the **distributed pull mode**, two models have been studied. In the first model, labelled Comb-NC, both choices are realized jointly. In the second model, labelled Sep-NC, choices are done separately to keep an exploration of several NSP paths.

As such, in the Comb-NC model only one path is explored. We found that, as illustrated in [ETICS-D4.4], the convergence of the algorithm does not suffer from this lack of exploration which is, on the contrary, beneficial from a runtime point of view. The corresponding MDP of this approach is defined as:

- Each state is a couple (next NSPs, locally chosen QoS, network availability) where the network availability is a discrete indicator of the network available resources (e.g. 0=congested, 1=80% used, 2=less than 80% used),
- Each action is a locally chosen QoS satisfying the requested QoS and the choice of the next NSP,
- The reward function is equivalent to the price of the proposed QoS if the end-to-end QoS offer respects the requested QoS and the customer accepts the proposal,
- The transition function is a combination of: the customer likeliness of acceptance, the customers' request law of arrival and the next NSP capacity to succeed.

In the distributed push mode, we modelled only the choice of acceptance of a request and the selection of the adequate combination of offers. Hence the MDP is quite similar to the one of the distributed pull mode except that the locally chosen QoS is replaced by the aggregated offer and the network availability by the aggregated availability.

In Reinforcement Learning (RL), when the transition function of an MDP is unknown, the so-called "model-free" algorithms should be preferred.

5.4.1.1.3.2. Reinforcement Learning Algorithms that we used:

Even if they have different properties, RL algorithms have a common framework detailed by algorithm 1. The scheme has three main steps within a finite or infinite loop. The choice and observation phases are highlighted using the same colours as in FIGURE 27 and FIGURE 28.

Algorithm 1 Generic learning algorithm

```
Initialize learning parameters
At each decision epoch $t$
        Select an action $a_t$ according to a formula
        Observe reward $r_t$ and new state $s_t$
        Update decision data according to a formula
        Update learning parameters
```

We focus on the Q-learning and Sarsa algorithms because of their "model-free" ability which make them particularly adapted to the inter-NSP offer negotiation where much behaviour is intractable.

The Watkins' Q learning algorithm [WD92]to learn optimal Q-values of each couple (state, action) at each decision epoch t according to a “Q-based” policy π . The Q-value function is defined as ensued:

$$Q_\pi(s,a) = \mathrm{E}\left[R_a^t \middle| s^t = s, a^t = a\right] = \sum_{s'\in S} P(s,a,s')\left[R(s,a,s') + \gamma \cdot Q_\pi(s',a')\right]$$

This recursive definition of the function is exploited by the Q-learning algorithm [WD92] to build a Q-table, without knowing the transition probabilities, but using the observed reward $R^t(s,a)$:

$$Q^{t+1}(s,a) = (1-\alpha^t(s,a))Q^t(s,a) + \alpha^t(s,a)\left[r^{t+1}(s,a) + \gamma^t \max_{a'\in A(s')} Q^t(s',a')\right]$$

where α is a learning rate determining the trade-off between exploration and exploitation. This latter also evolves at each decision epoch and the way it is actualized particularly impacts the convergence property of the algorithm.

Various kinds of "Q-based" policies explore the state/action space in order to learn the optimal policy. The recursive equation allows the system to reach a deterministic optimal policy. So, by definition the Q-learning have to be modified to learn an optimal mixed policy, equivalent to the mixed NE.

The SARSA (State-Action-Reward-State-Action) algorithm [SARSA] also uses the state-action function described above. It differs from Q-Learning algorithm in the update of the Q-values, and more specifically in the consideration of the future action. Q-values are updated as follows: $Q^{t+1}(s,a) = (1-\alpha^t(s,a))Q^t(s,a) + \alpha^t(s,a)\left[r^{t+1}(s,a) + \gamma^t Q^t(s',a')\right]$. The value used for the actualization is not the one of the action maximizing the future reward but the one of the "real" next action. This suggests some smooth modifications in Algorithm 2 to keep in memory the past action and reward when the update is performed.

Convergence: The Q-Learning algorithm is proven to converge to optimal Q-values under two assumptions, as demonstrated in [WD92] all couples (state, action) must be visited infinitely, and $\sum_{t=0}^{\infty} \alpha_t = \infty$, $\sum_{t=0}^{\infty} \alpha^2{}_t < \infty$.

This suggests that, in a finite-horizon MDP, the algorithm has been "trained" prior to its execution. The authors of [DM03] refined this proof demonstrating the existence of an upper bound according to how the learning rate α is updated: if it is updated using a polynomial function then the convergence time is polynomial, if it is updated using a linear function then the convergence time is exponential.

The authors of [SiJa00] provided a proof of convergence of the SARSA algorithm under the following assumptions: i) All couples (state, action) must be visited infinitely, and ii) the Q-based policy to choose actions is greedy at the limit (when $t \to +\infty$).

5.4.1.1.3.3. Learning Policies for NSP Choices

A Q-based policy is the way to select an action based on Q-values. Such policies must allow exploration of the environment (trying several actions to learn Q-values) but also exploitation (choosing frequently actions that maximize the expected reward). Initially, the agent has to learn, so the exploitation has priority. After convergence, Q-values are precise enough to be exploited, so exploitation has priority.

Greedy policy. The greedy policy selects always the action having the highest Q-value, such as $a_t = \arg\max_{a \in A_S} Q_t(s,a)$.

ε***-Greedy policy.*** This policy selects the action having the highest Q-value with a probability 1-ε, and a random action with probability ε. ε is initialized to a high value in order to encourage exploration. It decreases as Q-values become more precise. The policy is greedy at the limit.

Softmax policy. The softmax policy, defined as $P(a_t = a') = \frac{Q_t(s,a')}{\sum_{a \in A_S} Q_t(s,a)}$, selects an action with a proportional probability to its Q-value. This policy is not necessarily greedy at the limit.

Boltzmann policy. The Boltzmann policy selects an action using formula $P(a_t = a') = \frac{e^{Q_t(s,a')/\tau}}{\sum_{a \in A_S} e^{Q_t(s,a)/\tau}}$, where $\tau \in \Re_+^*$ governs the lag between action probabilities. When $\tau \to \infty$ lags between probabilities tend to 0 and the policy is nearly random and uniform. When $\tau \to 0$ lags between probabilities increase and the policy is nearly greedy. Usually, τ is initialized to a high value to encourage exploration, then it is decreased to tend to a greedy policy.

5.4.1.1.3.4. Implementation specification

This specification of the data model of the learning algorithm in order to be applicable either in centralized push or distributed push and pull was presented in the private deliverable D5.6[ETICS-D5.6], and will be made publicly available in Deliverable D5.8 [ETICS-D5.8].

5.4.1.2. Multi-Path Computation

Computing inter-NSP Multiprotocol Label Switching Traffic Engineering Label Switched Path (MPLS-TE LSP) through a pre-determined sequence of NSPs is quite straight as each Path Computation Element (PCE), using the Backward Recursive PCE- based Computation (BRPC), knows who is the next to be contacted in order to continue the computation. The optimality of the MPLS-TE LSP inter-NSP path depends strongly on the choice of the pre-determined sequence of NSPs on which the calculation works.

We proposed in ETICS a novel procedure allowing a forward discovery of multiple inter-NSP sequences and the computation of optimal MPLS-TE LSP inter- NSP path(s) over these NSPs sequences. The discovery operation and the constrained path computation phase are performed simultaneously starting from the source to the destination. Experimental evaluation showed that the proposed scheme is effective in terms of Traffic Engineering solution and protocol efficiency.

We defined the problem as Inter-NSP Multi-Constrained Optimal Path Over Multiple NSP Routes (ID-MCOP-MNR) problem, and proposed an algorithm to solve it. The details of the problem definition and the algorithm were described in [ETICS-D5.6] and will be made publicly available in the upcoming Deliverable D5.8 [ETICS-D5.8].

In order to tackle some of the issues inherent to the BRPC procedure, we also investigated another approach where several inter-NSP routes are explored and the pre-computation of those inter-NSP routes (as is done with BRPC) is avoided, in order to identify the e2e resources that would meet customer QoS requirements. This approach, called Forward Discovery PCE-based Computation (FDPC) was also described in D5.6 and will be made publicly available under deliverable D5.8.

5.4.2. Algorithms for NSP chain computation

In the Pull Model, NSPs exchange network capabilities in the form of Traffic Engineering routing information. Thus, each ASVR is capable of reconstituting the whole topology (at the hierarchical aggregate view) of the ETICS community of NSPs. Once this LSDB is built, it is easier to run on top of the topology a standard CSPF (Constraint Shortest Path First) algorithm to determine the optimal path based on the requested parameters.

As already mentioned, the ASVR not only embeds the OSPF-H-TE protocol, but also provides PCE functionalities (see figure below). This particular PCE, in addition to the BRPC support, is based around two kinds of Traffic Engineering Databases (TED):

- The standard TED is filled by learning the underlying network topology through the IGP-TE (here, OSPF-TE),
- The Hierarchical TED (H-TED) is fed by learning the inter-carrier topology through the OSPF-H-TE protocol.

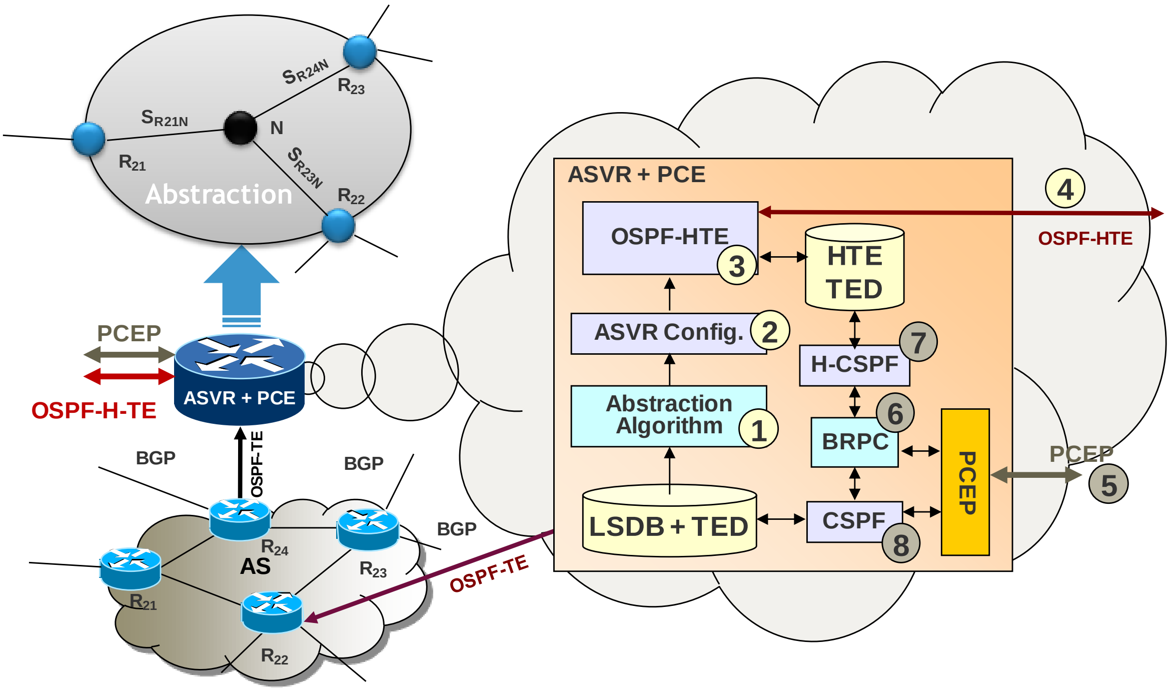


FIGURE 28: PCE MODULES INSIDE THE ASVR

The learning phase corresponds to the steps 1 to 4 in FIGURE 28 in order to fill the H-TED of the PCE. In the pull model, the SLA Manager trigs the PCE without specifying the AS Path (step 5). When an inter-domain request is received, the PCE starts the BRPC algorithm (step 6). This step needs an AS Path in order to know the next PCE in the chain to contact. To determine it, the PCE will use its H-TED and execute a CSPF on the hierarchical topology to determine the optimal AS path that satisfies the requested parameters (step 7 on the figure above). Once the AS path is computed, the first PCE will initiate the standard BRPC algorithm and forward the request to the next PCE (step 6 and 8). The request is processed Hop by Hop (i.e. PCE by PCE) until the destination.

Again, depending on the network capabilities provided by the different NSPs, the AS path computation could result into a more or less detailed vectors. In its simplest form, it could be composed only by the AS numbers, letting each PCE choose the PoI and the BGP routers inside the PoI. In its complex form, the BGP routers inside each PoI have been already selected.

Note that instead of letting the first PCE compute the AS path, it is also possible to fully distribute the computation on a hop by hop basis. In such case, each PCE only computes the optimal next AS in the chain in order to determine the next PCE to which the initial request must be sent. Then, the standard BRPC algorithm takes place. Note that in this way, the PoI (as well as the BGP routers inside this PoI) cannot be selected at this stage of the procedure. They are selected on the backward phase of the BRPC algorithm, i.e. when requests come back to the source. From a business perspective, it could be preferable to completely distribute the AS path selection. Nevertheless, from a technical point of view, the probability of rejection may increase when the AS path selection is done on a hop by hop basis with local criteria for the selection and without global objective of success. An NSP could also not select some PoI or BGP routers during the BRPC process.

CSPF algorithms have been extensively studied and numerous proposals are available today. However, the main principle consists in traversing a tree whose arcs are normalized with TE parameters. For that purpose, the CSPF algorithm starts to determine where the ingress and egress nodes are. Then, the CSPF searches the optimal path on the tree by first pruning all paths that don't match the criteria and then on the remaining paths selects the best one. In the case of PCE, it is possible to specify which criteria (e.g. bandwidth, delay, price …) must be optimized in the whole set of requested parameters through the PCEP protocol. This is possible for the PCC by specifying the objective function in the PCEP request.

As usual, each NSP is free to use / implement the CSPF that satisfies its criteria (both technical and business ones). Some CSPFs are more powerful that other ones, often at the expense of performance. So, as it is part of the PCE, the CSPF algorithm for the AS path selection is not specified.

5.5. ETICS SERVICE ACCESS POINT

Once the ETICS Core System is deployed in each NSP network and the initial learning phase is accomplished, the system becomes operational and could be triggered to request inter-carrier ASQ paths. The ETICS service access point is the interface through which ETICS customers can interact with the community of ETICS NSPs in order to get informed about available offers, as well as to request and order such offers.

5.5.1. INITIALIZATION PHASE:

Take the case of an ETICS customer that needs ETICS services to interconnect some site(s) (e.g. data centre, business site etc). In order to benefit from ETICS services, these sites must be first physically interconnected to the ETICS community of NSPs.

Therefore, prior to any use of the ETICS community ASQ paths, the sites need to be physically connected to the ETICS community (or one of its NSPs) through one of its points of interconnect (PoIs) (or PoEI for VPN services). Also, the customer, and in particular the user of the ETICS service access point, must know the names (following the ETICS namespace convention) of the PoIs/PoEIs to which the "sites" are connected.

As such the following steps are required in order for a site to take advantage of ETICS services:

- Physical interconnection of the site to the ETICS community
- The name of the PoI, and eventually the interface(s) on which the site is connected.

5.5.2. AVAILABLE INFORMATION

This interface should offer the following information to an ETICS customer:

- The regions that can be reached within the ETICS community together with the capabilities associated to these regions. The different region offers are described in Section 5.3.1.2.

- The points where it is possible to offer point to point (multipoint) network services such as business connectivity services.

- Optional: One can imagine a configuration in which the ETICS customers could as well have access to single-NSP PoI-to-PoI ASQ path offers as described in Section 5.3.1.1

5.5.3. Available actions

An ETICS customer can have one of two actions:

- Request an ASQ path offer: In return, the ETICS interface returns an offer.

- Order an ASQ path offer: If the offer satisfies the customer, the latter can order it.

5.5.4. Traffic identification negotiation

During the ordering phase, the ETICS customer and the ETICS community need to agree on the way customer traffic is identified at the entry of the ASQ path. This implies agreeing on the data packet headers that need to be matched at the entry of the ASQ path.

The PoI from which the ASQ path starts contains multiple network interfaces. The first step is to agree on which interface the traffic will be coming from. Then, a flow specification field must be agreed on. The flow specification field identifies which traffic should profit from the ASQ path. Traffic that does not correspond to this identification will not profit from the ASQ path. As a starting point, the fields that were used for matching in the openflow[14] specification 1.1 could be enough for our needs. These fields are the following:

Ingress Port	Ether source	Ether dst	Ether type	VLAN id	VLAN priority	IP src	IP dst	IP proto	IP ToS bits	TCP/ UDP Src port
TCP/ UDP Dst port	MPLS label	MPLS traffic class								

TABLE 1 FIELDS ON WHICH THE TRAFFIC CAN BE IDENTIFIED AT INGRESS POI

5.6. NETWORK PATH COMPUTATION

Path computation in ETICS architecture can be applied in different models and with different scopes, going beyond the pure network path computation where the objective is the calculation of a network route (in terms of network nodes or links hops) between two end-points, compliant with a set of path or metric restrictions and optimization criteria. In particular, in ETICS the path computation concepts and functionalities, traditionally associated to network information like domain topologies or TE metrics, can be enhanced and integrated with concepts related to the Business Support System / Operations Support System (BSS/OSS) area. This is a key feature to enable an effective interaction and cooperation between the service and the control plane in support of the computation of inter-carrier ASQ connectivity service offers and associated network paths.

The following sections present three architectural models for the path computation in ETICS, explaining how they can be applied to the ETICS architecture, their impact on the interaction between the functional modules and their internal procedures. Moreover, each section provides an initial analysis of the extensions required to support these new functionalities, both in terms of protocol modifications (e.g. proposing new

[14] The OpenFlow Switch Specification. Available at http://OpenFlowSwitch.org.

objects modelling the new concepts introduced by ETICS in the path computation) and upgrades of the internal procedures in the different architecture components.

5.6.1. SERVICE PCE: COMPUTATION OF INTER-CARRIER ASQ ROUTE OFFERS

In the ETICS architecture, the NSBP includes the functions related to the management and handling of network connectivity product offers, service specifications and service instances from an operational and business perspective. The creation of the inter-carrier ASQ network service offers on the NSBP side can be effectively supported through the interaction with an enhanced version of the traditional PCE, called Service PCE.

The Service PCE extends the traditional path computation framework to a new concept of PCE, which computes not only network routes, but also network connectivity offers with price information in inter-carrier scenarios. In this sense, the Service PCE can be considered as a converging point between service and control plane routing functions. In fact, it is able to mix heterogeneous information at the service and control plane (TE metrics, prices, customer profiles) in order to compute offers for assured-quality connectivity services across inter-carrier domains, while considering the source/destination endpoints, the traversed Inter-carrier interconnection points (PoIs), TE information, policy data, etc.

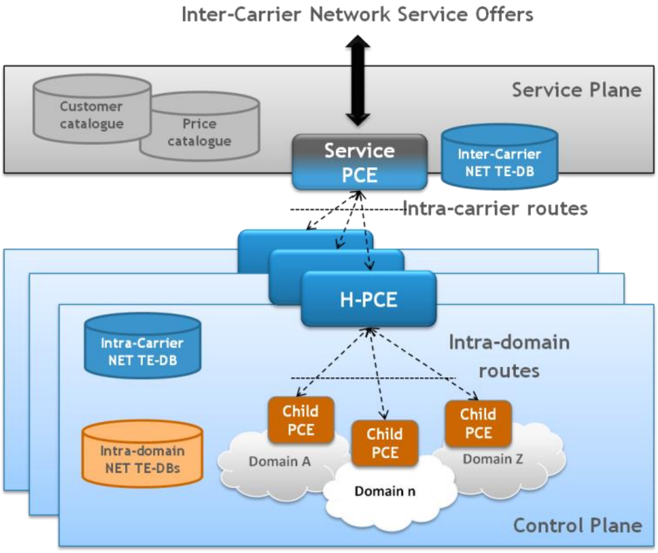


FIGURE 29: SERVICE PCE FOR INTER-CARRIER ASQ SERVICE OFFERS COMPUTATION

In a typical deployment, as shown in FIGURE 29, the Service PCE acts as a second-level parent PCE. The multiple ASes, each of them managed by a different carrier, are organized in multiple domains and adopt the hierarchical PCE model [IETF-DR-6] to compute their internal multi-domain network routes. The Service PCE combines TE indicators about the inter-carrier network topology, the intra-carrier routes obtained from the H-PCEs in each AS, together with service-layer information about prices and customers.

The automated built inter-carrier network connectivity offers can be then passed to the NSBP for the completion of the push or pull model procedures. In fact, in this model, a Path Computation Client (PCC) can be integrated in the NSBP to request route offer computation either on-demand (i.e. connectivity offer tailored to a specific customer request at the NSBP interfaces) or in-advance (i.e. pre-computation of a connectivity offer to be stored in the NSBP Product Catalogue). Therefore, the Service PCE can be deployed in both the ETICS fully centralized pull and push models described in section 4.6.

Some protocol extensions are required in the PCE communication Protocol (PCEP) in order to support route offer computations. The path computation request should be able to specify that the request itself is related not only to a network path, but also to the associated prices. On the response side, a new object should be defined to model the price information associated to a route offer. In particular this object should be able to describe the adopted price model (i.e. *pay-as-you-go* or *flat*), the currency type, the price value for data or time unit and the upper bound for which the offer is valid (e.g. max data volume or time length).

5.6.2. From an Inter-carrier ASQ Path Offer to an Inter-carrier ASQ Network Path

In the Push model, an inter-carrier ASQ path is computed entirely at the service plane and typically calculated based on considerations mainly related to the BSS/OSS area (e.g. SLAs between the carriers, customer profiles, etc.), while it does not take into account control plane indicators like the TE metrics or intra-domain network topology. Consequently, the computation result could provide just a list of loose hops, without specifying the full end-to-end network path. In this scenario, the control plane path computation becomes fundamental to provide the detailed path, not only within a single-NSP or intra-domain scope but also in the Inter-Carrier scope to refine EROs expressed as a list of ASes (i.e. NSP chain), e.g. defining the ASBRs or the interfaces at the PoIs.

The ETICS framework defines the concept of *Contract Identifier* (contract ID) associated to a route offer proposed from an NSP. The contract ID defines a set of information, related to the offer itself and potentially linked to BSS/OSS considerations elaborated during the creation of the offer. It typically has a local scope within the NSP and can be applied during the path computation associated to the network segment of the NSP.

Following this approach, an inter-carrier ASQ path offer defines a sequence of NSPs and, for each of them, the contract ID for their offer. The inter-carrier network path computation can follow the Backward-Recursive PCE-based Computation (BRPC) model [RFC5441], where the path computation request is forwarded along the NSP chain and the end-to-end path is computed composing the network paths of the intra-NSP domains. In each domain, the local PCE server extracts the associated contract ID and computes the path taking into account the data defined for the contract ID. The information associated to the contract ID, and consequently to the ASQ service offer, can define policies to be applied in the path computation, restrict the range of possible ASBRs, specify a set of TE parameters valid for service offer, etc. In order to preserve the internal topology confidentiality within the different NSP's domains, the path computation result could be expressed in form of path keys representing the path segment.

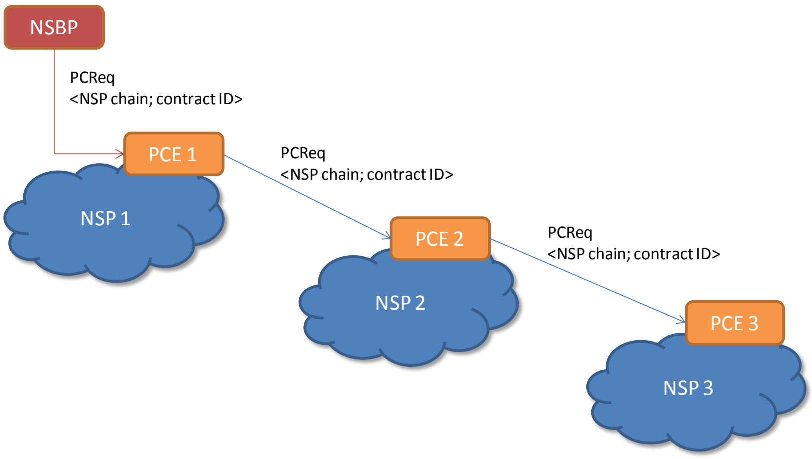


FIGURE 30: INTER-CARRIER ASQ PATH COMPUTATION IN BRPC MODEL (PUSH SCENARIO)

From an architectural perspective, the PCE servers at the control plane must be able to interact with a functional entity in charge of maintaining the correspondence between the contract IDs and the mentioned information associated to the offers. Moreover, it is expected that the PCE servers are able to apply some policies for the path computations, e.g. interacting with an external Policy Decision Point (PDP).

Also in this case, some protocol extensions are required at the PCEP level. In particular, the path computation request must be able to specify the contract ID. An option could be the definition of a dedicated object including the contract ID value. An alternative solution could be the encoding of the contract ID within the VENDOR-CONSTRAINT object as proposed in the IETF draft for the specification of proprietary constraints within path computation request [IETF-DR-5].

5.6.3. FROM AN NSP CHAIN TO AN INTER-CARRIER ASQ NETWORK PATH

As in the previous approach, in the Pull Model the inter-carrier network path computation takes as input a pre-computed NSP chain that, in this case, can include a pre-selection of BGP routers inside the PoIs to be traversed by the chain itself. This NSP chain can be computed adopting the H-TE procedures, where CSPF algorithms are executed over a high-level hierarchical topology (i.e. the inter-carrier network topology) that is learnt at the ASVR through the OSPF-H-TE protocol. The PCE co-located in each ASVR is then capable to run CSPF algorithms on top of the hierarchical Traffic Engineering Database (H-TED). This allows applying more global optimization functions related to the end-to-end path, instead of local policies or constraints to be applied at each single NSP.

Once the NSP chain is available, the PCE task is to compute a complete ERO given the constraint of the BGP routers sequence to be traversed. This constraint can be easily specified in the PCReq using the standard PCEP protocol defined in RFC5440, without requiring any additional extension. In particular, the BGP router sequence can be expressed through a list of hops in the IRO object, that describe the AS associated to each BGP router. The path computation procedure is triggered by the SLA Manager in the NSBP and is typically based on the same BRPC approach (see FIGURE 31) adopted in the Push Model and described in the previous section. For confidentiality, the PCRep could express the path in terms of Path Key Segments (PKS) instead of explicit ERO hops. During the signalling phase, at the entry node of the given NSP, the intra-domain explicit ERO will be retrieved through a PCReq to the local PCE.

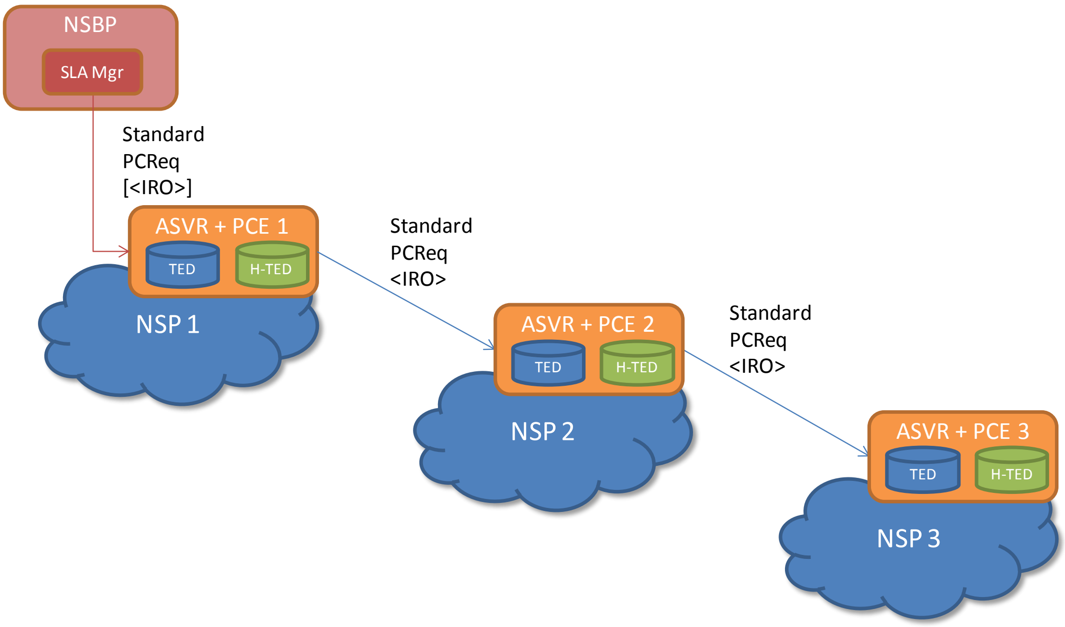


FIGURE 31: INTER-CARRIER ASQ PATH COMPUTATION IN BRPC MODEL (PULL SCENARIO)

Taking into account the presence of a pre-defined domain path, it is clear that, given a specific set of objective function and metrics to be minimized, a single complete ERO and a single offer will result from each path computation. In order to obtain more offers, different types of requests, still compliant with the pre-defined IRO, are required. Such requests could specify different optimization criteria, path constraints or parameters to be minimized or maximized.

5.7. INTER-CARRIER ASQ PATH PROVISIONING

The previously described blocks of the ETICS architecture return an inter-carrier ASQ path to the ETICS customer. If the customer accepts the inter-carrier ASQ path, it needs to be provisioned. Provisioning means enforcing a traffic identification scheme at the border routers involved in the PoIs of an ASQ path: on which incoming data packet headers each NSP in the ASQ path chain needs to take the forwarding (and ASQ enforcement) decision? This is exactly the scope of this section.

A first input to this problem was first provided within [ETICS-D5.6], and will be further enhanced in the upcoming detailed specification deliverable D5.8 [ETICS-D5.8]. This section gives an overview of the problem and the approach recommended by ETICS.

The section is two-fold. First, it deals with how the traffic can be identified at the border routers. It second deals with how such a traffic identification scheme can be enforced at the data plane level.

5.7.1. WHAT IS THE PROPER TRAFFIC IDENTIFICATION SCHEME?

5.7.1.1. Problem specificity for ETICS: destination-based routing is not sufficient

In today’s best effort Internet, traffic is identified and routed at AS Border Routers (ASBRs) based on the IP destination IP header field of packets. However, ETICS proposes a way to reach destinations according to specific needs. This proposed architecture implies that several flows to the same destination may follow different paths, depending on the contracts that the owners of these flows bought for their traffic delivery.

We illustrate this with the example of FIGURE 32. The **red** and **blue** flows have the same destination but do not follow the same path. **As a consequence, the classical "simple" destination-based routing identification scheme is no longer sufficient for the needs of ETICS.**

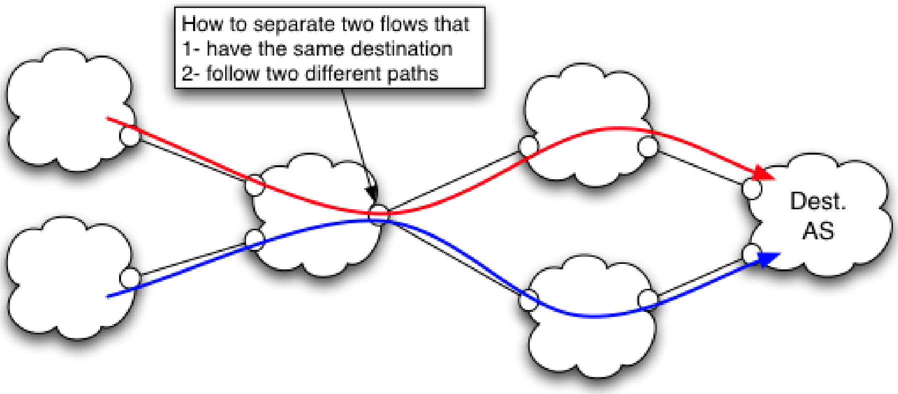

How to separate two flows that
1- have the same destination
2- follow two different paths
Dest.
AS

FIGURE 32: FLOW DIFFERENTIATION FOR A COMMON DESTINATION

5.7.1.2. Requirements of the traffic identification solution:

We think that an efficient identification scheme needs to be insensitive to the following issues:

- Forwarding scalability: The identification scheme must be scalable at the forwarding level. To give a figure about the current scale of Internet's destination-based routing, a typical Forwarding Information Base (FIB) in the Internet DFZ is about 450 000 lines.
- Policy compliancy: the identification scheme must not allow a domain to use a path that is not compliant with the path proposed by its neighbours or to steal (free ride) an ASQ path that is destined to another domain. We illustrate later in this section a global identification scheme that is not policy compliant.

5.7.1.3. Recommended ETICS traffic identification approach

In order to choose the most appropriate traffic identification approach, we first compare the different possibilities in terms of traffic identification.

5.7.1.3.1. Input to the traffic identification problem: the service access point

The input to the traffic identification problem is given during the phase in which the customer orders the inter-carrier ASQ path. During this phase (See the Service access point in Section 5.5 and in particular Sec. 5.5.4), the ETICS customer and the ETICS interface agree on the way the traffic is identified at the first point of interconnect. This results in a set of traffic identifiers that uniquely identify the traffic that needs to flow within the ASQ path (e.g. a set of source IP addresses belonging to the customer and a set of destination IP addresses that the inter-carrier ASQ path needs to reach). We call this identification the *original headers identification*. This identification needs to be performed at least at the entry of the ASQ path in order to know which traffic needs to follow the ASQ path.

5.7.1.3.2. Possible identification schemes

An inter-carrier ASQ path defines a network path through multiple NSP domains. The packets that were identified by the original headers identification need to follow this network path. Any traffic identification scheme at the PoIs should ensure that these packets follow this network path. Different ways are possible to ensure this goal.

5.7.1.3.2.1. Re-identification at each intermediate PoI (using original headers)

A naïve approach is to use the original headers as traffic identifiers along the entire path. In this configuration, the routing decision is taken, at each PoI, based on the original header identification. We illustrate this through a simple example in FIGURE 33. The figure shows the identification scheme on top of the generic inter-carrier ASQ path that concluded Sec.3. We assume in this figure that the traffic identification negotiation concluded to the following original headers identification: (S1…Sx) as source IP addresses (d1…dm) as destination IP addresses.

The naïve approach consists therefore in replicating this original identification header along the path at each PoI. As such, at each PoI, the router needs to match the packets against the original headers identification and take the routing decision to enforce the route (and the QoS) accordingly.

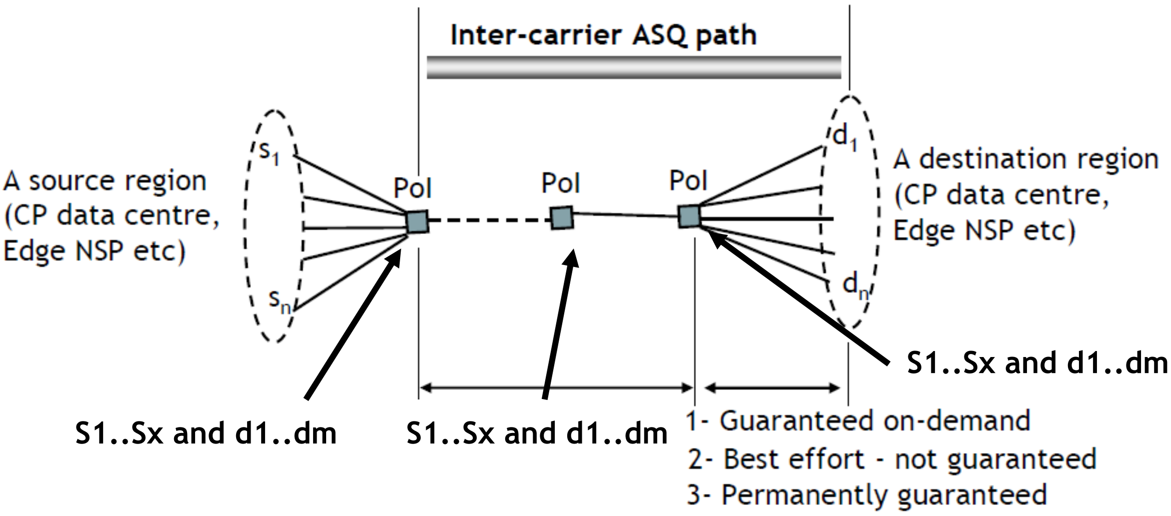


FIGURE 33: IDENTIFICATION USING ORIGINAL HEADERS AT EACH POI

5.7.1.3.2.2. Indirect identification: path identification:

We use FIGURE 34 to illustrate through a simple example the indirect identification approach (taking the same assumptions as FIGURE 33 as far as the result of the traffic identification negotiation step is concerned). This procedure consists in using a path identifier (or a tunnel id) instead of routing based on the original headers. The original headers, result of the traffic identification negotiation, are only used at the entry of the ASQ path (the first PoI). The traffic is then tunnelled (encapsulated) and is routed according to the tunnel/path identifiers.

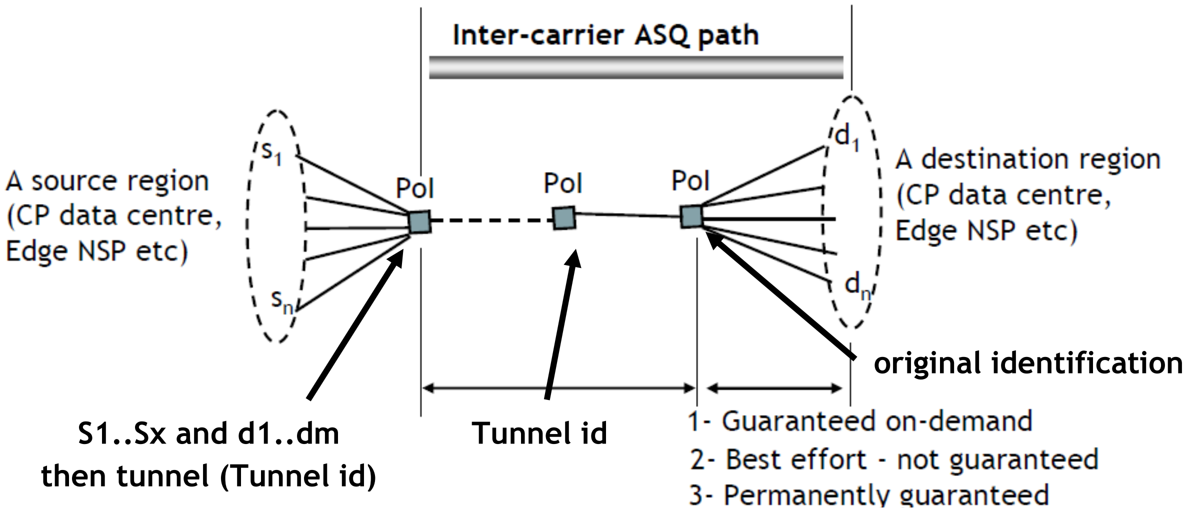


FIGURE 34: INDIRECT IDENTIFICATION/PATH IDENTIFICATION

Note that in both described scenarios (Section 5.7.1.3.2.1 and Section 5.7.1.3.2.2), the original header identification is reused at the last PoI before the traffic splits to reach the different hosts in the region.

In the indirect identification scheme described hereby, we distinguish between two extreme cases: End-to-End flat labelling and source-label stacking. The first is the least scalable in terms of forwarding; the second is the most scalable. However, it is not applicable in practice (for reasons that we will explain later). It has also severe policy compliance issues.

5.7.1.3.2.2.1. End-to-end flat labelling

In this case, the path from stub to stub is identified by the path identifier.

We have evaluated the number of potential routes available towards all the domains in the Internet in the worst case. Addressing all the paths of Internet (toward all destination ASes) a domain can use, the number of paths is in the order of magnitude of 10^7.

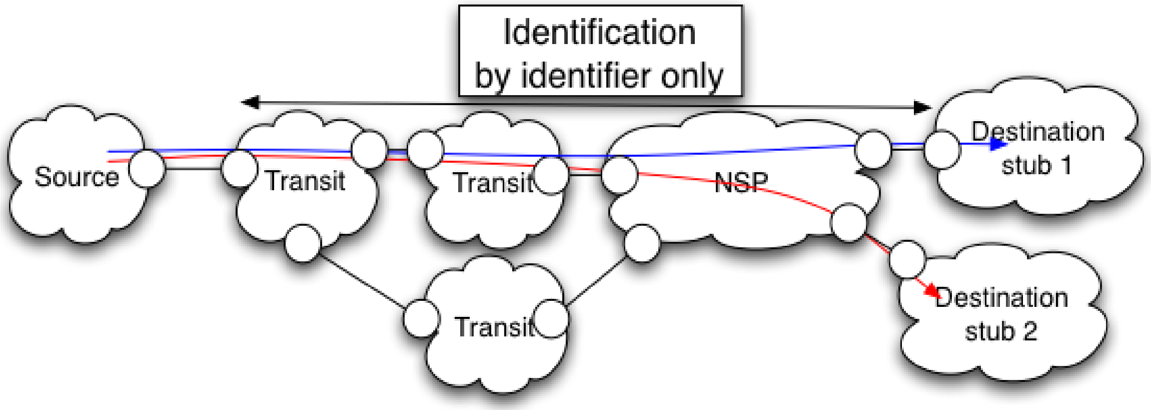


FIGURE 35: END-TO-END FLAT LABELLING

5.7.1.3.2.2.2. *Source identifier stacking*

In such an identification scheme, each domain aggregates all the flows that need to be carried to a common next domain. We must highlight that this scheme is an example, which we consider not being applicable in the ETICS context. Nevertheless it is a good way to underline issues that can be encountered.

FIGURE 36 presents an example of such a scheme. The source (or its provider) adds to the original packet a series of identifiers. Each identifier identifies one domain of the path. For instance, NSP-1 forwards the packets according to the external identifier (i.e. C) only and stripe off this identifier at its exit ASBR, just before forwarding the packet to the next AS (i.e. NSP-2). NSP-2 then receives the packets and forwards it according to the external identifier (i.e. B) only and so on.

This process is very close to the source routing [RFC791] but applies at the AS level.

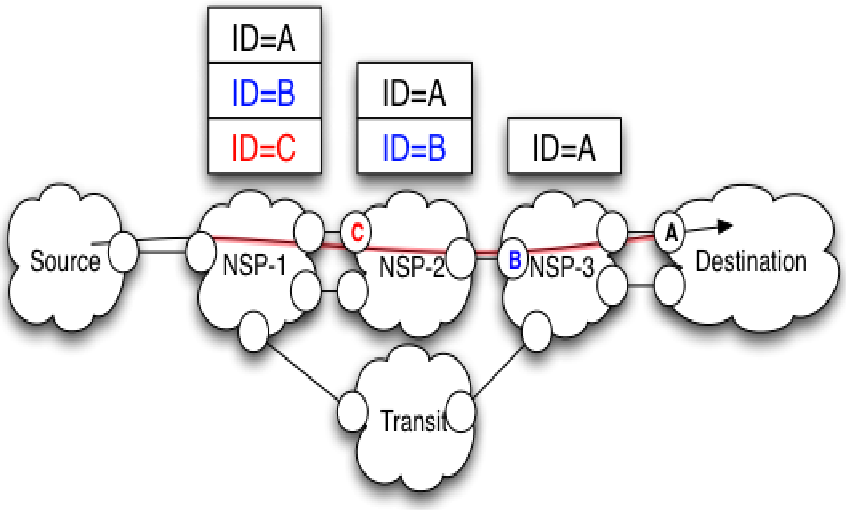


FIGURE 36: EXAMPLE OF SOURCE IDENTIFIER STACKING

This type of scheme resolves the FIB scalability issue as each AS would only need the same number of routes than the number of its neighbours.

For instance, an AS having 1 000 neighbours would only need 1 000 entries in its FIB.

Despite the fact that it is insensitive to the FIB scalability issue, this type of schemes encounters non-negligible issues:

- All the identifiers are inserted into encapsulation header(s) therefore increasing the payload of the packets.
- If formats of identification were not the same among the crossed domains (e.g. NSP-1 is MPLS based whereas NSP-2 is LISP based), the source domain would be obliged to deal with all these formats.

A domain may potentially use paths its neighbours did not advertise to him. An example of policy violation is shown in FIGURE 37.

- By knowing the **blue** and the **green** path (and their associated identifiers), the source AS, which could either be the real source or a transit AS, is able to deduce the **red** path, even if it has not been advertised. In order to prevent such type of violation, packet filtering associated to deep label inspection is necessary. This type of filtering would need to be performed at the entrance of each domain and is very costly (in term of computational cost). **Therefore, this scheme is not usable.**

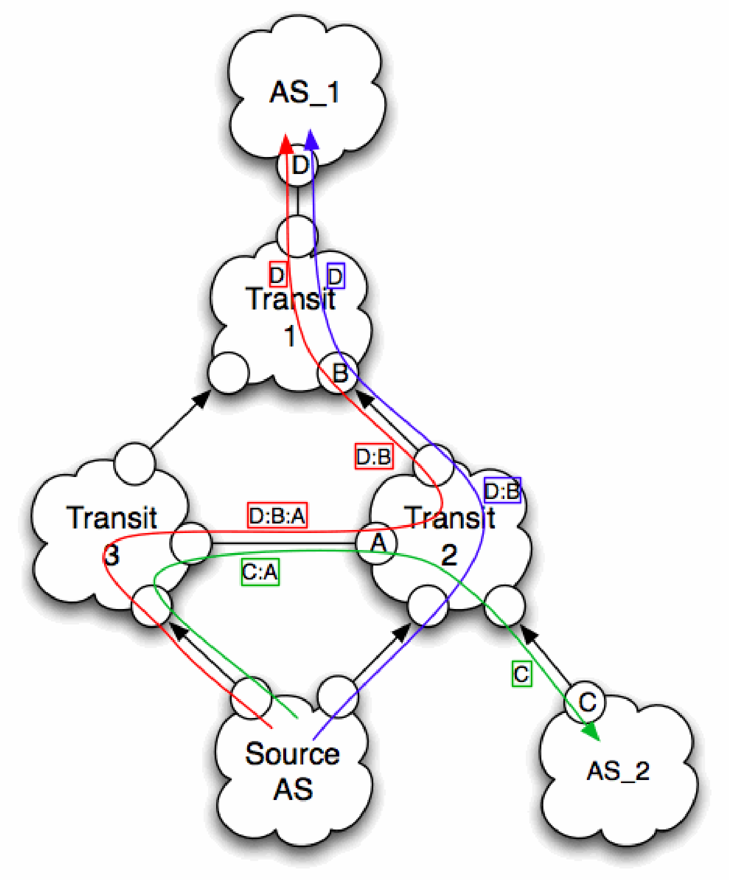


FIGURE 37: POLICY VIOLATION EXAMPLE

Note that for both indirect identification mechanisms, nearby the destination, at the exit of the aggregated part of the ASQ path, identification may be performed based on the original header of the packets.

5.7.1.3.2.2.3. Hybrid stacking:

This approach lies in the middle of the design space between end-to-end flat labelling and source stacking. In the transit networks, where the scalability issue is most likely to take place, there is substantial potential for the aggregation of different paths in common path subsections. Such aggregation may be performed thanks to an extra header. We will see in the next section how such an aggregation could be performed.

5.7.1.3.3. Traffic identification approach recommended by ETICS

Re-identifying the traffic using the original headers, as described in Sec.5.7.1.3.2.1, in not suitable for ETICS. This would in effect imply a high number of rules (policy based routing) that the border routers need to keep. Today's core routers mainly perform "simpler" tasks as destination-based routing or label switching. Therefore, we recommend to focus on the indirect traffic identification scheme, thus restricting the original headers identification use only at the edge (at the entry of the ASQ and at the exit before terminating on a region).

Therefore, the tempting approach for ETICS relies therefore on the following key points:

1. Original identification at the edge and packets encapsulation: packets are identified using the original headers (according to the traffic identification negotiation process) and are then encapsulated to enforce their path till the next intermediate PoI.
2. Path identification: The intermediate PoIs do not look at the original packet headers and rely only on the path identifier to take their routing and processing decision. Already existing labels or identifiers can be used as Path Identifiers. For instance, MPLS encapsulation path enforcement can use MPLS labels (20 bits per incoming/outgoing router interface) as the path id. One could also imagine that the IPv6 Flow Label (20 bits) could be used in the future for such a goal. Another possibility could be the use of the 64 lowest-order bits of /64 IPv6 addresses and LISP path enforcement. All these possibilities must only impact the extra header and not modify the user packet (including the original header identification). However, for the short term, it is realistic to focus on the MPLS label as path or tunnel identifier for identification at intermediate PoI.

As a conclusion, the indirect traffic identification scheme is the one recommended by ETICS. We saw three flavours of such a mechanism: the flat labelling, the source stacking and the hybrid approach. The flat labelling works fine except that it has scalability issues since routers need to keep a label for each ASQ path in the community. Source stacking, as we explained is not feasible because of policy compliance issues as well as because of the size of the headers (due to stacking). We think that the hybrid mode needs as such to be considered and evaluated. We leave this evaluation for future work and only give an illustration of how the NSPs in the community can aggregate paths to reduce the state that needs to be kept in the different routers.

5.7.1.3.3.1. Aggregation of ASQ paths in the hybrid mode;

Several ASQ paths may be aggregated, into a single and bigger ASQ path, by intermediate ETICS partners. The FIGURE 38 illustrates such an aggregation.

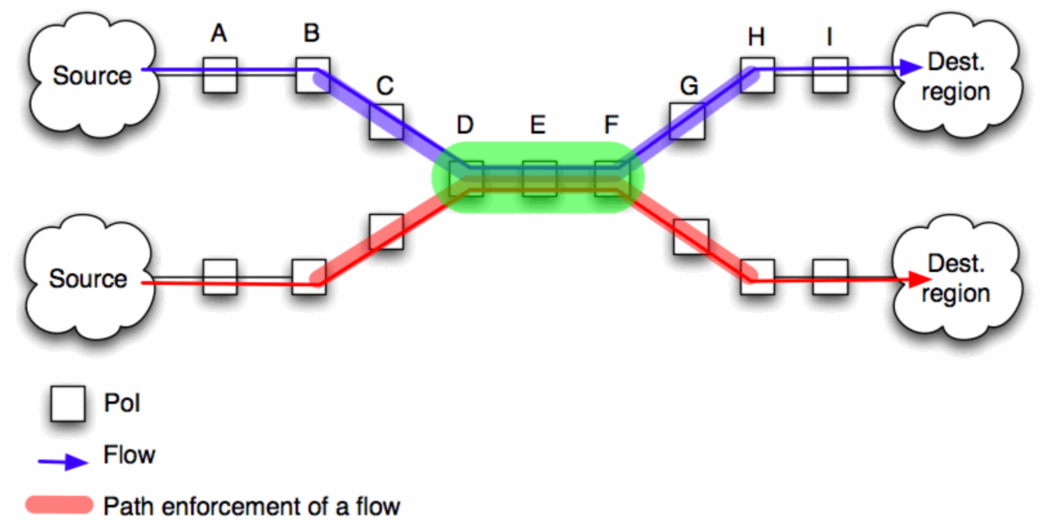


FIGURE 38: ASQ PATH AGGREGATION

More traffic identification points are used to aggregate ASQ paths:

1. Point D: At the aggregation point of ASQ paths. This aggregate is another ETICS ASQ path. It must therefore identify the flows, which are to be forwarded into this aggregate, and deduce another outgoing path-ID (i.e., Path-ID_ag). This identification is performed thanks to a FEC, which is comparable to the one at point B. This identification may be performed based on the already existing Path-ID (i.e., the one inserted by the PoI B) and/or on the original packet IP destination address. Taking into account the original IP destination address cannot be performed as it leads to scalability issues, as underlined earlier. We assume that all the flows following an ASQ path are aggregated in the same aggregate. Therefore it is possible to identify the corresponding ASQ paths by only taking into account the associated path-ID. Packets are then encapsulated another time and the Path-ID_ag is inserted into the extra header in order to forward the packet till the exit of this ASQ path aggregate.
2. Point E: This point only has to forward the traffic to the next PoI and is therefore equivalent to the previously analyzed points C and G. It is therefore better to identify the path based on the external path-ID (i.e., Path-ID_ag) only, which is used for the aggregation.

Once we converged on a traffic identification scheme, we next give an idea about how such an identification scheme could be enforced.

5.7.2. ENFORCING THE TRAFFIC IDENTIFICATION SCHEME: A DISTRIBUTED END-TO-END SIGNALLING MECHANISM (RSVP OVERLAY)

This section deals about how the traffic identification schemes described in the previous section could be enforced on the network.

Today, and mostly for security reasons, only the BGP protocol is allowed to run at the Peering Points. It is not thus possible to use other signalling protocols to set up routes that enforce the traffic identification scheme on the border routers. RSVP-TE, for instance, could not be used at the Peering Point to stitch or nest the tunnels. To overcome this limitation, one might think about setting up IPsec tunnels between BGP routers in order to let RSVP-TE cross inter-domain boundaries. Unfortunately, this only could work if all the NSPs are homogenous: this does not work if at least one NSP is not supported MPLS or MPLS-TE. So, exactly like we propose to use an IGP-TE in overlay mode (for the distributed pull), we propose to use RSVP-TE the same way. After endowing the ASVR with the PCE functionality, we need to augment it with the RSVP-TE support too. As a result, the ASVR is more or less a complete MPLS-TE router with all the TE features: IGP-TE, MPLS-TE, RSVP-TE and PCE. Like other RSVP-TE routers, the ASVR could be triggered to setup a TE tunnel. But, this time, instead of creating the TE tunnels into the real network, the TE tunnels will be setup between the ASVRs. In a standard router, the RSVP-TE *path* message determines the route of the tunnel, or simply follows the *path* provided in the Explicit Route Object (ERO) e.g. the one computed previously by the PCE. Each router that receives the *path* message checks if resources are available and marks them as pre-reserved before forwarding the *path* message to the next router. Once the egress router i.e. the tail end of the TE tunnel, receives it, it sends back the *resv* message to confirm the resource reservation and indicate which label the previous router must use. Hence, Hop by hop, every router receives the reservation confirmation and the inner label. It finishes configuring its routing LSP table and confirming the resources reservation. In case of failure, the router that does not have sufficient resources available sends back a *path error* message. From this point, the source sends periodically RSVP *path* messages to maintain the TE tunnel up.

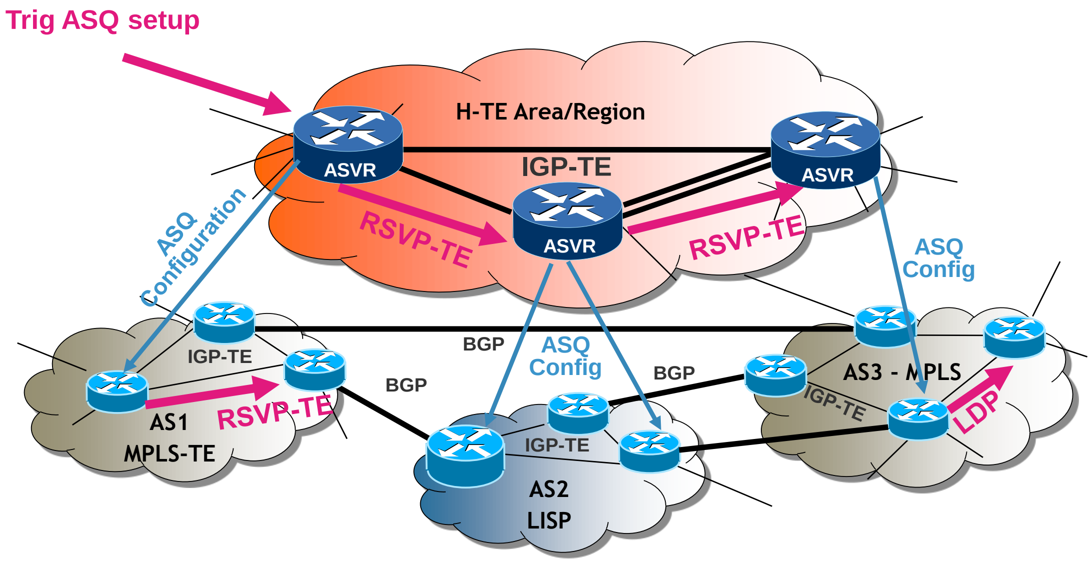


FIGURE 39: RSVP-TE IN OVERLAY MODE THROUGH ASVR

In the particular case of the pull scenario, instead of looking to its local resources (which has no sense as it has no data plane), the ASVR will trigger the ASQ enforcement into the underlying network. This configuration depends on the underlying technology used by the Operator. It could be a DS-TE tunnel, MPLS, MPLS-BGP-VPN, LISP, simple DiffServ router configuration etc. Regardless of the intra-NSP technology in use for ASQ path enforcement, the ASVR will forward the initial RSVP-TE path message to its neighbour NSP (ASVR) only once the part of the ASQ path is correctly configured into the network it controls. So, from the H-TE overlay network perspective, we would use RSVP-TE in the same way as in a standard network. The modification is purely internal to the ASVR: instead of looking to some internal resources, the ASVR will trigger the underlying network to check resources availability and enforce the QoS. Again, standard mechanisms are available for that purpose like Netconf or proprietary protocol could be used like OSS/BSS of the equipment providers. In case of failure, a path error message is sent back advertising the problem to the previous ASVR that could release their respective configurations. In case of success, the last ASVR sends a *resv* message to its predecessor with the Traffic Identifier that must be used at the PoI between the BGP routers. Several cases could occur:

- MPLS continuity: Disregarding if tunnels are TE or DS-TE aware or not, we could use an MPLS label between two BGP routers (more precisely between the interfaces of the two BGP routers) to identify the traffic at the PoI. So, the ASVR could transmit this label avoiding an end-to-end RSVP-TE message that performs the stitching or nesting of the different MPLS tunnels. We solved the problem mentioned previously and also save one step (i.e. the stitching or nesting is done after all sub MPLS tunnels are setup with another particular RSVP-TE message).
- Non MPLS to MPLS domain: In this particular case, the ASVR that controls the MPLS domain will send back a label that will not be usable by the non MPLS domain. In fact, the BGP router must be considered as the Ingress Label Edge Router (LER) for the TE tunnel. So, it must be configured with

a Forwarding Equivalent Class (FEC) that identifies the incoming traffic in order to label the correct IP packets i.e. the ones which must follow the ASQ path. To avoid impact on performance, we recommended using only the DiffServ Code Point (DSCP) as traffic identifier in conjunction to IP address prefix (source and/or destination). In all cases, the ASVR of the MPLS domain must be aware that the previous domain is not MPLS aware and so, could not provide traffic identified by a label. A Router Information LSA could be used to advertise what technology the underlying network is supporting.

- MPLS to non MPLS domain: This time, the ASVR has no label to propagate to the previous one. So, it is easier for the previous ASVR to understand that the next domain is not MPLS aware, and so, that the BGP router is the Egress LER of the TE tunnel. The configuration consists simply for the Egress LER to pop the label and deliver a standard IP traffic. Again, to increase performance of traffic identification, we recommend that the BGP router set the DiffServ Code Point of the IP packets output the tunnel before delivery them to the next domain. Then, ASVR could exchange the DSCP value instead the label.

So, by using RSVP-TE between ASVR we solve two problems: Stitching and Nesting of TE tunnels at the inter-domain and support of connection-less technologies. This gives us the possibility to maintain an end-to-end session coherency between the different operators as well as refreshing regularly the ASQ path. Of course, RSVP-TE timers could be largely increased at the ASVR level. We also inherit all MPLS-TE management features like:

- Protection and Fast-Reroute: it becomes easier to coordinate Fast-Reroute and protection between NSPs e.g. add a backup BGP router in the PoI, a backup PoI for the ASQ …
- Management of the ASQ path: Removal, modification will be easier,
- Path Key Sequence (PKS) could be used to protect privacy of NSP,
- Use MPLS OAM to monitor the ASQ even if it not completely composed by MPLS tunnels

5.8. MONITORING

In previous deliverables, network monitoring has been identified as important building block of the overall ETICS ecosystem, e.g. [ETICS-D4.2] and [ETICS-D5.2]. Moreover, the topic has been addressed by the test-bed design (D6.1 [D6.1] Section 2.7) where a network topology suitable to check aspects of the ETICS network monitoring has been presented.

In this section, the term “domain” is used with respect to network domains where “domain” means an adjacent part of the (Inter-)network infrastructure that is under the administrative control of one authority.

Note: For technical reasons, non-adjacent parts of the network infrastructure must be considered as different domains even if they were under administrative control of the same authority.

In the next few paragraphs, we describe which aspects of network monitoring in ETICS have already been addressed in previous deliverables. However, starting with Section 5.8.1, network monitoring is described in a self-contained way such that there is no need to check external references, even though they are provided throughout the text.

Network Monitoring has been discussed so far in [ETICS-D4.1] (D4.1 Sections 3.2.6 Monitoring & Measurement, D4.1 Section 5.3 End-to-End SLA and Network Monitoring, Annex C Network Monitoring) and in [ETICS-D4.2] (D4.2 Section 6: SLA Monitoring and Assurance Architecture). The latter section examines the ETICS monitoring system and related interfaces to the Control, Data and Management Planes, and thus it represents the main source of information for the present deliverable.

Furthermore, a first attempt to identify the interface points of the monitoring subsystem with the core ETICS system has been performed in D5.2 [ETICS-D5.2] in Section 7.2.

5.8.1. Purpose of the ETICS monitoring system

The purpose of network monitoring in ETICS is threefold:

(1) Verification of SLA fulfilment,

(2) Debugging ETICS installations and configurations, and

(3) Educational purposes.

Regarding (1): The market for AQ services will yield definitive information as to whether fully fledged monitoring systems like NMON or other less technical measures will represent the method of choice. However, since QoS-enabled transport services ("goods") are envisioned to be priced more aggressively, it is assumed that disputes between domain owners along the data path are hardly solvable without monitoring technology. Having said that, it is obvious that the monitoring solution must be carefully designed in order to keep monitoring costs low. This proposal seeks to take this into account. Moreover, NMON can serve as an input to estimate customer's Quality of Experience (QoE), and consequently countermeasures can be taken ***before*** the QoE reaches a lower threshold, which helps to avoid user churn. Mapping of QoS parameters (delay, jitter, etc.) and multimedia parameters (codec, motion level, etc.) into QoE levels, that is to say service quality as perceived by the end user, is indeed part of the ETICS framework. QoS parameters, which are the output of the NMON architecture, could provide input to the QoS-to-QoE mapping box.

Regarding (2): Especially during the set-up process of QoS enabled services, NMON can be used to precisely spot the points of failure, which results in faster time-to-market and lower initial costs.

Regarding (3): In the case where over-provisioning is not (economically) feasible, traffic behaviour is hard to predict. NMON in this respect can serve as enabler for the research area of inter-domain QoS, which still has encompasses a large research potential. Related work has already been studied and described in the [ETICS-D4.1], Annex A (State of the Art), which will not be replicated here.

A more detailed description of the NMON purposes and scenarios as well as details on the NMON terminology and methodologies are given in Section 5.3 and Annex C respectively of [ETICS-D4.1].

5.8.2. Metrics with high relevance for ETICS network monitoring

For several years, the measurement community has been defining metrics in order to perform measurement in networks, in particular in ITU-T (International Telecommunication Union Telecommunication Standardization Sector) and in the IPPM (IP Performance Metrics) working group of the IETF. The metrics can be classified into three categories:

- Local metrics: performed at the destination, they show mainly the variation of the service received.
- One-way end-to-end metrics: relevant for evaluating the network service or for performing measurements for non-interactive applications.
- Two-way end-to-end metrics: relevant for evaluating interactive applications or TCP-based transport services.

Within ETICS we will only focus on the main metrics relevant for evaluating the QoS (delay, jitter, bandwidth and loss) experienced in the network, either targeting the full end-to-end path or single interconnect links:

Local metrics

We will focus on the jitter, also called variation in packet delay (RFC 3393). When monitoring the network service, this metric gives an indication of the network state evolution (getting into congestion or not) since it shows the dynamics of queues in the network. Additional metrics such as packet reordering (RFC 4737) (for example, applications using the TCP transport protocol) may be considered.

One way end-to-end metrics

- One-way delay [RFC2679]: this is the transfer delay experienced by packets in the network. This is in general the most important metric when estimating the network service for non-interactive or non TCP-based applications since the symmetry of the network paths cannot be assumed.
- One-way loss [RFC2680]: this is the loss experienced by packets in the network and complements naturally the information given by the one-way delay.
- Available bandwidth [RFC5136]: this is the amount of unused bandwidth on a link.
- Obtained bandwidth: this is the throughput actually consumed by a given application.

Two way end-to-end metrics

The two-way delay, also called round-trip time (RFC 2681) is relevant when the return path has to be considered for the communication.

In order to evaluate the QoS obtained from the network, end-to-end monitoring of the relevant metrics is a priority for ETICS. On the one side, one-way metrics are more appropriate since they do not assume symmetry of the network (services), on the other side, they are noticeably more difficult to obtain since they usually require synchronization of probes and/or end hosts on both “ends” of the data path and always require some form of correlation of measurements performed at both ends.

5.8.3. Monitoring approaches

In the following subsections, network monitoring approaches which have been selected as options for monitoring of an ETICS network will be described. Selected out of a pool of alternatives given in Section 4.1.4.3 of [ETICS-D4.3], OAMMON and active/passive NMON will be introduced, which might be used in parallel as they are orthogonal to each other.

5.8.3.1. OAM monitoring

OAM (Operations, Administration, and Maintenance) is a general term used to describe the processes, activities, tools, standards, etc., involved with operations, administration, and maintenance activities mainly used in the context of computer networks. A set of standards, e.g. ITU-T Y.1731 [Y1731] and [IETF-DR-4], to name only some prominent ones, has already been defined. In the ETICS context we restrict ourselves to the OAM aspects of monitoring – i.e. determining QoS metrics – of computer networks.

This topic had been introduced later in the ETICS project and has not yet been described in previous architectural deliverables.

OAM standards, operating on Layer 2 of the OSI reference model [ISO-OSI], are already implemented in most of recent network equipment and can readily be used for Layer 2 (L2) monitoring purposes., e.g. Link Trace and Ethernet Loopback.

However, besides these advantages, OAM has some drawbacks: firstly, it suffers from scalability issues, and secondly, it lacks security mechanisms needed for an inter-operator use. For this reason, the OAM monitoring is used only for intra-domain purposes.

More precisely, OAMMON is the functionality for collecting information in order to monitor the network, for avoiding, useless OPEX expenses or downtime cost. OAMMON concepts will be applied for fast reaction and possible faults; it is not defined for specific Quality concepts. The figure below shows some points where the OAMMON can be applied:

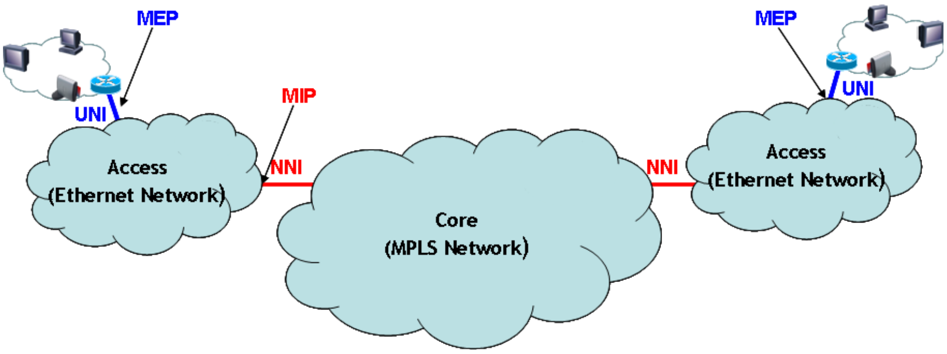


FIGURE 40: MEP AND MIP MONITORING POINTS

MEPs (Maintenance Association End Point) and **MIPs** (Maintenance Association Intermediate Point) are defined for covering the network as much as possible, considering, as important factors, the dimension and the performances. There are three techniques to implement OAMMON in the ETICS project:

5.8.3.1.1. OAMMON Continuity Check

Main goal of the OAMMON is the **monitoring**, by using a high frequency of distributed events, the connectivity between two or more points into a limited network portion. In this context, the OAMMON

Continuity Check is the feature that provides what is defined as “Pro-active Maintenance”. Each termination point of a connection sends an OAMMON Continuity Check Messages (OAMMON CC PDU) towards a list of “extremities”.

As soon as this dialogue (for whatever reason) fails, specific Alarms are sent by NEs informing that the “other end” connection is unreachable. This procedure has the great advantage that, working on the upper levels (layer 2) of the network, it may detect also un-connectivity due to congestion or any other reason that does not permit the detection at Physical Level (Layer 1).

5.8.3.1.2. OAMMON Link Trace Management

The **Link Trace Management** concept is the capacity to handle protocol information transmitted by a maintenance endpoint on the request of the administrator to track a hop-by-hop path to a destination maintenance endpoint. It is used to guarantee the wellness of vital connectivity data within a path. Furthermore, the **OAMMON Link Trace** is used for two main purposes:

1. **Path discovery**: understanding how and where the traffic flows (but it is not our case);
2. **Fault localization**: once determined that MEP does not reach a predetermined point into the network (using OAMMON CC, OAMMON Loopback, from customer report, etc.), the result is to exactly detect the point where the traffic drops.

5.8.3.1.3. OAMMON Loop Back Management

With the **OAMMON Loopback** the operator uses a MEP to ping another MEP or a MIP (if supported by NE).

This is useful in order to know whether a point of the path is reachable by a extremity of a connection.

The OAMMON PDUs, in **Loopback Management** context, are exchanged between maintenance endpoints to satisfy an administrative request to verify connectivity to a particular maintenance point. Loopback indicates whether the destination is reachable or not; it does not allow hop-by-hop discovery of the path. It is similar in concept to ICMP Echo (Ping).

5.8.3.1.4. OAMMON scenarios

OAMMON gives a great flexibility to the network operator, who can choose if the OAMMON will be enabled and in which context defining, e.g. MEPs and MIPs distribution, OAMMONCC vs Link Trace (the two are not mutually exclusives). In ETICS network the PCE is the point of the network that should react to an alarm retrieved by the network in order to reveal whether the privileged traffic needs to be further protected by for example rerouting it to a preferable alternative route or removing other traffic. We define the following scenarios:

5.8.3.1.4.1. Management-based PCE usage - Single PCE Path Computation

This is the simplest scenario, where the OAMMON Network Monitoring System (NMS) constructs the explicit path that it supplies to the head-end LSR using information provided by the operator. It consults the PCE, which returns a path used by NMS.

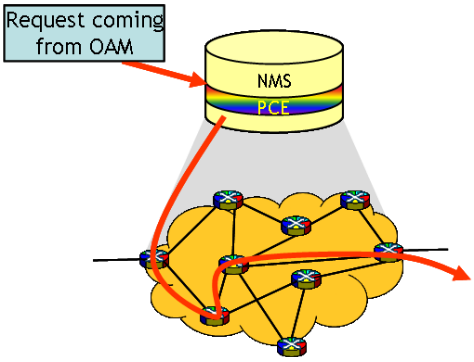


FIGURE 41: SINGLE PCE PATH COMPUTATION: NETWORK EXAMPLE

5.8.3.1.4.2. Management-based PCE usage - Multiple PCE Path Computation

In contrast to the previous scenario the head-end NMS makes a request to an external PCE. The returned path is completed (fully constrained) with respect to the local network, where the NMS has the complete knowledge and control. The NMS only knows the information about the external network that BGP (or similar protocols) provides to it.

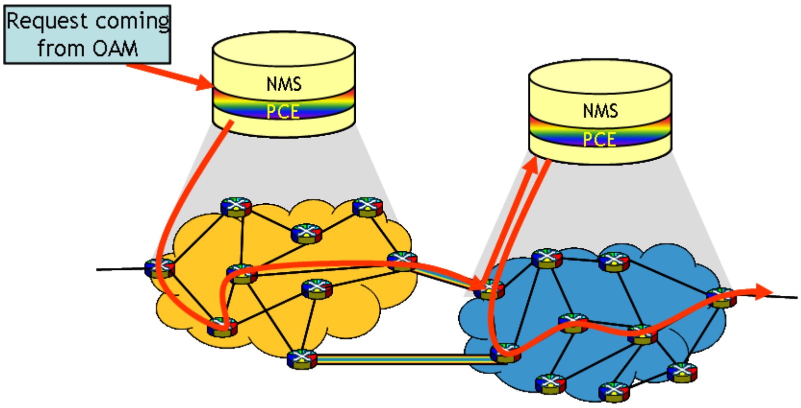


FIGURE 42: MULTIPLE PATH COMPUTATION: NETWORK EXAMPLE

5.8.3.1.4.3. Management-based PCE usage - Multiple PCE Path Computation with Inter-PCE Communication

In this scenario the OAMMON triggers the NMS with a specific service request. Multiple PCE path computation with inter-PCE communication involves coordination between distinct PCEs. In this scenario, the NMS network node or component that requests the computation makes a single request and receives a full or partial path response, but the response is actually achieved through the coordinated, cooperative efforts of more PCEs. In this model, all policy decisions may be made separately by each PCE based on computation information received by previous PCEs.

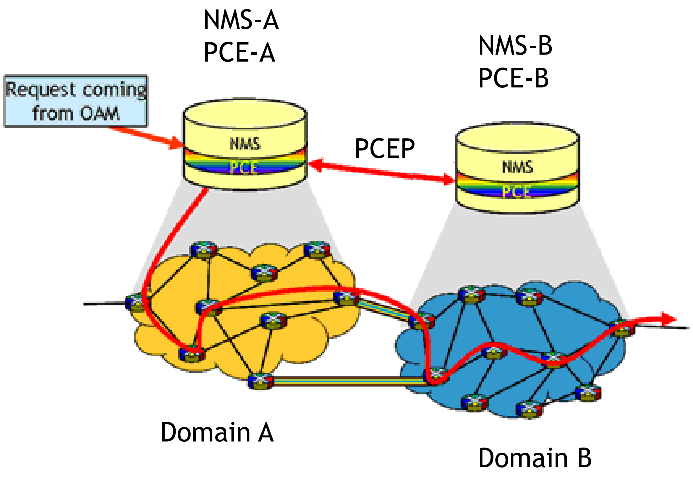


FIGURE 43: MULTIPLE PCE PATH COMPUTATION WITH INTER-PCE COMMUNICATION: NETWORK EXAMPLE

The schema presented in FIGURE 43 is a simple example of PCE inter-domain path computation. The domain A is managed by the NMS-A, having embedded its own PCE (PCE-A). The domain B is managed by NMS-B having embedded its own PCE (PCE-B) – please refers to the slide attached (new FIGURE 43).

The system starts after a new path configuration request coming from the OAM. The OAM have embedded the map of the NMS working in the interested domains and the related IP addresses. So, the OAM choose the NMS-A, sending to it a new inter-domain path configuration request. The NMS-A transform the request coming from the OAM in a PCEP request to the PCE (acting as PCC versus its own PCE). The PCE communicate with other PCEs (e.g. PCE-B) via PCEP. There is no communication and message exchange between NMS-A and NMS-B. The presence of two NMS is just to generalize the context: in other words the process of inter-domain path configuration can be started from NMS-A or NMS-B independently. As result of the path configuration request, both PCE realizes the best path computation inside the proper domain (producing no packets exchanged via PCEP).

As a final consideration, this last scenario can be considered the less probable because it implies that the internal network information (such as structure, composition etc.) is “published” to all the partners.

5.8.3.2. Autonomous monitoring methodology

Another option is to consider that each autonomous system will administrate its own monitoring architecture independently of the other systems. In this autonomous monitoring architecture, SLA violations on the end-to-end[15] path will be detected by end-to-end active or passive measurements (cf. FIGURE 44). Active measurements would usually be launched by the source end-user network (or by local measurements at the destination end user network, depending on the type of metrics).

Detection of SLA violations will set off a mechanism to access monitoring information at the autonomous systems (domain) level in order to locate QoS degradations (see FIGURE 45). The access to the monitoring data will be provided by the NSPs. They may send their measures in response to a solicitation from the monitoring system or authorize other domains to access the data. This latter solution asks for a hierarchical architecture which allows finding and transferring only this data, which is needed for a certain

[15] In this context, “end-to-end” means from access NSP to access NSP, not to the User Equipment

request/measurement. This is necessary because a full transfer of all monitoring data would produce way too much traffic (cf. [ETICS-D4.2] Section 6.4).

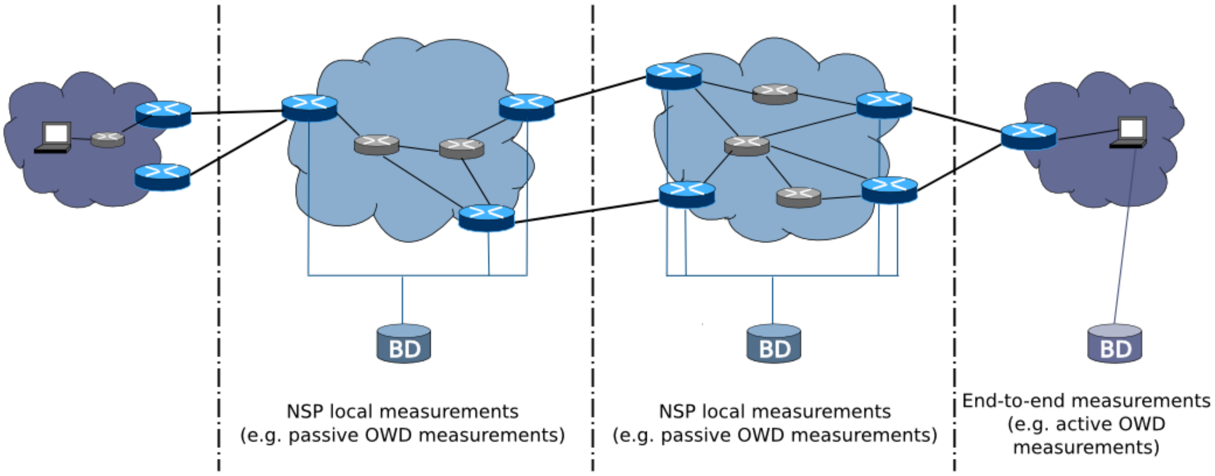


FIGURE 44: GENERIC AUTONOMOUS MONITORING ARCHITECTURE

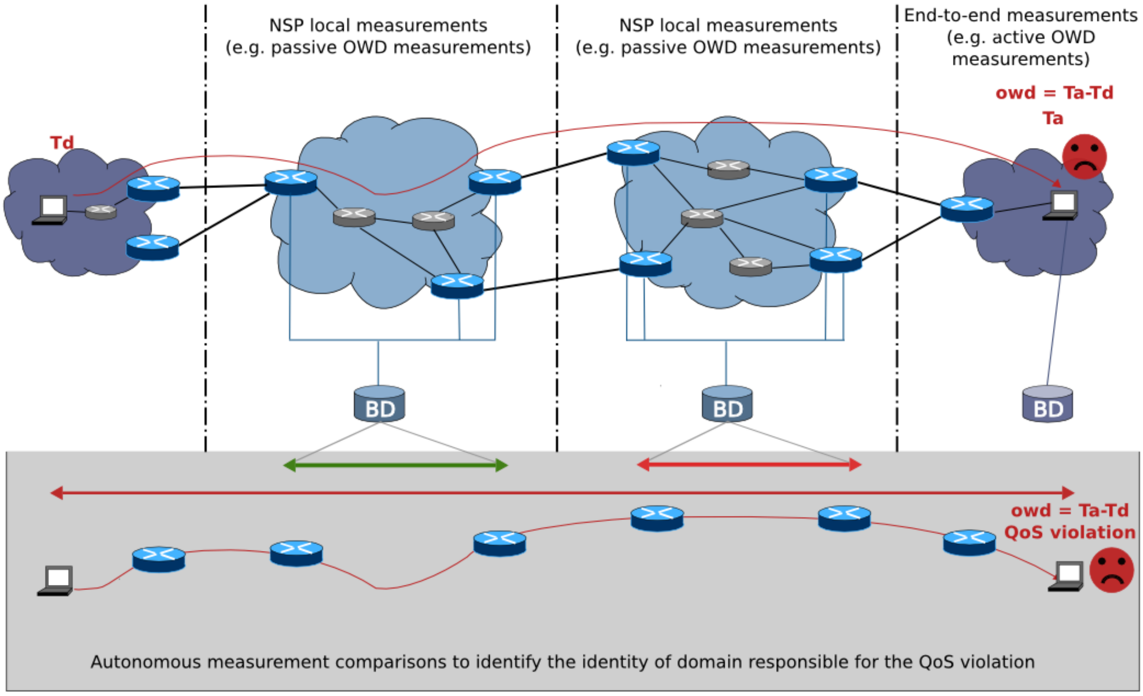

NSP local measurements
(e.g. passive OWD measurements)
NSP local measurements
(e.g. passive OWD measurements)
End-to-end measurements
(e.g. active OWD
measurements)
owd = Ta-Td
Ta
Td
BD
BD
BD
owd = Ta-Td
QoS violation
Autonomous measurement comparisons to identify the identity of domain responsible for the QoS violation

FIGURE 45: QOS DEGRADATION LOCATION IN CASE OF A SLA VIOLATION IN THE AUTONOMOUS MONITORING ARCHITECTURE

The hierarchical monitoring architecture utilising this autonomous monitoring methodology is called "NMON" and is described as follows:

Functionality of the NMON system must carefully be balanced against cost thereof, so the overall monitoring system may not be too complex. This hierarchical approach described here is simple and straightforward and fits well to the autonomous monitoring methodology. Moreover, scalability issues and security/confidentiality concerns can also easily be addressed. This approach is meant to implement the ETICS monitoring system quickly and with reasonable effort.

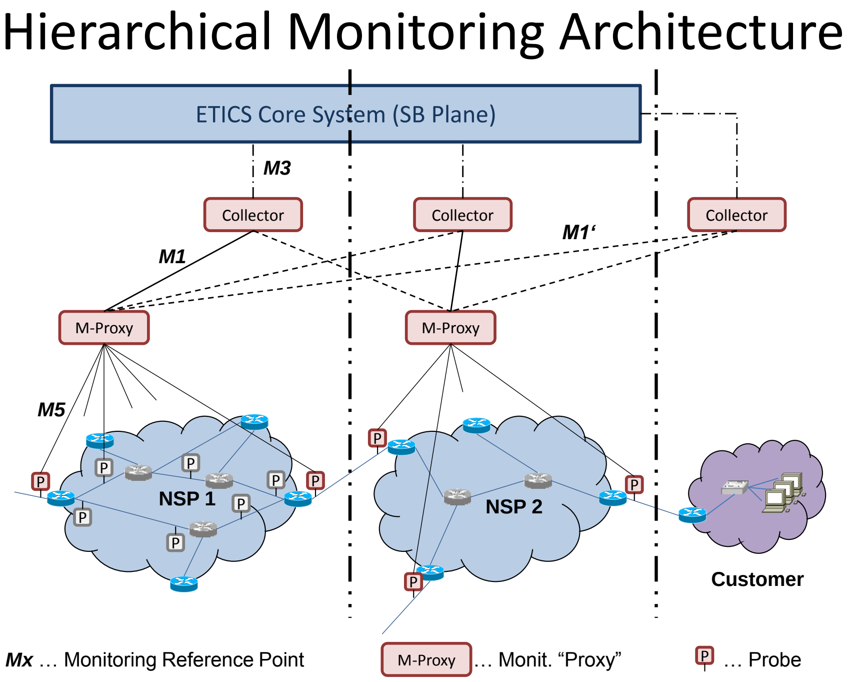


FIGURE 46: HIERARCHICAL MONITORING ARCHITECTURE

FIGURE 46 shows the basic functional entities of the hierarchical approach. The probes are deployed in the network (physical layer) at least on any TDP (cf. Section 5.3.1.1) of any (monitored) link in both directions. An operator may also decide to use the system for monitoring the network internally and to that end deploy probes also within the network, which is demonstrated for *NSP1* in FIGURE 46 (greyish probe symbols).

The M-Proxy (Monitoring-Proxy) function acts as point-of-contact for the other monitoring functions. For availability and/or load-balancing purposes, more than one instance of the M-Proxy function could be deployed per domain. Other functions, such as caches and databases for storing data where applicable (e.g. link-based performance data), could be co-located with the M-Proxy function.

The collector function is part of the front-end and is used to collect, correlate, and evaluate the data retrieved from the probes. The collector function also needs to talk to the ETICS ("non-monitoring") system (via M3) to retrieve necessary data such as routes of established ASQ paths and SLA/SLS information. None, one, or more instances of the collector function can be deployed to either

- NSP domains, including the end customer's domain, or,
- A number of trusted third parties (monitoring service provider).

Which scenario shall be used for the final ETICS architecture is subject to further studies, but is certainly influenced by the extent trust relationships are established between actors. However, the decision where to deploy instances of the collector function does not change the reference points.

M1 is the reference point between the Collector and the M-Proxy function within the same domain. *M1'* is basically the same reference point as *M1*, but improved by some authentication / encryption mechanisms needed for inter-domain, untrusted relations. *M1/M1'* protocols should support forwarding and redirection of requests/connections for availability and load-balancing reasons.

M5 is the reference point between a probe and the M-Proxy function, which is always within the same domain.

Besides *M3*, the main function of which has been described previously, the main function of reference points *M1*, *M1'*, and *M5* is to transport information for configuring the probe and to transfer monitoring data from the probe towards the collector function. For all reference points shown, functions and protocols will be specified in WP5.

NMON covers two mechanisms: active and passive monitoring, both of which are developed on top of the same architecture as shown in FIGURE 46. Active- and passive monitoring have in common that they both work on OSI layer 3 (in contrast to OAM monitoring) and both use hashing in some way. This means that packets are captured at different points on the path and are correlated afterwards by means of hashes. To this end, monitoring information is retrieved by the collector function through monitoring proxies (M-proxy in FIGURE 46). The difference of active and passive NMON is that in the former case packets are inserted for the sole purpose of getting QoS values, while in the latter case only packets of the user's application traffic are captured without injecting additional packets into the network.

Active NMON:

Active NMON is based on active probe packets (i.e. packets injected into the network), which are inserted at the ingress edge and sent through the network towards the egress edge for the sake of monitoring the QoS that they receive. QoS parameters that can be monitored by active mechanisms include, but are not limited to, end-to-end availability, delay and available bandwidth.

At the ingress edge, the emission timestamp is embedded in the payload of active probe packets; at egress edge the reception timestamp is logged. End-to-end delay can be computed at egress edge or at a collector level as the difference between the emission timestamp and the reception timestamp.

One-way latencies measurement has been standardized by OWAMP (One-Way Active Measurement Protocol [RFC4656]). Available bandwidth can be obtained by post-processing the one-way latencies of a series of active probe packets. Different methods exist for doing so, an example of those methods being *Forecaster* [NeHa]. End-to-end availability is a straightforward by-product of active one-way latency monitoring, as any packet which is not received at the egress node is logged as dropped.

In ETICS one-way metrics (availability, latencies, available bandwidth) must be monitored from ingress node to egress node but also per NSP since it is important to locate the faulty NSP in case of SLA violation. As in the case of Active Flow based monitoring, active probe packets are also marked at the ingress node such that they can be identified and evaluated by equipment on the path. To this end, a "magical value" and a sequence number are inserted into the payload of probe packets. The magical value identifies a monitoring job, and the sequence numbers identify the successive packets in the train of active probe packets that correspond to a monitoring job.

Observe from this discussion that there is a lot in common between Active NMON and Active flow based monitoring. The difference between them stems from the behaviour of intermediate equipment on the path. In the case of active flow based monitoring, the intermediate pieces of equipment are active. They include a timestamp (i.e. the timestamp of that packet at intermediate equipment) into the payload of active probe packets. At the egress node an active probe packet payload consequently carries a sequence of timestamps, and the contribution of each NSP to the global latency can be computed as the difference between two of those successive timestamps. On the contrary, in Active NMON all intermediate pieces of equipment are passive. They recognize active probe packets by a series of features among which are the magical value and sequence number (mandatory), as well as port number and protocol (optional). They log the timestamps of active probe packets into a memory efficient data structure that we call sketch and which is based on different hash functions. The timestamp of active probe packets can be retrieved by a collector that is able to send monitoring requests to ingress nodes, egress nodes and intermediate check points through the monitoring proxies.

Passive NMON:

The ETICS passive monitoring system is also based on hashing: for every packet at an observation point (cf. FIGURE 46) a hash value is calculated from the packet headers (fields which might change *en route* must be masked out) and part of the payload of the packet. The same hash function must be applied at all observation points. Consequently, if one packet passes several observation points, it is always gets assigned the same hash value. In this way, the packet can be "tracked" on its way through the network by finding the same hash value at different observation points. The hash value, along with the time-stamp when the respective packet was captured, is stored directly on the probe or on a fast storage which is attached using a high-speed I/O interface.

In modern networks with transfer rates of 10GBit/s and above, it is not feasible to store all hash ⇔ timestamp values, because this would consume too much storage space. Therefore, a filter based on the hash value will be put in place. Of course, in order for this to work, the same filter must be applied at all observation points. In this way, it is possible to store only a portion of the {hash, timestamp} pairs in order to reduce storage space and I/O demand.

The actual calculation of traffic metrics is done by the Collector function (cf. FIGURE 46), which retrieves the timestamps from the probes via the M-Proxy function. The M-Proxy primarily acts as contact point for all monitoring data requests. After verifying access rights, the Collector might be redirected directly to the probe or storage where to retrieve the data from in order to shorten the path the monitoring data takes (this is not shown in FIGURE 46). Furthermore, the M-Proxy function can be deployed several times in a network for load balancing and redundancy reasons.

Please note that in FIGURE 46, the Collector function is shown in the operator's networks and also in the end-user's network. This is to show that the Collector function could be located at all those places in the network but of course this is not necessary (e.g. not every end-user will have an instance of the Collector function). Another option, which is preferred by some of the NSPs in the ETICS consortium, is to have the Collector function located inside a trusted third party.

As an additional privacy measure, the following algorithm shall be applied when verifying that an SLA has not been violated: Whenever an SLA shall be verified for a given time period (e.g. the last hour), the end-to-end (or edge-to-edge) metrics are evaluated and verified. If no violation is detected, this result is reported and the measurement is finished. If, on the contrary, a violation is detected, then – and only then – per-domain performance metrics are retrieved and domain(s) which have violated the SLA are identified. This algorithm also supports the scenario of deploying a trusted third party which performs the measurements.

It is expected that the transfer of monitoring data on reference points M1, M1', and M5 (cf. FIGURE 46) will consume a considerable amount of bandwidth. In order to facilitate this, an optimized, binary protocol will be proposed, the definition of which is currently work in progress. Existing protocols, such as IPFIX [RFC5101], have been considered, but they usually result in too much overhead, translating to a much higher implementation effort. Furthermore, the protocol can be redefined at a later point in time, if necessary.

It is assumed that some network operators will be quite reluctant to let additional traffic be injected into their networks only for measurement purposes. The passive network monitoring methodology accommodates this attitude perfectly. Consequently, it is assumed that the passive monitoring solution could be adopted at much more ease than solutions which foresee that traffic is injected into the network. Therefore, passive monitoring well supports fast practical deployments of the ETICS network monitoring solution.

Another substantial advantage of passive monitoring is that the actual user's traffic is measured which factually forms a "proof" of the service quality which has been delivered, while in the case of active monitoring newly generated packets are used, which might not experience the same treatment by network equipment as the users' traffic. Furthermore, with passive monitoring, it is a lot harder for a network operator to manipulate the performance evaluation results[16], because packets of the unmodified user traffic are picked for evaluation, so it not possible to provide "measurement packets" with preferential treatment.

On the negative side, it must be mentioned that very high-performance measurement equipment must be used, both in terms of CPU power and in terms of storage space directly at the probe or at least at the same site.

5.8.4. INTERFACING NETWORK MONITORING WITH THE CORE ETICS SYSTEM

5.8.4.1. Overview

In contrast to the core ETICS system (i.e. the part of the architecture which is not directly concerned with measurements), which can be regarded as a set of closely interrelated functions serving a common architectural purpose, the ETICS monitoring system is designed as a flexible building block that is only loosely connected to the rest of the system. The reason for such a design is threefold: Firstly, monitoring systems should in general be as independent as possible from the core system which is under measurement, in order avoid the distortion of the obtained results. Secondly, NSPs may prefer their individual selection and deployment of monitoring systems, which are not imposed by the ETICS framework. Thirdly, keeping the monitoring subsystem highly flexible and generic guarantees the highest flexibility and adaptability for all potential future measurement tasks. Fourthly, concerning maintainability

[16] Even though we assume that a certain level of trust is established between members of the ETICS consortium.

of the ETICS system a replacement of monitoring approaches through revised components may be considered in the future, which should not affect other components.

Measurements can be used and integrated from various sources, e.g. OAM monitoring (cf. Section 5.8.3.1) and NMON (cf. Section 5.8.3.2).

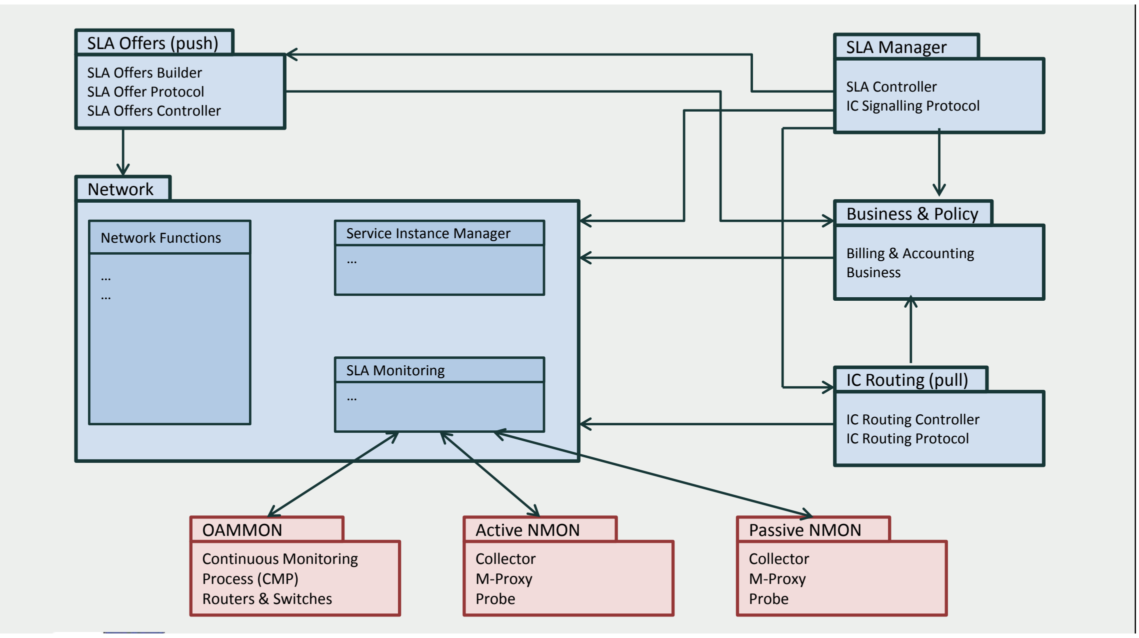


FIGURE 47: INTERFACE TO THE CORE ETICS SYSTEM

In FIGURE 47, the overall architecture of the ETICS core system (blue) and the monitoring sub-systems (red) are shown: The "SLA Monitoring" instance gets the information about contracted SLAs from the "SLA Controller". Whenever the "SLA Monitoring" instance detects a new contracted SLA, it automatically triggers the monitoring process. This process is subject to implementation and configuration specifics, which are out of the scope of this document. All information that is necessary to perform the measurements has to be provided through the SLA. Most notably, the ingress- and egress points of the traffic as well as the description of all Points of Interconnect (PoI) in conjunction with the Traffic Delivery Points (TDP) on the path comprising the mechanisms and traffic identification methods used at the respective TDPs are required.

OAM monitoring, as already explained, is self-contained within one operator's network. In this case, an instance of a "Continuous Monitoring Process" (CMP) functional entity is connected to the "SLA Monitoring" instance. The CMP instance coordinates retrieval of OAM monitoring information from the networking equipment.

In the case of NMON, instances of the Collector function are connected to the "SLA Monitoring" instances. In contrast to OAM monitoring, NMON spans multiple operators, ideally from edge to edge of the traffic path. To this end, additional (ETICS ASQ) routing information for all paths of all monitored SLAs will be needed. It is envisioned that this practically spans the whole Internet. Routing information within the core ETICS system is stored within the "IC Routing Controller" which is connected to "SLA Monitoring". The latter one must be able to pass this information to NMON and optionally to other measurement modules requiring this information.

The measurement results are reported back to the "SLA Monitoring" instance, where the measurement results are checked against the contracted SLAs. If a contract violation is detected, all pertinent information (e.g. faulty domain, violated QoS parameters, etc.) is provided to the "SLA controller" in order for the latter to react and take appropriate decision to correct the problem.

Finally, measurement results could be used by an NSP to adjust their SLA offers or Network Capabilities, in particular when measurements serve to fulfil the traffic matrix.

In the following sections, the communication between the ETICS core system and the monitoring sub-systems is described. As shown in FIGURE 46 this interface is named "M3". In the next section, this interface will be described in an abstract way; i.e. the messages (behaviour) and the information elements are described, but without giving specific encoding. Therefore, the description is given here in D4.4 instead of WP5/T5.1.

5.8.4.2. Communication – Behaviour

In this section, the communication (proposed message flow) for M3 is given. Depending on the monitoring sub-system used, different message flows apply.

5.8.4.2.1. Passive NMON

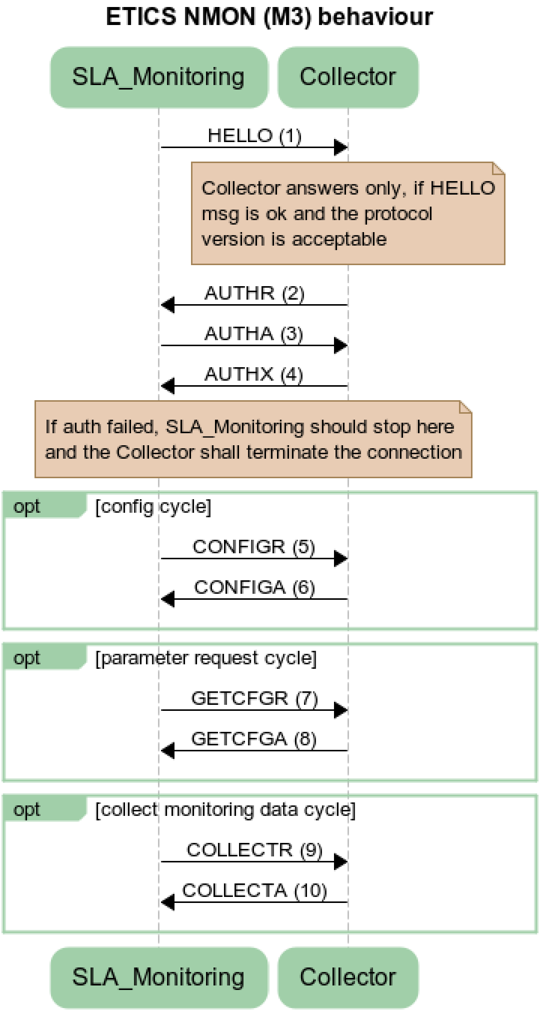

ETICS NMON (M3) behaviour
SLA_Monitoring
Collector
HELLO (1)
Collector answers only, if HELLO msg is ok and the protocol version is acceptable
AUTHR (2)
AUTHA (3)
AUTHX (4)
If auth failed, SLA_Monitoring should stop here and the Collector shall terminate the connection
opt
[config cycle]
CONFIGR (5)
CONFIGA (6)
opt
[parameter request cycle]
GETCFGR (7)
GETCFGA (8)
opt
[collect monitoring data cycle]
COLLECTR (9)
COLLECTA (10)
SLA_Monitoring
Collector

FIGURE 48: MESSAGES: BASIC AND PASSIVE NMON

FIGURE 48 shows the messages for authentication, Probe set-up and requesting passive NMON information:

Initially, the Collector accepts connections from an SLA_monitoring instance to which the Collector only answers, if the HELLO message is well formed and the protocol version is acceptable. Otherwise, the Probe shall not send respond at all. This is a simple, yet effective first protection against scanning attacks: if the Server (Collector) does not reply, the attacker cannot free the resources used for the attack (TCP connection, states) until a certain time-out has elapsed, thus making the attack more difficult and resource consuming.

After successful authentication, other information exchange cycles, like a config cycle or a collect-monitoring-data cycle may be triggered by the SLA_monitoring instance.

During the implementation work in WP5, it became obvious that the probe needs to know about the ASQ paths that have been set-up on the respective connection. Especially in the MPLS-case essential information is not contained in the MPLS header information, as it is available as meta-information of the MPLS control-plane, which must be forwarded to the probe. This shall be done by means of a configuration-cycle.

On the other hand, e.g. after a reboot of the ETICS business plane elements, there must be a way to retrieve the configuration of a probe. This is achieved by means of the parameter-request cycle.

Finally, after a probe has been configured and running for a while, measurements can be done. In the case of passive NMON, this is done by issuing a collect-monitoring-data cycle.

More information about the transported information elements is given in Section 11.6.1.

5.8.4.2.2. Active NMON

AUTHENTICATION

Conversely to the Passive NMON case, triggered by evident security issues it is necessary to establish a authentication procedure between SLA monitoring and the monitoring proxy before any other kind of communication is accepted.

The authentication procedure will likely be based on public key cryptography. Each SLA monitoring instance (a “prover”) will have a pair of keys: a public key which is openly accessible on the network and a private key which should always be kept secret (integrity may be corrupted by attackers).

There exists a relation between the public and the private keys. This relation is based on a one way function or equivalently on a mathematical problem which is known to be NP-complete. This means that it is practically impossible for an intruder to guess the value of the private key from the value of the public one. Classical examples of one way functions are based on the mathematical problem of the discrete logarithm or the factorization of a big integer as a product of two prime integers.

For authenticating itself the instance of SLA monitoring (the “prover”) will have to convince the monitoring proxy (the “verifier”) that SLA monitoring knows the value of the private key that corresponds to the public one.

In public cryptography authentication methods are known as zero knowledge protocols. A zero knowledge protocol makes it possible for a "prover" to convince a "verifier" that it knows a secret (the private key) without revealing that secret.

Whatever the one way function the protocol is based on, any zero knowledge protocol can be summarized as a similar succession of steps that we are going to summarize:

(i) First of all the prover selects a random mask and keeps it secret but makes a commitment on the value of this mask by transmitting to the verifier the image of the random mask through the one-way function. Doing so it becomes impossible for the prover to change the random mask afterwards.

(ii) The verifier then selects a random bit (0 or 1) and communicates its value to the prover. This is a way of challenging the prover.

(iii) Depending on the value of the random bit the prover has to communicate either the value of the random mask of step (i) or a combination of this random mask and the private key to the verifier. The verifier then checks that the prover does not lie by computing the image of the value he has just received through the one way function and comparing it either to the value received at step (i) or to a combination of this value and of the public key.

As we have just explained the authentication of SLA monitoring can be summarized as a succession of three steps as it is displayed on FIGURE 48. This is true whatever the specific public key cryptography authentication protocol selected. In order not to duplicate work the authentication procedure will be similar for active NMON and for passive NMON.

INFORMATION NECESSARY TO PERFORM THE MEASUREMENTS

Active NMON is based on active probing. This means that some traffic is injected into the network for the sake of monitoring the Quality of Service that this traffic receives. Traffic injection is performed in such a manner that the QoS received by probe packets is representative of the QoS of the traffic of interest (ASQ good). The QoS received by probe packets is compared by the monitoring collector to the technical parameters of the SLA.

Active NMON makes it possible to check the conformity of the SLA with respect to three different metrics which are: the delay (both end-to-end and hop-by-hop), the probability of packet losses, and the end-to-end available bandwidth.

All information that is necessary to perform the measurements is provided by SLA monitoring to the monitoring collector/controller. At least the following information must be provided:

- For each involved NSP:

    - NSP unique identifier
    - Text description (optional)

- For each SLA:

    - SLA unique identifier

- SLA valid dates
  - beginning date (optional)
  - end date
- SLA path
  - Ordered list of ASBRs on the SLA path, for each ASBR
    - IP address of the node
    - IP address of the measurement equipment corresponding to the node
    - NSP identifier of the NSP to which the node belongs
    - Role [terminal/transit] (Optional)
- traffic identification methods used at traffic delivery points
- QoS parameters
  - end-to-end maximum delay
  - maximum delay per hop
  - maximum probability of losses
  - minimum bandwidth requirement

5.8.4.2.3. OAMMON

In this sub-section, the interface between the OAMMON block and the SLA monitoring block of the ETICS-CORE ARCHITECTURE (OAMMON-IF) is described, as already shown in FIGURE 47.

It is responsible for the resources discovery, the processing and continuous monitoring of SLA resources. Although the exact implementation will depend on the devices, we propose here a set of message and workflow defining the abstract messages to enable the OAMMON to operate. The main functionalities offered by the OAMMON-IF are the following:

- **Synchronization**: It is used to retrieve the availability of the SLA objects and resources.
- **Process**: allows the configuration of OAM resources through the mechanisms provided by the OAMMON.
- **Monitoring**: Exchange of monitoring parameters related to failures and performances of the SLA resources.

The table below explains the messages associated with each of the functionalities.

Functionality	Message	Direction	Description
Synchronization	SLA resource query Event notification	OAMMON ←→ ETICS-CORE-ARCH	Request to retrieve info about SLA resources (SLA resource discovery). Event notifications from ETICS-CORE-ARCH to OAMMON are also considered

Process	Configuration command for available SLA resources	OAMMON → ETICS-CORE-ARCH	Request for committing available SLA resources
	SLA resources commissioning notification	OAMMON ← ETICS-CORE-ARCH	Notification of commit resources request
Monitoring	SLA resource status query	OAMMON → ETICS-CORE-ARCH	Request to retrieve monitoring information about the status of a SLA resource
	SLA resource status notification	OAMMON ← ETICS-CORE-ARCH	Asynchronous notification of a failure in the SLA resource

The OAMMON interface considers two kinds of workflows to implement its supported basic functionalities: common workflows and specific workflows.

Common workflow:

The common workflow shows details about the setup of the OAMMON manager and the ETICS-CORE-ARCHITECTURE agent connection for the session and authentication management. Initialization and termination functionalities are performed each time a new connection between the OAMMON and ETICS-CORE-ARCHITECTURE interfaces is needed.

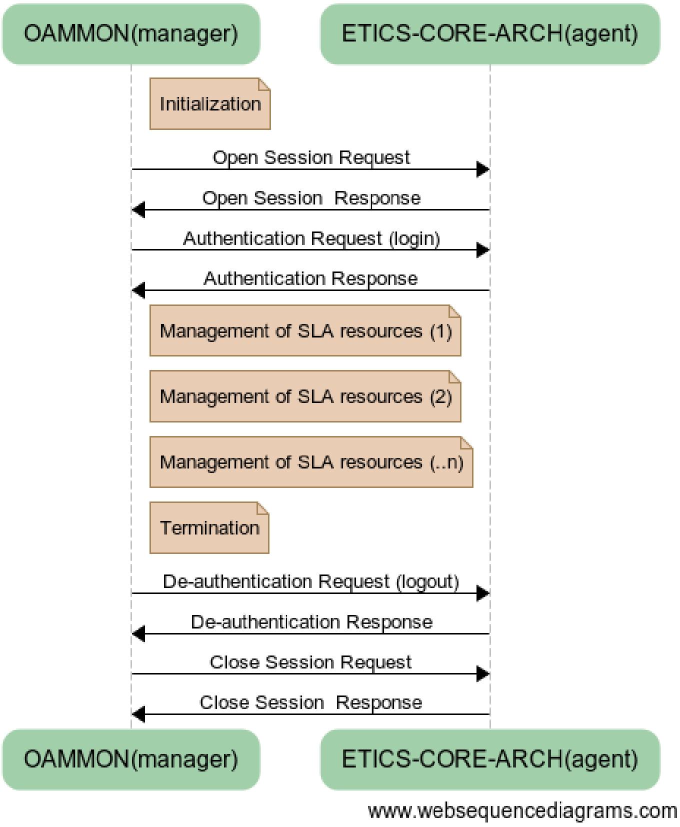

OAMMON(manager)
ETICS-CORE-ARCH(agent)
Initialization
Open Session Request
Open Session Response
Authentication Request (login)
Authentication Response
Management of SLA resources (1)
Management of SLA resources (2)
Management of SLA resources (..n)
Termination
De-authentication Request (logout)
De-authentication Response
Close Session Request
Close Session Response
OAMMON(manager)
ETICS-CORE-ARCH(agent)
www.websequencediagrams.com

FIGURE 49: OAMMON COMMON WORKFLOW

Specific workflow:

The specific workflow shows details about the SLA management between the OAMMON (manager) and the ETICS-CORE-ARCHITECTURE (agent) for the monitoring of the SLA resources. Synchronization, Process and Monitoring functionalities are performed inside the authenticated session described by the common workflow.

With the terms “Generic Request”/“Generic Response” we mean one of the functionalities related to the management of SLA resources (Synchronization request/response, Process request/response, Monitoring request/response).

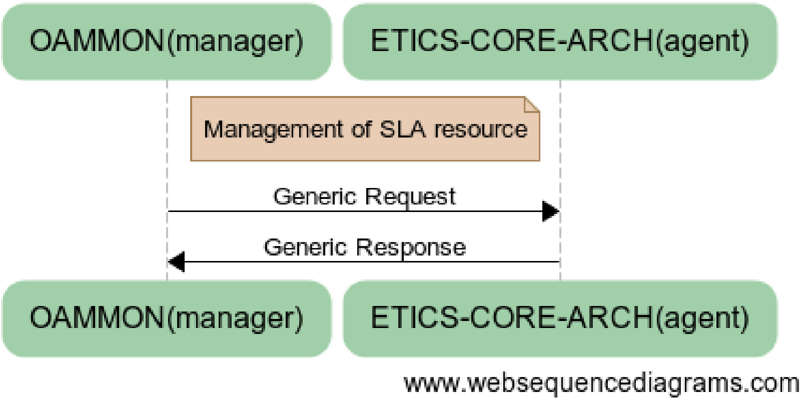


Figure 50: OAMMON Specific workflow

6. EXTENSIONS OF ETICS CORE-SYSTEM ARCHITECTURE

This section is dedicated to extensions to the ETICS core-system architecture. While the previous section has mainly focused on aggregate resource ASQ paths especially for the NSP-to-NSP use case, the present section provides functional extensions to the architecture. These extensions are optional for the deployment of the ETICS architecture and may be used for enhancing the QoS for specific use cases.

As a starting point in Section 6.1 a framework called Service Enhancement Functional Area (SEFA) containing functional extensions for enabling additional services going beyond the initial scope of ETICS, e.g. enabling session services, is given. Such a functional extension is called Service Enhancement Function (SEF).

In a similar light, the Congestion Exposure (ConEx) approach is presented in Section 6.2 as an exemplary Service Enhancement Function (SEF) enabling capacity sharing on aggregate resources. Beyond the indicated capacity sharing usage, congestion pricing may also be envisioned as further ConEx use case, which could well correlate to the economic investigations in [ETICS-D3.3] and [ETICS-D3.5] (congestion pricing, Paris Metro Pricing, willingness-to-pay, etc.).

Beyond that, the business/enterprise service context is specifically targeted in Section 6.3 by investigating the realisation of Virtual Private Networks (VPNs) in the ETICS context.

6.1. SERVICE ENHANCEMENT FUNCTIONAL AREA AND SERVICE ENHANCEMENT FUNCTIONS

Going beyond aggregate-level inter-carrier ASQ paths, Section 5.2 of [ETICS-D4.2] aimed at highlighting how ETICS can more holistically enable end-user ASQ connectivity session handling triggered over the “vertical” *E7* reference point and the subsequent connectivity session handling over the E1 reference point. Moreover, the descriptions in [ETICS-D4.2] included how these capabilities are related to and interact with the management of inter-carrier ASQ paths. As a result, an extension to the ETICS reference architecture was presented that addresses inter-NSP session handling and Service API based interaction with Information Service Providers (InfSP).

In order to provide a more generic extension of the ETICS architecture the so called Service Enhancement Functional Area (SEFA) was introduced in [ETICS-D4.3]. SEFA enriches the basic ETICS architecture and provides the base for individual as well as specific value added functions. This could be for example application service related QoS/QoE monitoring or application specific service quality differentiation in conjunction with application service control functionalities (e.g. media adaption), and services on top of the ASQ paths. Hence, the SEFA architecture can be used to enrich a broad range of ETICS services that are indicated in Section 3.3.

6.1.1. Motivation

Before refining the concepts related to SEFA the motivation for the introduction and reiteration of SEFA in the present deliverable is illustrated by means of a railway metaphor:

- The roadbed, rails and points are figuratively the network with its functionalities.
- The stitching of several railway segments which enables a specific train connection from location A to B are figuratively the ASQ path.
- The passenger or lading which needs to be conveyed / handled is figuratively the application.
- The type of train (which has direct relations to passengers or lading as well as to the rails), e.g. high speed train, subway train, freightliner, etc. is figuratively the Service Enhancement Functional Area (SEFA), i.e. it is the glue which closes the gap / builds the "bridge" between specific (application / network) service demands and an existing network architecture in order to enable a new and/or enhanced service (use case).
- The several components of the train, e.g. wagons, seats, chassis, rails, etc. are figuratively the several Service Enhancement Functions (SEFs), i.e. several SEFs (as part of the SEFA) build the glue between the service and an existing network architecture in order to fulfil the specific service demands and thus, to enable the specific service.

Similar to the chosen scenario, in AQ interconnection the requirements raised by particular use cases may substantially differ. While a train will be designed and build / re-build on the requirements of the convey scenario, the underlying infrastructure (railroads and also type of train) is more generic, i.e. corresponding to the more generic nature of aggregate resource ASQ paths built on top of an existing infrastructure. Consequently, the intention of the SEFA and SEFs is to derive use case specific requirements of all involved actors as well as to describe the additional needed functionalities of all involved entities in relation to an existing architecture. Such functionalities could be in the fashion of, e.g. application service related QoS/QoE monitoring entities and/or derivation of context information in order to enhance application service quality delivery to customers in conjunction with e.g. application service control functionalities, such as media adaptation. Moreover, such functionalities could be some kind of use case related policy and accounting entities in order to enable e.g. charging and/or business models. In the context of developing a new ecosystem, such as ETICS, the description of use case specific SEFs could help to structure the design of the overall ecosystem architecture. Based on the resulting and granular set of specific SEFs an overview of the needed ecosystem capabilities can be provided.

6.1.2. High-level concepts

The ETICS framework described in deliverable D2.2 [ETICS-D2.2] includes multiple actor roles, such as edge network service provider (Edge NSP), transit network service provider (Transit NSP), transport network service provider (Transit NSP), information service provider (InfSP) and business customer (BC), as shown in FIGURE 19.

The NSP actor roles, presented in FIGURE 19, are closely related to the network plane of the ETICS architecture, described in D2.2 and D4.2. The network plane is subdivided in different sub network planes including Network Data Plane, Network Control Plane, Network Management Plane and Network Service and Business Plane. Functions within the network planes are performed by the ETICS NSPs roles for delivering and managing network or connectivity services. Besides the network plane, an application plane exists. The InfSP actor role is closely related to the application plane. The functions in the application plane are performed by the InfSP role for the purpose of delivering and managing application services. However, the application plane is not in the focus of ETICS.

The main focus of ETICS lies on the network plane. More specifically, the ETICS framework is related to establishing aggregate level resource ASQ paths for inter-carrier E2E QoS services. However, besides that ETICS has also identified additional business opportunities by enriching the basic ASQ paths and services in the network and application planes by means of so-called service enhancement functions which can be used to realize added value services that go far beyond the basic ETICS ASQ goods. The intention of ETICS is not to investigate or develop the added value aspects in the whole dimension, rather than to consider and to enable added values by means of the ETICS framework. FIGURE 51 a) presents the location of the service enhancement functional area in relation to the application and network plane in the ETICS architecture.

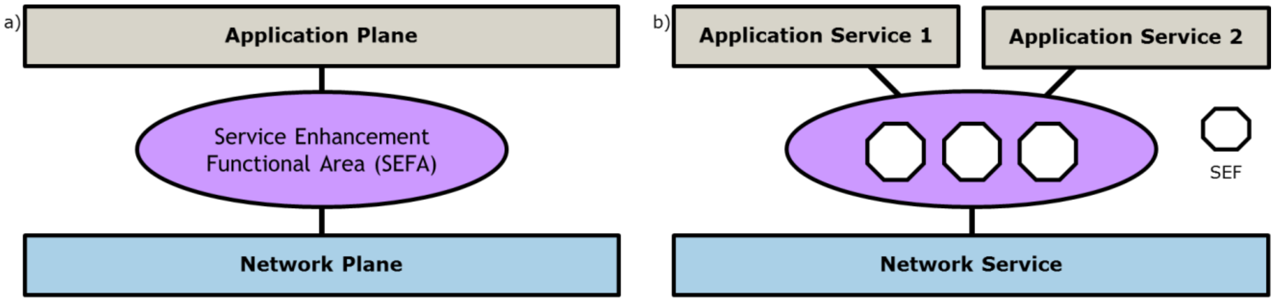


FIGURE 51: LOCATION OF SERVICE ENHANCEMENT FUNCTIONAL AREA IN RELATION TO APPLICATION AND NETWORK PLANE

The Service Enhancement Functional Area can be understood as an abstract "container" of service enhancement functions and is hence an extension of the high level ETICS architecture in order to enable additional functionalities / services. The SEFs interacts between InfSP, customer and ETICS provider (NSPs). In other word, the SEFA defines the place to enrich network plane and application plane functionalities by means of the service enhancement functions (SEFs), as shown in FIGURE 51 b).

In addition to that it can be stated that there is no restriction regarding the functionalities of SEF and their plane where they belong. As a consequence SEFs can also be application functions and / or network service functions. An example for that in the network plane could be for instance resource admission functions, routing function or capacity sharing functions. An application plan SEF could be for instance session handling, session specific accounting and billing functionalities.

6.1.3. SERVICE ENHANCEMENT FUNCTION (SEF)

A specific added value service is represented / instantiated by means of an individual service enhancement function. This SEF gathers and combines information or parameters in order to create the specific added value service or to provide information / parameters for external added value services on top of ETICS architecture or ETICS internal services. The SEF supports both the application services and ETICS services (e.g. ASQ). The generation of added value is based on the interaction between SEF, network, application, device and customer. Moreover, each specific use case / added value service requires an individual SEF implementation and specific parameter sets.

According to the definitions of the data, network control, management planes and network service business plane in [ETICS-D2.2] and [ETICS-D4.2], a high-level definition of the Service Enhancement Functional Area (SEFA) is provided in the following:

- SEFA is an abstract area of the ETICS framework that extends the basic ASQ paths towards specialized added value services.
- These added value services can be used to produce / enrich higher layer application services as well as network services.
- To achieve this purpose the SEFA interacts with other ETICS planes (data, control, network service business and management) in order to gather, trigger, combine and control added value service specific parameters and actions.
- The specific added value service is represented / instantiated by means of an individual Service Enhancement Function (SEF) inside the SEFA.
- SEFA is an abstract functional area that contains all possible SEFs.

In that context the Service Enhancement Function (SEF) can be understood as follows:

- The SEF represents a specific added value service or function realized by means of an individual implementation.
- In many cases the different stakeholders of a certain SEF must have special SEF relationships / agreements (regarding this SEF) that are outside the scope of the ETICS architecture.
- Besides that, it is assumed that the SEF may have value added service specific interfaces to application services.
- In general the SEF consists of three components:
    - Logic: The entirety of functionalities that characterize and form the specific Service Enhancement Function.
    - Interfaces:
        - Internal SEFA interface: Communication between different SEF instances within the SEFA of one or multiple actors.
        - External SEFA interface: Communication between SEF/SEFA instances and information/parameter providing external units or systems, e.g. RACS [ETSI-1] or NASS [ETSI-2].
- Multiple SEFs can interact with each other and / or can be combined to more complex added value services.

There exists a wide range of different use case scenarios being of interest for ETICS (see Section 3.3). Each use case scenario has its own requirements on the network architecture. This means, each use case scenario needs a specific set of functionalities as well as interfaces which interconnect the functionalities in order to setup the specific use case. The intention of the introduced Service Enhancement Functional Area (SEFA) approach is to help describing use case specific services with their required specific functionalities. As such the present section aims at presenting a structural framework, i.e. “cookbook”, for formulating SEFA specifications for particular use cases.

Please, take note that the SEFA approach neither represents a specific use case nor specifies a special set of functionalities. The intention of the SEFA approach is to systematically enable use-case specific (extended) architecture descriptions on top of the ETICS core system architecture, e.g. realising ASQ connectivity based on ASQ paths.

6.1.4. THE SEFA COOKBOOK

The SEF approach describes (network / application) services with focus on a specific use cases and the intention is to extend the ETICS architecture without changing the original ETICS objectives. Moreover, SEF(A) does not influence the implementation nor integration of ASQ related services that are described in Chapter 4.4 before. Subsequently, the notion of layers will be used in order to explicitly distinguish between individual service descriptions. The aim is to help structuring the design of network service and application service architectures built on top of a generic ASQ path-centric ETICS architecture. Whenever setting up a new service, some changes (e.g. introducing new parameters) or extensions in the network architecture have to be done in order process the new service. In the SEFA/SEF approach this new functionality XYZ is represented by means of the specific Service Enhancement Function with an abstract notation (SEF-[XYZ]). As a result, the SEF-[XYZ] can be seen as a complement of the existing architecture.

The proposed SEFA “cookbook” is divided in the following six focal points:

0. Goal of SEFA use case:

In the run-up to the SEF use case description based on the SEFA cookbook the goal of the SEF use case is described from a high level point of view. Moreover, the in involved actors and their roles in this specific use case are described.

1. Prerequisites of SEFA design:

The SEFA approach takes a given network architecture with its implemented or theoretical specified functionalities as prerequisites. Such network architecture could be, e.g. the ETICS core architecture or a deployed and existing network architecture of a network operator (*Note: The SEF / SEFA principle can be applied to any existing network architecture which are described in, e.g. TMF/IPsphere, RACS, NASS frameworks.*). In the context of ETICS, the ASQ paths will serve as “base layer” for the services established on top of them, i.e. use case specific SEFs being realised as specific SEF-layer on top of ASQ paths.

2. Naming of SEF-layer:

Each SEF-layer has its own name depending on the specific use to be realised. For example, if the aim is to setup a service, such as the “Graceful Denial of Service” (GDoS) (for details we kindly refer to [ETICS-D4.2]) the name of the SEF-layer will be SEF-“Graceful Denial of Service” (SEF-GDoS). This step is only naming of the SEF-layers but this step helps to distinguish between use cases when describing the requirements of the different use cases. Thus, in the case of realising several use case scenarios there will be several SEF-layers. Each SEF-layer is independent from each other and is related to the base-layer. The benefit of this layered approach is that mixing up of requirements and / or functionalities of multiple use cases will be avoided.

3. Requirements on SEF-layer:

The use case specific requirements on network service functionalities or application service functionalities are described in this paragraph. The requirements can comprise different topics, e.g. traffic identification, accounting, session handling, application information, network capabilities, customer and/or policy information and so on. Rather than providing a technical solution in this paragraph the needed capabilities are described.

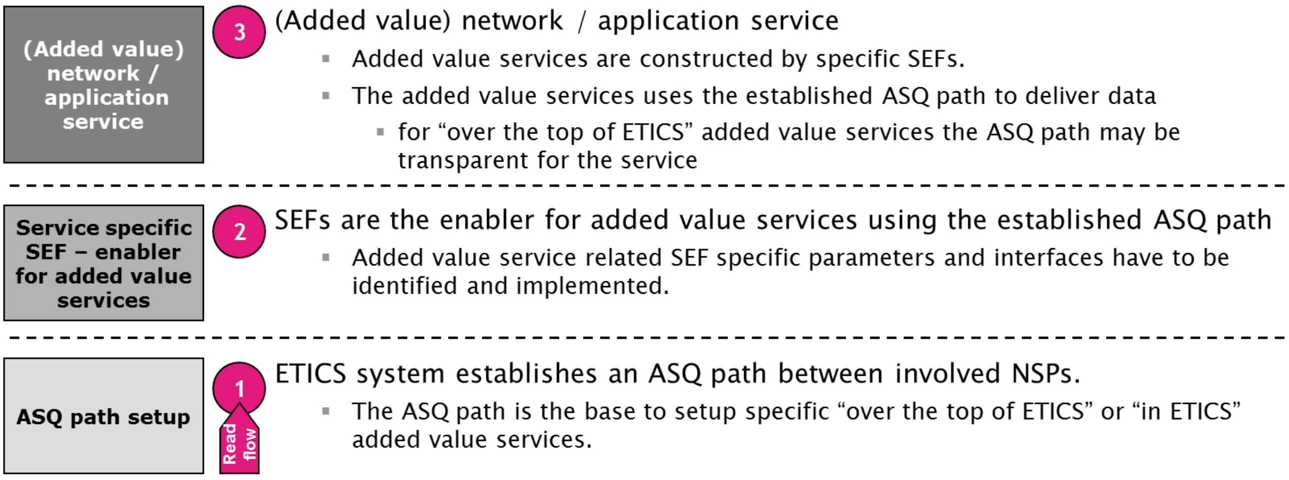
4. Description of SEF-layer:

The “Description of SEF-layer” is divided into two phases. In the first phase SEF-layer internal high level functional elements are derived from the use case specific requirements described in “Focal point – Requirements on SEF-layer”. In the second phase the derived high level functional elements are mapped to the base-layer in order to specify the functionalities of the specific SEFs. This mapping process minds the capabilities of the base-layer without interfering with the architecture or functionalities of the base-layer. Only interfaces to transmit parameters between SEF-layer and the base-layer are specified. As a result, each SEF-layer describes only the additional functionalities which are needed to realise the aimed use case on-top of the base-layer. Several SEFs can be specified in the SEF-layer, which are distinguished through the name of the location where the specific SEF will be utilized and the functionality the SEF will perform. For example, in the SEF-GDoS scenario [D4.3] a SEF which performs accounting on the point of attachment (PoA) of the edge network service provider will be named as SEF-GDoS- Point of Attachment – Accounting (SEF-GDoS-PoA-Accounting).

5. Realisation of SEF-layer:

This paragraph aims at describing how the specific SEFs of a SEF-layer can be realized. For this purpose existing frameworks, e.g. PCC, RACS, NASS or other frameworks will be reviewed. Whenever an existing framework already provides the functionality (or at least parts of it) of a specific SEF, this framework should be foreseen (and used) to realise this specific SEF. Otherwise a new description of functionalities should be given. It is very likely that the design of framework conform SEFs eases the integration of these SEFs in the network architecture of network service provider.

Based on the methodology introduced by the SEFA cookbook the ETICS framework is able to describe new and / or enhanced functionalities and / or services which are not foreseen in the generic ETICS framework that focuses on the aggregate-level ASQ path composition and establishment process. Examples for such added value services are, e.g. Graceful Denial of Service, managed connectivity service, bidirectional traffic delivery, etc. Due to the SEFA cookbook the setup is divided into three steps. These successive three steps to setup SEF-based (added value) services are shown in FIGURE 52. In step 1 the ETICS system establishes an ASQ path between the involved NSPs. This step works identical to non-SEFA-supported use cases. The service specific SEFs enable to setup new and / or enhanced services using the pre-established ASQ path in step 2. The added value service itself is realised in step 3 and uses the underlying ASQ path to transmit the data, while the transmission logic including the functioning of SEFA and the predefined ASQ paths is hidden to the application service and network service. In the case of an external and independent SEF-based service (i.e. a *non-integrated service*) on top of an ASQ path, the ASQ path is transparent for the service. ETICS providers may however also aim at providing SEF-based service being aware of certain ASQ path related parameters (i.e. *integrated services*). This services may provide SEF service related interfaces to the ETICS planes, e.g. to network, control and/or management plane.


(Added value) network / application service
3
Service specific SEF – enabler for added value services
2
ASQ path setup
1
Read flow

(Added value) network / application service

- Added value services are constructed by specific SEFs.
- The added value services uses the established ASQ path to deliver data
  - for "over the top of ETICS" added value services the ASQ path may be transparent for the service

SEFs are the enabler for added value services using the established ASQ path

- Added value service related SEF specific parameters and interfaces have to be identified and implemented.

ETICS system establishes an ASQ path between involved NSPs.

- The ASQ path is the base to setup specific "over the top of ETICS" or "in ETICS" added value services.

FIGURE 52: SUCCESSIVE STEPS FOR SETTING UP SEF-BASED SERVICES

Keeping the SEFA cookbook methodology and FIGURE 52 in mind, a simplified way of illustrating the use case specific SEFs in relation to the (ETICS ASQ path) "base-layer" can be derived. FIGURE 53 shows the coexistence of use case specific SEFs in the ETICS framework. Moreover, in FIGURE 53 is sketched that not each SEF has a direct relationship with the ASQ path (for example SEF-GDoS).

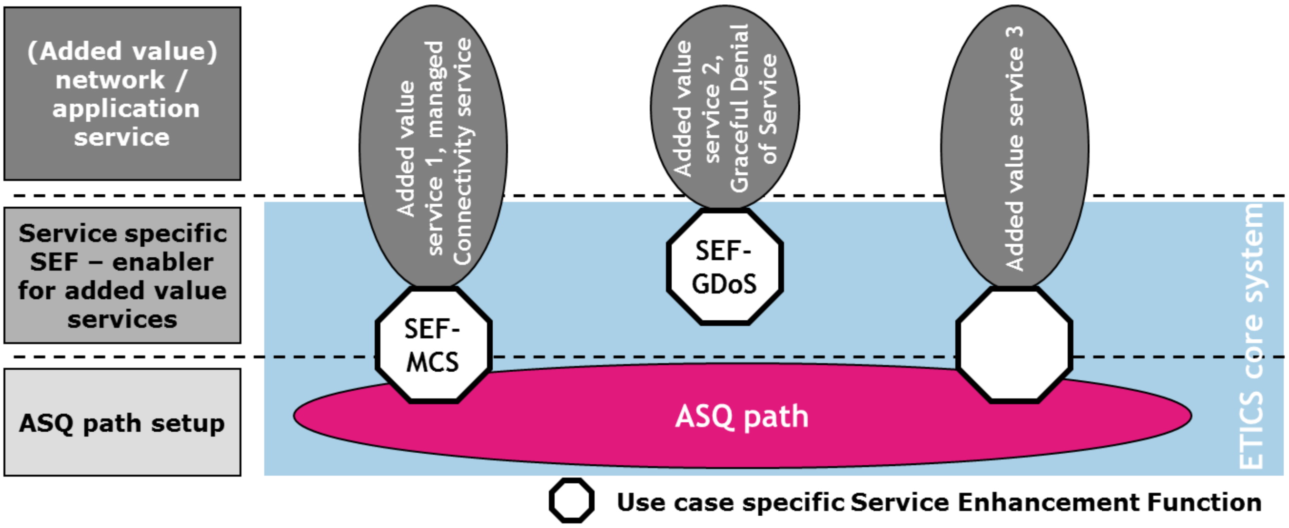

(Added value) network / application service
Service specific SEF – enabler for added value services
ASQ path setup
Added value service 1, managed Connectivity service
Added value service 2, Graceful Denial of Service
Added value service 3
SEF-MCS
SEF-GDoS
ASQ path
ETICS core system
Use case specific Service Enhancement Function

FIGURE 53: COEXISTENCE OF USE CASE SPECIFIC SEFS IN THE ETICS FRAMEWORK

FIGURE 54 illustrates a high level view on several use case specific SEF layers in relation to the ETICS reference model and reference points. Each SEF layer represents a specific SEF use case and hence, a specific (added value) network and / or application service.

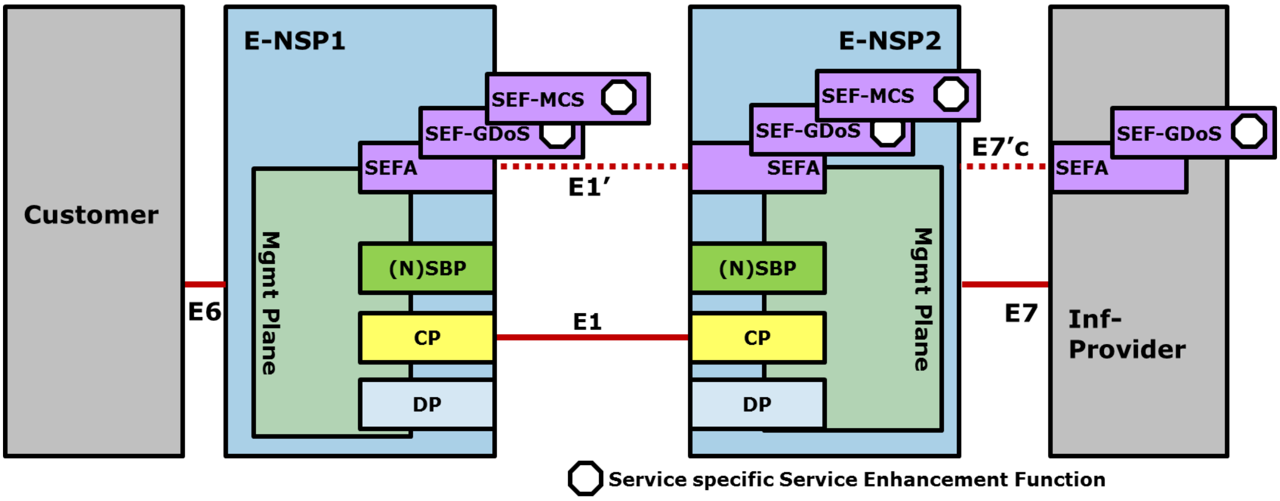


FIGURE 54: SEFA LAYERS AND ITS USE CASE SPECIFIC SEFS IN THE ETICS REFERENCE MODEL AND REFERENCE POINTS

In the following sub-section two SEF use cases will be described.

6.1.5. Enhanced ETICS use cases based on SEFA

By means of the generic ETICS core system architecture and additional functionalities of SEFA new (added value) use cases can be derived. Such use cases can represent an added value for all ETICS actors (content provider, business customer, content delivery network (CDN), network service provider (NSP), and end customer). This SEFA extension can even be seen as potentially beneficial for 3rd party providers such as video streaming platforms.

In the remainder of this section two SEFA based use cases, i.e. the “Graceful Denial of Service (GDoS)” (also see Section 11.7.1) and “Session based connectivity service” (also see Section 11.7.2) will be introduced as examples for additional functionalities to be added to the generic ETICS architecture in order to create new added value services.

These two use cases differentiate between each other regarding the need for interfaces with the underlying ASQ path and the information they request from different actors of the ETICS community:

A) **Non-Integrated SEFA service** – as example “Graceful Denial of Service (GDoS):

For this kind of SEFA based service no direct interaction with the ASQ path establishing entities may be needed. The content provider assumes that there is some kind of abstract network transport that is able to deliver the content to the residential customer in an acceptable quality, without having any information regarding the underlying ASQ path and its parameters. For more details, we kindly refer to the annex Section 11.7.1.

B) **Integrated SEFA service** – as example “managed connectivity service”:

For this kind of SEFA based service the service establishing party has to have some knowledge about (parameters of) a dedicated ASQ path that may be already pre-established or signalled. For example it may be realistic to assume that providers pre-establish an ASQ path based on demand forecasts and afterwards use SEFA in order to utilise this link in the end customer business. The notion “integrated” has been chosen in order to underline the close cooperation between the ASQ path creating instances and the SEFs that is needed to derive service specific ASQ path information in order to construct the specific added value service. For more details, we kindly refer to the annex Section 11.7.2.

This kind of SEFA based session-connectivity service can also be used in the context of VPN services where the VPN end-side requests information regarding available bandwidth and/or VPN parameters from the network service provider instance (e.g. PCE).

6.2. CAPACITY SHARING BASED ON CONGESTION EXPOSURE (CONEX)

Congestion Exposure (ConEx) is an Internet Protocol (IP) extension that is currently under standardization by the Internet Engineering Task Force (IETF). With ConEx, information on the congestion level of an end-to-end network path is provided by the sender for further usage in the network. This information can be used for e.g. traffic management, accounting or enforcing resource sharing policies. Keeping the extended ETICS architecture including the SEFA approach in mind, ConEx can be described as a dedicated SEF use case for capacity sharing that provides an additional control structure within one ASQ traffic service. In order to be harmonised with the naming of SEF-layers the SEF layer providing ConEx functionality will in the following be named as Service Enhancement Function – Congestion Exposure (SEF-ConEx).

The SEF-ConEx is applicable for services with ConEx-enabled end nodes but can also improve overall service experience with significant amounts of legacy nodes. There are many cases where several end systems or users share a link to a common destination with varying capacity demands, see section 6.2.3. If this shared links becomes a bottleneck, the distribution of this scarce capacity amongst the contributing end systems can become critical for user experience and SLA conformance. SEF-ConEx provides the ability to police capacity sharing according to the service provider’s policies. The (SEF-ConEx) policing functionality needs to be provided by network elements upstream of any limiting bottleneck, so the SEF-ConEx can be instantiated by the edge NSP alone. This per-user policing functionality can be parameterized using the interface E7’. To ensure that the end nodes implement ConEx correctly and auditing functionality is proposed with the ConEx protocol (see section 6.2.1). The (SEF-ConEx) auditing elements would be perfectly placed close to the receiver, downstream of all potential bottlenecks. If the bottleneck is most likely in the edge provider’s domain, the audit element can also be provided by the edge provider. Please note that the audit is not necessarily needed for every service instance but it needs to be statistically present in order to effectively discourage infringements.

6.2.1. THE CONGESTION EXPOSURE PRINCIPLE

In ConEx the congestion signal is based on either loss or Explicit Congestion Notification (ECN) [RFC3168]. ECN is a TCP/IP extension that allows intermediate network nodes (i.e. routers) to mark packets in the IP header instead of dropping them in case of congestion. To mark packet at the bottleneck link before the queue overflows, an Active Queue Management Scheme must be implemented in the node like e.g. Random Early Detection (RED) which marks packet with a certain probability depending on the queue length.Therefore, such routers need to implement an Active Queue Management (AQM) mechanism to realize the early marking before the queue overflows und thus packets need to be dropped. In the Transmission Control Protocol (TCP) the receiver will feedback the loss and ECN information in its acknowledgements. Thus using ECN, both end points of a TCP connection have a view on the congestion level of the end-to-end path while intermediate network nodes can just see some of the congestion information, more precisely information regarding bottlenecks upstream on the path. Please note, that the term congestion is always used meaning loss or ECN marks only, and neglects other signs of congestion, such as increased delay (delay is captured by the monitoring approaches, while packet loss is difficult to be measured via monitoring solutions).

In ConEx, the sender feeds this information on *whole-path* congestion back into the network with the next transmission in a connection. Thus this provides an estimation of the expected congestion level on the path to all nodes on the path. Now each node on the path can see the whole path congestion as well as the already experienced congestion of traversed nodes. By subtraction of the experienced congestion from the whole-path congestion, the expected rest-of-path congestion can be calculated as well.

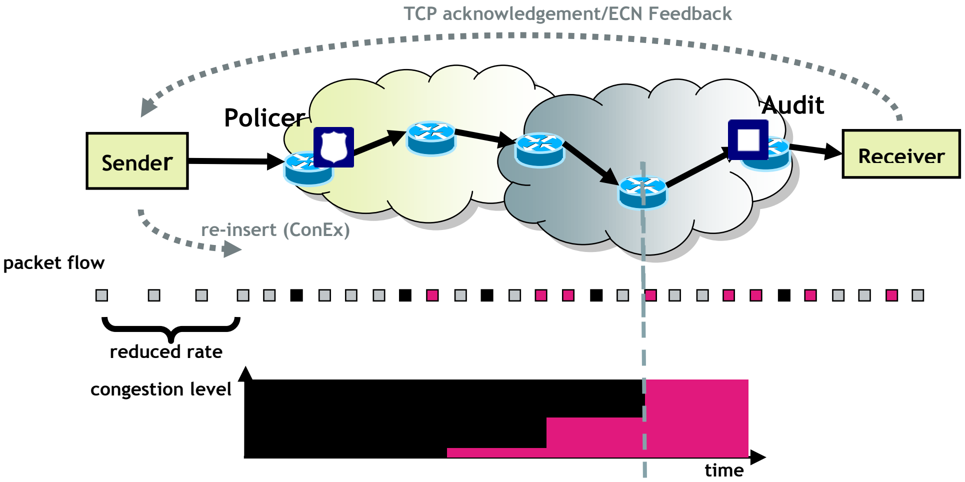


FIGURE 55: CONGESTION EXPOSURE PRINCIPLE

In FIGURE 55 an ECN-capable transmission between a sender and a receiver traversing two interconnected networks is shown. The sender sends packets, which eventually ECN-marked, to the receiver, shown in red, by one or more intermediate networks nodes that are currently facing full queues. The receiver feeds this congestion information back to sender. Based on the information the sender will perform any regular reaction to congestion as it would do anyway e.g. reducing the sending rate based on the implemented TCP congestion control algorithm. Furthermore, the sender will mark one packet with the ConEx re-feedback marking (black) for each congestion feedback signal and thus each ECN-marked packet that was seen at the receiver. The number or percentage of ConEx-marked packets is the same over the whole path and can be seen by all network nodes while the number or percentages of ECN-marked packets can grow over the path with every congested link. Assuming mostly ECN-markings and no packet loss, at the end of the path the number of ConEx- and ECN-marked packets should be the same (with a delay of one Round-Trip Time (RTT) for the ConEx information). Thus at this point of the path, it can be monitored if the sender has sent sufficient ConEX-markings and thus declared the level of congestion honestly to the network. To realize this controlling functionality a separate network device is needed which is called *audit*.

Now information about the current congestion level is available at all nodes along the network path. In [IETF-DR-9] several use cases are discussed. One use case is to introduce traffic management at the network ingress to control congestion later on the network path. Therefore a policer is installed at the first hop of each user. This policer will limit the number of ConEx markings per user and thus provide an incentive to avoid congestion if possible. E.g. background transmissions can easy decrease its sending rate or even stop sending and resume at later point of time in case of congestion. The advantage of such a ConEx-based policer is that limitations are only applied if congestion occurs and thus some kind of policing is needed while other policers, e.g. rate limiters, enforce restrictions even if the network is not fully utilized.

6.2.2. CONEX PROTOCOL MECHANISM

The ConEx protocol will be standardized in the IETF for IPv6. Members of the ETICS consortium largely contributed and will be contributing further in this standardization effort including a Linux-based prototype implementation. [IETF-DR-10]

As the ConEx congestion signal is based on loss and ECN, the number of lost or marked packet or even better the number of lost or marked bytes needs to be provided by the receiver. The acknowledgement mechanism in TCP in general only provides a signal that a data segment is missing by sending duplicate ACKs. It does not tell which further segments are missing or how many. To address this problem and thus avoid spurious retransmissions, the Selective ACKnowledgement (SACK) extension can be used. The SACK mechanism is largely deployed today and e.g. supported by nearly 90% of webservers [KNT12]. A similar problem applies to the ECN feedback mechanism in TCP. This mechanism provides only one feedback signal per RTT. Thus if more than one marking is received within one RTT, this will not be noticed by the sender. Due to the oscillating behaviour of TCP congestion control it is not uncommon that several markings occur in the same RTT. Currently there is some effort in the IETF driven by members of the ETICS consortium to develop a more accurate ECN feedback. Various approaches have been proposed and discussed that either re-use the existing ECN bits in the TCP header or use additional fields (e.g. a TCP option) or unused header space to realize the more accurate ECN feedback [IETF-DR-5, IETF-DR-6, IETF-DR-7].

The ConEx information itself is carried in the IPv6 Destination Option Extension Header. In the ConEx Destination Option (CDO) four bits are defined [IETF-DR-11]. One bit announces if the IP packet is ConEx-enabled while the other three provide the actual congestion signal. Besides the two bits that are needed to reflect the number of lost or ECN-marked bytes separately, there is one more signal – the so-called *credit*. This signal should be used by the sender to provide a (worst-case) estimation of the expected congestion in order to simplify the audit functionality as the ConEx re-feedback signal is delayed by one RTT. In order to allow precise byte-wise congestion accounting, the ConEx marks are multiplied with the size of the packet carrying the marks.

In order to mark the right number of bytes with the respective ConEx re-feedback signal, a ConEx sender has to keep state on how many bytes have been congestion marked or lost [IETF-DR-8]. A network node can monitor the fraction of the ConEx-marked IP packets and their payload length to reconstruct the congestion level over a certain time period.

6.2.3. THE CONEX-BASED CAPACITY SHARING SEF

The ETICS architecture provides Assured Service Quality (ASQ) traffic services to enterprise or other business customers as well as to information services providers (e.g. Communication providers or Over-The -Top providers). In the following scenario we consider a bidirectional service using Best Effort in the

PoI2Region part. This can be a number of residential end users that have an assured bandwidth services to certain content providers' content (that are not directly connected to their access NSP) either in a permanent way or on-demand for a limited amount of time (e.g. watch an HD/3D movie). The service used does not have fixed service requirements per user, e.g. an adaptive and thus variable codec is used, but it is providing an ASQ traffic service for all its users together to avoid disruptions of other services. That means within the ASQ traffic service there are still a number of end users competing on the available capacity. While the content provider does not know how to distribute the available capacity correctly and fairly between its consumers, each consumer does have knowledge on how important a certain transmission is at a certain point of time. Therefore, weighted fair queuing at the ASQ path entry does not provide an appropriate solution. The perceived Quality of Experience (QoE) of the users as a whole would increase if every user would back off less important transissions in case of congestion (that means, if the demand of the content provider's costumers exceeds the capacity provided by the ASQ service). Nevertheless, for each single user there is no incentive to do so. A SEF-ConEx for policing ConEx-based capacity sharing can provide this incentive and enforce non-corporative users not to impair the service of other users.

This incentive is implemented by a per-user ConEx policer, as shown in FIGURE 56, which only reacts in case of congestion. Other than e.g. a rate-limiter, a ConEx policer does also allow any user to consume of the whole capacity of the ASQ traffic service if e.g. no other users are present. Thus ConEx can provide a better user experience within an ASQ service class if different competing parties use the same service. This is an enhancement of the Quality of Experience on top of the ASQ service provided by the ETICS framework.

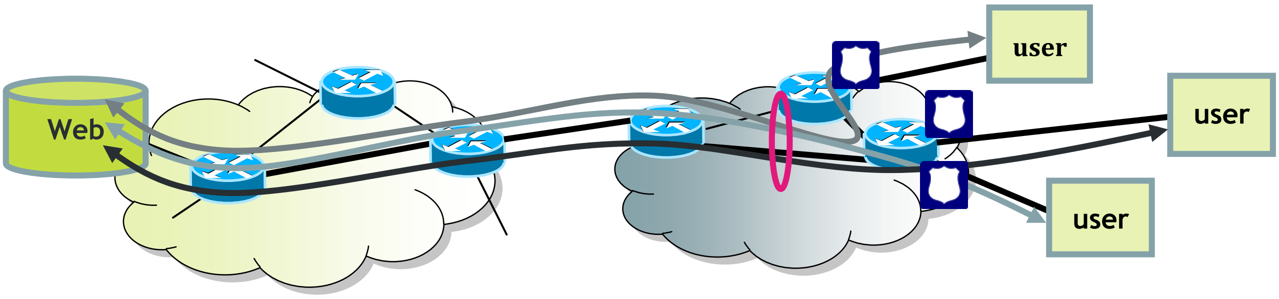


FIGURE 56: SCENARIO FOR BIDIRECTIONAL SERVICE USAGE WITH PER-USER POLICING OF RESIDENTIAL COSTUMERS

The parameterization of the ConEx-based policers implement the providers policies for this service and are crucial for proper functioning of ConEx. These parameters are defined by the Information Service Provider and transmitted to the respective NSP (typically the Edge NSP) via the E7 interface (see FIGURE 19). The exact architecture and parameterization of ConEx-policers are therefore focus of ongoing research in ETICS.

6.2.4. Low Latency Service

A similar scenario can be regarded when having an assured quality interactive communication service to certain destination end-hosts (e.g. HD Videoconference, telepresence, HD audio). For such interactive services it is important to achieve low latency between the communication end points. One problem for latency are large buffers which get filled up frequently by any transmission using a loss-based congestion control algorithm such as used in today's TCP algorithms. Smaller buffers on the other hand would increase the loss rate and/or decrease throughput. To avoid packet loses and maintain small queue fill sizes, ECN can be used. A low marking threshold for the AQM will keep the average queue fill size low but will still be able to buffer small bursts if needed. If small queue fill sizes are a crucial goal, the appropriate reaction on ECN markings has to be enforced. While keeping the necessary state at any potential bottleneck, i.e. at any queue, is expensive, ConEx provides the necessary information at all places in the network. Therefore, for ConEx-enabled traffic ConEx policing can ensure that the end hosts will react properly to the ECN congestion signal and thus do not fill up the queue completely. Today's loss-based congestion control algorithms are not able to utilize a link if the buffer is too small. Furthermore, interactive communication services have different requirements on congestion control like a smoother sending rate or a possibility to interact with adaptive coding mechanisms. The whole field of RTP Media Congestion Avoidance Techniques (rmcat) that support low latency requirements are currently investigated in the IETF and will be further regarded in ETICS.

6.3. ENTERPRISE/BUSINESS CONNECTIVITY: VIRTUAL PRIVATE NETWORK (VPN) SERVICES

In the following, we identify the major issues pertaining the realization of inter-NSP VPNs by the ETICS architecture and describe the architectural elements and functional components required for the provision of such services. We start by providing a high level description of different VPN service use cases of increasing complexity, corresponding to the different stages of the ETICS deployment process i.e., from bootstrapping to full deployment. Then we proceed with a more detailed presentation of the required architectural elements for the realization of these scenarios paying particular attention to the specificities of each service scenario.

6.3.1. VPN USE CASES

6.3.1.1. Use case I (VPN Access)

In the first and simplest use case, we consider the case of a particular Single-NSP services i.e., ETICS bootstrapping scenario involving two Edge-NSPs. As shown in FIGURE 57, a VPN service has already been instantiated by E-NSP A for a certain BC. Initially, all VPN sites are located exclusively at E-NSP A. The target of the ETICS based service is to allow the extension of the existing VPN by further enabling the inclusion of a remote site located at E-NSP B. The remote site may simply be a remote site of the BC, but more complex scenarios can be envisioned, where the remote site belongs to a cloud service provider. In this use case, no PE functionality is provided by E-NSP B; the remote site is to be given remote access to a PE at E-NSP A. In practice, this translates to the establishment of a CE-PE relationship on top of an ASQ path, and therefore it does not constitute an inter-NSP VPN service. It is noted that similar functionality could be achieved by allowing the direct interconnection of the remote CE to the rest of the BC sites at Edge NSP A. Nevertheless, this would change the already established VPN model by necessitating the establishment of multiple direct interconnection paths thus limiting the flexibility of establishing various VPN topologies (e.g., hub-and-spoke, see also Section 11.2) while also increasing the complexity of the overall architecture incurring scalability concerns.

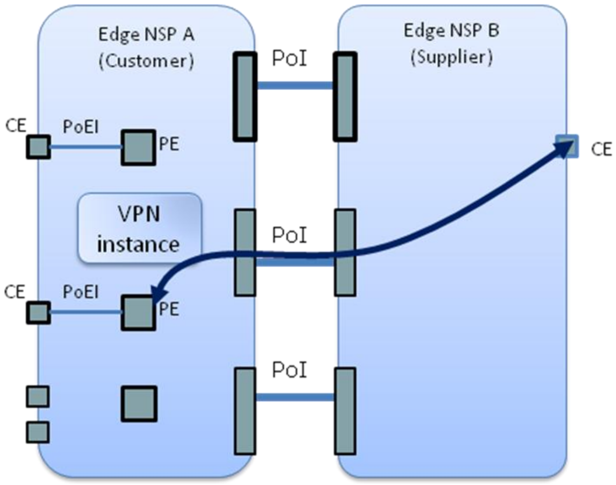


Figure 57: VPN service – Use case I

6.3.1.2. Use case II

In a second use case, we consider the use of PE functionality in both edge NSPs. In this case, a remote site CE establishes an ordinary connection with a local (at NSP B) PE, which then handles VPN routing. Basically, the VPN service in this case can be realized by any of the three aforementioned options of RFC 4364.

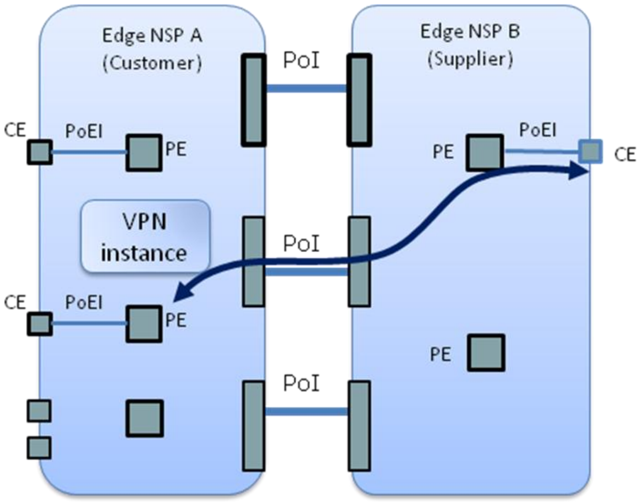


Figure 58: VPN service – Use case II

6.3.1.3. Use case III

In a third use case we again consider cases where VPNs are based on bilateral agreements between Edge-NSPs (i.e., no transit NSPs) but this time more than two Edge-NSPs are involved. As discussed in Section 11.2, these results in the most complicated scenario, where the selection of the VPN topology plays a significant role in the overall utilization of resources, both on a PE level (i.e., VPN routing state) and on a network level (i.e., bandwidth consumption).

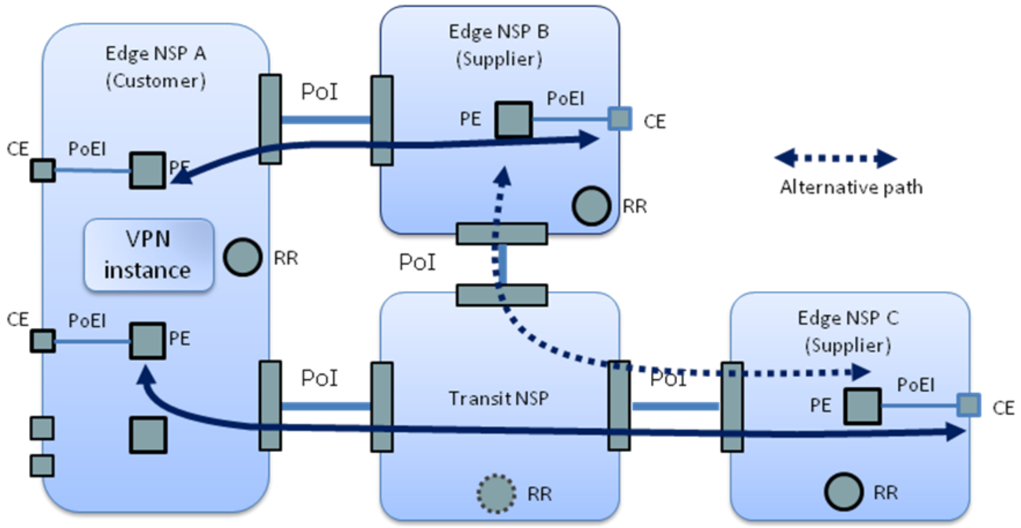


Figure 59: VPN service – Use case III

6.3.2. Traffic aggregation levels

The provision of VPN services necessitates first the assessment of the traffic aggregation levels considered in the context of the ETICS architecture, along with the expected size of the inter-carrier VPN market. Here, traffic aggregation levels refer to service level granularity i.e., the mapping between basic ASQ paths and individual VPN service offerings. As explained in the following, aggregation levels have a direct impact on the scalability and complexity of the entire architecture with respect to VPN service provisioning. This assessment is firstly made regardless of the specific VPN functionality that must be supported (i.e., establishing VPN routing state), and therefore applies to all types of business connectivity services. Then we delve into the details of VPN services.

In the most straightforward case, VPN services can be considered as an extension of the ETICS services with the entire set of service negotiation and establishment procedures being carried out on a per BC level. This would practically entail the inclusion of the PoEIs[17] in the calculation of the exact ASQ paths. From the service establishment point of view, and more specifically with respect to the H-TE technical approach, this inclusion is not expected to induce any overhead in the process of computing the NSP chains as it does not affect the size of the AS level graph. However, the calculation of the exact traffic paths (i.e., the BRPC algorithm) must take place on top of larger topologies exactly because of the inclusion of the PEs. This is expected to lead to an increase of convergence time. Moreover, and most importantly, employing the BRPC algorithm on a per-VPN service request basis would further induce control plane overheads, due to the large number of expected VPN service instances. These overheads relate to the associating signalling functionality and state established at the various networking entities. More specifically, establishing a separate ASQ path per VPN service is expected to present scalability issues with respect to the routing and forwarding functionality. The provisioning of a separate ETICS service per VPN would necessitate the establishment of separate LSPs per VPN service, regardless of the potential path overlaps, thus imposing a significant overhead on the routing and forwarding substrate.

[17] At this point it suffices, for simplicity reasons, to identify PoEIs as PE-CE pairs (see Section 2.6).

Given these scalability concerns we consider an alternative deployment path where VPN services are delivered based on already established ASQ paths. These goods may be based on PoI-to-Region and/or PoI-to-PoI ETICS service primitives. For use case I, we foresee a solution based on PoI-to-Region services. In this case, a PE at the Edge NSP supporting a VPN instance is designated as a member of a PoI participating to a PoI-to-Region service instance. The Region then refers to areas in the other Edge NSP where a remote client site may reside. For use cases II and III we employ PoI-to-PoI services and aim to take advantage of the highly potential existence of common, inter-NSP path segments among multiple VPN service instances (note that in this case the existence of PEs at all interconnected Edge NSPs facilitates the aggregation of multiple VPN instances i.e., a PE typically supports multiple VPN sites). The traffic of each VPN service instance can occupy only a portion of already available aggregate traffic pipes while the associated control plane overheads are amortized across the entire set of VPN service instances.

6.3.3. Mapping VPN Services to Existing ASQ Paths

Aggregating multiple individual VPN service instances imposes the requirement for *mapping* these instances to already established ASQ paths. In turn this poses the following functional requirements:

- *ASQ path discovery*. This requires maintaining some form of state regarding already established ETICS services and can be accomplished at the Service Facilitator level. For each established ASQ path (identified by a unique GSID) the Facilitator maintains the following information:

  (i) *NSP chain*. This information can be maintained in the form of list of IP addresses of ASBRs (i.e. an Explicit Route Object in the PCE terminology) or a list of AS numbers. It is first used to determine whether the ASQ path is suitable for a specific request in terms of interconnected end-points. If the ASQ path terminates at a PE (or set of PEs), then the IP address of that PE can be also maintained for the same reason (see also Section 6.3.4.1). In general, more detailed information can be further included describing the exact service provisioning path. This information can be provided to the Service Facilitator by the local ASVR.

  (ii) *Pricing*. In accordance to Section 6.3.4.1.

  (iii) *QoS*. For the case of bandwidth, information refers to the residual bandwidth available at the ASQ path as required to determine whether the ASQ path can support a requested VPN service instance. This information is available at each ASVR (D5.2) and can be therefore communicated to the Service Facilitator so as to keep track of the end-to-end residual bandwidth of a certain ASQ path. Moreover, currently available techniques for establishing hierarchical LSPs [RFC 4206] further allow the maintenance of this information at the aggregation points (e.g., ASBR^{A}_{1} in FIGURE 60).

- *Admission control.* Having identified an ASQ path for the interconnection of the BC sites, the ETICS architecture should assure that there are enough resources to satisfy the QoS requirements for a specific VPN service instance. This task can be accomplished by the Service Facilitator based on the available QoS information described above. However, this functionality only suffices to verify resource availability across an ASQ path and does not refer to the network segments beyond the ingress/egress nodes of the ASQ path e.g., a path connecting a CE to an egress ASBR. To further extend admission control to this network segment we can utilize the signalling functionality provided by SEFA with the distinction of applying this solution on a per VPN service level and not on a per session level.

- *Traffic identification-level mapping.* Having identified an ASQ path that can support a specific VPN service instance, the ETICS architecture must proceed in a mapping between the packets belonging to the VPN and the existing ASQ path. This is crucial to realize the benefits of traffic aggregation as discussed above. In order to accomplish this, VPN service provisioning shall be based on the routing and forwarding state available for the selected, already established aggregate ASQ path. This is realized by means of tunnelling as shown in FIGURE 60.

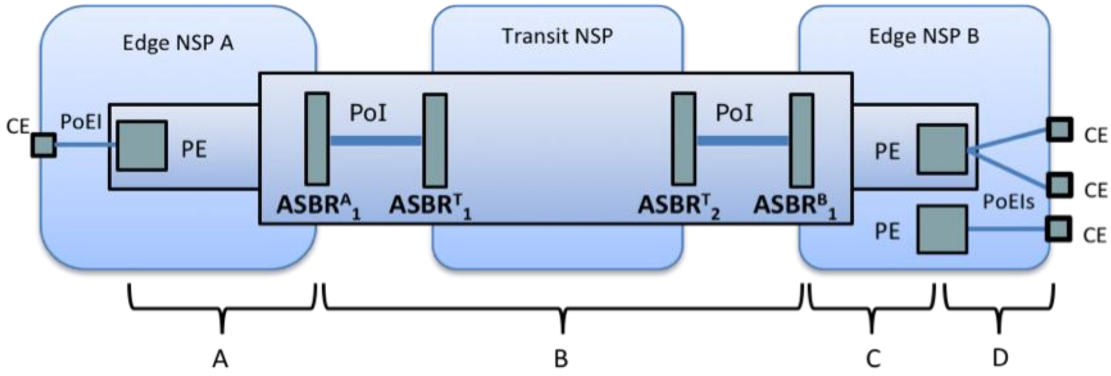


Figure 60: Traffic identification

Targeting at the provisioning of BGP/MPLS L3VPNs, we assume here a homogeneous MPLS-enabled network environment, employing LSPs for the forwarding of traffic from PE to PE. For each established aggregate ASQ path we assume that the required forwarding state has been established so as to transmit a packet from the ingress point to the egress point (segment B in FIGURE 60). Then, the forwarding of traffic between PEs can take place by means of LSP aggregation [RFC 4206]. This translates to RSVP-TE signalling between the interconnected PEs. The Head-end (ingress ASBR) of the already established ASQ path (ASBR$^{A}_{1}$ in the example of FIGURE 60) is responsible for adjusting the value for the available bandwidth of the respective path. At the same time, the respective Path message is neglected in the intermediate ASBRs, therefore avoiding the establishment of per tunnel forwarding state. At the tail-end (egress ASBR) of the ASQ path (ASBR$^{B}_{1}$ in the example of FIGURE 60) the message is again utilized to establish the forwarding state until the egress PE. As a result one forwarding entry per tunnel (i.e., PE-PE pair) is created at the ingress and egress ASBRs of an existing ASQ path.

In the case of Use case I (VPN Access) we further consider a more simplified approach, in which CEs interconnect to the remote PEs using IPSec tunnels. In this case, we also have to enable the mapping between these tunnels and the aggregate ASQ paths. In the Point-to-Region direction The solution to this issue is based on the overall solution provided in the domain of traffic identification for PoI-to-Region services and is currently shaping.

6.3.4. SERVICE PLANE FUNCTIONALITY

In the following we describe the basic service plane functionality for the establishment of a VPN service instance.

6.3.4.1. Offers representation

Each VPN service offering is represented by three distinct elements:

(i) *Service footprint*. This is similar to the case of Point-to-Region service in that a service offering provides information about the network areas that can be served. For the simplest use case of VPN Access, the PoI-to-Region service primitive can be used to indicate the service footprint. In the case of VPN use cases II and III, the differentiation comes from the fact that the VPN service terminates at a PE. Therefore service footprint refers to potentially established PoEIs corresponding to the CEs and PEs deployed in the network. This results in the requirement for representing both the network locations that can be served, so that a service offering can be assessed in the context of a specific request for a set of CEs to be interconnected, and the specific PEs that shall participate in service provisioning. However, VPN services are by definition based on the interconnection of privately addressed nodes. To overcome this limitation we make the assumption that each potentially interconnected VPN site, as well as each PE, hold a public IP address. Service footprint is then expressed by sets or ranges of public IP addresses that can be served by deployed PEs in an analogy to Point-to-Region services. Portal users can therefore check whether a certain BC site can be served. A certain site location is then mapped to the corresponding PE providing the service. This information is statically defined by Edge-NSPs and it is not displayed in the ETICS portal i.e., it is only used for instantiating the VPN service.

(ii) *QoS guarantees*. This refers to the offered QoS for this service e.g., Bandwidth in Mbits/s, Jitter, availability, etc. In this case of VPN services, this is about Permanently Guaranteed service level (see Section 3.4.1).

(iii) *Price*. Flat rate per QoS vector. This corresponds to the usual monthly recurring charging schemes applied for VPN services. A service initiation (activation) fee can also be imposed for the establishment of the service (see D3.5).

6.3.4.2. Service composition process

In the following we provide the set of steps followed for the establishment of a VPN service, with a particular focus on the specificities of this particular service. We focus on the Distributed Pull scenario.

(1) ETICS portal is populated with VPN service offers/capabilities. Each offer is represented as described in Section 6.3.4.1.

(2) A service request is submitted. This can be performed either by Edge-NSPs acting on behalf of BCs (e.g., Scenario I, Section 6.3.1.1), or by the BC themselves. The request specifies the locations of the sites that are going to be interconnected in the VPN. This is done in the public IP address space as discussed in the previous section.

(3) NSP chain computation. Prior to engaging in this procedure, the indicated IP addresses in the submitted request are used to infer the PEs that are to be interconnected. These PEs shall form the end-points for the VPN service, and therefore are the ones that are to be interconnected by the ETICS architecture. Typically, NSP chain computation shall degenerate to the *ASQ path discovery* process described above. If no existing ASQ path can serve a certain request, then the already available NSP chain computation procedures apply.

(4) Path computation. This procedure starts once an NSP chain has been defined/selected, and it may be obsolete in case detailed path information is already available.

(5) Admission control. As described in 6.3.3, in the case a suitable existing ASQ path is discovered the availability of the requested resources shall be verified.

(6) Traffic shaping. Once exact services paths have been defined and the availability of resources has been verified, the interconnected PE devices are configured with the appropriate traffic policing parameters corresponding to the QoS parameters of the service. For use case I, traffic shaping parameterization is foreseen for remotely interconnected CEs.

(7) Communicating NSP chain to the Route Reflectors (RR) of Edge-NSPs. This step is optional and targets at taking advantage of the knowledge of the NSP chain to limit the signalling overhead for the communication of the VPN routing state i.e., RRs may have established BPG sessions with multiple RRs at different domains, only a subset of which may participate a specific VPN service instance; in this case knowledge of the NSP chain can be used to avoid forwarding routing state to RRs not participating the specific service instance. It is noted that this does not affect the regular use of Route Targets.

(8) Establishing BGP sessions between Edge-NSP RRs and between these RRs and PEs supporting the VPN (when the sessions are not already available due to previous VPN service instances).

(9) Establishing VPN routing state. This includes currently well-established practices regarding the following:

    a. Establishing Route Target values at PEs and RRs.

    b. Advertising VPN routing information. This includes the CE-to-PE, PE-to-RR, RR-to-RR and RR-to-PE interactions.

    Depending on whether step (6) is supported the corresponding BGP Update messages are circulated only among the RRs of the participating NSPs.

6.3.4.3. L2VPN provisioning

The previous paragraphs deal with VPN solutions relying on BGP. An alternate solution to L2VPN based on BGP is L2VPN based on the Label Distribution Protocol (LDP). Provider Edges (PEs) use LDP signalling to establish end-to-end Multi-Segment Pseudowires (MS-PWs). These PWs themselves are tunnelled across each domain in MPLS-TE LSPs that can be setup with MPLS-TE signalling and routing protocols (RSVP-TE and OSPF-TE).

This L2VPN use case benefits from the pull and push model to discover the PEs on the end-to-end ASQ path. The Path Computation Element is the key element in both pull and push models.

In the pull model, the service plane, after the ordering of a L2VPN, queries a hierarchical PCE to select on demand the appropriate domains to render the service according to the target service level agreement (SLA). The entry and exit ASBRs of each domain represent the switching PEs (S-PE).

In the push model, the NSPs declare a priori their offers which are then combined to create end-to-end offers. A contract identifier represents an instance of such an end-to-end offer within the scope of a SLA. Once the PCE server associated with the ingress domain receives a network path computation requests from the service plane, it extracts the contract identifier and queries a server to obtain relevant pieces of information such as the acceptance criteria, the TE parameters, the ASBRs in the local domain, the next PCE server address.

6.3.5. VPN Access Issues

In the case of the *VPN Access* use case, there is the need for CEs to advertise their private routing information to the PE so that reachability by the rest of the VPN sites is established. Considering the fact that in this VPN service scenario CEs and PEs reside in different Edge-NSPs, it is not possible to have a direct, physical interconnection between them. This functionality can therefore be realized with the establishment of BGP session between PEs and remote CEs.

Another issue regarding this service relates to the ability of a PE to determine the ingress attachment circuit for packets sent by different remote. This is required for the selection of the appropriate VRF. The use of the source IP address (CE) is an option here.

6.3.6. Compatibility issues

Of particular importance toward the adoption of the ETICS architecture for the support of inter-NSP VPNs, is the degree of compatibility of the ETICS specific functionality with already deployed protocols and established practices. Potential incompatibility issues would obviously disrupt the currently established VPN service provision landscape obstructing a transition towards an ETICS enabled architecture. The selection of the described used cases, as well as the architectural features presented, target exactly this concern.

On a high level, the selected traffic aggregation levels result in a hierarchal design, where VPN functionality is significantly simplified. Inter-NSP forwarding issues are dealt with by the ETICS architecture, simplifying VPN specific procedures to the level of simply establishing Hierarchical LSPs, i.e., a standardized, non ETICs-specific functionality. On another level, the use of the NSP chain knowledge is only considered an optional feature that allows taking advantage of the ETICS architecture to reduce the associates control plane overheads. Again, the protocols used for the exchange of VPN routing information remain intact.

7. SCALABILITY

The present section collects a series of investigations on the scalability of the ETICS architecture. The section starts by identifying the size of the problem, i.e. the scope that ETICS needs to take into account. On this basis Network Service and Business Plane and Control and Data Plane issues are independently assessed. Subsequently, the proposed monitoring solutions are inspected for their suitability for the defined problem scope. Orthogonally, this section closes with a brief outlook on expected network efficiency gains regarding the capacity dimensioning.

7.1. SIZE OF THE PROBLEM

The scale [ETICS-D5.4] of the problem can be identified by the number of NSPs and their interconnection topology (and/or the number of considered networks or ASes), the size of the considered NSP networks, the scale of the offer dimension (impacts the push and the path selection), the scale of customer ASQ path requests dimension (the ability of NSPs to process ASQ path requests).

A detailed measurement-based study in [ETICS-D5.4] identifies 6500 transit networks among the 40 000 ASes of today's Internet. This gives us an idea of the number of the NSPs that would form ETICS communities (at least for the PoI to PoI offers on which the path computation and the service composition is performed). Among these 6500 networks, 70% of them have a degree of less than 10 and 98% of them of less than 100. The average number of networks crossed by a data connection in the Internet is between 3 and 4. These figures have been used, when possible, by the partners to generate topologies in the simulation studies of [ETICS-D5.4].

In this section we propose an IPv4 Internet Tomography study at the AS level granularity. More precisely we provide measurements about various kinds of Internet "size indicators" related to:

- Autonomous Systems, such as:
    - The number of observed ASes depending of their nature
    - The number of observed Inter AS links
    - The number of observed prefixes originated by ASes
    - An estimation of Customer Cones size of ASes
- Public Internet Exchange Points, such as:
    - The number of identified IXPs
    - The number of observed ASes by an IXP
    - An estimation of the maximum inter-AS links that can be established by ASes in IXPs
    - An estimation of the maximum number of intra-AS links that can be established by AS in IXPs

The motivation of this study is to get an idea of the order of magnitude of the problem size if a solution such as ETICS would be deployed in today's IPv4 Internet. This could help better understanding whether it would scale or not and could provide inputs for the performance and dimensioning studies.

The study is based on:

- BGP routing tables collected in February 2012 over 244 Autonomous Systems by the RouteViews and the RIS projects. These tables enable us to build a map of the IPv4 Internet at the AS granularity and to understand which AS originates from which prefixes. However it must be reminded that such analysis suffers from inherent limitations, namely the topology obtained is known to be incomplete as for instance it is only possible to observe peering links (free-settlement peering) if tables are collected on one of the two peers or on one of their customers. Furthermore, we are at most only able to say whether two ASes are connected or not, but not how many eBGP sessions exist between them.
- Delegated files provided by the five Regional Internet Registries (ARIN, RIPE, APNIC, LACNIC, AFRINIC). These files enable us to assign countries to AS, therefore we can geo locate them in continents, make our measurements continent by continent and identify potential regional trends. It should be noticed that while an AS does not have necessarily a geographical reality (an AS can land in many countries and continents), in the facts many of them do (universities, companies, local ISPs or hosters, etc.)
- A list of public IXPs and their members built upon public information found on the web. We use these information to better understand the ability of AS to inter connect with one another and estimate the number of intra and inter AS links that can be established. Even if this list is not exhaustive, we strongly believe it is quite representative. But it must be kept in mind that we only focus on public IXPs and that we do not take into account private facilities as a solution such as ETICS is the more likely to take place into public facilities such as today's IXPs.

In the further sub sections we present the results of our various measurements.

In ANNEX 11.5.4 the results on AS and link distribution of the Internet are presented based on various measurements. These analyses show that the distribution of ASes is not the same in the continents and most of links are located in EU and North America.

This suggests that ETICS deployment can be started at regional level where such kind of QoS needs and gains can be economically valid. The next paragraph gives an overview and the technical conclusions.

7.1.1. Conclusions on Size of the problem

As a conclusion, *all these results* tend to prove that the order of magnitude of the problem is quite reasonable on average as most ASes have low values for most indicators and because the AS that would correspond to NSPs in ETICS are quite few (less than 10% of all ASes).

Moreover the solution could be regionalized if needed making it possible to consider deploying such a solution as the one proposed by ETICS at a large scale in today's Internet.

However, the results also show that for some ASes among those which would correspond to NSPs, indicators values can be very high, especially customer cone size. Such very specific behaviours must therefore be taken into account in the design of the solution to ensure they can be handled and that they would not prevent the system from scaling.

7.2. NETWORK SERVICE AND BUSINESS PLANE SCALABILITY

According the approach in Section 5.3.1 and 5.3.2 two main groups of processes can be identified: the offers (the push model) and the capabilities (pull) of the ETICS NSPs. Both these groups have been treated in D4.3 chapter 5 from a point of view of dimensioning of ICT infrastructures and protocols. In substance it was identified that the storage areas (pre-compiled offer, ASQ requests), queues at any actor role domain and CPU will scale with the size of the ETICS community, at least with a service and business plane adopting Web REST technologies.

7.2.1. OFFER EXCHANGE (PUSH MODEL)

The scalability analysis has to be developed for the various configurations occurring in the push model (centralized and distributed) and to the specific type of PoI-to-PoI configurations.

1) Scalability of the offer dissemination/publication/storage/update.

2) Scalability of the algorithms for offer computation

Concerning **1)** the NSBP analysis made in D4.3 has identified no scalability issues in the offer process made at REST web technology even if collaboration is done through a portal (fully centralised) or between NSPs in the distributed collaborations. Both NSP customers and offer coordinators should in advance make the dimensioning of IT resources (storage, queue, CPU, etc.) according to the dimension of the ETICS communities.

Concerning **2)** the results are also described in D5.4:

The purpose of the simulations has been to **compare** various algorithms for the key ETICS feature of **NSP path selection** according to the published offers (PULL mode) or capabilities (PUSH mode). These algorithms apply to various modes of the architecture. For the PUSH centralised mode, learning algorithms have been proposed to solve the issue faced by the Service Facilitator when combining the published offers.

The simulations have been conducted for three algorithms for offer computation for NSP networks up to 1000 and offers up to 1000. The results in terms of period to finish the computation are not diverging and such time is in the order of tens of seconds (max. 22 sec) which is acceptable for a SLA phase that is comparable to the provisioning phase for aggregated network services. Despite some differences between push and pull models, these results generally apply to both.

In addition to the above considerations, as we have seen in previous sections, ETICS ASQ paths can be of two types: PoI-to-PoI and PoI-to-region. Thus, it is relevant to analyse for these cases possible scalability issues as follows.

Please take note that the analysis on the SBP on PoI-to-PoI ASQ path offers made in D4.3 has identified no scalability issues in the offer process made at REST web technology. This will not be further reviewed in this deliverable.

7.2.2. On Demand Offer process (PULL)

As explained earlier, in the On Demand Scenario, the NSPs are not publishing offers in advance. Instead, offers are computed according to the SLA request and published to others NSPs. Compared to the PUSH scenario, the number of offers exchanged or notifications of offer updates between NSPs are limited and do not increase too fast when the number of NSPs grows. But, the service composition implies, at least, to propagate the SLA request between selected NSPs, with the aforementioned drawbacks:

- imposing additional crank back mechanisms and more time required to obtain a result from the service composition process,
- or flooding the SLA request to all NSPs at the same time, which consumes more bandwidth and path computation power.

The simulation results in terms of time to finish the computation are not diverging and such time is in the order of tens of seconds (around 20 sec) which is acceptable for an SLA phase that is comparable to the provisioning phase for aggregated network services.

The runtime required by the distributed pull mode is significantly less than under the distributed push mode essentially because of the optimization of the learning algorithm Comb-NC which avoids a full exploration of the NSP topology [ETICS-D5.4].

For the NSP provider, the ETICS service and business plane infrastructure has to be dimensioned as well to survive to the following scalability problems: receiving several pull offer requests from different NSP customers. In this case REST queuing and storage of requests have to be dimensioned and cleaned up periodically.

7.2.3. Scalability of Monitoring subsystems (Service & business plane aspects)

Please refer to the dedicated sections on control and data plane (cf. Section 7.4 - 7.6) as we did not foresee any service and business plane scalability issues for Monitoring.

7.3. CONTROL AND DATA PLANE

7.3.1. On Demand Offer process (PULL) - HTE

Concerning the PULL model, the relevant scalability factor corresponds to the flooding time between all the H-TE ASVRs. Indeed, as using OSPF-TE[18] to convey the Network Capabilities, the stability of the solution depends on the time necessary for all ASVRs in the H-TE area to acquire the knowledge of all NSPs within the ETICS community[19]. The convergence time must be lower than the frequency of Network Capabilities

[18] The study remains valid for IS-IS-TE

[19] We assume that the H-TE area is equal to the ETICS community in term of NSPs invlove in the PULL model

announcements in order to obtain a stable scenario. If the convergence time is for example larger than 1s and the frequency of Network Capabilities update less than 1s, ASVRs will never get a stable vision of the AS topology.

To assess the scalability of the H-TE solution, we simulate the Network Capabilities (i.e. the Link State Advertisement messages in OSPF-H-TE) exchange between ASVRs. For that purpose, we start from the Internet Tomography results to simulate various network topologies, in terms of the number of ASes and consequently the number of ASVR's (we take the hypotheses of 1 NSP = 1 ASVR). The H-TE mechanism just performs an LSDB (Link State Database) synchronisation as no SPF (Shortest Path First) computations are done by the ASVR. SPF computations are only performed on request by the PCE collocated with the ASVR. Thus, the flooding time in the H-TE area could be summarized to the propagation time of the LSA between the ASVRs. But, instead of a standard network topology, the topology formed by the ASVRs has a particularity regarding the links between the ASVRs. In fact, the propagation delays between ASVR's are quite large (between 10 to 100 ms) with respect to the transmission time of the traffic on the link between two routers in a standard network. These delays are due to the fact that ASVR's are connected through dedicated tunnels (IP GRE or IP sec) that spawn across many routers. As a real example, the ETICS testbed where some ASVR's have been deployed, give a range of delays between 20 and 50 ms.

7.3.1.1. Simulation Conditions

We have selected a certain number of hypotheses to run our simulation:

- Topologies are obtained with BRITE [FFF99][MLM01] as topology simulator. BRITE parameters are fixed to correspond to the Internet Tomography analysis in particular the interconnection degree,
- 10 runs with 10 different topologies per point have been used during simulation,
- The distance provided by BRITE between two nodes are used to compute the delay between two ASVR with the formula: distance/10 in ms to achieve a range comprise between 10 and 100 ms,
- Bandwidth for the link connection between ASVR is fixed to 10Mbit/s or 100 Mbit/s,
- Processing time of OSPF message is considered to be negligible.

We have defined two notions regarding the topology:

- The "Slave" router is defined as the utmost one. This corresponds to the OSPF diameter of the considered network,
- The "Master" router is defined as the most connected one. This corresponds to the largest number of neighbours for this given router.

7.3.1.2. Simulation results

First of all, we begin by starting up all ASVRs within the H-TE area randomly in a given interval (Figure 6). Due to some limitation of the NS-3 simulator (mostly in term of memory consumption), we have not been able to simulate a topology with more than 200 ASVRs, i.e. more than 200 ASs.

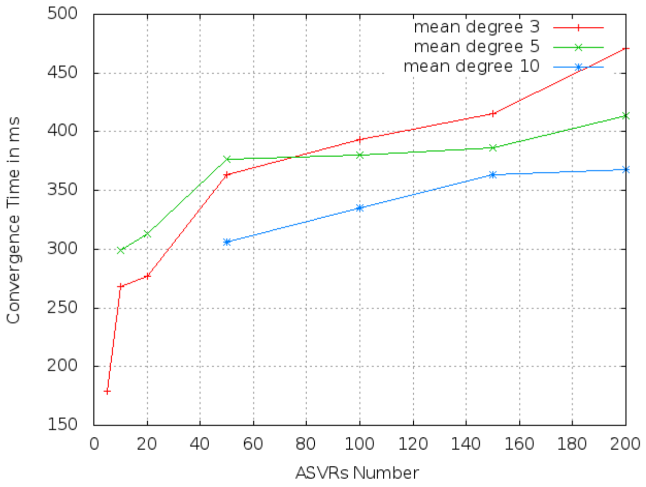


FIGURE 61 : CONVERGENCE TIME WHEN ALL ASVRS ARE STARTED UP IN A GIVEN INTERVAL

Convergence time is always lower than 500 ms for interconnections of degree 3, 5 and 10. For interconnections of degree 10, the simulation must start at a minimum of 50 nodes. Lower values could conduct to a topology too close to a full mesh that it is not suitable and not representative to a real AS topology. We could remark that the convergence time decreases when the interconnection degree increases and that the curves are asymptotic regarding the number of ASVRs.

In the second run of the simulation, we begin with a stable LSDB synchronization between all N-1 ASVRs and start up one ASVR. This corresponds more or less to what happen in real life: you add a new ASVR in an already running H-TE area. This time, we could run simulation up to 500 ASVR i.e. 500 ASes without encountering simulation problem with NS-3 (as opposed to the simulation duration when getting all ASVRs up). The figure 7 gives the results when we start up the “Slave” ASVR (i.e. the utmost one) and Figure 8 gives the results for the “Master” ASVR (i.e. the most connected one).

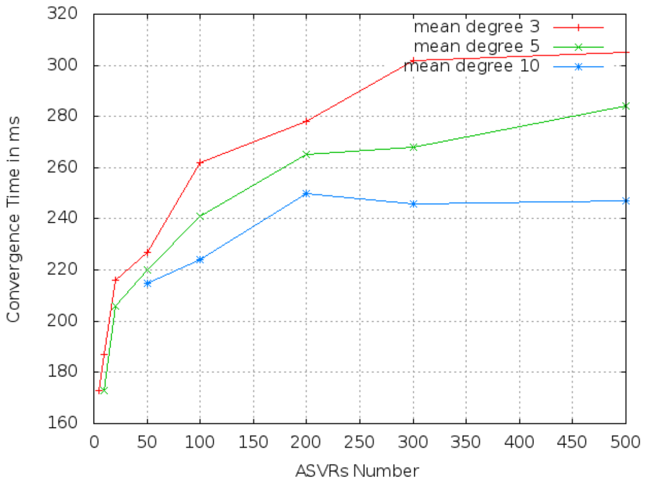


FIGURE 62 : FLOODING TIME WHEN FIRE UP THE “SLAVE” ASVR

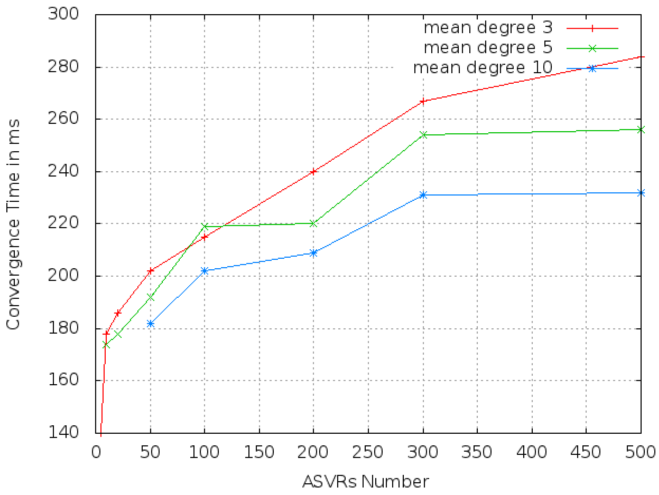


FIGURE 63 : FLOODING TIME WHEN FIRE UP THE "MASTER" ASVR

In both cases, the convergence time is lower than 320 ms disregarding the interconnection degree. Convergence time is also inversely proportional to the interconnection degree. Curves are asymptotic regarding the number of ASVRs. Again, like in the first round of the simulation, for an interconnection degree = 10, the minimum number of ASVRs must be 50.

The last runs concern the advertisement of subnet within the H-TE area. We consider again a stabilized H-TE area where all ASVRs have synchronized their LSDB. The number of messages is directly extrapolated from the customer cone size of the Internet Tomography analysis. Figure 9 gives the results when advertising a range of subnet comprise between 10 and 1000.

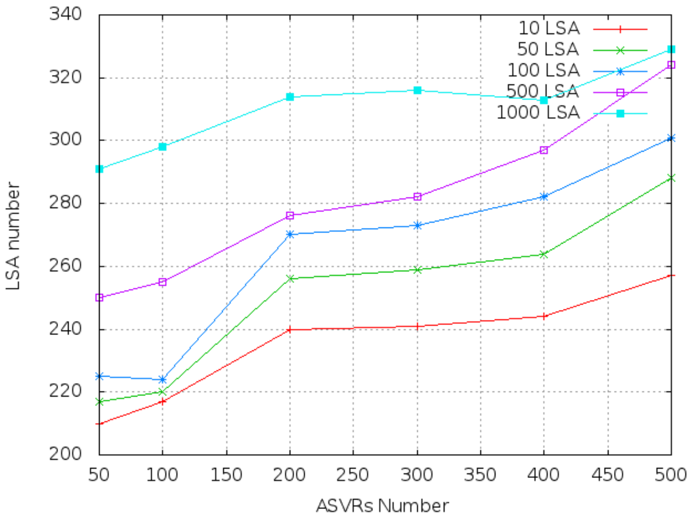


FIGURE 64 : FLOODING TIME OF SUBNET WITHIN THE H-TE AREA

The convergence time remains lower than 350 ms and seems not to follow an exponential curve.

7.3.1.3. Size of messages and Database

To complete the convergence time, we have also computed the size of the messages as well as the size of the database in the worst case to determine the upper limit of the system. In H-TE model, the abstraction algorithm describes the AS topology with pseudo-nodes:

- A nucleus (the ASVR itself),
- All ASBRs which are located at the border of the AS in the Point of Interconnect (PoI)

Interconnected by unidirectional pseudo-links:

- Pseudo-links between the Nucleus and ASBRs,
- Inter-AS links that connect an ASBR to foreign ASs (i.e. the BGP links)

Then, network prefix advertisement must be added to this model. Looking to the particular OSPF implementation of H-TE, the size of Opaque LSA messages is given below. We obtain per ASBR:

- 1 x Inter-ASv2 Opaque LSA = 1 x 216 bytes
- 2 x pseudo-links Opaque LSA = 2 x 216 bytes
- n x Summary LSA = n x 36 bytes

The total size of messages, and thus, the size of the LSDB is given by the following equation:

$$\sum_{i=1}^{n} LSA_{Sum}(i) + \sum_{i=1}^{m} LSA_{interAS}(i) + 2 \times \sum_{i=1}^{m} LSA_{TE}(i)$$

Where LSA_{sum}, $LSA_{inter\text{-}AS}$ and LSA_{TE} correspond respectively to the network prefix announcement, the advertisement of NGP links and the pseudo links between the ASVRs and the ASBRs. N denotes the number of subnets announced by the ASVRs and m the number of ASBRs in the AS. This gives a size of 36n + 648m bytes. For 10 000 subnets and 50 ASBRs, we reach ≈ 400k bytes. We could consider that the size of messages and corresponding LSDB per AS is quite low and completely scalable.

7.3.1.4. Conclusion

The simulation has provided a convergence time under 500ms in the worst condition (large number of ASVRs and low interconnection degree). However, in this configuration the interconnection degree increases proportionally with the number of ASes as observed with the Internet tomography analysis. We could easily achieve an ETICS Community of 250 ASes (i.e. 500 ASVRs with 2 ASVRs per AS for redundancy) with margin of progression regarding the asymptotic evolution of convergence time.

The H-TE database size per ASVR is quite small and increases in 0(n) with the number of ASVRs. Again, for the limit of 500 ASVRs we reach in the worst case ≈ 200 Mbytes LSDB size per ASVR and message volume exchange within the H-TE area. This figure may be strongly decreased as not all ASes will announce 10 000 network prefixes and 50 ASBRs. In particular Edge NSPs are quite small regarding Transit NSPs as observed by the Internet Tomography analysis.

The evolution of convergence time versus the router numbers is asymptotic and not exponential. This is due mostly to the proportional increase of interconnection degree as well as the fixed size of maximum OSPF network diameter (we use the same rectangular area in BRITE to generate the topologies). In fact, the convergence time depends mostly on the interconnection degree and the propagation delay between ASVRs.

As main conclusion, we can affirm that the PULL model is largely scalable regarding the targeted ETICS community. Of course, apply H-TE mechanism to the Internet scale is another challenge that may appear to be outside of the general scope.

7.3.2. ASSESSING THE SCALABILITY OF THE TRAFFIC IDENTIFICATION SOLUTION

The scalability of the traffic identification solution relates mainly to the number of ASQ paths that NSP Border routers need to handle. The more ASQ paths the system is able to handle, the more it is scalable.

ETICS has explored two main identification schemes: original headers identification (Sec. 5.7.1.3.2.1) and the indirect identification (Sec 5.7.1.3.2.2). The original headers identification is less scalable giving the way routers today handle policy based routing. Policy based routing is in fact less scalable than destination-based routing and longest prefix matching. To give an idea about the number of destination routes that current routers handle in destination-based routing, a typical Forwarding Information Base in the Internet DFZ is today around 450,000 lines.

Now for the indirect identification, more investigation is needed in order to assess its scalability. We identified two “theoretical” variants of this identification scheme, the source stacking (Sec. 5.7.1.3.2.2.2) and end-to-end flat labelling (Sec. 5.7.1.3.2.2.1). The first is the most scalable since each border router needs to handle as many identifiers as the number of border routers in its domain (or neighbours). However, it is prone to policy compliance issues, and it imposes a high overhead because of the increased header size due to stacking. The second is the least scalable because each border router will keep one different identifier per ASQ path: it needs therefore to route according to as many identifiers as ASQ paths that go through it, which can be important for a core router. We proposed the possibility of using a hybrid mode in which the labelling is flat and is aggregated whenever possible; we plan to investigate its scalability in the future.

7.4. PASSIVE NMON SUB-SYSTEM SCALABILITY

In the course of the ETICS project, monitoring sub-systems (active- and passive Network Monitoring (NMON) and OAMMON) have been developed. In this section and its sub-sections, scalability considerations thereof are presented. While the focus of Active- and Passive NMON is on inter-provider QoS measurements, the focus of OAMMON is on intra-provider QoS measurements.

These sub-systems have already been described in Section 5.8.3 – “Monitoring Approaches”; please see there for further details.

7.4.1. INTRODUCTION

Even though based on a known concept, the majority of the passive NMON has been newly developed to the better extent. Therefore, Requirements as specified in ETICS D2.2, section 6 have been taken into account to the largest extent possible.

The basic principle of the passive NMON sub-system is based on correlating packet’s hashes observed at different points in the network, which means that the concept is packet-based and as such is extremely resource consuming and must scale with the high bit-rates of modern backbone networks. As a consequence, it is necessary to tune and optimize all parts of the overall system.

Therefore scalability considerations have been taken into account at all levels of the passive NMON sub-system, namely on the

- architectural,
- protocol,
- algorithmic, and at the
- hardware level.

Side note: most of the optimizations have also been translated into the prototype implementation (➔ WP5) in order to be able to draw better conclusions of the system performance and costs in a real-world deployment.

7.4.2. SCALABILITY MEASURES ON THE ARCHITECTURAL LEVEL

Scalability measures on this level are of utmost importance because once instances of the system are deployed, it is extremely difficult (read: cost intensive) to change something on this level. As a consequence, the fundamental architecture has been developed early in the project to allow enough time for reviewing.

The basic idea here is to have a “Proxy Layer” which serves various purposes like: access control, scalability, redundancy, and traffic engineering.

For more details about measures on this level refer to the ANNEX, Section 11.8.1.1.

7.4.3. SCALABILITY MEASURES ON THE PROTOCOL LEVEL

Scalability on the protocol level basically boils down to an efficient implementation, the ability to transfer a huge number of packet data within reasonably sized chunks, and the ability to handle a sufficiently high number of requests in parallel.

Another property, which could also be interpreted as scalability issue, is the extensibility; i.e. the ability to add further information elements to messages or other types of messages to the protocol after it has initially been specified.

Both of these issues are discussed in more detail in the ANNEX, Section 11.8.1.2.

7.4.4. SCALABILITY MEASURES ON THE ALGORITHMIC LEVEL

Some measures taken regarding algorithms are merely “performance” optimizations, which are described in more detail in [ETICS-D5.7] (to be releases, sub-section on “library u_nmon”). However, we can also argue that such measures are – at least to some degree – related to scalability, because they help to scale up the system to a higher number of users / requests per second / etc. ***without*** the need to deploy more hardware, bandwidth, or other resources.

The most prominent optimizations are:

- Use multi-threading
- Introduce Filters
- Optimize the (resource consuming) correlation (algorithm)

These measures are described in more detail in the ANNEX, Section 11.8.1.3.

7.4.5. Scalability measures on the hardware level

While a solution with standard hardware for data rates up to 10 Gbps seems to be absolutely feasible (probably with a hash-filter applied); for even higher rates, support of specialized hardware will be needed. The reason for this is that the packets can only be filtered out (by the hash-filter) AFTER the hash has been calculated. Even though current multi-threaded hardware does not have issues calculating the hash for even higher bit-rates (thanks to our multi-threaded algorithms), it seems that only the process of reading the packets from the line and moving it into user-space already touches hardware limits of standard hardware. This problem might be solved in the future, but a conservative assumption is, that support of specialized hardware is needed. There are several options for this:

1. Specialized capture cards and probes, which are already commercially available (e.g. Endace DAG-cards and probes)
2. The probe functionality could be integrated into high-bandwidth routers.

A detailed analysis of both options is out of scope of the project, but the second option seems quite feasible because the algorithms of the capture-part of the probe are fairly simple and should not be a problem for implementation into hardware.

7.4.6. Summary and Conclusions for the passive NMON sub-system

The most important conclusion we can draw from our analysis and the prototypical implementation of the passive NMON is that passive network monitoring indeed is feasible with the restrictions mentioned for the scalability on the hardware level. For details about the implementation and the performance thereof, please refer to [ETICS-D5.7] (to be released; sub-section on packet capturing).

Our research shows, that the bottleneck of the overall system will be either the network card (driver) or the storage system. While the former problem seems to stem from software implementations and thus is expected to be volatile, the latter one is a more general restriction, but current generation storage systems should already be able to cope with the load.

We conclude that for links with a bandwidth above 1 or 2 Gbps, it is necessary to apply a hash-filter, effectively only using a fraction of packets for the analysis. This will add a small statistical uncertainty, but by applying statistical algorithms, the influence to the absolute value of the measured metrics will be minimal (due to a rather high sample size, worst case one out of eight for 10 Gbps).

An optimized protocol (will be described in [ETICS-D5.8], Section on “EMONIT (M1, M5) Protocol Specification”) helps to reduce bandwidth consumption (savings of 60% and more are possible), reduce CPU utilization, and speeds up monitoring requests and consequently supports scalability.

Overall, we conclude that the developed passive monitoring system can scale up to the bandwidth of currently deployed networks and such which will be deployed in the near future.

7.5. ACTIVE NMON SUB-SYSTEM SCALABILITY

We examine the scalability of active NMON in the terms of Capacity consumption, Memory and computing resources consumption and Deployment cost.

7.5.1. CAPACITY CONSUMPTION:

(4) **Information collection**
For each measurement job, only end-to-end results of performance metrics (delay, bandwidth, loss, ...) are collected from the receiver nodes. This means that the number of messages is proportional to the number of jobs, which is acceptable. However, when SLA violations occur, the collector must probe intermediate nodes for their measured transit times. As violations are supposed to occur rarely, the network may easily support a transitory traffic (the load is proportional to the number of transit domains in the path of the SLA), if the congestion level allows it.

(5) **Control traffic**
At the beginning of each job, the controller transmits job ADD messages to senders, receivers, and capture points. This traffic is supposed to be light-weight and the channels of communication between controller and measurement nodes can be another off line network unused for the QoS traffic.

(6) **Measurement traffic**
Our active measurement algorithms while aiming to have a good accuracy of the measurement of performance metrics, never overloads the network and try to keep the measurement traffic friendly with the data traffic.

7.5.2. MEMORY AND COMPUTING RESOURCES CONSUMPTION

All difficult computation tasks are done centrally at the controller who is supposed to have enough memory and computation power to handle hundreds of jobs in parallel. However, storing intermediate results at capture points allows us to save network capacity while sacrificing some of their storing capacities. In our implementation, we use an intelligent data structure called CMS that allows keeping the memory consumption constant while ensuring a collision-less insert and lookup of measurement information.

7.5.3. MEMORY AND COMPUTING RESOURCES CONSUMPTION

The cost can be estimated as a factor of the number of machines deployed for the purpose of the monitoring:

- Senders and receivers can be any of the existing devices at the edge networks. Their number is dependent of how far (how close to final users) the system will monitor the SLAs.
- The number of capture point machines can be reduced by using only one at each PoP and by applying a port mirroring strategy to monitor traffic transiting at one interface of an ASBR at a given moment and then switch to another interface for another job.

7.6. OAMMON SUB-SYSTEM SCALABILITY (CONTROL AND DATA PLANE ASPECTS)

In the ETICS architecture the scalability of the OAMMON depend on a number of factors:

- Access speed of the OAMMON system to the ETICS core architecture
- Response times of the monitored services
- Monitor poll interval
- Number of monitor threads running concurrently
- Time between monitoring requests
- Number of profiles

The ETICS-CORE-ARCHITECTURE determines the monitoring service response times. In other words the ETICS-CORE-ARCHITECTURE controls concurrent threading and the time between monitoring responses. It is possible to control the scalability and the performance of the whole infrastructure managing the number of profiles and the monitor polling intervals when defining the monitoring application using OAMMON.

Access speed of the OAMMON system to the ETICS core architecture

In LAN monitoring environments, OAMMON and ETICS-CORE-ARCH run over high-speed networks. Therefore the response times of services are low, typically less than one second. The poll interval is also affected by OAMMON data logging functions, which involve a higher number of disk access operations.

Response Times in Remote Monitoring Environments

In remote monitoring environments, where service monitoring run over a WAN or an Internet connection, the response times are less predictable, and it is the network response time, instead of monitor performance, that limits the poll interval. For this reason, to enhance the OAMMON scalability is needed to use polling intervals that allow enough time for responses to be received before the next monitoring request starts.

Monitoring poll interval

The poll interval determines the rate at which the OAMMON performs monitoring requests on a service. It’s possible to define the poll interval using the following formula as a guide:

Minimum poll interval = number of profile elements × average response time / maximum number of threads running concurrently

The number of profile elements configured for a monitor determines the total number of monitoring requests performed by the OAMMON. The average response time varies according to the access speed to the ETICS core architecture, so when the poll interval is selected, it is needed to choose an appropriate value to the application environment. In the worst case, if the application fails to respond then the average response time will be the timeout value set for the monitored element.

7.7. VPN SERVICES

Mapping VPN service instances to established aggregate ASQ paths constitutes a first architectural choice towards a scalable design (see Section 6.3.2). However, still some aspects of the described VPN service provisioning framework affect the scalability of the overall architecture.

The first one relates to the dimensioning of the network with respect to the deployment and use of PEs. A large number of PEs, though increasing the associated CAPEX/OPEX, increases service footprint. However, in the specific context of the ETICS supported VPN services, a large number of PEs is unfavourable in terms of control plane overheads i.e., the number of tunnels established between pairs of PEs (see Section 6.3.3), increases with the number of PEs in the network. These tunnels impose both signalling and state overheads for the entire architecture.

Moreover, scalability concerns are raised in the case of the *VPN Access* service type, for increased numbers of interconnected CEs. The concerns here relate to the number of BGP sessions that must be established and maintained in order to convey VPN routing information across the NSP borders, as well as to the total number of tunnels required for the respective (data plane) traffic.

VPN Gateways

Taking into account these scalability concerns we consider the use of a limited set of PEs within the ETICS architectural context. We term these PEs as *VPN Gateways* to signify their role in the ETICS architecture as the components enabling the VPN routing functionality, acting at the same time as the entry points in the ETICS architecture. VPN Gateways act as aggregators of VPN traffic reducing the associated tunnelling overheads as already discussed.

7.8. NETWORK EFFICIENCY GAINS BY TRAFFIC DIFFERENTIATION

In this section we investigate and quantify the gain in terms of network efficiency or capacity usage when considering various traffic differentiation scenarios as compared with no traffic differentiation. We will first give an introduction being then followed by summary of key findings. Further elaborations can be found in Section 11.8.2.

7.8.1. INTRO AND OBJECTIVES

The analysis is performed by a modelling approach based on the waiting time distribution of an M/G/1 non pre-emptive queue. The analysis does not go into details on relative vs. strict admission control and how this can relate to the different traffic classes deployed.

There are two main settings. One is where the required link capacity is calculated given the load, and the other is where the link capacity is given and the possible load is calculated, both given a set of constraints.

While ETICS is focusing on the inter-carrier aspect it is still relevant to also include investigations of settings that are more oriented toward access, aggregation and backhaul segments of the interconnected networks. It is important to note that while relatively speaking the greatest value of traffic differentiation is typically on these network segments there are several advantages of maintaining an end-to-end QoS traffic class marking, even when using heterogeneous QoS network design approaches in the different segments of the networks. For instance, in the core of a network, a common practice is the mapping/de-mapping operations of QoS classes onto PHB classes. In details, during the mapping stage (ingress), a larger set of QoS traffic classes are mapped onto a fewer number of per-hop-behaviour (PHB) classes. Instead, during the de-mapping stage (egress), PHP classes are re-mapped on QoS traffic classes taking care to not overwrite original class of service. By ensuring end-to-end traffic class semantics new end-to-end business models can be supported.

Recognizing that for some long-distance links there are specific needs for traffic differentiation and protection that perhaps is better handled by tunnel-oriented network technologies, while still maintaining the end-to-end QoS traffic classes for the assured quality IP packets. The future trend toward applying dynamically provisioned virtual links or ASQ paths can imply that it becomes important for the customer NSP to rate-limit at his ASQ path entry point (at the PoI) in order to control the usage of the ASQ path according to the agreed SLA. Here again, it is important to control the service levels for the different types of traffic and the trade-offs between investments, running costs and service level, considering the competitive environment.

In addition to the improved network efficiency and service level control in events of small or moderate congestion, whether transient (seconds) or longer time intervals, the even more important cases are in the event of failures. In these rather rare cases sever congestion may occur and the traffic differentiation approach becomes a very important tool for matching the value of the traffic and the protection level and resources invested for the traffic in the different segments of the network.

It is also recognized that a typical end-customer want traffic differentiation for its own benefit. Consider for instance a DSL line where a typically case today can be that the customer is experience trouble by some traffic interfering negatively with his other traffic. In such cases there can be a benefit to enable, in agreement with the customer, and to deploy access differentiation policies. One use case here can be to temporally demote sustained bitrate flows to a lower priority class (e.g. kids streaming low value content) in order to improve the QoE for interactive traffic (parents connected by VPN to work). A more efficient approach, by avoiding expensive deep packet inspection in the access router or service gateway, is to enable and maintain traffic classification end-to-end across interconnected networks as initiated and controlled by the end-user. Again, we note that the PHB differentiation for the interconnected traffic may be achieved by fewer classes than what is deployed for the access and aggregation.

While the so-called “least effort” (LE) or background (BG) traffic concepts can be deployed based on local edge NSP broadband products and policies we do not expect the BG traffic class as such to be operational for the interconnection and long-haul backbone in the coming years as this would raise many issues related to business models as well as net-neutrality. However, in a longer time horizon there can also be benefits of introducing and supporting cheaper traffic classes also at the inter-carrier level so that this can benefit the customers that would like to utilize a background traffic class for his benefit.

We also observe as the more aggressive deployments of fibre and LTE mobile access progress the strain on the aggregation, metro and backhaul network segments increases. This is partly due to the “all-you-can-eat” nature of TCP. Again, traffic differentiation can be an important tool for the NSPs to have better control of the service levels and QoE, and for facilitating a more gradual network investment approach knowing that the network infrastructure is deployed in an efficient way.

The modelling methodology used in these studies is presented in the Annex Section 11.8.2.1.

7.8.2. MAIN FINDINGS

More detailed key results of this study are presented in the Annex Section 11.8.2.2. In the following a brief summary of results is given:

Our analysis has shown that the delay requirement is a critical and sensitive number that directly impact the needed capacity. By increasing the value of the maximum allowed delay for the traffic class with the strongest delay requirement (and similarly also for the other traffic classes) a significant additional "gain" can also be achieved for the loads below 500 Mbs. Hence, by supporting differentiation one can have a greater freedom in selecting the exact requirements for each class.

Below we summarise and highlight the main findings of the examples that have been analysed:

1. Traffic differentiation will have a positive effect for capacity of real interest;
    a. like in the range 5Mbit/s-500 Mbit/s for typical access links/routers and
    b. 50Mbit/s-5Gbit/s for typical core links/routers.
2. Most of the gain is obtainable with two classes, however additional gain is possible to obtain by introducing more than two traffic classes. The effect of having more than two traffic classes is seen mainly for lower capacity
    a. like less than 50 Mbit/s for access links/routers
    b. and less than 500 Mbit/s for core links/routers.
3. The magnitude of the differentiation capacity gain is significant for a quite broad range of capacity and has its maximum of approximately 1.75 for scenarios I and II, while for scenario III the maximum capacity gain is approximately 1.5.

Allowing more relaxed delay requirements in particular for links below 500 Mbs some additional gain can be achieved.

8. CONCLUSIONS

The present deliverable has proposed a final ETICS solution aiming at providing suitable means for rethinking the present interconnection paradigm at several axes: in particular, ETICS aims at supporting NSPs in regaining sufficient control over IP interconnection traffic flows without conflicting with current deployments in practice. By deliberately focusing on aggregate traffic resources ("(big) pipes") not only scalability problems are addressed, but also flexibility is promoted in the provisioning of quality differentiated services, i.e. Assured Quality (AQ) services as in the focus of ETICS. Such (big) pipes may be seen as fundamental architectural basis for enabling new Business Models in globalized, heterogeneous, and interconnected networks based on better control and a better differentiation between service demands. On top of aggregate traffic resources, the realisation of quality-differentiated, session-based services need to be handled in order to facilitate the utilization of pre-established pipes. Altogether, ETICS is aimed at providing tailored solutions for improving cooperation among and beyond NSPs (e.g. with enterprise customers) capable of providing high quality interconnection services.

By taking heterogeneous needs into account, the ETICS architecture has unified a comprehensive set of concepts trying to address needs relevant to different actors, rollout phases, and technological backgrounds. While the method of choice focuses on interconnected network services with quality guarantees and fully automated processes, the present deliverable also well accommodates for early rollout phases and trials. This has been reflected by aligning the rollout of individual components and capabilities to the time axis, i.e. a roadmap concept drawing an evolution path from minimalistic better than best effort services for the very early phases to full deployments. This is complemented by clearer architectural recommendations regarding individual components.

Subsequent paragraphs will provide a more detailed summary of the stated architectural work:

The present deliverable intentionally focuses on the ETICS service management component with strong links to the Interconnected Networks, Service Enhancement Functional Area (SEFA), Virtual Private Network (VPN), and charging of session & application services as illustrated in the big picture of FIGURE 9. Proposed as tradable AQ Services in overlay to existing network infrastructure, the generic and flexible ASQ paths allows accommodating for a great diversity of different use cases and further extensions. The proposed ETICS architecture is agnostic to AS-specific technological decisions, which is required for reaching a global scope for AQ service demand. In addition, the integration of existing network technologies has been specifically taken into account. This basically relates to ETICS's fundamental assumptions to not exclude connection-oriented or connection-less technologies in ASQ interconnection.

Besides the recognition of interesting use cases for premium session-aware services in interconnected networks, an intentional focus has been laid on "(big) AQ pipes", i.e. aggregate resource ASQ paths, serving as fundamental basis for more fine-grained utilisations. In particular, the present architecture has the capability to support Point-of-Interconnect to Point-of-Interconnect (PoI-to-PoI; and the special case of Point-of-Enterprise Interconnect (PoEI)) - as well as PoI-to-Region services on the basis of generic ASQ paths. By the assistance of SEFA, more light-weight session-aware services may be released on top of PoI-to-Region services, which may positively impact the scalability of the solution relative to other end-to-end ASQ alternatives.

SEFA has specifically been presented as flexible mechanism for realising advanced functionalities on top of generic ASQ paths. Exemplarily, Graceful Denial of Service (GDoS), session-based connectivity service, and Congestion Exposure-based (ConEx) capacity sharing have been illustrated. In this light, ConEx may also be seen as potential mechanism to realise congestion pricing for AQ services in the future.

Offers and network capabilities have to be exchanged in order to establish a particular ASQ path. Offers may be exchanged between customer and supplier on the basis of a “ready to wear” (push) or “made to measure” (pull) paradigm. In both cases, a series of network resources spanning one or more NSPs are traded, while the computation of ASQ paths may be organised in various degrees of centralisation. As a result, the cross-product of architectural options has yielded a multitude of deployment scenarios given in [ETICS-D4.3].

Originating from an extensive set of architectural options, the present work has moreover been able to draw a clearer picture on architectural recommendations – on the one hand regarding alternatives fitting to exogenous circumstances and on the other hand regarding a more detailed working out of technical differences (e.g. SLA lifecycles). Out of a set of possible deployment scenarios, a few have been extracted being regarded to be especially suitable for automation. In a subsequent evaluation, a possible stepwise roadmap of ETICS’s rollout from simple bootstrapping solution to the final roll out of a fully advanced system has been given. Through an investigation of manual, only bilateral, and also PoI-to-Region based rollouts, the bootstrapping phase essentially focuses on simplicity. .

Availing ourselves of ETICS’s governance assumption, the architecture does not restrict which actor can provide the composition of network services, i.e. the construction of ASQ paths provided by network resources from different NSPs. In relationship, also a non-discriminatory solution has been developed which enables the buying, purchasing, and reselling of network resources of every actor in the ecosystem as long as the technical ETICS facilities are deployed. Nevertheless, in the presentation an intentional focus on NSP-to-NSP has been set due to a closer immediate relevance.

Special emphasis has also been lead on developing a monitoring architecture with clear interfaces to the main ETICS architecture. For this purpose, two orthogonal monitoring approaches have been proposed: OAM monitoring and centralised monitoring (in active and passive flavour). While OAM monitoring may be regarded as extension of known intra-domain monitoring mechanisms, the other newly designed monitoring concepts have been specifically tailored to the needs of an inter-NSP context. Nonetheless, both solutions provide essentially required assistance in efficiently validating provisioned interconnection network services against purchased offers from other NSPs. Monitoring may also be seen as valuable tool for improving the monetization of ASQ paths or for establishing an internal quality management for interconnection services. For bootstrapping phases, monitoring solutions may be omitted to keep the required investments low, however this may require a certain level of trust between cooperating partners.

Originating from three main service types relevant for charging, i.e. PoI-to-PoI, reachability (charged for reaching a region, i.e. PoI-to-Destination Region), and tunnel services (e.g. VPN), the final architecture in addition also takes various charging modalities into account (also see [ETICS-D3.5]). While the present architecture essentially builds on directional charging, preferably Sending Party’s Network Pays (SPNP), more advanced Initiating Party's Network Pays (IPNP) mechanisms may be realised on top of it for PoI-to-PoI and reachability services. Tunnel services may in parallel be directly constructed with IPNP due to their bilateral nature. Due to the high importance of tunnel services - especially in the enterprise service context, e.g. regarding premium corporate video conferences - a special work stream has elaborated the realisation of VPN services on top of an existing ASQ path infrastructure. Thus, VPNs may be regarded as economically attractive special case.

While the final ETICS architecture may be very extensive, capable, and flexible, for further work issues or related work context please refer to Section 9.

9. FURTHER WORK

While the collection of architectural work in ETICS culminating in the present deliverable have been able to address a wide range of issues in the field of AQ network interconnection, some features have not been addressed so far. Therefore, the present section will conclude this deliverable by drawing the attention to some untouched fields in the context around ETICS. While some details may be elaborated in the detailed specification given in [ETICS-D5.8], we would like to emphasise that most of the given items for further work go beyond the focus of ETICS. Nevertheless, all of them may be regarded to be important for a successful transition towards a new Internet ecosystem and the associated shift in technologies for cooperating on quality assured Internet services.

Amongst other we subsequently list some important puzzles to solve in the future:

- **PoEI**: The holistic integration of PoEI services in the architecture may require a deeper investigation around required special processes (e.g. regarding billing), and requirements and capabilities of enterprise customers.
- **Congestion pricing**: While congestion pricing may be well known from literature, its realisation in the context of ETICS, e.g. by the help of ConEx, may have to be further analysed.
- **Advanced functionalities / session-services**: In addition to the exemplary illustration of three advanced functionalities being realised by the help of SEFA, more SEFs may be of interest for practical realisation. Moreover, a deeper investigation of technologies for realising session-services on top of ASQ paths may be investigated.
- **Extended passive monitoring**: The implemented passive monitoring sub-system may be extended in the future by adding further functionality, modification or addition of algorithms, and especially the development of hardware support, i.e. the integration of monitoring functionality in the routers. Beyond that, the feasibility of illustrated approaches may further proven by practical deployments. In particular, we would also like to highlight the potential (future) extensions of the NMON monitoring approach discussed in Section 11.6.2.

10. BIBLIOGRAPHY

[AE03] Allman, M., Eddy, W.M., and Ostermann, S., “Estimating loss rates with TCP,” SIGMETRICS Perform. Eval. Rev. 31, 3, 12-24, December 2003.

[BV02] Benko, P., Veres, A., “*A Passive Method for Estimating End-to-End TCP Packet Loss,*” Proc. IEEE Globecom 2002, Taiwan.

[AM06] Amante, S., et al., *Inter-provider QoS MIT Communication Future Programme*, 2006.

[BGPMEA] Huston, G., *BGP Routing Table Analysis Report*, Web site: http://bgp.potaroo.net/.

[Djarallah11] Djarallah, N.B., Le Sauze, N., Pouyllau, H., Lahoud, S., Cousin, B., „*Distributed E2E QoS-Based Path Computation Algorithm over Multiple Interdomain Routes*,“ 3PGCIC 2011: 169-176, 2011.

[DM03] E. Even-dar and Y. Mansour, “*Learning rates for Q-learning,*” Journal of Machine Learning Research, 2003.

[DP09] Djarallah, N.B., Pouyllau, H, “*Algorithms for SLA composition to provide inter-domain services,*” Integrated Network Management 2009: 460-467

[ETICS-1.6] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D1.6 – ETICS technological Roadmap v2*, 2012 (internal).

[ETICS-D2.1] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, Deliverable D2.1 – *Current business models and services; scenarios for the future; high-level requirements – How can the future Internet look like?*, May 2010.

[ETICS-D2.2] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D2.2 – Business and technical requirements for future network architectures*, January 2011.

[ETICS-D2.3] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D2.3 – Business and technical requirements for future network architectures – Final version*, December 2012.

[ETICS-D3.2] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D3.2 – Potential business models analysis and requirements*, January 2011.

[ETICS-D3.3] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D3.3 – Financial/Economic Dynamic Analysis*, December 2011.

[ETICS-D3.4] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, Deliverable D3.4 – Business, legal and socioeconomic impacts, May 2012.

[ETICS-D3.5] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D3.5 – Final business models analysis*, expected 2012.

[ETICS-D4.1] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D4.1 – End-to-End service specification template*, November 2010.

[ETICS-D4.2] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D4.2 – ETICS architecture and functional entities high level design*, June 2011.

[ETICS-D4.3] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D4.3 – Revision of ETICS Architecture and Functional Entities,* January 2012.

[ETICS-D5.2] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.2 – ETICS Draft Detailed specification of the inter-carrier service delivery system*, December 2011.

[ETICS-D5.3] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.3 – First release of selected components for the Inter-Carrier WP5*, 2011.

[ETICS-D5.4] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.4 – Simulative Assessment on ETICS intercarrier Service Delivery solution*, 2011.

[ETICS-D5.6] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.6 – Detailed specification of selected components for the Inter-Carrier Service Delivery for first and second SW releases*, 2012.

[ETICS-D5.7] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.7 – Final release of selected components for the Inter-Carrier Service Delivery, Expected 2013.*

[ETICS-D5.8] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D5.8 – Detailed specification of ETICS components for the Inter- Carrier Service Delivery*, expected 2013.

[ETICS-D6.1] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Deliverable D6.1 – ETICS Testbed Specification and Implementation*, May 2011.

[ETICS-M5.5] Economics and Technologies for Inter-Carrier Services (ETICS), INFSO-ICT-248567, *Milestone M5.5 – Simulators ready*, 2011.

[ETSI-1] European Telecommunications Standards Institute (ETSI) Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN), *Resource and Admission Control Subsystem (RACS): Functional Architecture*, Technical Report ETSI ES 282 003 v3.4.2, 2010.

[ETSI-2] European Telecommunications Standards Institute (ETSI) Telecommunications and *Internet converged Services and Protocols for Advanced Networking (TISPAN): Network attachment subsystem (NASS)*, ETSI ES 282 004 v3.4.1, March 2010.

[FFF99]. Faloutsos, M., Faloutsos, P., and Faloutsos, C., "*On Power_Law Relationships of the Internet Topology,*" ACM Computer Communication Review Cambridge, MA, September 1999.

[HB06] Howarth, M. P. et al., "*End-to-end quality of service provisioning through inter-provider tra-c engineering,*" Computer Communications, 2006.

[HB75] Hilgard, E.R., and Bower, G.H., "*Theories of Learning,*" Prentice-Hall, 1975.

[IETF-DR-1] Internet Engineering Task Force (IETF), Network Working Group, *Advertisement of Multiple Paths in BGP – Internet Draft*, Work in Progress, ed. Walton, D., Chen, E., Retana, A., September 2011.

[IETF-DR-2] Internet Engineering Task Force (IETF), *Network Best Practices for Advertisement of Multiple Paths in BGP – Internet Draft*, Work in Progress, ed. Uttaro, J., Van den Schrieck, V., Francois, P., Fragassi, R., Simpson, A., Mohapatra, P., October 2010.

[IETF-DR-3] Internet Engineering Task Force (IETF), Network Working Group, *North-Bound Distribution of Link-State and TE Information using BGP – Internet Draft*, Work in Progress, ed. Gredler, H., Medved, J., Farrel, A., Previdi, S., September 2011.

[IETF-DR-4] Internet Engineering Task Force (IETF) Draft Y.1731-07, MPLS-TP OAM based on Y.1731, ed. Busi, I., and van Helvoort, H., 2011.

[IETF-DR-5] Internet Engineering Task Force (IETF), TCP Maintenance and Minor Extensions Working Group, *Problem Statement and Requirements for a More Accurate ECN Feedback – Internet Draft*, Work in Progress, ed. M. Kuehlewind, R. Scheffenegger, October 2012.

[IETF-DR-6] Internet Engineering Task Force (IETF), TCP Maintenance and Minor Extensions Working Group, *More Accurate ECN Feedback in TCP – Internet Draft*, Work in Progress, ed. M. Kuehlewind, R. Scheffenegger, July 2012.

[IETF-DR-7] Internet Engineering Task Force (IETF), TCP Maintenance and Minor Extensions Working Group, *Accurate ECN Feedback Option in TCP – Internet Draft*, Work in Progress, ed. M. Kuehlewind, R. Scheffenegger, July 2012.

[IETF-DR-8] Internet Engineering Task Force (IETF), Congestion Exposure Working Group, *Congestion Exposure (ConEx) Concepts and Abstract Mechanism – Internet Draft*, Work in Progress, ed. Mathis, M., Briscoe, B., October 2012.

[IETF-DR-9] Internet Engineering Task Force (IETF), Congestion Exposure Working Group, *ConEx Concepts and Use Cases – Internet Draft*, Work in Progress, ed. Briscoe, B., Woundy, R., Cooper, A., July 2012.

[IETF-DR-10] Internet Engineering Task Force (IETF), Congestion Exposure Working Group, *TCP modifications for Congestion Exposure – Internet Draft*, Work in Progress, ed. Kuehlewind, M., Scheffenegger, R., May 2012.

[IETF-DR-11] Internet Engineering Task Force (IETF), Congestion Exposure Working Group, *IPv6 Destination Option for ConEx – Internet Draft*, Work in Progress, ed. Krishnan, C., M. Kuehlewind, M., Ucendo, C., September 2012.

[IETF-IPPM] Internet Engineering Task Force (IETF), IP Performance Metrics Group, Available: http://datatracker.ietf.org/wg/ippm/

[IMS] 3rd Generation Partnership Project, IP Multimedia Subsystem (IMS), [Online]. Available: http://www.3gpp.org/article/ims

[IMS-2] GSM Association (GSMA), GSMA, IMS Profile for Voice and SMS, IR.92, v3.0, 2012.

[ISO-OSI] International Organization for Standardization (ISO), Information technology – Open Systems Interconnection – Basic Reference Model: The basic model, ISO 7498-1, 1991.

[IPSPH-R1] IPsphere 1.0, IPsphere Framework Technical Specification (Release 1). June 2007.

[IPSPH-TR158] [IPsph-TR158] TR158, *IPsphere Framework, General Requirements and Technical Architecture*, Release 2.0.

[IR.34] GSMA, "IR.34 - Inter-Service Provider IP Backbone Guidelines," 2008.

[Kl76] Kleinrock, L., "*Queueing Systems, Volume II: Computer Applications*," New York, John Wiley & Sons, 1976.

[KLM96] Pack Kaelbling, L., Littman, M. L., and Moore, A.P., "*Reinforcement Learning: A Survey,*" Journal of Artificial Intelligence Research, 1996.

[KNT12] Kühlewind, M., Neuner, S. and Trammell, B., *On the state of ECN and TCP Options on the Internet*, accepted at PAM'13, March 2013.

[MLM01] A.Medina, A.Lakhina, I.Matta and J.Byers, "*BRITE: Boston University Representative Internet Topology Generator,*" Web access: http://cs-pub.bu.edu/brite/index.htm, March 2001.

[MMMR08] Marco Mellia, Michela Meo, Luca Muscariello, and Dario Rossi. 2008. Passive analysis of TCP anomalies. Comput. Netw. 52, 14, pp. 2663-2676 October 2008.

[NeHa] Neginhal, M., Harfoush, K., and Perros, H., "*Measuring Bandwidth Signatures of Network Paths*," In Proceedings of the 6th International IFIP-TC6 Conference on Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet (NETWORKING'07), Springer, 2007.

[PMAM98] Paxson, V., Mahdavi, J., Adams, A., Mathis, M., "*An Architecture for Large-Scale Internet Measurement,*" IEEE Communications, 1998.

[RFC791] Internet Engineering Task Force (IETF), RFC 791, Internet Protocol: DARPA Internet Program – Protocol Specification, 1981.

[RFC2234] Internet Engineering Task Force (IETF), Network Working Group, "*RFC 2234 – Augmented BNF for Syntax Specifications: ABNF,"* ed. Crocker, D., Overell, P., 1997.

[RFC2679] Internet Engineering Task Force (IETF), Advanced Network & Services, *RFC 2679 – A One-way Delay Metric for IPPM*, ed. Almes, G., Kalidinidi, S., and Zekauskas, M.,, September 1999.

[RFC2680] Internet Engineering Task Force (IETF), Advanced Network & Services, *RFC 2680 – A One-way Packet Loss Metric for IPPM*, ed. Almes, G., Kalidinidi, S., and Zekauskas, M., September 1999.

[RFC3168] Internet Engineering Task Force (IETF), Network Working Group, *RFC 3168 – The Addition of Explicit Congestion Notification (ECN) to IP*, ed. Ramakrishnan, K., Floyd, S., and Black, D., September 2001.

[RFC3357] Internet Engineering Task Force (IETF), IP Performance Metrics Working Group, *A One-way Loss Pattern Sample Metrics*, RFC 3357, ed. R. Koodli, R. Ravikanth, August 2008.

[RFC4206] Internet Engineering Task Force (IETF), Network Working Group, RFC 4206 – Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic Engineering (TE), ed. Kompella, K., Rekther, Y., October 2005.

[RFC4271] Internet Engineering Task Force (IETF), Network Working Group, *RFC 4271 – A Border Gateway Protocol 4 (BGP-4)*, ed. Rekhter, Y., Li, T., and Hares, S., January 2006.

[RFC4655] Internet Engineering Task Force (IETF), *RFC 4655 – A Path Computation Element (PCE)-Based Architecture*, Farrel, A., Vasseur, J.-P., Ash, J., August 2006.

[RFC4656] Internet Engineering Task Force (IETF), Network Working Group, *RFC 4656 – A One-way Active Measurement Protocol (OWAMP)*, ed. Shalunov, S., Teitelbaum, B., Karp, A., Boote, J., and Zekauskas, M., September 2006.

[RFC4761] Internet Engineering Task Force (IETF), Advanced Network & Services, *RFC 4761 – Virtual Private LAN Service (VPLS) Using BGP for Auto-Discovery and Signaling*, ed. Kompella, K. and Rekhter, Y., January 2007.

[RFC5101] Internet Engineering Task Force (IETF), Network Working Group, *RFC 5101 – Specification of the IP Flow Information Export (IPFIX) Protocol for the Exchange of IP Traffic Flow Information*, ed. Claise, B, January 2008.

[RFC5136] Internet Engineering Task Force (IETF), Network Working Group, *RFC 5136 – Defining Network Capacity*, ed. Chimento, P., Ishac, J., February 2008.

[RFC5441] Internet Engineering Task Force (IETF), Network Working Group, *RFC 5441 – A Backward-Recursive PCE-Based Computation (BRPC) Procedure to Compute Shortest*

Constrained Inter-Domain Traffic Engineering Label Switched Paths, ed. JP. Vasseur, R. Zhang, N. Bitar and J.L. Le Roux, 2009.

[RFC6049] Internet Engineering Task Force (IETF), *RFC 6049 – Spatial Composition of Metrics,* ed. Morton, A, Stephan, E., January 2011.

[RFC6534] Internet Engineering Task Force (IETF), IP Performance Metrics Working Group, *Loss Episode Metrics for IP Performance Metrics (IPPM),* RFC 6534, ed. N. Duffield, A. Morton, J. Sommers, May 2012.

[RKS06] Rewaskar, S., Kaur, J. and Donelson Smith, F., "*A passive state-machine approach for accurate analysis of TCP out-of-sequence segments,*" SIGCOMM Comput. Commun. Rev. 36, 3 , 51-64, July 2006.

[S00] Savage, S. "*Sting: A tool for measuring one way packet loss,*" in Proceedings of IEEE INFOCOM '00, Tel Aviv, 2000.

[SBDR08] Sommers, J., Barford, P., Duffield, N., and Ron, A., "*A geometric approach to improving active packet loss measurement,*" IEEE/ACM Trans. Netw. 16, 2, 307-320, April 2008.

[SE03] Seitz, N. International Telecommunication Union (ITU), *Recommendation Y.1541 and Y.1221- A Basis for IP Network QoS Control and Traffic Management,* Available: http://www.itu.int/ITU-T/worksem/qos/presentations/qos_1003_s5p1_pres.ppt

[SARSA] Rummery, G.A., Niranjan, M., Technical note, "*On-Line Q-Learning Using Connectionist Systems*".

[SiJa00] Singh, S.P., Jaakkola, T. , Littman, M.L. and Szepesvari, C., „*Convergence Results for Single-Step On-Policy Reinforcement-Learning Algorithms,*" Machine Learning, 38(3):287{308, 2000.

[SPT94] Sastry, P.S., Phansalkar, V.V., and Thathachar, M.A.L, "*Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games With Incomplete Information,*" IEEE Transactions on Systems, Man, and Cybernetics, 1994.

[Ta91] Takagi, H, "*Queueing Analysis, Volume 1: Vacation and Priority Systems, Part 1*," Amsterdam, North-Holland, 1991.

[TJDB06] Tesauro, G., Jong, N. K., Das, R., and Bennani, M. N, "*A hybrid reinforcement learning approach to autonomic resource allocation,*" In ICAC '06: Proceedings of the 2006 IEEE International Conference on Autonomic Computing. IEEE Computer Society, 2006.

[Valancius-tiers] V. Valancius, C. Lumezanu, N. Feamster, R. Johari, and V. V. Vazirani, "*How many tiers? pricing in the internet transit market*," in ACM SIGCOMM, pp. 347–356, 2011.

[WC96] Wang, Z., and Crowcroft, J., "*Quality-of-Service Routing for Supporting Multimedia Applications,*" IEEE Journal on Selected Areas in Communications, vol.14, num. 7, 1996, pages 1228-1234.

[WD92] Watkins, C.J.C.H., and Dayan, P., Technical Note: "*Q-Learning,*" Journal of Machine Learning Research, 1992.

[XB05] Xiao, J., and Boutaba, R., "*QoS-aware service composition and adaptation in autonomic communication,*" IEEE Journal on Selected Areas in Communications, 23, 2005.

[Y1541] International Telecommunication Union (ITU), *Recommendation ITU-T Y.1541* (11/2011) – Network performance objectives for IP-based services, 201.

[Y1731] International Telecommunication Unit (ITU), *ITU-T Y.1731 – OAM functions and mechanisms for Ethernet based networks*.

[3GPP-1] 3rd Generation Partnership Project (3GPP), *Quality of Service (QoS) concept and architecture*, TS 23.107 v11.0.0, ETSI, September 2009

[3GPP-2] 3rd Generation Partnership Project (3GPP), *End-to-end Quality of Service (QoS) concept and architecture*, TS 23.207 v11.0.0, ETSI, September 2009

[3GPP-3] 3rd Generation Partnership Project (3GPP), *IP Multimedia Subsystem (IMS); Stage 2*, TS 23.228, v5, ETSI, 2006.

11. ANNEX

11.1. CHANGES OVER D4.3

Changes:

- Section 2:
    - New section on Virtual Private Networks (VPNs) state-of-the-art and inter-carrier QoS metrics
    - Focus on interrelation of related technologies with ETICS, e.g. interrelating IPX and IMS usages with ETICS, as well as worked out advantages and disadvantages of approaches
- Section 3:
    - New section on definitions, and buyer and supplier scenarios
    - Integration of use cases related to the Service Enhancement Functional Area (SEFA), and more detailed illustration of the required inter-carrier ASQ paths
- Section 4:
    - New visions and goals section presenting the big picture of the ETICS solution
    - Updates integrated from [ETICS-D5.6] regarding basic/automated mode and SLA lifecycle
    - New section on buyer and supplier scenarios associated with the ETICS architecture
    - New overview on rollout roadmap being further detailed in [ETICS-D2.3]
    - Deployment scenarios incorporating updates triggered from [ETICS-D3.5]
- Section 5:
    - New monitoring architecture interfacing with the rest of the system
    - Harmonisation of contents with [ETICS-D5.6]
- Section 6:
    - Revamped presentation of the Service Enhancement Functional Area (SEFA) and Service Enhancement Functions in the form of a cookbook being illustrated with the help of two concrete examples (see details of the examples in the annex)
    - New contributions von VPN services, as well as Congestion Exposure (ConEx).
- Section 7:
    - New section on scalability of individual components
- Annex

- ○ Additional related works on inter-provider QoS metrics & verification, as well as details on VPN topologies
- ○ New Sections on VPN Topologies and inter-provider QoS metrics and verification
- ○ New SEFA use cases illustrated on the basis of the SEFA cookbook
- ○ New contribution on loss measurements & congestion identification
- ○ New complementary material on scalability, and network efficiency and gains
- ○ New ETICS Glossary

11.2. RELATED WORK: VPN TOPOLOGIES

An important aspect of VPN services relates to the definition of VPN topologies. We use the term *VPN topology* to refer to the graph created by connecting the VPN sites that have direct communication. Typical topologies are: (i) *full mesh*, where all sites are directly interconnected to each other, (ii) *hub and spoke*, where a single site (hub) is directly connected to all other sites, and communication between any pair (or set) of sites is established through the hub node. (iii) *multi-hub and spoke*, where multiple nodes act as hubs, enabling the communication between subsets of the sites.

FIGURE 65 shows an example of a hub and spoke topology; with Edge NSP 2 having the role of the hub i.e., the PE serving this VPN instance at Edge NSP 2 concentrates the routing information from all other sites. This has obviously a significant impact on the size of the routing state maintained at each PE, as there is a single PE that is required to maintain all VPN routes. This provides a significant scalability advantage over full mesh topologies. At the same time however, this topology structure has also an effect on the resources required to support the requested QoS. This can be illustrated with the example of Edge NSPs 3 and 5 which can establish a direct E1 relationship. In this case, the corresponding traffic between these domains consumes the resources on the interconnection of E-NSP 2 and E-NSP 3, as well as those on the interconnection of E-NSP 3 and E-NSP 5. Apparently, a careful planning of the topology is required so as to balance the scalability benefits/costs.

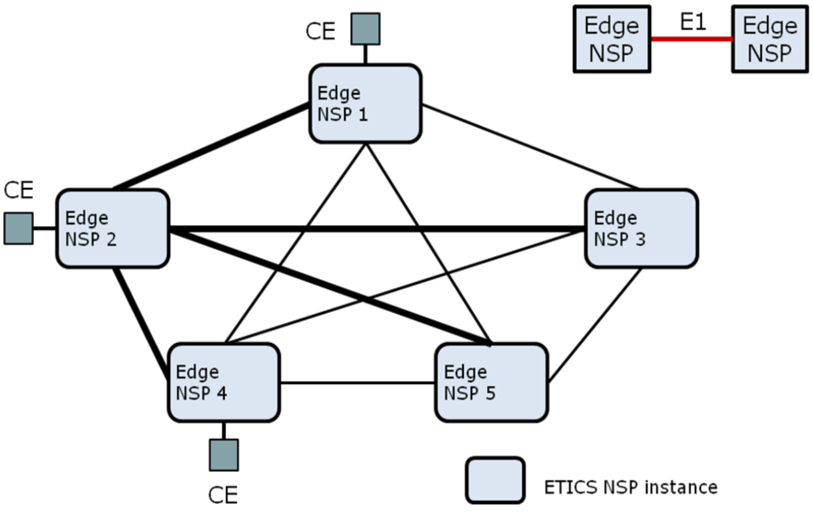


Figure 65: A hub and spoke VPN topology

11.3. RELATED WORK: INTER-PROVIDER QOS METRICS & VERIFICATION

This section discusses the MIT Communication Future Programme recommendations for the definition and verification of performance metrics to support inter-provider QoS [AM06]. This set of guidelines largely leverages on the specifications agreed at both IETF and ITU standards bodies, more specifically those within the IETF IPPM working group [IETF-IPPM] and ITU [Y1541]. More details on measurements' report methodology and effects of routing on measurements can be found in [AM06].

- The ITU-T [Y1541] specifies six classes 0-5 of service, in Table 2.1 below we report them together with their associated QoS parameters given by:
    - IPTD – IP packet transfer delay
    - IPDV – IP packet delay variation or jitter
    - IPLR – IP packet loss rate
    - IPER – IP packet error rate

Class 0 is the class with the most stringent QoS guarantees and Class 5 with the least.

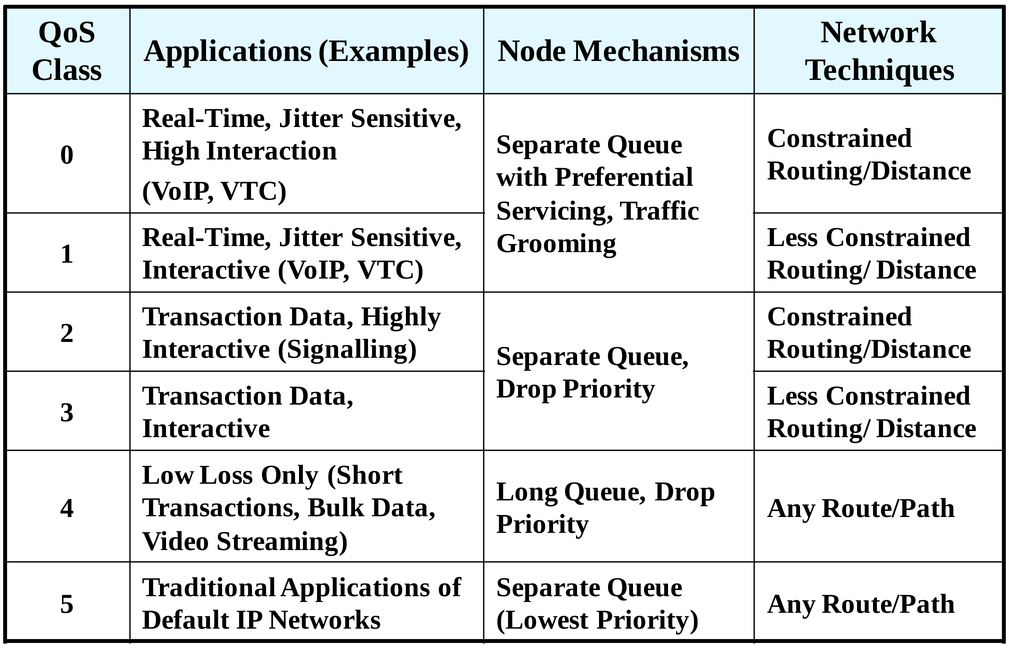
- A possible association between [Y1541] Classes, Diffserv per-hop-behaviour, and Y.1221 transfer capabilities is given in Table 2.2
- Examples of applications that the ITU-T [Y1541] classes can support, together with implications on node and network routing behaviours are summarised in Table 2.3
- Recommendations in [AM06] with regards to the definition of IPTV, its bounds and reporting methodology are given in Table 2.4
- Recommendations in [AM06] with regards to the definition of IPDV, its bounds and reporting methodology are given in Table 2.5
- Recommendations in [AM06] with regards to the definition of IPLR, its bounds and reporting methodology are given in Table 2.6

Network Performance Parameter	Nature of Network Performance Objective	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
IPTD	Upper bound on the mean IPTD	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	Upper bound on the 1-10^{-3} quantile of IPTD minus the minimum IPTD	50 ms	50 ms	U	U	U	U
IPLR	Upper bound on the packet loss probability	$1*10^{-3}$	$1*10^{-3}$	$1*10^{-3}$	$1*10^{-3}$	$1*10^{-3}$	U
IPER	Upper bound	$1*10^{-4}$					U

Table 2.1 [Y1541] recommended classes of service

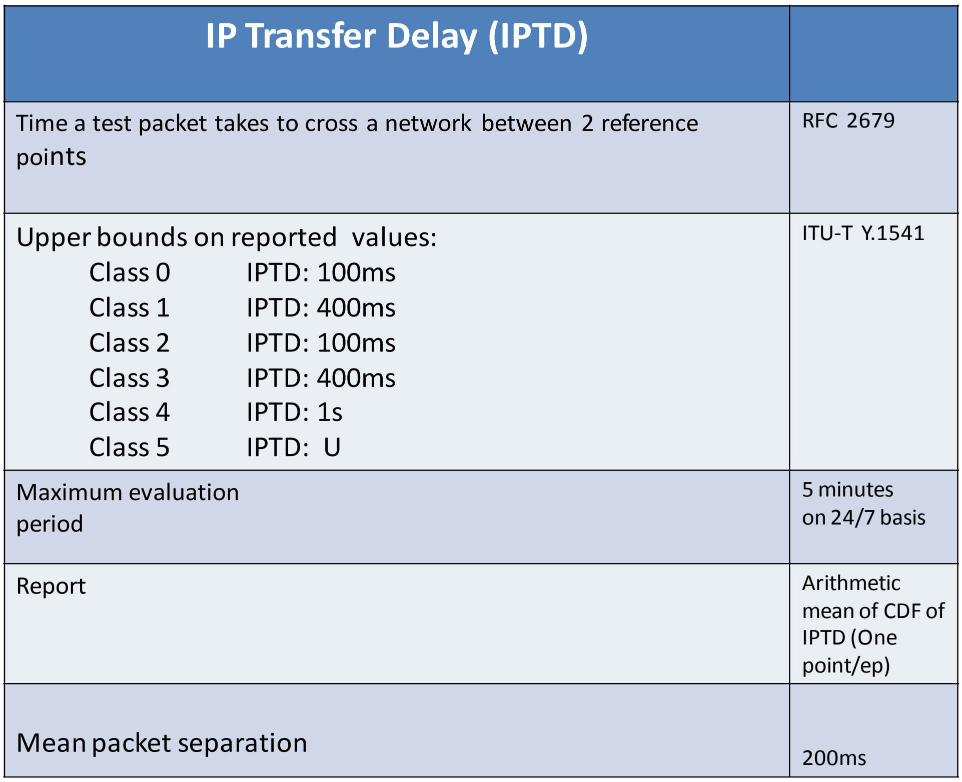
Y.1221 transfer capability	Associated DiffServ PHB	IP QoS Y.1541 class
Best Effort (BE)	Default	QoS Class 5 (Unspecified)
Statistical Bandwidth* (Modified to Limit Delay)	AF	QoS Classes 2,3,4
Dedicated Bandwidth (DBW)	EF	QoS Classes 0 and 1

Table 2.2 Possible classes association



QoS Class	Applications (Examples)	Node Mechanisms	Network Techniques
0	Real-Time, Jitter Sensitive, High Interaction (VoIP, VTC)	Separate Queue with Preferential Servicing, Traffic Grooming	Constrained Routing/Distance
1	Real-Time, Jitter Sensitive, Interactive (VoIP, VTC)		Less Constrained Routing/ Distance
2	Transaction Data, Highly Interactive (Signalling)	Separate Queue, Drop Priority	Constrained Routing/Distance
3	Transaction Data, Interactive		Less Constrained Routing/ Distance
4	Low Loss Only (Short Transactions, Bulk Data, Video Streaming)	Long Queue, Drop Priority	Any Route/Path
5	Traditional Applications of Default IP Networks	Separate Queue (Lowest Priority)	Any Route/Path

Table 2.3 [Y1541] classes' application, node and network routing behaviour [SE03]



IP Transfer Delay (IPTD)	
Time a test packet takes to cross a network between 2 reference points	RFC 2679
Upper bounds on reported values: Class 0 IPTD: 100ms Class 1 IPTD: 400ms Class 2 IPTD: 100ms Class 3 IPTD: 400ms Class 4 IPTD: 1s Class 5 IPTD: U	ITU-T Y.1541
Maximum evaluation period	5 minutes on 24/7 basis
Report	Arithmetic mean of CDF of IPTD (One point/ep)
Mean packet separation	200ms

Table 2.4 Summary of recommendations in [AM06] for IPTD

IP Delay Variation (IPDV)	
Difference between IPTD of test packet and packet with lowest IPTD in evaluation interval	RFC 3393
Upper bounds on reported values Class 0 IPDV: 50ms Class 1 IPDV: 50ms Class 2 IPDV: U Class 3 IPDV: U Class 4 IPDV: U Class 5 IPDV: U	ITU-T Y.1541
Maximum evaluation period	5 minutes on 24/7 basis
Report	• 99th percentile of CDF of IPTD-IPTD(min) (One point/ep) or alternatively • 99th percentile of CDF of IPTD-IPTD(min) plus number or packets with large delay (Two points/ep)
Mean packet separation	200ms

Table 2.5 Summary of recommendations in [AM06] for IPDV

IP Packet Loss Ratio (IPLR)	
Packet loss is ratio between uncorrupted packets arriving at dest. reference point and those leaving at send reference point. A packet is lost if its IPTD > T_{max} (T_{max} = 3 sec ITU-T Y.1540)	RFC 2680
Upper bounds on reported values Class 0 IPLR: 1×10^{-3} Class 1 IPLR: 1×10^{-3} Class 2 IPLR: 1×10^{-3} Class 3 IPLR: 1×10^{-3} Class 4 IPLR: 1×10^{-3} Class 5 IPLR: U	ITU-T Y.1541
Maximum evaluation period	5 minutes on 24/7 basis
Report	Packet loss probability (accurate at 0.1%) (One point/ep)
Mean packet separation	200ms

Table 2.6 Summary of recommendations in [AM06] for IPLR

Measurements should be preferably be carried out by active probing, which consists in generating ad-hoc packets of the same size of the class that is being measured.

11.4. USE CASES: ACTOR-BASED ANALYSIS

For each of the ETICS ecosystem actors multiple ways of using the ETICS inter-carrier services are possible.

The focus of ETICS is on assured quality (AQ) network services traded between the NSPs (the NSP-to-NSP network services) in order to establish inter-carrier ASQ paths across the different NSPs. Use cases which are focusing on the NSP-to-NSP ASQ path level can be understood as generic ASQ path use cases targeting to establish an assured service quality (ASQ) infrastructure. Based on such an ASQ path infrastructure the NSPs can offer specialised transport services towards their final customers (e.g. content provider, business customer, neighbouring NSPs etc.), as for instance managed quality connectivity services. In order to realise such “managed connectivity” use cases that are based on or “on-top-of” the generic ASQ infrastructure additional functionalities have to be implemented and deployed. These additional functionalities may also have more or less impact at the NSP-to-NSP interactions in order to enable and support/enhance such end-to-end ASQ inter-carrier connectivity as well as other services on-top of the ASQ path.

In order to provide a more generic extension of the ETICS architecture the so called Service Enhancement Functional Area (SEFA) was introduced in [ETICS-D4.3], which enriches the basic ETICS architecture and provides the base for individual as well as specific value added functions and services on top of the ASQ paths, as for instance establishing and handling of managed connectivity services that are carried by (or “on-top-of”) the infrastructure level ASQ paths. An example of such an enriched use case, which considers connectivity service as such, as well as how to manage these connectivity services, is explained in further detail in Section 6.1.

The following paragraph deals mainly with generic ASQ path use cases and illustrates how different actors of the ETICS community can benefit in different ways from the ETICS architecture. It should be mentioned that these use cases represent only a subset of all use cases that become possible with the ETICS framework. Most of these cases have already be defined and discussed in Deliverable D1.6, however some improvements and extensions have been incorporated here.

The actors content provider, business customer, content delivery network (CDN) and network service provider (NSP) shown in FIGURE 3 can benefit from the generic ASQ path inter-carrier service, such as:

- The actors filling out the customer side of the ETICS architecture (content provider, business customer and content delivery network) are able to request ASQ inter-carrier service from the ETICS community.

- An NSP is able to request ASQ inter-carrier service from the ETICS community.

11.4.1. CONTENT PROVIDERS

Content providers can request inter-carrier services for one of the following reasons:

11.4.1.1. Assured-quality destination-based Internet transit

Today, when content providers do not use the services of a CDN, they buy bundled connectivity services to reach the entire Internet. Such a service is called the transit service, its pricing is volume-based. An Internet transit service is a-best-effort reachability service to all potential destinations that are connected to the Internet.

In the context of inter-carrier ASQ services, content providers might want to buy a *destination-based assured quality transit service*, that is an ASQ path to a given set of destinations within one (or multiple) access ISP(s). This set of destinations within one (or multiple) access ISP(s) is what we call a *region*. Without loss of generality, we will focus on the

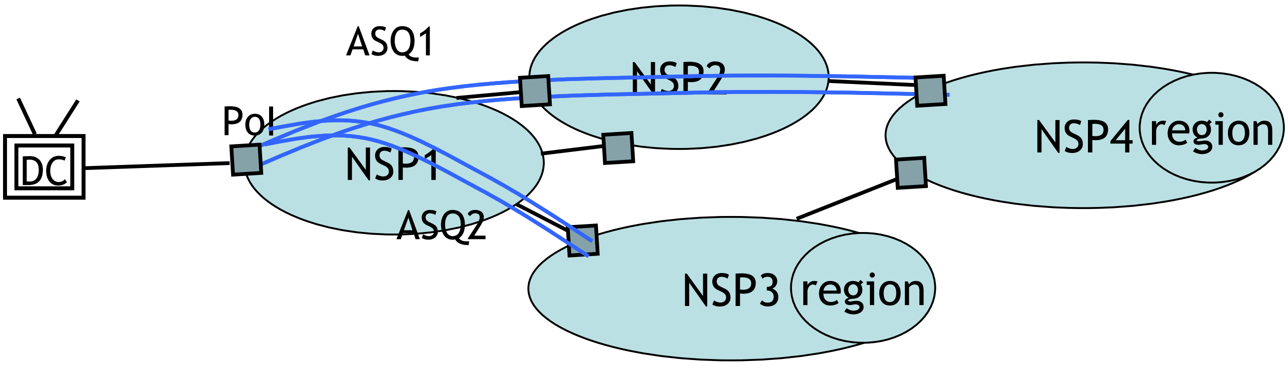


FIGURE 66: ENHANCED INTERNET SERVICE FOR CONTENT PROVIDERS

Such tiered or destination-based transit pricing has been coined by Valancius et al. [Valancius-tiers] as an alternative to the current bilateral transit agreements. The study led by Georgia Tech showed the benefits of tiered (destination-based) transit pricing. However, they recommended to limit the pricing to only three tiers: roughly one low transit price for close destinations, a medium price for “medium” destinations, and a higher price for farther destinations. Their argument is that having as many tiers as destinations would be hard to manage in terms of contracts to handle, and having only three tiers could already provide near optimal benefit for transit providers. The ETICS framework could complement this by:

- Having as many tiers as CPs request. This was enabled by the automation of negotiation which simplifies the management “complexity” by automating the negotiation. (Note: The automation of the establishment of such ASQ paths themselves (as ETICS is describing) is anyway an important step to lower the cost of ASQ path production and hence ASQ traffic.)
- Adding quality guarantees to the tiered destination-based transit and extend it to the inter-carrier case.

In fact, in destination-based pricing as proposed by [Valancius-tiers], the transit provider has only control on the traffic that transits within its own network and has no guarantees therefore on the end-to-end paths. To illustrate this, in Figure 2, a destination-based transit provided by NSP1 to the Data Centre on the left cannot have guarantees on how the traffic will flow in NSP2, NSP3 and NSP4. However, an inter-carrier ASQ path to the different regions in NSP4 and NSP3 would give such end-to-end guarantees. (Note: Since the network parts of an ASQ path are operated by the NSPs themselves a real guarantee can only be given by an NSP for the own network part. For the cascading scenario all the “following” NSPs operate their own networks and the “sending” NSP can only trust them. Real end-to-end guarantees can only be given by the ETICS consortia as a whole.)

It is important to note that the traffic in such ASQ paths goes through two phases. (1) A first phase in which it is aggregated (e.g. through NSP1 and NSP2 for ASQ1, and through NSP1 for ASQ2). (2) A second phase, at the ingress point of the last access NSP, in which it is “disaggregated” to reach the different destination hosts within a region. Although it is relatively easy to give traffic delivery guarantees to the aggregated level (1) (e.g. A given aggregate bandwidth and a given delay), it is harder to quantify such guarantees for the “disaggregated” traffic (2) as the traffic will split to reach the different destinations.

Therefore, we distinguish between two possibilities depending on which guarantees are provided for steps (1) and (2):

11.4.1.1.1. *Have the ASQ path guarantees only on the aggregate level*

In this first possibility, the inter-carrier path is assured only till the ingress of the last access NSP. The **remaining part of the access NSP is therefore provided in Best effort**. Although somewhat counter-intuitive, such service can be interesting for content providers even if the last part of the ASQ path is best effort. In fact, similar intra-domain services are sold today in the form of “speed-ups” where a content provider asks its transit provider to “boost” the quality of its traffic delivery within its domain and let the rest of the path (access) go best effort.

11.4.1.1.2. *Also provide guarantees for each destination in the region, and not only to the aggregate level.*

An example for such a service is to provide a guarantee to have X Mbps and Y ms of delay at worst for each destination (or for Z% of the destinations).

Note that a content provider may need only an inter-carrier ASQ path between only one of its servers and one distant host. However, as we will explain in Sec. 7, for scalability reasons, we will only focus on aggregated traffic. It will be required to build such individual connectivity sessions between two hosts.

Addressing and traffic identification considerations:

1/ All the hosts in the destination set can benefit from the ASQ path

2/ The NSPs need to make sure that only traffic originated by the content provider goes through the considered ASQ paths. This might be enabled either through tunnelling or through appropriate *traffic identification* at border routers within PoIs.

11.4.1.1.3. *Definitions:*

We finish this section by clearly setting the definitions of a few terms introduced in this subsection which we will use through the rest of this deliverable.

We call a set of end hosts a *region*. A region can be addressed by an IP prefix or a set of IP addresses.

We call the part of the ASQ path where the traffic of multiple end hosts of a region goes through the same path the *aggregated traffic*. The aggregated traffic is typically traffic from a PoI to a PoI.

We call the traffic that is related to an individual host or to an individual application within an individual host an *individual session*.

11.4.1.2. ASQ path filled on demand (Individual-sessions-aware ASQ paths that are meant to be filled on-demand):

Content or application providers can demand an ASQ path to a large set of *potential* destinations, called *region*, with only a limited subset of users within the region that can use it simultaneously. In such a service, the ASQ connectivity is granted only upon the demand of the end user.

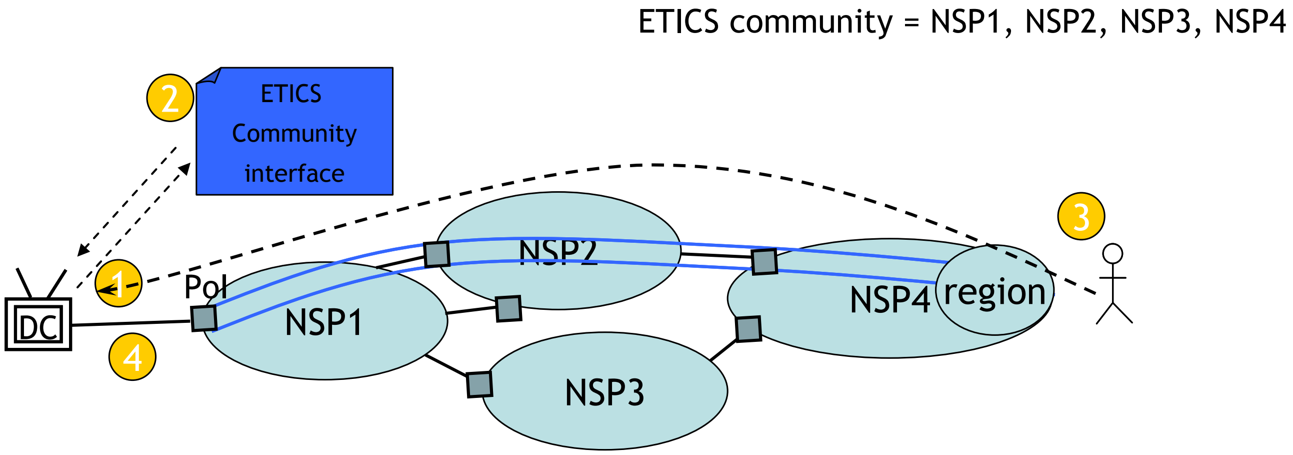

ETICS community = NSP1, NSP2, NSP3, NSP4
2
ETICS Community interface
3
1
PoI
DC
4
NSP1
NSP2
NSP3
NSP4
region

FIGURE 67: ASQ PATH TO A REGION WITH SESSIONS (MICRO FLOWS) HANDLING

Such an ASQ path is similar to the one in Section 11.4.1.1.2 except that the association of individual sessions' traffic to the aggregated part of the ASQ path is done in a dynamic way on-demand. We use Figure 3 to illustrate such an ASQ path and how it could be used by the content provider. In a first step (1), the content provider (depicted by the "TV" on the left side of the figure) requests a flow-aware ASQ path to the destination region in NSP4 (from an ETICS interface as we will see later). The flow-aware ASQ path can specify the maximum number, say N, of destinations or flows that can profit from the ASQ path simultaneously, as well as the characteristics of each individual session. Thanks to the ETICS mechanisms (that we will detail later), the ETICS system provides the content provider with the requested ASQ path (2). At this point, the ASQ path is "empty" and does not contain traffic. The filling of the ASQ path (saying which flow- dest,src address etc- goes inside the aggregated ASQ path)is be done upon the demand of an individual user. In fact, an individual user (a destination in the region) would first connect to the content provider's portal and then ask for an assured quality content delivery (3). In a fourth step, the content provider will associate the individual user's flow to the flow-aware ASQ path.

Note that although the individual sessions are expected to be very dynamic (upon an end user demand for a content), the big ASQ PoI-to-region path would be much less dynamic. They are negotiated only once. The content provider can ask for upgrading them depending on the users' demands.

We will see later in this document that a new functional block of the ETICS architecture (called Service Enhancement Functional Area (SEFA) could be used to facilitate establishing such services for handling individual sessions).

Challenges:

Traffic identification considerations: One of the main challenges of such a (potential) service, that the ETICS project needs to solve (if it sees an interest in such a service), is the association of the end user's micro flow with the big pipe (the flow-aware ASQ path).

11.4.2. Business customers

Business customers will basically request inter-domain ASQ paths to interconnect their sites or data centres. The services that business customers can expect from the ETICS inter-carrier services are:

Inter-carrier Layer 2 connectivity services with (or even without) assured quality.

- point-to-point services (layer 2 over ETICS)
- point-to-multi-point services ("VPLS" [RFC4761] over ETICS)

Inter-carrier Layer 3 connectivity services with (or without) assured quality.

- Point-to-point services
- Point-to-multi-point services

An example for an Inter-carrier Layer 3 point-to-point connectivity service is a so called "(Low-end) managed business connectivity service".

11.4.3. Network Service Providers

11.4.3.1. Individual-sessions-aware ASQ paths that are meant to be filled on-demand

Edge NSPs can request ASQ paths to reach regions in other edge NSPs. Such services can be used to provide Assured Quality host-to-host telco communication services. Similarly to the service in Sec.11.4.1.2, such service needs to be coupled with the ability to add individual sessions to the *aggregated traffic* in the ASQ path. Therefore, they are a kind of individual-sessions-aware ASQ paths that are meant to be filled on-demand.

One possible operating mode for such a service is the following. Upon the demand of an edge NSP, say NSP1, to reach a region within a distant edge NSP, say NSP2, an "empty" ASQ path is set between two NSPs. Upon the demand of an end-user 2 from NSP2 to have an assured quality communication with another end-user 1 within NSP1, the individual session that is relative to the two hosts 1 and 2 is added to the aggregated traffic part of the ASQ path. Once the communication is over, traffic related to this individual session is "removed" from the aggregated traffic ASQ path.

Challenges:

Such services have the same traffic identification concerns and challenges as those of Sec.11.4.1.1.3.

Moreover, our description presented a simplified version of the problem. In reality, please note that in order to have an assured quality host-to-host communication, two ASQ paths (in the two directions, traffic termination and traffic origination) need to be set. One goes from NSP1 to NSP2 and another one goes from NSP2 to NSP1.

Finally, the business model and the money flows for such service need to be clarified. In particular, who would pay for such a service? Edge NSP1 or Edge NSP2? Or both? And how would the revenues of the host-to-host communication be shared between the two edge NSPs and the transit NSPs (if any) that provide the aggregated traffic part of the ASQ path?

Please note that this is not the only way to enable host-to-host communication services. Another way to do it is to have a third-party (telco service provider/ application content provider) provide such services. In this case, the telco provider is responsible for building and buying such a service, using one of the previously described services. The idea is that this third party buys the inter-carrier ASQ paths that are necessary for the communication set up and retails them to end users in the form of communication services. It can act either as a hub/relay for the traffic between the end hosts or as a facilitator/enabler for the communication set up.

11.4.3.2. Assured-quality destination-based Internet Transit

Similarly to content providers in Sec. 11.4.1.1, edge NSPs can request assured-quality destination-based internet transit to reach either (1) destination content provider networks or (2) destination edge NSPs.

11.4.4. Content distribution Networks and Cloud Service Providers (?)

In the context of CDN interconnect; different CDNs might have punctual interconnection needs with guaranteed quality to operate transfers of content between CDNs.

However, a study in Deliverable D3.3 [ETICS-D3.3] concluded that this is not a viable use case for ETICS.

11.4.5. Residential end user

Residential end users may need inter-carrier services for one of the following reasons:

- Have Assured Quality connectivity services to certain content providers' content (that are not directly connected to their access NSP) either in a permanent way or on-demand for a limited amount of time (e.g. watch an HD/3D movie).
- Have Assured Quality connectivity services to certain destination end-hosts in order to have assured quality communication services (e.g. HD Videoconference, telepresence, HD audio).

Note that the two above mentioned services can be implemented thanks to inter-carrier services requested by edge NSPs (Section 11.4.3) or content providers (Section 11.4.1.2). However these ETICS actors (Edge NSP, Content provider) will request a "whole" ASQ path between ingress and egress edge NSP of the residential customers resp. content provider. In order to support a more fine granular service level as for instance a session-based connectivity service, additional mechanisms and functionalities have to be implemented and used. (As we will see later, **we assume that, for scalability reasons, it is not realistic to establish inter-carrier paths upon such individual sessions demands of residential end users**. Services for residential end-users will require the pre-establishment of inter-carrier ASQ paths like those of Sec. 11.4.3 and Sec. 11.4.1.2.)

The needed functionality for such a (more fine-granular) individual session-related connectivity based on the generic ASQ path service can be provided by an extension to the core ETICS architecture framework – the SEFA. This approach is illustrated and described in more detail in Section 6.1 below.

Challenges:

For individual session-based connectivity services the major challenge consists in identification and separate handling (where needed) of individual sessions within the overall ASQ path.

Open issues:

Decide who sells such services to the end user?

The content provider or the NSP? A priori, the first would be bought from the content provider. The second would be bought either from the NSP or from a third party: a telco provider.

11.5. "ORIGINAL" ETICS ACTOR ROLE MODEL

For completeness and for reference to previous deliverables (Del2.2 and Del4.3) the following ETICS actor role model illustrations are shown. See the main section regarding the updated ETICS actor role model (FIGURE 19). The first illustration, (bilateral cascading mode) originally also considered specifically transport NSP roles and network services. While this can be of relevance and value due to time constraints, the ETICS solution did not further explore the E4 and E5 reference points. However, the generality of the E0 reference point do allow to also specifically handle "sub-IP" transport network services based on the path computation mode.

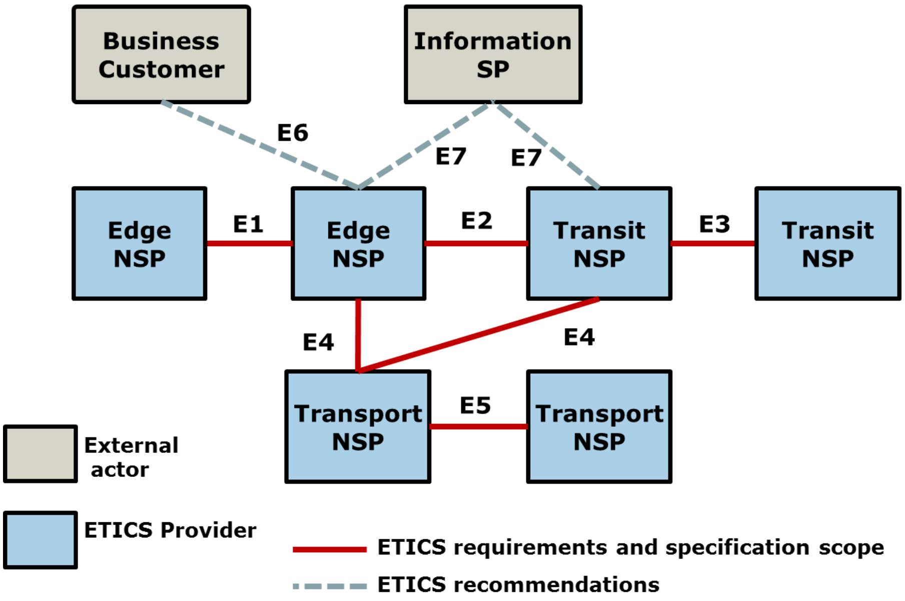


FIGURE 68 ETICS ACTOR-ROLE MODEL (BILATERAL CASCADING MODE)

The "original" actor role model for the path computation mode is shown below. Here, and as noted in the main section above, the so-called primary NSP[20] can maintain a multi-domain (inter-NSP) topology view and perform an inter-NSP-level composition and stitch together element network services as offered by the NSPs in the subordinate role[21].

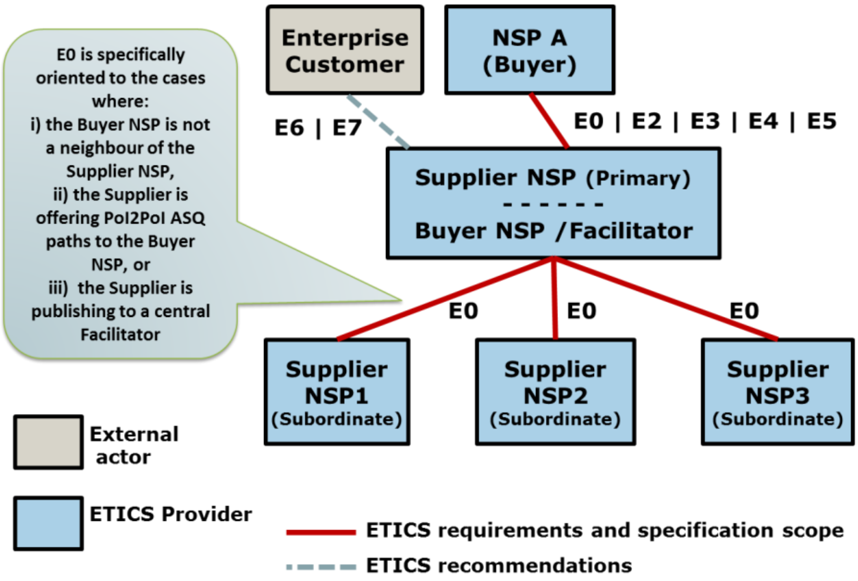

E0 is specifically oriented to the cases where:
i) the Buyer NSP is not a neighbour of the Supplier NSP,
ii) the Supplier is offering PoI2PoI ASQ paths to the Buyer NSP, or
iii) the Supplier is publishing to a central Facilitator
Enterprise Customer
NSP A (Buyer)
E6 | E7
E0 | E2 | E3 | E4 | E5
Supplier NSP (Primary)
Buyer NSP /Facilitator
E0
E0
E0
Supplier NSP1 (Subordinate)
Supplier NSP2 (Subordinate)
Supplier NSP3 (Subordinate)
External actor
ETICS Provider
ETICS requirements and specification scope
ETICS recommendations

FIGURE 69 ETICS ACTOR-ROLE MODEL (ASQ PATH COMPUTATION OPTION)

11.6. MONITORING

This section provides some annex material on monitoring for the Section 5.8. In particular, the information elements for the passive NMON communication, as well as potential extensions of the NOMON system are discussed.

11.6.1. PASSIVE NMON COMMUNICATION ELEMENTS

In the following, more details about the transported information elements is given:

HELLO – Message:

Protocol-ID: some „magic" value identifying the protocol. The Collector (server) uses this to determine, if the client connecting is a valid SLA_Monitoring instance.

Protocol-version: a numerical value identifying the protocol version the client wants to use.

AuthCap: Client authentication capabilities; i.e. the authentication methods the client can support

AUTHR – Authentication Request:

Auth-Meth: the authentication method the client shall use; e.g. username-password, or digest-authentication

[20] IPsphere use the notion of "Administrative Owner".
[21] IPsphere use the notion of "Element Owner" for this role.

Authr-Params: the parameters used for the authentication request; for digest-authentication, this would contain the challenge

AUTHA – Authentication Answer:

Autha-Params: the parameters used for the authentication answer; for digest-authentication, this would contain user-name, CNONCE, response

AUTHX – Authentication Result:

Submit the result of the authentication cycle.

Error-Code: indicating which problem occurred, shall include a code indicating "no error"

Error-Text: a textual description of the problem

CONFIGR – Configuration Request:

AddReplace: A bool value indicating, if the probe configuration shall be extended by this configuration, or replaced

Configuration: A configuration-set for the probe. This could e.g. be some XML formatted data element. The detailed specification is implementation dependent and outside the scope of this project.

CONFIGA – Configuration Answer:

Submit the result of the CONFIGR message.

Error-Code: indicating which problem occurred, shall include a code indicating "no error"

Error-Text: a textual description of the problem

GETCFGR – Get Configuration Request:

Request the configuration of the probe; no parameters (the configuration is always sent as a whole).

GETCFGA – Get Configuration Answer:

Submit the result of the GETCFGR message.

Error-Code: indicating which problem occurred, shall include a code indicating "no error"

Error-Text: a textual description of the problem

Configuration: The configuration-set of the probe. This could e.g. be some XML formatted data element. The detailed specification is implementation dependent and outside the scope of this project.

COLLECTR – Collect-Data Request:

This message starts the actual data collection for a passive NMON sub-system.

StartTime: the start-time of the time-span for which QoS metrics shall be calculated (usually given in seconds since epoch, i.e. seconds since 1st Jan 1970)

StopTime: the end-time of the time-span for which QoS metrics shall be calculated; seconds since epoch as for StartTime

PropertyFilter: the property-filter indicates which packets shall be selected for the measurements. The filter shall be given in conjunctive normal form and the following properties shall be supported:

- Network (layer)3 protocol
- Source- and destination IP address (IPv4 and IPv6)
- Transport (layer 4) protocol
- Transport layer address (i.e. port number)
- MPLS Labels (stacked)

Example for a filter in conjunctive normal form:

((IP-src == 86.59.24.130) OR (IP-src == 86.59.24.140)) AND (port-src == 80)

COLLECTA – Collect-Data Answer:

Submit the result of the COLLECTR message.

Error-Code: indicating which problem occurred, shall include a code indicating "no error"

Error-Text: a textual description of the problem

Result: a vector, giving for each second[22] in the requested time-span at least: average delay, delay variation, number of matched packets, number of unmatched packets, number of observed negative time-intervals, estimated bandwidth.

Other metrics might be added to the result-vector depending on future requirements.

Note on Information Elements content:

In this abstract, high-level description of the protocol, some information elements which are needed for a real implementation might be missing. Their definition is implementation specific and shall be added by the respective protocol description.

11.6.2. Discussion of potential (future) extensions of NMON

Using the measurements collected by the NMON system, several QoS metric can be derived. Bandwidth, delay and loss are the most interesting QoS metrics to draw conclusions about the quality experienced by the application or user. The available bandwidth and delay can be measured instantaneously at one point of time or even for one packet/probe. To achieve estimation about the mean available bandwidth or average delay, samples can be taken over a certain period of time.

This section focuses on loss measurements as loss measurement provides a slightly harder problem. One loss event is monitored by detecting if a certain packet was delivered or lost. This information for one packet or for one point of time does not provide any useful information regarding the Quality of Experience. To estimate an average loss rate it is not enough to sample single probe as the lost packets

[22] Other granularities than one second are possible, but we recommend one second

might be missed. Thus loss can either be measured directly at the loss point, if known, by counting all dropped packets, or, if two measurement points are available, the traffic volume before and after the loss point can be compared. This simple measurement does only provide an average loss rate but cannot provide any information on the loss pattern. Depending on the application the average loss rate might not be enough, e.g. if FEC is used the number of subsequent lost packets might be more important than the average loss rate. Moreover, NMON provides per packet data but does not guarantee that all packets are captured and thus does not provide an accurate loss measurement. In the following section we survey passive and active loss measurement methods mostly based on TCP which can be performed in addition to NMON monitoring.

In the loss pattern, effects of congestion control and application behaviour are observable. Future work will be investing these loss patterns to create a loss model of today's Internet traffic. The goal is to identify conditions of congestion by comparing loss measurements with the loss patterns generated by such a model. This information can be used to improve network measurements and failure detection. A first study of loss patterns in typical Internet usage scenarios is presented in [KNT12].

11.6.2.1. Passive Loss Estimation

For TCP traffic, header information can be utilized for retransmission detection and thus loss estimation. At the network edge, where forward (data packets) and backward (ACK) can be seen, one measurement point is sufficient to detect loss. At the sender-side retransmissions are estimated by detecting duplicated sequence numbers in the TCP header that at the same time as a different id field. SACK, DSACK or the recognition of duplicated ACK pattern can help to estimated spurious retransmission and thus improve the loss estimation. Spurious retransmissions often occur when small flows lose the last packets of a flow. Measuring at receiver-side, a hole in the sequence numbers can be detected. To distinguish missing packets from reordering and thus packets that will arrive later, the timestamp (in the TCP timestamp option) can be used if present. Alternatively a maximum delay threshold must be chosen. Only if a packet with the missing sequence number cannot be seen within the time frame defined by the threshold, it is accounted as loss. For measurement points in the network, both approaches need to be combined as the loss can either occur before or after the measurement point.

Several heuristics for passive retransmission detection and thus loss estimation have been proposed in the literature [BV02, RKS06]:

- LEAST: Loss Estimation AlgorithmS for TCP [AE03]
- TSTAT: TCP STatistic and Analysis Tool [MMMR08]

We implemented a measurement algorithm similar to those in the literature to perform a first loss measurement study. The decision diagram is shown in FIGURE 70.

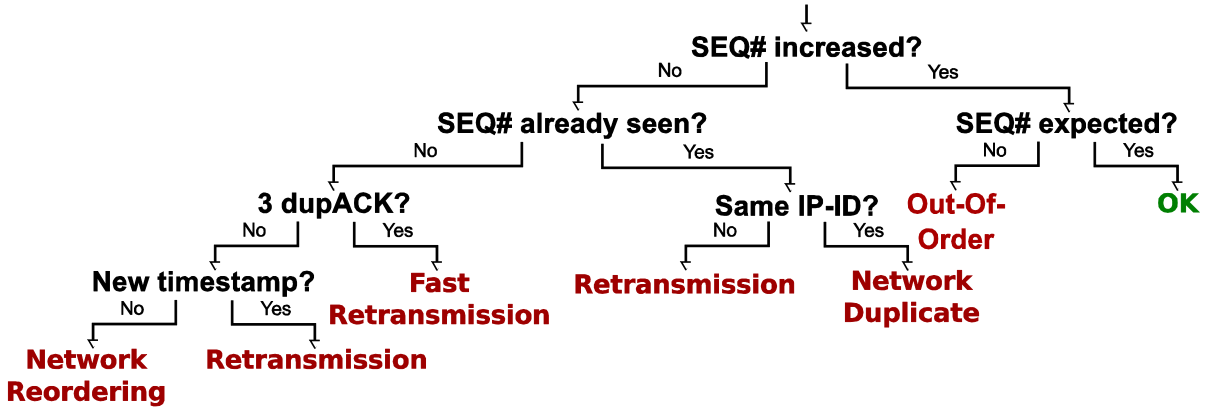


FIGURE 70: DECISION DIAGRAM FOR TCP LOSS/RETRANSMISSION ESTIMATION

11.6.2.2. Active Loss Measurements

While passive measurement can only be performed when there is actual user traffic, active measurements aim to continuously capture the network state.

Active measurements induce measurements traffic in the network. To avoid that this measurement traffic will influence the network state and thus the measurements results, active measurements aim to keep the measurement traffic low. Other than e.g. delay measurements, which usually have relatively stable measurements value over large time period, loss usually only occurs for short times and potential provide strongly varying measurement values.

While the goal is to measure packet loss experienced by other flows, the problem with active measurement methods is that they might either introduce own losses (if too often/too much load) or miss losses (if too few/too rare). Proposed tools try to modulate the measurement traffic such that the traffic volume induced is low but the losses can still be captured accurately. Poip (Poisson Ping) send UDP packets with Poisson modulated intervals and fixed mean rate [PMAM98]. BADABING sends sequences (two or three fixed-size probes of back-to-back packets) with geometric distributed intervals [SBDR08]. Sting sends specified number of TCP packets over raw socket and counts acknowledgements. By skipping the first sequence number it is ensured that the receiver will send one ACK per received packet. This allows a two-way loss measurement with a sender-side only tool [S00].

11.6.2.3. Loss Metrics

The IP Performance Measurements working group of the IETF has defined several metrics to describe loss rates and loss pattern:

- One-way Packet Loss Metric [IETF-RFC2680]
  - Type-P-One-way-Packet-Loss
  - Type-P-One-way-Packet-Loss-Poisson-Stream
  - Type-P-One-way-Packet-Loss-Average (loss rate)

While the Type-P-One-way-Packet-Loss describes only a binary value which is 0 if not loss occurred and one if so, the Type-P-One-way-Packet-Loss-Poisson-Stream is a sequence of time and Type-P-One-way-Packet-

Loss values pairs. Type-P-One-way-Packet-Loss-Average gives the average of all Type-P-One-way-Packet-Loss values of a Type-P-One-way-Packet-Loss-Poisson-Stream.

- One-way Loss Pattern Sample Metrics [IETF-RFC3357]
    - Type-P-One-Way-Loss-Distance-Stream
    - Type-P-One-Way-Loss-Period-Stream

These metrics extend a loss sequence by the distance and period which is the number of received packets between two losses and respectively the number of lost packets in a row.

- Statistics [IETF- RFC3357]
    - Type-P-One-Way-Loss-Noticeable-Rate
    - Type-P-One-Way-Loss-Period-Total
    - Type-P-One-Way-Loss-Period-Lengths
    - Type-P-One-Way-Inter-Loss-Period-Lengths

The noticeable rate is calculated based on a minimum threshold for the loss distance, as e.g. application using FEC can conceal some errors. Type-P-One-Way-Loss-Period-Lengths describes the number of loss periods while Type-P-One-Way-Loss-Period-Lengths describes the number of packets lost in one loss period. Type-P-One-Way-Inter-Loss-Period-Lengths gives a distance between two loss periods. This metric can be used to a define loss periods that have a small number of received packet between each other as belonging to one burst loss.

In [IETF-RFC6534] also the loss episode duration, loss-free episode duration and loss episode frequency are defined. Challenges like defining a loss threshold, clock synchronization and determining the packet size are discussed in [IETF-RFC3357].

When interpreting these metrics, effects of congestion control and application behaviour have to be taken into account. Congestion control algorithms themselves periodically induce overload to probe for available bandwidth. Therefore herein a burst loss is defined as an event consisting of all losses occurring in a TCP connection within one RTT starting from the first loss. Thus the number of losses within a burst strongly depends on the used congestion control scheme and current share of the bottleneck capacity. For standard TCP congestion control and a greedy source, periodic bursts of (more or less) constant size (1-4 packets/burst) are expected.

Moreover, a loss burst also counts the number of losses within a loss episode for aggregated traffic at one network node. Counting these events provides a new metric which captures packet loss in a congestion-control aware way, as losses occurring within a single RTT will be treated as a single event by TCP.

11.6.2.4. Future Work

This first analysis of loss measurement and loss pattern is motivated by the goal of identifying conditions of congestion based on the measurements of burst losses. This is a hard problem as a large number of effects might influence the loss pattern. In addition to effects of the bandwidth probing of TCP congestion control, application behaviour influences the loss pattern as well. E.g. YouTube traffic shows large burst losses even though TCP congestion control is used, as the application sends the traffic in blocks with only very short idle times.

As a next step to identify condition of congestion, a loss model of the application and used congestion control needs to be created. This is currently work in progress. The loss model currently focuses on three common classes of Internet activity – web browsing, download, and YouTube – to study their loss patterns individually. Based on this analysis, one will be able to compare loss measurements with loss patterns generated by such a model. Thus, situations, where the measured loss pattern (strongly) differs from the model, can be identified as potential problem cases that need to be further investigated. Furthermore, based on a more advanced loss model, one also might be able to distinguish shared congestion, self-congestion (bottleneck used by only one flow) and loss due to slow start overshoot from actual bit errors or losses caused by problems in the network. This is future work. As a further extension, spatial and temporal correlation of per-flow burst loss measurements, collected at the network edge, can be used for network tomography and locating of error sources.

11.7. SEFA BASED USE CASES

Complementary to the SEFA cookbook being introduced in Section 6.1.4, the subequent subsection will in detail discuss two exampary use cases taken from 6.1.5, i.e. Graceful Denial of Service and Managed Connectivity Service using pre-established ASQ paths.

11.7.1. SEFA BASED USE CASE 1 – GRACEFUL DENIAL OF SERVICE

In the following the realisation of an "Over the top of ETICS" use case, the "Graceful Denial of Service" (GDos), will be described by means of the SEFA cookbook.

0. Goal of SEF use case:

The goal of this specific SEF use case "Graceful Denial of Service" is to provide an information and / or engaged signal for IP-based services to mass marked customers. FIGURE 71provides a high level view on the graceful denial of service SEF use case scenario which is enabled by means of considering the available network capabilities of the mass market customer and the application service requirements on the network performance in the mass market customer access network.

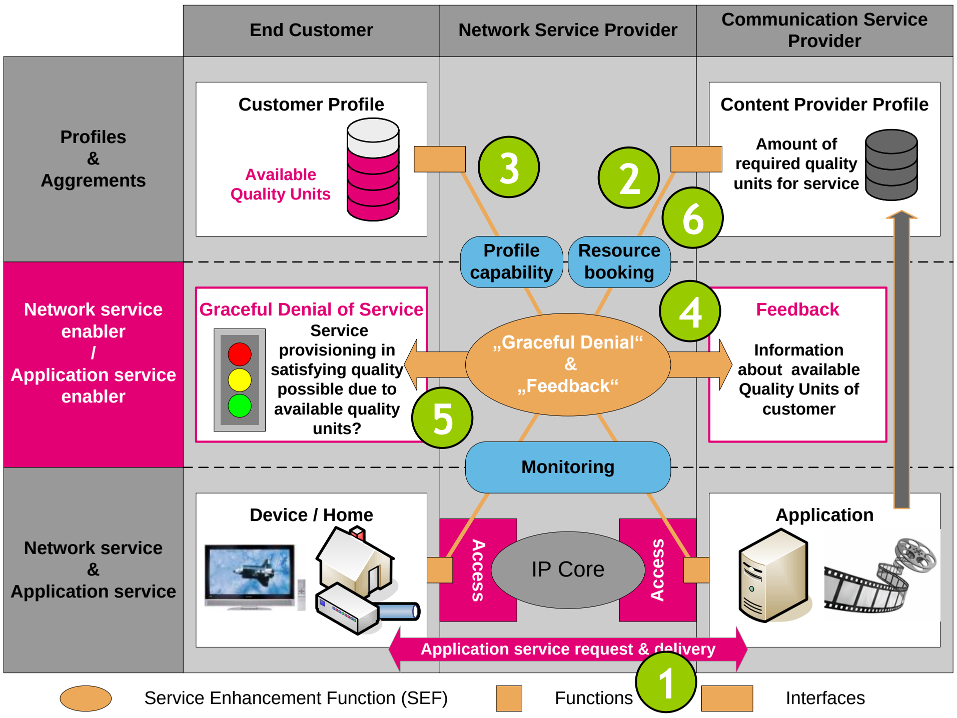

End Customer
Network Service Provider
Communication Service Provider
Profiles & Aggrements
Network service enabler / Application service enabler
Network service & Application service
Customer Profile
Available Quality Units
Content Provider Profile
Amount of required quality units for service
Profile capability
Resource booking
Graceful Denial of Service
Service provisioning in satisfying quality possible due to available quality units?
„Graceful Denial“ & „Feedback“
Feedback
Information about available Quality Units of customer
Monitoring
Device / Home
Application
Access
IP Core
Access
Application service request & delivery
1
2
3
4
5
6
Service Enhancement Function (SEF)
Functions
Interfaces

FIGURE 71: SEF USE CASE GRACEFUL DENIAL OF SERVICE

The high level description of the “GDoS” added value service is given as follows:

1. Customer requests application service, such as IPTV, in high quality from communication service provider (CmSP).
2. CmSP sends required amount of quality units needed for providing requested application quality to the network service provider (NSP). (Instead of quality units also a parameter which represents the needed network access bandwidth may be used for this added value service.)
3. Service Enhancement Function (SEF) requests available quality units (or access bandwidth parameters) from customer profile.
4. SEF - Graceful Denial of Service evaluates capabilities of the customer profile, such as available quality units and provides feedback to the CmSP.
5. CmSP informs customer about the state of service provisioning capabilities, such as the requested service is provideable in requested quality or not.
6. CmSP carries out resource ordering from the NSP in the case of user confirms start of application service.

1. Prerequisites of SEFA design:

In the context of ETICS a multi-provider network architecture is assumed consisting of two Edge NSPs and zero to N Transit providers. The Edge NSPs connect either the content provider or the mass market end customer to the network of this GDoS scenario. In the case that the NSPs are acting as an ETICS community the existence of an ASQ path could be assumed between the Edge NSPs. Since this scenario is an example for an “over the top of ETICS” SEFA use case the ASQ path is not necessary for the realisation for this SEF-GDoS use case.

FIGURE 72 shows the involved actors which have to implement SEF-GDoS related instances. (Note: The functionality of the corresponding SEF instances of different actors may vary according to the specific role of the actor in the added value service.)

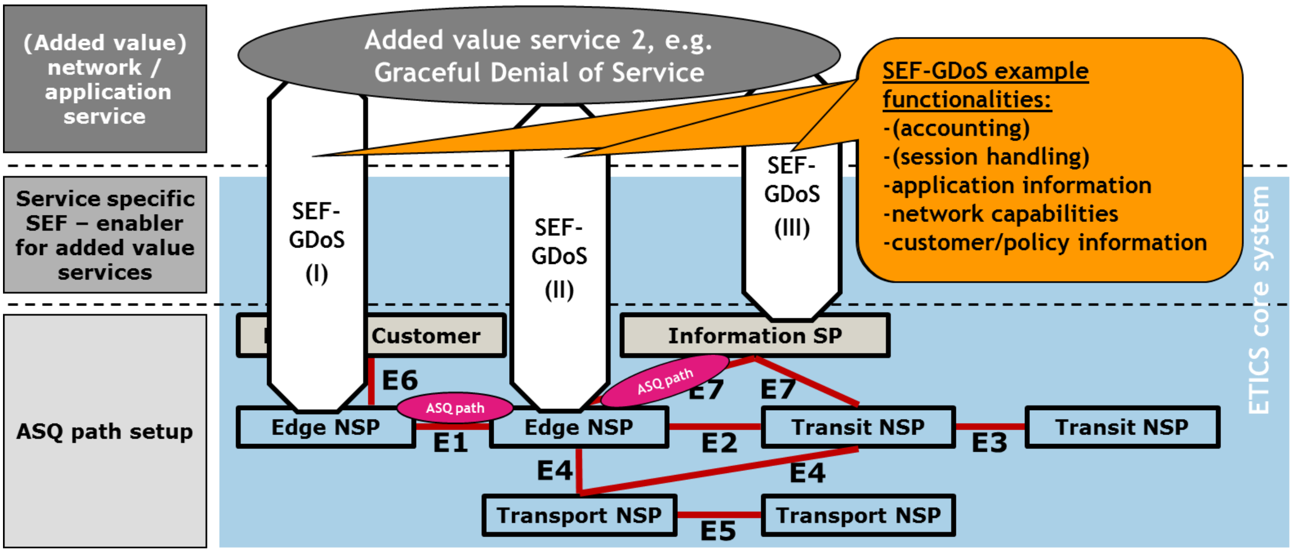


FIGURE 72: ACTORS OF THE GRACEFUL DENIAL OF SERVICE SEF USE CASE

2. Naming of SEF-layer:

As already introduced above the notation of SEF-GDoS has been chosen in order to highlight that this added value service “Graceful Denial of Service” is realised by the specific Service Enhancement Function SEF-GDoS.

3. Requirements on SEF-layer:

The added value service “Graceful Denial of Service” has the intention to realise a signal/feedback to the mass market customer in the case that some constraints (e.g. network resources, customer credit points) allow the delivery of the requested application service, e.g. video service only with limited quality and hence, reduced quality of experience.

The SEF-GDoS added value service involves several actors, as depicted in FIGURE 72 and has requirements on each actor in order to derive the engaged signal. The requirements on each actor depends on its role in the use case, which is described in the following:

- Edge NSP of mass market customer:
  - Connects mass market customer to the Internet
  - Transport of application data from information service provider to mass market customer.
  - Provides information about available access bandwidth of the connected mass market customer.
- Edge NSP of information service provider
  - Connects information service provider to the Internet

  - Transport of application data from information service provider to mass market customer.
  - Provides information about available access bandwidth of the information service provider.
  - GDoS signal/feedback performing instance
    - In the use case SEF-GDoS, the orchestrating SEF instance requests and gathers all needed information from the other SEF-GDoS instances in order to derive the feedback to the mass market customer. Without loss of generality it can be assumed that this orchestrating SEF instance is located in the network of the content provider connecting edge NSP.
- Information service provider
  - Offers application service (e.g. video service) to the mass market customer
  - Provides application information about required network capabilities to deliver the requested (video) service in customer satisfying quality.
  - (Customer / policy information could be used in order to derive statements regarding the status of the customer account that may be of importance for high quality content delivery.)
- Mass market customer
  - Request application service from information service provider.
- GDoS signal/feedback performing instance
  - In the use case SEF-GDoS, the orchestrating SEF instance requests and gathers all needed information from the other SEF-GDoS instances in order to derive the feedback to the mass market customer. Without loss of generality it can be assumed that this orchestrating SEF instance is located in the network of the content provider connecting edge NSP.

4. Description of SEF-layer:

According to the general structure of the SEFA cookbook this “Description of SEF-layer” is divided into two phases. In the first phase SEF-layer internal high-level functional elements are derived from the use case specific requirements. In the second phase the derived high level functional elements are mapped to the base-layer in order to specify more fine granular the needed functionalities of the specific SEF-GDoS instance.

In order to realise this SEF-GDoS added value service the different actors have to implement specific SEF-GDoS instances in their network. These actor role specific SEF-GDoS instances are depicted in FIGURE 72 by means of numbering I, II and III. Keeping in mind the requirements described in “3. Requirements on SEF-layer” of the SEFA cookbook, the functionality of these instances depends on the role of the actor, which is described in the following with focus on the high-level functional elements of the first phase:

a.) SEF-GDoS (I) – Mass market customer Edge NSP:

- The high-level functionality of SEF-GDoS (I) consists in providing information about the available bandwidth within the mass market customer access network.

b.) SEF-GDoS (III) – Information service provider:

- The high-level functionality of SEF-GDoS (III) consists in providing information about the required network capabilities for the application (e.g. video service) the mass market customer wants to use and also the information that a certain mass market customer has requested this application service. With this trigger information towards SEF-GDoS (II) the overall added value service creation for GDoS is initiated. SEF-GDoS (II) will provide SEF-GDoS (III) with "engaged signal" information which has to be forwarded to the mass market customer in order to signal the expected application service quality to the customer.

c.) SEF-GDoS (II) – Information service provider Edge NSP:

- The high-level functionality of SEF-GDoS (II) consists in providing information about the available bandwidth to the information service provider network. Besides that, SEF-GDoS (II) is the orchestrating SEF for the added value GDoS service and provides the "engaged signal" towards the information service provider, who "forwards" this information to the mass market customer. In order to provide the GDoS added value service, SEF-GDoS (II) interacts with SEF-GDoS (I) and (III) over SEF-GDoS specific interfaces (e.g. SOAP), gathers all information and derives the "engaged signal" feedback according to the SEF-GDoS internal logic.

FIGURE 73 provides a high-level view of SEF-GDoS functionalities and ETICS actors as well as the involved network elements. Moreover, the ETICS interfaces E1' and E7' illustrate the interaction between SEF-GDoS instances (I), (II) and (III).

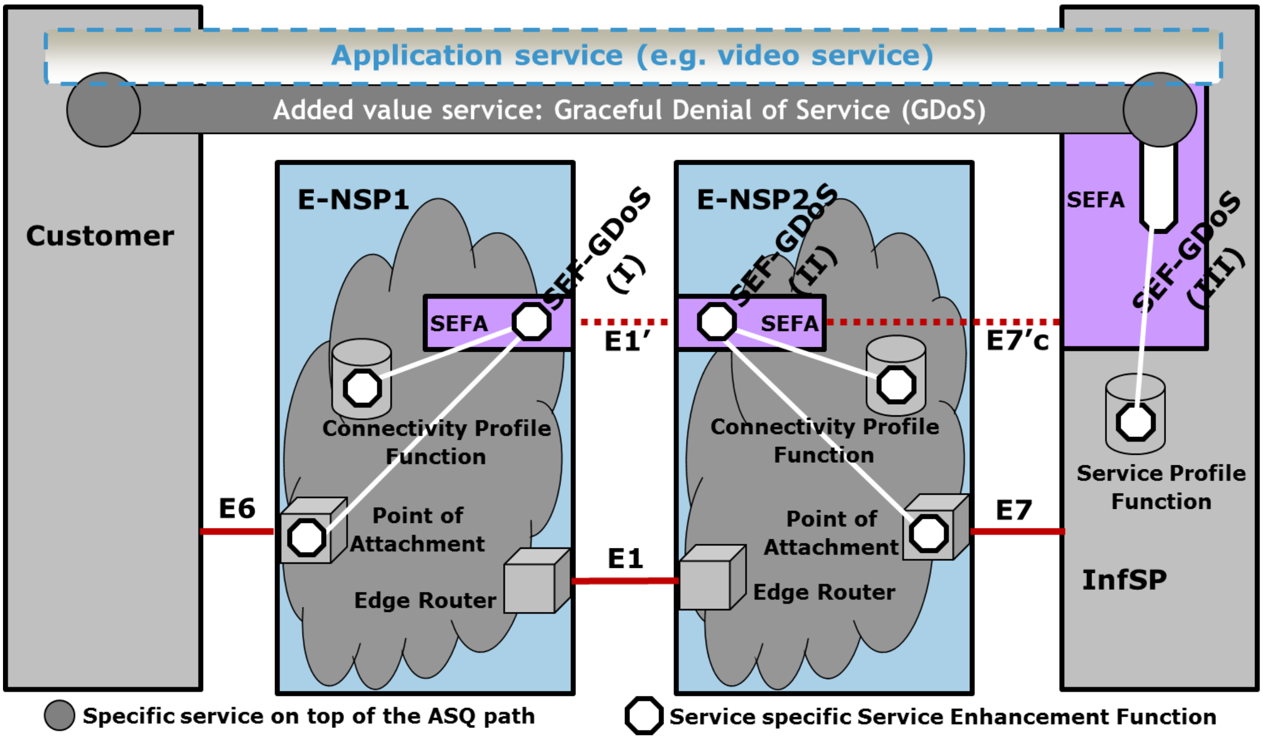

Application service (e.g. video service)
Added value service: Graceful Denial of Service (GDoS)
Customer
E-NSP1
E-NSP2
SEF-GDoS (I)
SEF-GDoS (II)
SEF-GDoS (III)
SEFA
SEFA
SEFA
E1'
E7'c
Connectivity Profile Function
Connectivity Profile Function
Point of Attachment
Point of Attachment
Edge Router
Edge Router
E6
E1
E7
Service Profile Function
InfSP
Specific service on top of the ASQ path
Service specific Service Enhancement Function

FIGURE 73: HIGH-LEVEL VIEW OF INVOLVED NETWORK ELEMENTS IN THE SEF-GDOS USE CASE

In the following, as part of the second phase, the high-level functional elements are mapped to the base-layer in order to specify more fine granular the needed functionalities of the specific SEF-GDoS instance and network elements:

a.) SEF-GDoS (I) – Mass market customer Edge NSP:

- In order to get the information about available bandwidth within the mass market customer access network, SEF-GDoS (I) interacts with the Connectivity Profile Function of the mass market Edge NSP and requests the needed parameters:
    - User access network profile information (e.g. IP address, quality units, etc).
    - Mass market customer access line parameters (e.g. used bandwidth, available bandwidth etc.)
- The Connectivity Profile Function is some kind of meta database for the SEF-GDoS and contains all information that has been provided from different entities of the mass market Edge NSP as for instance
    - Edge NSP customer data base,
    - System that provides information about the network capabilities in the mass market customer access network. (The RACS could be a candidate for such a system.)
    - Point of attachment, etc.

  using own service provider specific interfaces and data base requests (e.g. SOAP, mySQL).

b.) SEF-GDoS (III) – Information service provider:

- In order to gather the network capability requirements specified by the application service, SEF-GDoS (III) interacts with the Service Profile Function of the information service provider and requests the SEF-GDoS related parameters:
    - Information about the application the mass market customer has requested (e.g. video ID, service ID, etc.).
    - Related network capabilities requirements of this application (e.g. bandwidth).
- The Service Profile Function can be understood as an pre-filled database containing information service provider related application service and customer data.
- SEF-GDoS (III) “forwards” the “engaged signal” information to the mass market customer.

c.) SEF-GDoS (II) – Information service provider Edge NSP:

- In order to get the information about available bandwidth within the information service provider access network, SEF-GDoS (II) interacts with the Connectivity Profile Function of the Information Service Provider Edge NSP and requests the network access characteristics towards the Information Service Provider.

- The Connectivity Profile Function in SEF-GDoS (II) is once more some kind of meta database for the SEF-GDoS and contains all information that has been provided from different entities of the Information SP Edge NSP as for instance
  - Edge NSP information provider data base (containing for example used address ranges, agreements regarding bandwidth, SLAs and number of allowed flows etc.)
  - System that provides information about the network capabilities in the information provider access network. (The RACS could be a candidate for such a system.)
  - Point of attachment, etc.

  using own service provider specific interfaces and data base requests (e.g. SOAP, mySQL).
- SEF-GDoS (II) is the orchestrating SEF for the added value GDoS service and provides the “engaged signal” towards the information service provider.

A refinement of FIGURE 73 containing the different planes of the ETICS architecture and their interaction with SEF-GDoS is illustrated in FIGURE 74.

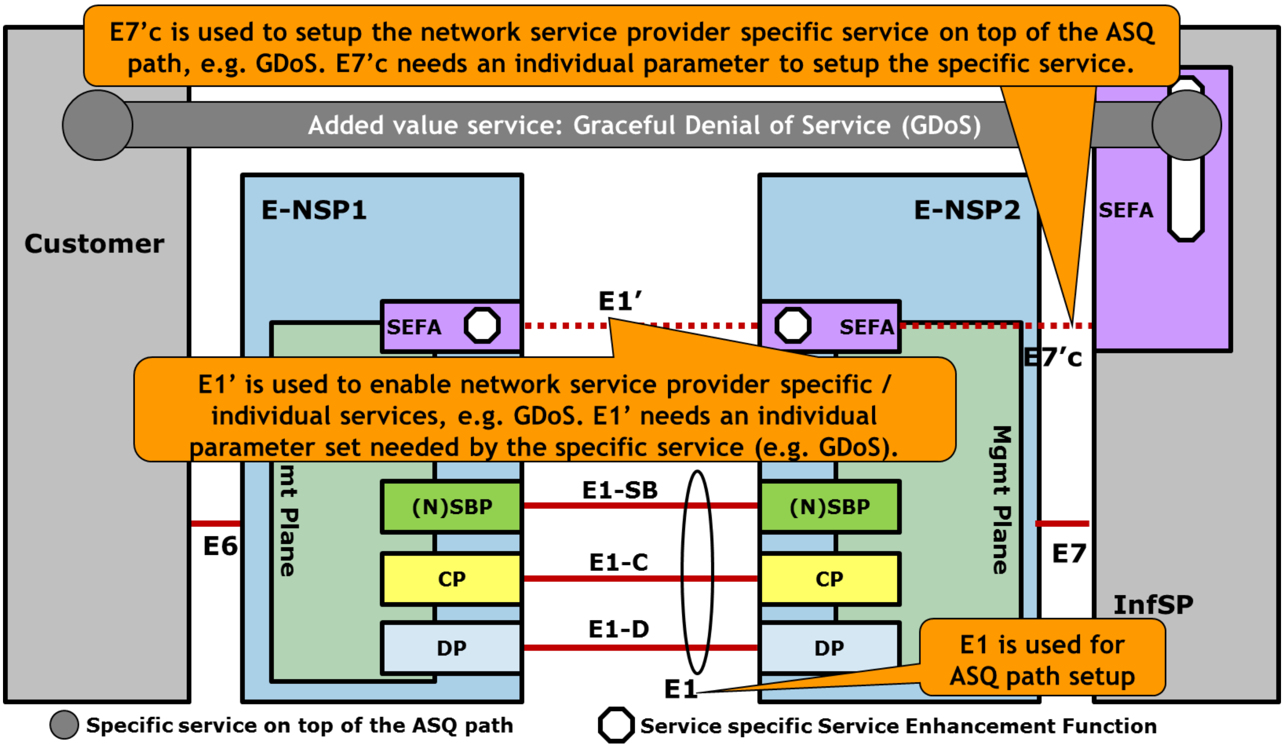

E7’c is used to setup the network service provider specific service on top of the ASQ path, e.g. GDoS. E7’c needs an individual parameter to setup the specific service.
Added value service: Graceful Denial of Service (GDoS)
Customer
E-NSP1
E-NSP2
SEFA
SEFA
SEFA
E1’
E7’c
E1’ is used to enable network service provider specific / individual services, e.g. GDoS. E1’ needs an individual parameter set needed by the specific service (e.g. GDoS).
Mgmt Plane
Mgmt Plane
(N)SBP
(N)SBP
E1-SB
CP
CP
E1-C
DP
DP
E1-D
E1
E6
E7
InfSP
E1 is used for ASQ path setup
● Specific service on top of the ASQ path
○ Service specific Service Enhancement Function

FIGURE 74: GDOS USE CASE SPECIFIC SEFS IN THE ETICS REFERENCE MODEL AND REFERENCE POINTS

The used interface E1’ of edge NSP 2 to request access network capabilities of the mass market customer from edge NSP1 comprises the following parameters:

- user_IP

The used interface E1’ of edge NSP 1 to reply access network capabilities of the mass market customer to edge NSP 2 comprises the following parameters:

- user_IP, available_bandwidth, used_bandwidth, level_best_effort, ASQ_info

The used interface E7' of information service provider to transmit network capability requirements of the application service to edge NSP 2 comprises the following parameters:

- user_IP, user_nsp_ID, required_bandwidth, InfSP_IP, InfSP_passphrase

The used interface E7' of edge NSP 2 to provide the Graceful Denial of Service information to the information service provider comprises the following parameters:

- user_IP, service_possible, ASQ_info

5. Realisation of SEF-layer:

The realisation of the SEF-GDoS use case has been implemented by means of web-services. The more detailed description of the Graceful Denial of Service implementation comprising the several SEF modules and interfaces is given in D5.3 section 4.1.1.1.

11.7.2. SEFA BASED USE CASE 2 – MANAGED CONNECTIVITY SERVICE USING PRE-ESTABLISHED ASQ PATHS

In Chapter 7 of this document so called ETICS bootstrapping scenarios are discussed, starting with some very simple NSP-to-NSP interconnection scenarios. While there are many interesting problems related to inter-network routing, a very simple bilateral business relationship between only two network service providers (NSPs) is enough to enable assured service quality (ASQ) traffic exchange to take place.

An exemplary realization of such a simple bootstrapping scenario is for instance when a business customer of one ETICS NSP buys ASQ connectivity to a particular business customer in the network of another ETICS NSP, and vice versa under the assumption that IPNP ("Initiating party pays principle") is used.

In the following the realisation of such a bootstrapping scenario - "Managed connectivity service using an ASQ path – IPNP charging" use case - will be described by means of the SEFA cookbook.

0. Goal of SEF use case:

The goal of this specific SEF use case "Managed connectivity services using an ASQ path – IPNP charging" has the intention to apply the IPNP charging principle for an ETICS ASQ service, as shown in FIGURE 75.

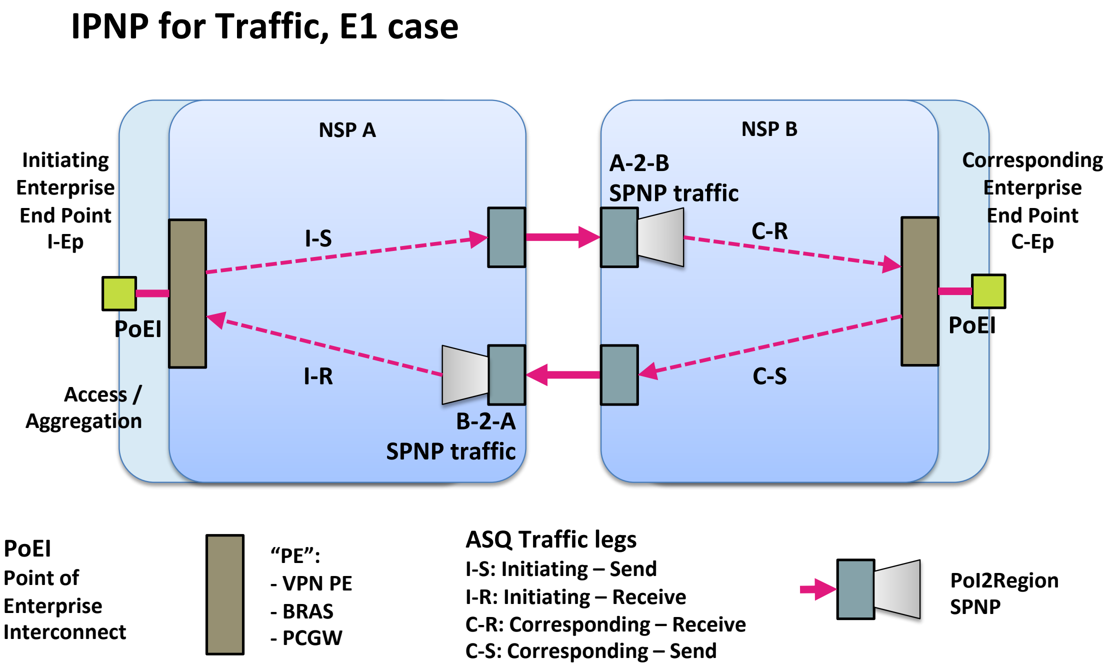

IPNP for Traffic, E1 case
NSP A
NSP B
Initiating Enterprise End Point I-Ep
Corresponding Enterprise End Point C-Ep
A-2-B SPNP traffic
B-2-A SPNP traffic
I-S
I-R
C-R
C-S
PoEI
PoEI
Access / Aggregation
PoEI Point of Enterprise Interconnect
"PE": - VPN PE - BRAS - PCGW
ASQ Traffic legs
I-S: Initiating – Send
I-R: Initiating – Receive
C-R: Corresponding – Receive
C-S: Corresponding – Send
PoI2Region SPNP

FIGURE 75: MANAGED CONNECTIVITY SERVICE USING AN ASQ PATH –IPNP CHARGING SCENARIO

1. Prerequisites of SEFA design:

For this example of a managed connectivity service scenario it is assumed that 2 Enterprise locations are connected to two different Edge NSPs. Between the PoEIs of these 2 Edge NSPS 2 directed ASQ paths have been established for instance in order to realize some kind of high quality communication between the two Enterprise locations.

In order to ease the billing for the needed ETICS connectivity services the enterprise wants to apply the Initiating Party Network Pays (IPNP) charging principle–compensating the SPNP charges in case where the initiating Enterprise location wants to pay the bill for both directions of traffic (Initiating Location => Corresponding Location and Corresponding Location => Initiating Location). In order to achieve this a per instance of the ASQ Managed connectivity Service has to be deployed in each Enterprise PoEI. (Note: The inter NSP ASQs are billed on a SPNP base between the both Edge NSPs.)

Since the PoEI charging has to be realized for certain destinations / regions and hence specific ASQ parameters a knowledge of the used ASQ path / SLA information is needed (e.g. amount of transmitted bytes, Source / Destination addresses, Price information etc.) and this use case can be considered as example of an "in ETICS" SEFA use case.

FIGURE 76 shows the involved actors which have to implement SEF-MCS related instances. (Note: For this symmetric example of Edge NSPs the needed functionality in both PoEIs is identical. The only difference exists in the fact that one of the NSP SEF realizations has to cover also the "orchestrating" SEF functions.)

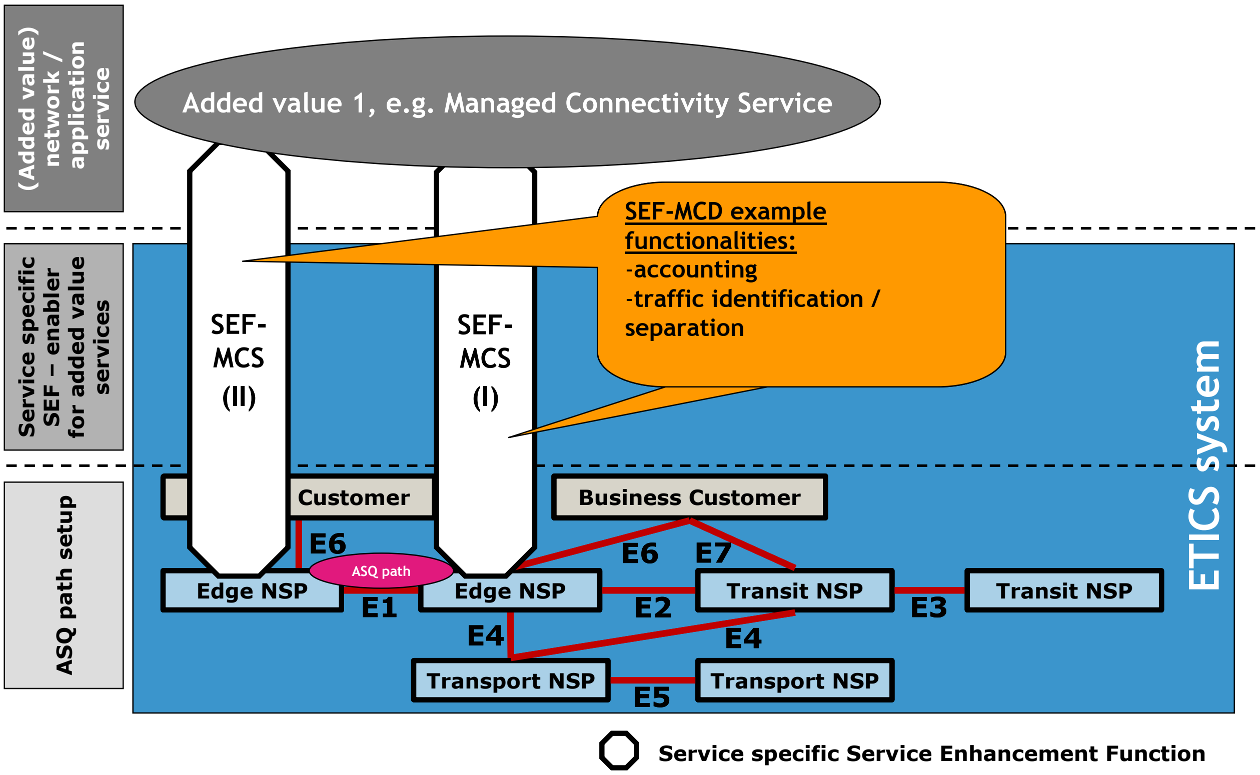

(Added value) network / application service
Service specific SEF – enabler for added value services
ASQ path setup
Added value 1, e.g. Managed Connectivity Service
SEF-MCS (II)
SEF-MCS (I)
SEF-MCD example functionalities:
-accounting
-traffic identification / separation
ETICS system
Customer
Business Customer
E6
E6
E7
ASQ path
Edge NSP
Edge NSP
Transit NSP
Transit NSP
E1
E2
E3
E4
E4
Transport NSP
Transport NSP
E5
Service specific Service Enhancement Function

FIGURE 76: ACTORS OF THE MANGED CONNECTIVITY SERVICE SEF USE CASE

2. Naming of SEF-layer:

As already introduced above the notation of Service Enhancement Function – Managed Connectivity Service (SEF-MCS) has been chosen.

3. Requirements on SEF-layer:

The added value service “Managed connectivity services using an ASQ path – IPNP charging” has the intention to realise an IPNP charging for the two ASQs of a Managed connectivity service between two Enterprise locations that are connected to different NSPs.

As already illustrated within the SEF-GDoS use case above, there exist different requirements on the different actors of the usage scenario that are described in the following:

- Edge NSP of Enterprise location I-Ep (IPNP “paying party”):
  - Connects Enterprise location I-Ep to the Managed connectivity service
  - Transport of managed connectivity service data.
  - Provides accounting information about I-Ep transmitted data belonging to the ASQ path (I-S => C-R).
  - SEF-MCS orchestrating functionality for realizing the IPNP principle
- Edge NSP of Enterprise location C-Ep (IPNP “Invited party”)
  - Connects Enterprise location C-Ep to the Managed connectivity service
  - Transport of managed connectivity service data.

  - Provides accounting information about C-Ep transmitted data belonging to the ASQ path (C-S => I-R).

- Enterprise customer / End point
  - No SEF functionality needed for this specific SEF use case (SEF-MCS).

4. Description of SEF-layer:

In order to realise this SEF-MCS added value service the different actors have to implement specific SEF-MCS instances in their network. These actor role specific SEF-MCS instances are depicted in FIGURE 76 by means of numbering I and II. The functionality of these instances depends on the role of the actor, which is described in the following with focus on the high-level functional elements of the first phase:

a.) SEF-MCS (I) – Edge NSP of Enterprise location C-Ep:

- The high-level functionality of SEF-MCS (I) consists in providing information about the amount of transmitted C-S (and if needed also received C-R) bytes belonging to the corresponding managed connectivity service ASQ path(s).

b.) SEF-MCS (II) – Edge NSP of Enterprise location I-Ep:

- The high-level functionality of SEF-MCS (II) consists in providing information about the amount of transmitted I-S (and if needed also received I-R) bytes belonging to the corresponding managed connectivity service ASQ path(s).
- Besides that, SEF-MCS (II) is the orchestrating SEF for the added value service “MCS using IPNP” and gathers the complete accounting statistics of both involved NSPs in order to derive the IPNP bill and make sure that the Inter-NSP SPNP bills are treated correctly. In order to provide the MCS added value service, SEF-MCS (I) interacts with SEF-MCS (II) over SEF-MCS specific interfaces (e.g. SOAP), gathers all information and derives the “IPNP statistics” feedback to both NSPs according to the SEF-MCS internal logic.

FIGURE 77 provides a high-level view of SEF-MCS functionalities and ETICS actors as well as the involved network elements. Moreover, the ETICS interfaces E1’ illustrate the interaction between SEF-MCS instances (I) and (II).

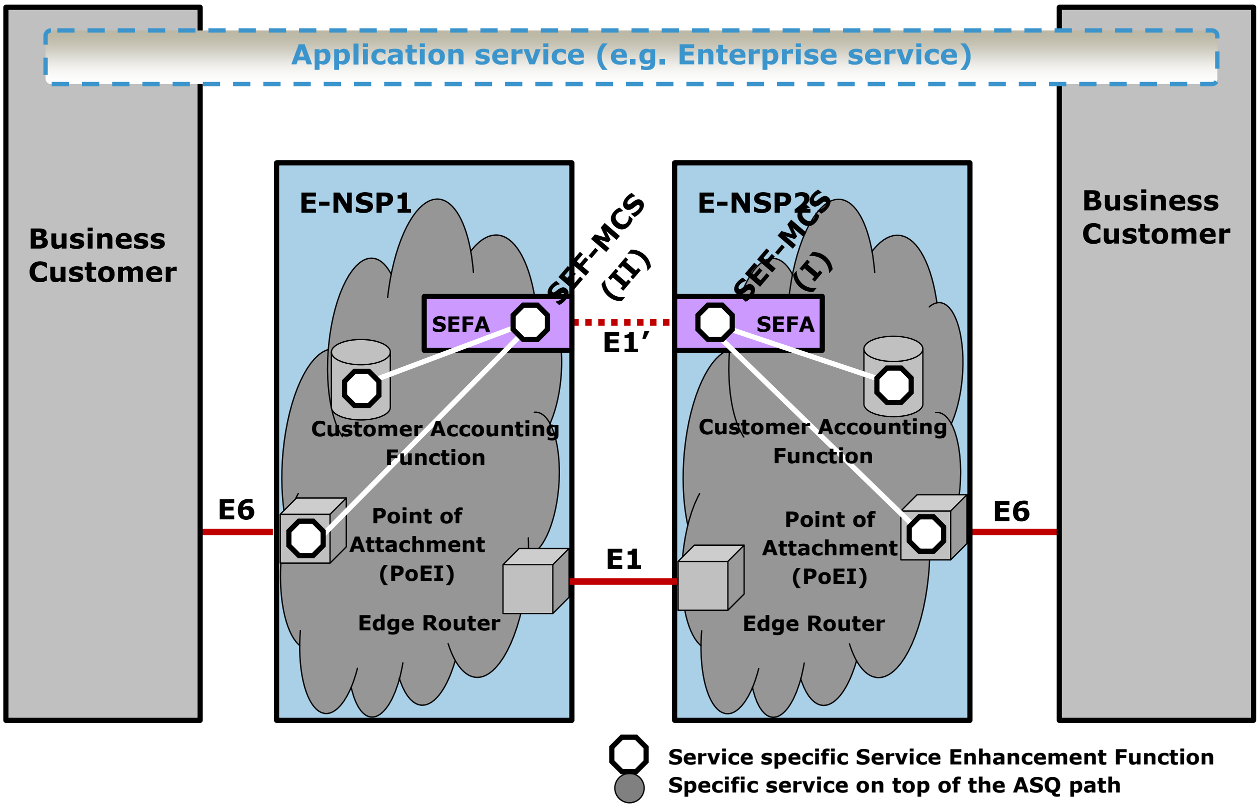

Application service (e.g. Enterprise service)
Business Customer
E-NSP1
SEF-MCS (II)
SEFA
E1'
E-NSP2
SEF-MCS (I)
SEFA
Customer Accounting Function
Point of Attachment (PoEI)
Edge Router
E6
E1
Customer Accounting Function
Point of Attachment (PoEI)
Edge Router
E6
Business Customer
Service specific Service Enhancement Function
Specific service on top of the ASQ path

FIGURE 77: HIGH-LEVEL VIEW OF INVOLVED NETWORK ELEMENTS IN THE SEF-MCS USE CASE

In the following, as part of the second phase, the high-level functional elements are mapped to the base-layer in order to specify more fine granular the needed functionalities of the specific SEF-MCS instance and network elements:

a.) SEF-MCS (I) – Edge NSP of Enterprise location C-Ep (IPNP "invited party"):

- In order to get the information about the amount of transmitted C-S bytes, SEF-MCS (I) interacts with the Customer Accounting Function inside the PoEI of Edge NSP B and requests the needed parameters. The Customer Accounting Function is some kind of statistics database for the SEF-MCS and contains all information regarding the send / received bytes belonging to the corresponding ASQs of the managed connectivity service.
    - Edge NSP customer data base,
    - System that provides information about the send and received network traffic in the NSP B network with respect to a given combination of Source and Destination addresses. (The NASS could be a candidate for such a system.)
    - Corresponding Point of attachment (PoEI)
- SEF-MCS (I) has also to provide an interface (E1') towards SEF-MCS (II) in order to send the gathered information to the orchestrating SEF-MCS component.
    - This SEF functionality can be implemented using service provider specific interfaces and data base requests (e.g. SOAP, mySQL).

- Another interface towards the ETICS portal (ETICS data base) is needed in order to look up the needed ASQ / SLA related information for the involved ASQ paths of the managed connectivity service.

b.) SEF-MCS (II) – Edge NSP of Enterprise location I-Ep (IPNP "Initiating party")

- SEF-MCS (II) has to provide the equal functionalities as SEF-MCS (I).
- In addition to that the orchestrating unit of SEF-MCS has to be performed:
  - Gathering of ASQ SLA specific information (e.g. price, data volumes, communication end points, duration of connections etc) for both used ASQs of the managed connectivity service.
  - Compiling traffic statistics for both used ASQs.
  - Deriving of overall costs that have to be paid according to IPNP by the "initiating party" of the managed connectivity service.

    (Note: Additional functionality can be added depending on the concrete SEF-MCS use case, e.g. traffic balance for SPNP based on the statistics of involved enterprise endpoints.)

5. Realisation of SEF-layer:

A realisation of the SEF-MCS use case could also be implemented by means of web-services, enhanced with needed functionality for requesting traffic statistics inside the different involved PoEIs.

Another realisation of SEFA-based / SEFA-related additional functionality realised on top of generic ASQ paths is capacity sharing, which will be illustrated in more details in Section 6.2.

11.8. SCALABILITY

This section will present complementary material for Section 7.

11.8.1. PASSIVE NMON SUB-SYSTEM SCALABILITY DETAILS

11.8.1.1. Architectural Level

As already mentioned, the fundamental architecture has been developed and presented to all partners early in the project in order to have enough time to find insufficiencies of the concept. The result of the discussions has already been presented in detail in Section 5.7 ("Monitoring") and the rest of this section summarizes this architectural view with a focus on scalability concerns.

The basic architecture of the NMON sub-system is the same for active and passive monitoring solutions: it follows an autonomous methodology (which means: every provider deploys – and is responsible for – their own monitoring equipment) in a hierarchical architecture.

Hierarchical in this sense means that monitoring requests are directed towards the Monitoring Proxy (M-Proxy) Function (instances of which can be replicated within a provider's network) rather than directly towards the probes which would not scale well (cf. FIGURE 46).

The collector function is used to collect, correlate, and evaluate the data retrieved from the probes. The collector function also needs to talk to the ETICS core system (via M3) to retrieve necessary data such as routes of established ASQ paths and SLA/SLS information. It is open where instances of the Collector Function shall be deployed, specifically there is no need to deploy them inside the customer's network and above this, for scalability reasons, it is also possible to replicate instances of the Collector Function.

The introduction of the "Proxy Layer" is the main source for a scalable system on the architectural level because by replicating instances of the M-Proxy function, it is possible to integrate "redirect functionality" into the protocol and in this way implement load-balancing and traffic engineering functionality. Not only could monitoring requests be redirected to another proxy, probably nearer to the probe, but monitoring traffic could also be directed to an alternative path through the network to the respective probe.

We argue that with this concept in place, possibly utilizing replication of the monitoring data, much like it is done nowadays for CDNs, the architecture of the network monitoring sub-system is sufficiently scalable.

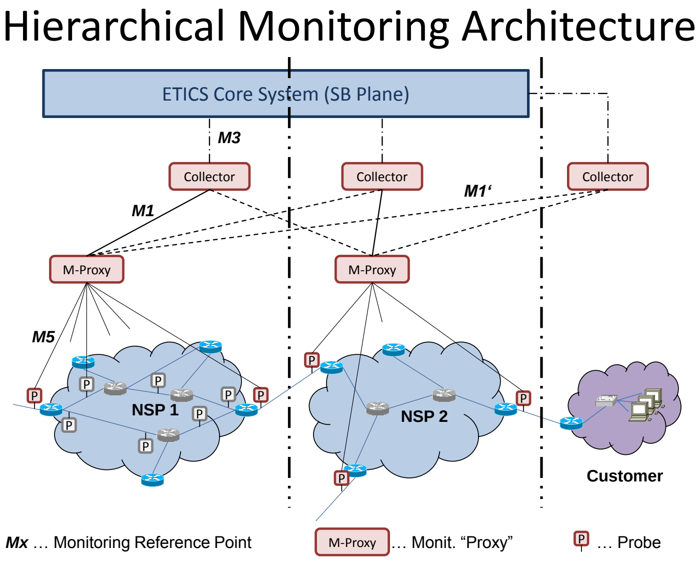


FIGURE 78: HIERARCHICAL MONITORING ARCHITECTURE

11.8.1.2. Protocol Level

We have met the efficiency criterion by specifying an optimized, binary protocol for the transfer of monitoring data between the Probe and the Collector Function (usually via the M-Proxy Function) to avoid the overhead of the nowadays very popular text-based protocols which can easily add an overhead of several hundred per cent (!) to the goodput data. Furthermore, there is no need to uselessly convert data back and forth from binary- to text representation (for transferring it) and then back to binary again at the receiver, thus saving CPU cycles. Taking into account that packet information of several million packets needs to be transferred per monitoring request in a real-world deployment, this indeed makes a tremendous difference.

In our proposed protocol, we group together packet information of a certain number of packets (currently 64K) into a chunk of reasonable size (in this case max. 8 MB) which can be assembled in memory. The advantage of this concept is threefold:

a) Preparing a chunk instead of single packets improves efficiency,

b) Less operating system calls are needed, and

c) The OS can better optimize the transmission of bigger junks, both in the stack and on the network level (e.g. by utilizing jumbo frames).

We considered using a lightweight compression, which could probably increase overall performance, but

a) This would need some further research and trials (for which we don't have enough resources), and

b) We expect the performance improvement to be moderate because the binary data already has much higher entropy than text which lowers the efficiency of compression.

Above that, in order to meet the extensibility criterion, all relevant messages have hooks in the grammar to be extended with further information elements and the protocol version is exchanged at startup such that even the message flow can be changed while remaining compatible with existing implementations. A protocol following this recommendations, named ETICS monitoring information transfer (EMONIT), has already been specified in ABNF (cf. [RFC2234]) and prototypically implemented in the course of WP5 work and is described in more detail in [ETICS-D5.8], sub-section "EMONIT (M1, M5) Protocol Specification".

11.8.1.3. Algorithmic Level

The more prominent optimizations in this category are:

Multi-threading:

One of the core techniques used to help in scaling up the system to higher packet rates is multi-threading. Currently, the hardware architecture tends to offer more and more processing cores per CPU package. Consequently, the focus of the optimization was on developing algorithms supporting multi-threading. At the same time, the developed algorithms get along with as little memory usage as possible, because especially with multi-threaded algorithms, the memory access quickly becomes a bottleneck.

All developed algorithms are multi-threading enabled where possible and meaningful, which are:

- Capturing packets ("capture-thread")
- Disk-access: read / write packet chunks ("reader / writer threads")
- Protocol ("protocol client / server threads)
- Collector[23]

Filtering:

On high bit-rate links, standard hardware[24] is not able to cope with the extremely high, worst-case packet rates. E.g. on a only 1 Gbps link, worst-case packet rates of up to 1.4M (millions!) of packets/second could occur. Therefore, it is inevitable to filter out some packets, which must be done based on the hash because we must keep the same packet at all observation points, and we can only calculate QoS metrics between pair-wise matching packets (in our system, the same packet always has the same hash, that's the key

[23] The Collector in the prototype is not multi-threaded, but the algorithm developed allows for a multi-threaded implementation.

[24] For cost reasons, standard hardware (in contrast to specialized hardware like capture cards) is our target

strategy we use). By applying the same rules on all observation points, we make sure to always keep the same packets. As this filter is based on the hash, this is called the ***hash-filter***.

When a Collector wants to perform a measurement, it usually only wants to do so on a sub-set of packets, e.g. belonging to an ASQ-path. To avoid the transfer of unnecessary data, a filter is applied at the probe, selecting only the information of requested packets. This filter is based on properties of the packet, like IP-address and/or MPLS-label or the like, and therefore it is called ***property-filter***.

Correlation:

Correlation is the process of finding the packet data of the same packet from different observation points and calculating QoS metrics (and probably extracting other information) of matching pairs. At such high packet rates, this is a quite challenging task!

The key point of our developed algorithm works by splitting the requested time-span into smaller time-spans (like 1 second) and perform optimized correlation operations on those. This is described in detail in [ETICS-D5.7], sub-section on “library u_nmon” in enclosing section about implementations. In the mentioned deliverable we also show, by analysing the used sub-algorithms, that the solution scales well with respect to increasing packet rates and also with respect to larger time-spans to be analysed.

Side-note about Collectors:

The correlation task, which is done inside the Collector, is very processing power demanding. As already explained in section 7.4.2 (scalability on an architectural level), instances of the Collector can easily be replicated such that there is no scalability issue in this respect.

11.8.2. Network Efficiency Gains

This section complements the material presented in Section 7.8, which will concentrate on the used modelling methodology as well as on delay distributions with/without priority classes.

11.8.2.1. Modelling Methodology

In this section the modelling methodology will be explained in more detail. The main idea behind is the assumption that it is possible to classify the traffic into traffic classes and handle them differently to differentiate among them. One of the main differentiation metrics in communication networks is the QoS requirements for the different traffic classes. By exploiting these differences it is possible to dimension the network resources so that real gains are achievable. Before we discuss any examples we shall state the modelling foundation in more depth.

At most we divide the traffic into four traffic classes named here:

- Real time (RT)
- Non real time (NRT)
- Best effort(BE)
- Background (BG)

Several other options for differentiation may also be possible. It is well known that the access part of the network will benefit most from differentiation mainly because the performance gains will be largest here. It is however not necessary to have a high number of differentiating traffic to take out the differential potential. Often a two class differentiation option could be sufficient, i.e. the differentiation is done according to traffic with high delay requirements or not.

For real time (RT) traffic the end-to-end delay and delay variation is usually the most important QoS parameter because for this traffic types the original bit-stream has to be recreated at the receiver site. Since the variable part of the end-to-end delay breaks down to the variable delay/waiting time in each router/multiplexer it will be beneficial to control the delay/waiting time for each router/multiplexer in a particular end-to-end path to control the entire end-to-end delay for RT traffic.

If no differentiation is implemented all traffic will get the same treatment in the network nodes and hence the strongest delay requirements must yield for all traffic, while if differentiation option is implemented only the RT traffic needs to comply with the strongest delay requirements while the other traffic classes may have much weaker requirements.

The way we compare statistical multiplexing with or without differentiation, is done by considering the two generic router/multiplexer model briefly described below. As a reference model we shall take the case without differentiation and compare it with scenarios where differentiation is performed. By making comparison of these two multiplexing models we may quantify the possible capacity gains by introducing differentiation.

Multiplexing without differentiation

When differentiation option is not used all traffic will be served as a single class and the strongest QoS requirements will determine the maximum offered load.

The router/multiplexer model without any form of differentiation is taken as single server queue where all the traffic is served in an FCFS (First Come First Served) manner as shown in FIGURE 79 below.

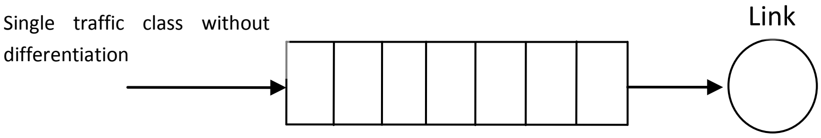


FIGURE 79 MULTIPLEXING MODEL WITHOUT CLASSES

Multiplexing with differentiation according to P priority classes

One of the most effective differentiation methods is obtained by assigning priority to the traffic classes. On the IP level this is done on a per packet basis. In practice the priority will be of non pre-emptive type, since pre-emption of an on-going packet transmission will not be effective.

In the general case we consider differentiation based on a non pre-emption P priority traffic classes regime as shown in FIGURE 80 below.

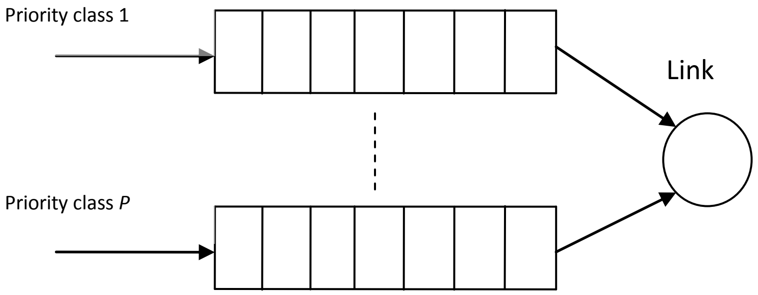


Figure 80 Multiplexing model with P priority classes

At most we consider cases with the four traffic classes defined in the introduction namely: RT (real time), NRT (non-real time), BE (best effort, basic) and BG (best effort, background) where the priority is given by the order of the line-up.

Modelling assumptions

The dimensioning modelling is based on the following assumptions:

- The traffic is classified into P priority classes.
- Packets arrive according to Poisson processes for all the priority classes and the arrival rates for priority class p are λ_p; $p = 1,..,P$.
- The link capacity is C (given in bits/sec).
- The packet length distribution is identically for all priority classes with mean P_L (given in bits).
- The offered traffic may be written $B_p = \lambda_p P_L$ and the corresponding loads are $\rho_p = {}^{B_p}\!/_{C}$ and further we also assume a stable system, i.e. we require that $\rho = \sum_{p=1}^{P} \rho_p < 1$.

Hence, with these definitions we have the mean service times for packets as $b = \mu^{-1} = {}^{P_L}\!/_{C}$ (and where μ is the corresponding service rate).

Dimensioning based on delay requirements

For dimensioning purposes we set up the following delay requirement: The waiting time for IP-packets of the p-priority class over a generic (access or core) router/multiplexer should be less than d_p (ms) for more than $1-\varepsilon_p$ part of the packets. We shall apply the complementary waiting times distribution to determine the required capacity for both the cases; with and without priority.

For differentiation to have any effect the delay requirements have to be different for different priority classes, i.e. by allowing a much looser delay requirement for the lower priority classes.

Dimensioning without differentiation

With the assumption made we may use the CDF of the waiting time distribution of an M/G/1 queuing model to find the required capacity. Without any differentiation we have to use the strongest delay requirements for all the traffic, i.e. d_1 and ε_1. The CDF of the waiting time in an M/G/1 queue is on the form $W^C(t) = F(\frac{t}{b}, \rho)$ where b is the mean service time and $F(x, \rho)$ is the PDF of the waiting time of the M/G/1 queue with (mean service time taken to unity without any priority) and load ρ. Hence the required capacity C_{np} (no priority) is given as

$$C_{np} = C(B, d_1, P_l, \varepsilon_1)$$

where $B = \sum_{p=1}^{P} B_p$ is the total offered traffic and $C = C(B, d, P_l, \varepsilon)$ is the solution of the functional equation $F({}^{Cd}\!/_{P_l}, {}^{B}\!/_{C}) = \varepsilon$.

Similar if the capacity C is given and we would rather like find the maximum traffic that may be carried without breaking the delay requirements; then the sought bandwidth is found as:

$$B_{np} = B(C, d_1, P_l, \varepsilon_1)$$

where $B = B(C, d, P_l, \varepsilon)$ now is the solution of the functional equation $F({}^{Cd}\!/_{P_l}, {}^{B}\!/_{C}) = \varepsilon$.

Dimensioning based on differentiation with P priority classes

The capacity dimensioning is based on the waiting times for priority class p in an M/G/1 non pre-emptive priority queuing model with P priority classes, where we also assume the priority classes have identical service time distribution. For this model the CDFs of the waiting times is given in terms of functions related to the following expressions:

- The PDF if the waiting times for the highest priority classes is given by $W_1^C(t) = \frac{\rho}{\rho_1} F(\frac{t}{b}, \rho_1)$ where $F(x, \rho_1)$ PDF of the waiting time in a M/G/1 queue with (mean service time taken to unity without any priority) and load ρ_1.
- Similar the PDF if the waiting times for the lower priority classes written as: $W_p^C(t) = \frac{\rho}{\rho_p^+} F_2(\frac{t}{b}, \rho_{p-1}^+, \rho_p)$ for $p = 2, \ldots, P$, and where $\rho_p^+ = \sum_{j=1}^{p} \rho_j$ and further $F_2(x, \rho_1, \rho_2)$ is the corresponding CDF of the waiting time for the of low priority class in a two priority M/G/1 queue with (mean service time taken to unity) and high priority load ρ_1 and low priority load ρ_2 respectively.

The required capacity C_{pr} (with priority) may be found as the smallest possible value of the capacity C so that all the delay requirements for all the priority classes are fulfilled, i.e.

$\frac{B}{B_1} F(\frac{Cd_1}{P_l}, \frac{B_1}{C}) \le \varepsilon_1$ and $\frac{B}{B_p^+} F_2(\frac{Cd_p}{P_l}, \frac{B_{p-1}^+}{C}, \frac{B_p}{C}) \le \varepsilon_p$ for $p = 2, \ldots, P$ and where $B_p^+ = \sum_{j=1}^{p} B_j$.

The solution of this optimisation problem above may be written as

$$C_{pr} = \max\{C_1, C_2, ..., C_P\}$$

where $C_1 = C(B_1, d_1, P_l, \varepsilon_1 \frac{B_1}{B})$ is and $C_p = C_2(B^+_{p-1}, B_p, d_p, P_l, \varepsilon_p \frac{B^+_p}{B})$ for $p = 2,\ldots,P$ and where the functions $C(B, d, P_l, \varepsilon)$ is defined implicit as the solution of $F(Cd/_{P_l}, B/_C) = \varepsilon$ as above and further $C_2(B_1, B_2, d, P_l, \varepsilon)$ is defined implicit as the solution of $F_2(C_2 d/_{P_l}, B_1/_{C_2}, B_2/_{C_2}) = \varepsilon$

Similar if the capacity C is given and we would rather like find the maximum traffic that may be carried without breaking the delay requirements a similar approach is possible. Suppose further that the percentage distribution among the different traffic classes; f_j is given, i.e. we have $B_j = f_j B$ (where $\sum_{j=1}^{P} f_j = 1$), and we would maximize the total offered traffic B so that all the delay requirements are fulfilled (for all the priority classes). Then the maximum offer traffic B_{pr} may be found as the smallest offered traffic B so that all the delay requirements for all the priority classes are fulfilled, i.e.

$F(\frac{Cd_1}{P_l}, \frac{Bf_1}{C}) \le \varepsilon_1 f_1$, and $F_2(\frac{Cd_p}{P_l}, \frac{Bf^+_{p-1}}{C}, \frac{Bf_p}{C}) \le \varepsilon_p f^+_p$ for $p = 2,\ldots,P$ where $f^+_p = \sum_{j=1}^{p} f_j$. The solution of this optimization problem is then given by:

$$B_{pr} = \min\{B_1, B_2, ..., B_P\}$$

where $B_1 = \frac{1}{f_1} B(C, d_1, P_l, \varepsilon_1 f_1)$, and $B_p = B_2(C, f^+_{p-1}, f_p, d_p, P_l, \varepsilon_p f^+_{p\,1})$ for $p = 2,\ldots,P$ and

where $B = B(C, d, P_l, \varepsilon)$ (as above) is the solution of the functional equation $F(Cd/_{P_l}, B/_C) = \varepsilon$ and further $B_2 = B_2(C, f_1, f_2, d, P_l, \varepsilon)$ is the solution of the functional equation $F_2(\frac{Cd}{P_l}, \frac{B_2 f_1}{C}, \frac{B_2 f_2}{C}) = \varepsilon$.

In the numerical examples we have taken the packet length to be exponentially distributed and the actual form of the function for this case is given in the Annex 11.8.2.3.

Differentiation gains

To be able to quantify the gain by introducing differentiation we define some of the interesting performance measures. One important parameter is the maximum loading for the multiplexer under the different options:

$$\rho \max_{np} = \frac{B}{C_{up}}$$

for multiplexing without differentiation and

$$\rho \max_{pr} = \frac{B}{C_{pr}}$$

for multiplexing with differentiation. The two last expressions give the maximum load as function of the overall traffic B, however, similar expression are also possible when the capacity C is the free variable:

$$\rho \max_{np} = \frac{B_{np}}{C}$$

for multiplexing without differentiation and

$$\rho \max_{pr} = \frac{B_{pr}}{C}$$

for multiplexing with differentiation.

An even more interesting measure is to compare the ratio of capacity without and with the differentiation option for quantifying the capacity gain as:

$$DiffGainCapacity = \frac{C_{np}}{C_{pr}}$$

Similar we may also define the traffic gain as the increase in total offered traffic by applying differentiation relative to the offered traffic without any form of differentiation, i.e. we take:

$$DiffGainTraffic = \frac{B_{pr}}{B_{np}}$$

Observe the two expressions for the gains are not equal. While the first expression is a function of the traffic the latter have the capacity as the independent variable and hence the expressions are by definition different.

11.8.2.2. Key results

By the modelling approach described in the previous subsection different scenarios are possible and can be analysed. The numbers of traffic classes may be arbitrary, however we will limit to at most four traffic classes as described in the introduction namely:

- Real time (RT)
- Non real time (NRT)
- Best effort (BE)
- Background (BG)

Before describing some of the scenarios considered, we shall first discuss the order of magnitude of the QoS requirements that are essential for the differentiation to be effective. The differentiation gains are possible to be obtained due to the rather huge differences in the delay requirements for RT traffic compared to elastic type of traffic. While the overall delay for RT traffic like conversational voice and video must be limited to just a few hundred millisecond end-to-end, elastic traffic may accept end-to-end delay of several seconds. Since the end-to-end packet delay adds up of variable waiting times in each multiplexer plus a fixed part that is traffic independent (transmission and propagation), the end-to-end delay requirements will be satisfied by setting appropriate waiting time bounds for each multiplexer. To fulfil the end-to-end delay requirements we propose to distinguish between typical core routers and access routes delay requirements i.e. allowing the access requirements by a decade. This may be argued by the fact that the numbers of core router hops end-to-end will be larger than the numbers of access router hops. Below we consider three scenarios with different distributions between the traffic classes.

The scenario I is the base scenario and includes the following for the parameter settings:

- 10% Real time (RT),
- 20% Non real time (NRT),
- 50% Best effort (BE) and
- 20% Background (BG)

Scenario II and III are variations around this base scenario. In scenario II, we start removing the BG traffic (but keeping the same relative size of the other traffic classes as in the Scenario I. In Scenario III, RT and NRT traffic is treated as on class and therefore, in this case, we end up with only two priority classes. We have defined two delay requirements; the strongest ones will correspond to high capacity core links while the more relaxed delay requirements will fit best for access links. For all the scenarios we have taken the mean packet length to be 10 kbits.

TABLE 1 SCENARIO I: DIFFERENTIATION WITH FOUR TRAFFIC CLASSES

	Real Time (RT)	**Non real time (NRT)**	**Best effort (BE)**	**Background (BG)**
Traffic distribution	10%	20%	50%	20%
Delay requirements for core routers	0.1ms	0.5ms	1.0ms	5.0ms
Delay requirements for access routers	1.0ms	5.0ms	10ms	50ms
Percentage of traffic that may break the requirements	1%	1%	1%	5%

TABLE 2 SCENARIO II: DIFFERENTIATION WITH THREE TRAFFIC CLASSES

	Real Time (RT)	**Non real time (NRT)**	**Best effort (BE)**
Traffic distribution	12.5%	25%	62.5%
Delay requirements for	0.1ms	0.5ms	1.0ms

core routers			
Delay requirements for access routers	1.0ms	5.0ms	10ms
Percentage of traffic that may break the requirements	1%	1%	1%

TABLE 3 SCENARIO III: DIFFERENTIATION WITH TWO TRAFFIC CLASSES

	Real Time (RT) and Non real time (NRT)	Best effort (BE)
Traffic distribution	37.5%	62.75%
Delay requirements for core routers	0.1ms	1.0ms
Delay requirements for access routers	1.0ms	10ms
Percentage of traffic that may break the requirements	1%	1%

FIGURE 81-FIGURE 85 below summarise the numerical results base on the three differentiation scenarios chosen. The first figure give the required total capacity needed to carry the traffic for all three scenarios in addition to the case without any differentiation, (the upper curve in FIGURE 81).

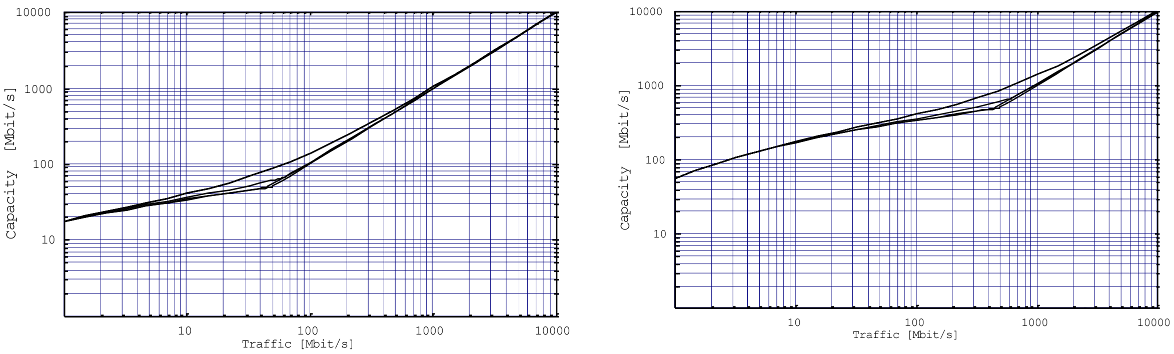


Figure 81 Required capacity as function of traffic, upper curve no differentiation, then two classes (scenario III) and three classes (scenario II), and lowest curve four classes (scenario I). Left figure for access link and right core link.

As expected we observe that scenario I gives the lowest capacity while the case without differentiation give the highest capacity. This figure shows the potential the differentiation options have and we observe that the range where differentiation is effective is somehow limited, like in the range 5Mbit/s-500 Mbit/s for typical access links/routers and 50Mbit/s-5Gbit/s for typical core links/routers. However, these intervals are well in the range of what is interesting for typical networks.

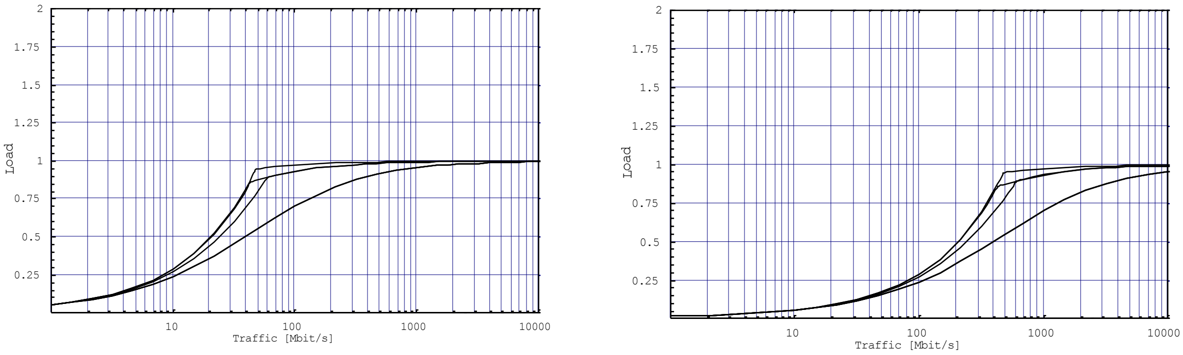


Figure 82 Maximum loading *as function of traffic, lower curve no differentiation, then two classes (scenario III) and three classes (scenario II), and upper curve four classes (scenario I). Left figure for access link and right core link.*

FIGURE 82 we give the corresponding maximum loading as function on the offered traffic. Sa expected we observe the differentiation have a quite large effect in the range of capacity mentioned above. Another observation is that the difference between the scenarios is not that big. However, scenario I (with four traffic classes) will give the highest utilization and therefor will give highest gain. This is even more clearly seen in FIGURE 83 where the actual magnitude of the gain is depicted.

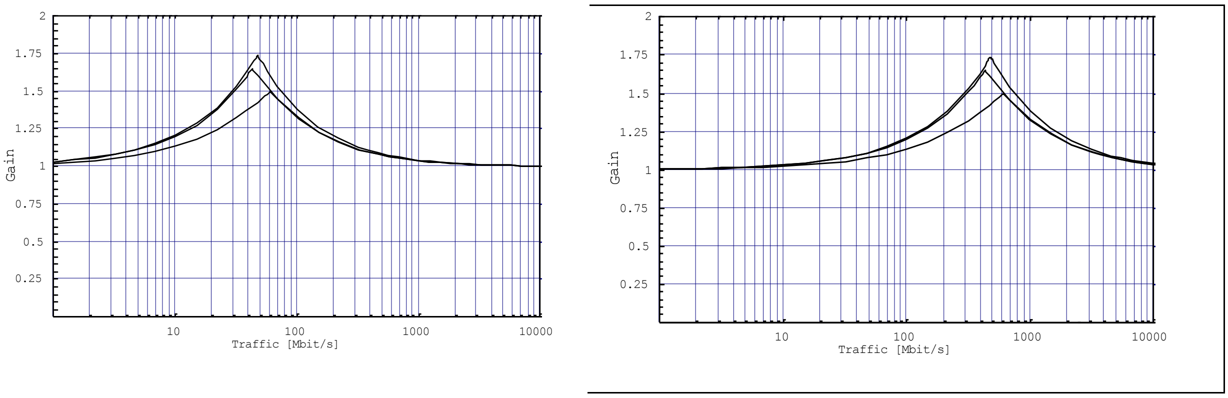


Figure 83 Capacity Gain *as function of traffic, lower curve two classes (scenario III) and three classes (scenario II), and upper curve four classes (scenario I). Left figure for access link and right core link.*

We see that scenario I and II almost give the same gain, while the two class option (scenario III) will have notable lower gain for low traffic i.e. in the range less than 50Mbit/s for access links and less than 500Mbit/s for core links.

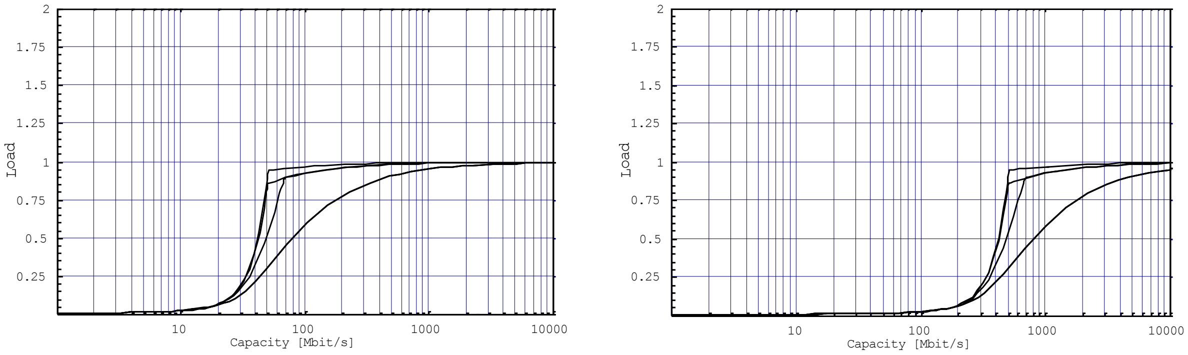


Figure 84 Maximum loading *as function of link capacity, lower curve no differentiation, then two classes (scenario III) and three classes (scenario II), and upper curve four classes (scenario I). Left figure for access link and right core link.*

If we consider the similar measures in terms of link capacity rather than offered traffic (FIGURE 84 and FIGURE 85) we observe quite similar performance however we observe that the gain is higher in terms of offered traffic (as function of the capacity) than seen above where the gain was measured in terms of capacity (as function of offered traffic). We also observe the difference in gain by scenario I and II the two traffic class case (scenario III) where the difference is notable for the lower bitrates, (see FIGURE 85).

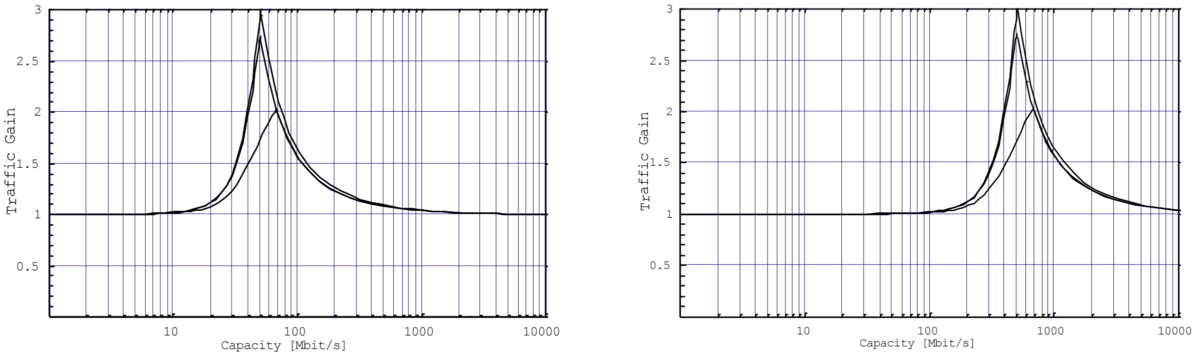


Figure 85 Traffic Gain *as function of link capacity, lower curve two classes (scenario III) and three classes (scenario II), and upper curve four classes (scenario I). Left figure for access link and right core link.*

11.8.2.3. M/M/1 case

In this annex we shall briefly discuss how to obtain the delay distribution for the M/G/1 queue with non-preemptive priority with two priority classes to obtain the functions $F(x,\rho)$ and $F_2(x,\rho_1,\rho_2)$ used for in the differentiation modeling. The method used is to invert the Laplace transforms. Explicit expression for the Laplace transform is found in the literature; see for instance the books of Takagi [Ta91] or Kleinrock [Kl76].

For the M/M/1 case the functions $F(x,\rho)$ and $F_2(x,\rho_1,\rho_2)$ may be found by inverting the corresponding Laplace transforms we find:

$$F(x,\rho)=\rho e^{-(1-\rho)x}$$

Further we may write $F_2(x,\rho_1,\rho_2)$ in terms of a geometric weighted sum of Bessel functions (of second type) by $S(\alpha,\beta)=\sum_{k=0}^{\infty}\beta^k I_k(2\alpha)$

$$F_2(x,\rho_1,\rho_2)=e^{-x(1+\rho_1)}\left(\rho\frac{1-\rho_1}{\rho_2}S(x\rho_1^{1/2},\rho_1^{1/2})+\frac{\rho^2-\rho_1}{\rho_2}S(x\rho_1^{1/2},\frac{\rho}{\rho_1^{1/2}})-I_0(2x\rho_1^{1/2})\right)$$

and where we take $\rho=\rho_1+\rho_2$. We may also express $F_2(x,\rho_1,\rho_2)$ in terms of trigonometric integrals

$$J(\alpha,\beta)=\frac{1-\beta^2}{\pi}\int_{\theta=0}^{\pi}\frac{e^{2\alpha\cos\theta}}{1-2\beta\cos\theta+\beta^2}d\theta$$

giving the following expression of $F_2(x,\rho_1,\rho_2)$:

$$F_2(x,\rho_1,\rho_2)=\frac{\rho^2-\rho_1}{\rho_2}H(\rho^2-\rho_1)e^{-x(\frac{1-\rho}{\rho}\rho_2)}+$$
$$\frac{1}{2}e^{-x(1+\rho_1)}\left(\rho\frac{1-\rho_1}{\rho_2}J(x\rho_1^{1/2},\rho_1^{1/2})+\frac{\rho^2-\rho_1}{\rho_2}J(x\rho_1^{1/2},\frac{\rho}{\rho_1^{1/2}})-(1-\rho)I_0(2x\rho_1^{1/2})\right)$$

A third expressions for $F_2(x,\rho_1,\rho_2)$ may be found in terms of integrals of type

$$I(\alpha,\eta)=\sqrt{\frac{2}{\pi}}\int_{z=0}^{\sqrt{8\alpha}}\frac{e^{-z^2/2}}{(1+\frac{\eta}{8\alpha}z^2)\sqrt{1-z^2/8\alpha}}dz$$

which gives the following expression of $F_2(x,\rho_1,\rho_2)$:

$$F_2(x,\rho_1,\rho_2)=\frac{\rho^2-\rho_1}{\rho_2}H(\rho^2-\rho_1)e^{-x(\frac{1-\rho}{\rho}\rho_2)}-\frac{1}{2}(1-\rho)e^{-x(1+\rho_1)}I_0(2x\rho_1^{1/2})+$$
$$\frac{1}{4}\frac{e^{-x(1-\rho_1^{1/2})^2}}{\sqrt{\pi}\sqrt{x\rho_1^{1/2}}}\left(\frac{\rho}{\rho_2}(1+\rho_1^{1/2})^2I(x\rho_1^{1/2},\frac{4\rho_1^{1/2}}{(1-\rho_1^{1/2})^2})-\frac{1}{\rho_2}(\rho_1^{1/2}+\rho)^2I(x\rho_1^{1/2},\frac{4\rho\rho_1^{1/2}}{(\rho_1^{1/2}-\rho)^2})\right)$$

All of the expressions for $F_2(x,\rho_1,\rho_2)$ are well suited for numerical calculations by applying numerical methods to calculate the integrals $J(\alpha,\beta)$ or $I(\alpha,\eta)$ above by.

Moreover the following uniform asymptotic expansion of the integral $I(\alpha,\eta)$ yields for all $\eta \geq 0$:

$$I(\alpha,\eta) \sim \sum_{k=0}^{\infty} \frac{(2k-1)!!}{k!} \frac{1}{(16\alpha)^k} I^k(\alpha,\eta) \text{ where}$$

$$I^k(\alpha,\eta) = (-1)^k \left(\frac{8\alpha}{\eta}\right)^k \left(F(2\sqrt{\frac{\alpha}{\eta}}) - \sum_{l=0}^{k-1} (-1)^l (2l-1)!! \left(\frac{\eta}{8\alpha}\right)^l \right) \text{ and}$$

$$F(x) = \sqrt{\pi} x e^{x^2} erfc(x)$$

11.8.3. Internet Tomography

A total number of 39346 AS were observed from the routing tables collected. We divided these AS into four categories:

- Single-homed AS: The AS which only one neighbour and this neighbour is a provider.
- Multi-homed Non Transit AS: The AS which has many neighbours and these neighbours are all providers.
- Multi-homed Transit AS: The AS which has many neighbours and is at least provider of one of them which is a single homed AS or a multi-homed non transit AS.
- Multi-homed Transit Transit AS: The AS which has many neighbours and is at least provider of one of them which is a multi-homed transit AS.

The table below shows the repartition of the number of AS by nature and continent.

Type	Total	North America	Europe	South America	Oceania	Africa	Asia	Antartica	unkown
Single-homed AS	**13942**	4876	4269	233	459	240	1291	13	2561
Multi-homed Non Transit AS	**18945**	7921	5975	374	382	186	1875	14	2218
Multi-homed Transit AS	**3401**	1014	1508	74	92	72	382	0	259
Multi-homed Transit Transit AS	**3058**	925	1274	102	97	68	428	0	164
Total	**39346**	**14736**	**13026**	**783**	**1030**	**566**	**3976**	**27**	5202

TABLE 4: AS REPARTITION BY NATURE AND CONTINENT

It appears first that continents are not equals, most AS are located in Europe and North America. Considering categories, less than 10% of AS are multi-homed transit transit AS (3058 out of 39346) which is interesting as these AS are the more likely to be the NSPs which would take part to the ETICS community. Moreover most AS (around 80%) are non-transit AS and could therefore be integrated into the community not as NSPs but as regions

When counting AS-AS links, we obtain that most inter AS links are established by AS located in Europe or North America, respectively 98616 and 72566 over 216470 observed links. This confirms these two markets are far bigger than the others. What is more it also appears that links are established in a regional manner. Table 2 shows indeed that most links are established between AS in located in the same continent, no matter the continent considered.

	Europe	North America	South America	Asia	Oceania	Africa	Antarctica	Unknown	Total
Europe	80514	9871	336	2547	198	705	73	4372	98616
North America	9871	56720	997	2108	417	181	5	2267	72566
South America	336	997	4126	26	1	2	0	1593	7081
Asia	2547	2108	26	14000	226	101	0	1548	20556
Oceania	198	417	1	226	2818	3	0	247	3910
Africa	705	181	2	101	3	960	0	132	2084
Antarctica	73	5	0	0	0	0	0	0	78
Unknown	4372	2267	1593	1548	247	132	0	1420	11579
Total	98616	72566	7081	20556	3910	2084	78	11579	216470

TABLE 5 : AS LINKS REGIONALIZATION

This result suggests that markets can be tackled regionally. Therefore, a regional deployment of ETICS (if it was considered or needed) would not be a drawback.

The results of FIGURE 86 show that Non-Transit AS have a low degree, 100% of single homed AS have a degree of 1 (as expected) and 70% of multi-homed AS have a degree of 2. For most Transit AS, degree is no more than 10 (90% of the cases) and even for transit transit AS, very few AS have a large amount of adjacencies, less than 10% have a degree greater or equal to 100.

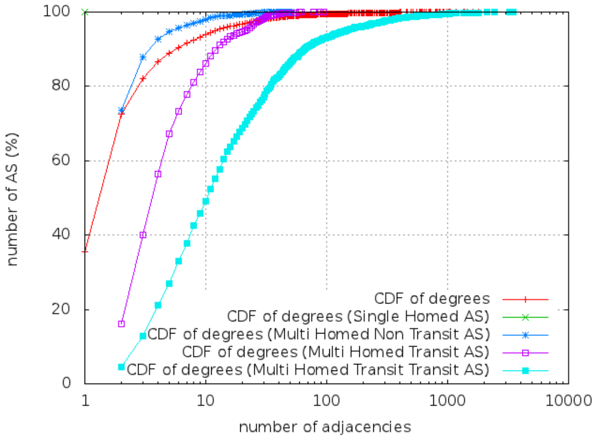


FIGURE 86 : NUMBER OF ADJACENCIES

As far as public IXPs are now considered, we identified 304 of them, most of them in Europe and North America as illustrated in the table below.

	Total	North America	Europe	South America	Oceania	Africa	Asia	Antartica	unkown
Public Internet Exchange Points	**304**	72	146	19	14	11	42	0	0

TABLE 6 : IXPS REPARTITION BY CONTINENT

2382 AS were present in these IXPS over a total of 7366 interconnections (some AS are connected in more than one IXP). The distribution of AS presence, illustrated on Figure 2, shows that presence is quite limited.

40% of AS are only connected in one IXP and less than 10% of AS are connected in more than 10 IXPs.

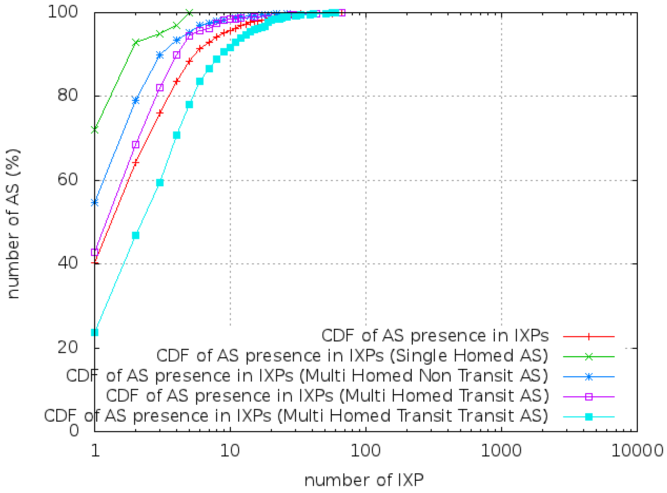


FIGURE 87 : PRESENCE OF AS IN IXPS

The detailed presence of AS in public IXPs is given in Table 7 in a continent by continent manner and depending on the nature of the AS. We also provide an estimation of the number of intra AS links that can be established[25] given the presence of AS in various IXPs. We obtain, based on our data, that at most 31155 intra AS links could be established.

	Type	**Total**	**North America**	**Europe**	**South America**	**Oceania**	**Africa**	**Asia**	**Antartica**	**Unkown**
Single Homed AS	Number of AS	**100**	13	46	0	11	3	7	0	20
	Maximum number of intra AS links	**69**	6	42	0	13	0	3	0	5
Multi-Homed Non Transit AS	Number of AS	**758**	196	423	5	21	10	50	2	51
	Maximum number of intra AS links	**4378**	2762	1451	3	33	4	43	1	81
Multi-Homed Transit AS	Number of AS	**489**	110	293	4	22	9	35	0	16
	Maximum number of intra AS links	**5514**	3733	1542	1	76	9	87	0	66
Multi-Homed Transit Transit AS	Number of AS	**924**	227	502	19	35	13	113	0	15
	Maximum number of intra AS links	**21194**	8537	8347	92	535	83	1629	0	1971
Total	**Number of AS**	**2271**	**546**	**1264**	**28**	**89**	**35**	**205**	**2**	**102**
	Maximum number of intra AS links	**31155**	**15038**	**11382**	**96**	**657**	**96**	**1762**	**1**	**2123**

TABLE 7 : ESTIMATION OF INTRA AS LINKS

Finally, investigating on inter AS links now, the results presented by Figure 3 show there are quite few AS per IXP. 60% of IXPs have less 10 members and less than 1% have more than 100 members.

[25] An AS connected in N IXPs can establish N! / (2*(N-2)!) intra AS links, the maximum number of intra AS links is the sum of this number for all AS we observed in IXPs.

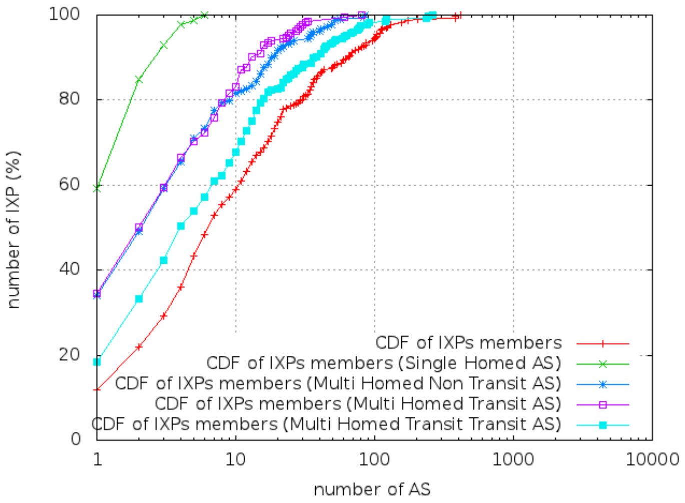


FIGURE 88 : NUMBER OF IXPS MEMBERS

Moreover estimating the number of potential inter AS links that could be established[26] we obtain that at most 438276 inter AS links could be established by AS present in IXPs.

Type	**Total**	**North America**	**Europe**	**South America**	**Oceania**	**Africa**	**Asia**	**Antartica**
Maximum inter AS links*	**438276**	88038	327565	407	5540	630	16096	0

TABLE 8 : ESTIMATION OF INTER AS LINKS

11.8.4. ORIGINATED PREFIXES AND CUSTOMER CONES

We now investigate on originated prefixes and customer cone size to get a better idea of the order of magnitude of the number of regions that would be needed to be announced in the system.

[26] In an IXP containing N AS N! / (2*(N-2)!) inter AS links can be established, the maximum number of inter AS links is the sum of this number for all IXPs we identified.

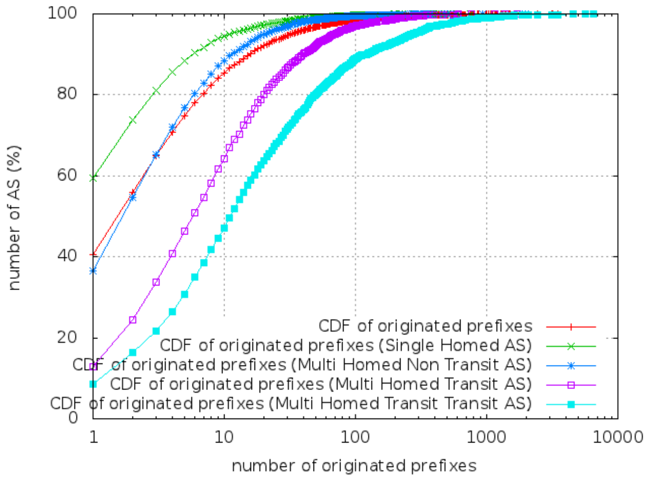


FIGURE 89 : NUMBER OF ORGINATED PREFIXES

FIGURE 89 illustrates that many AS only originate one prefix, 40% on average and up to 60% for single homed AS. Therefore, mapping prefixes to origin AS makes sense most of the time. Moreover, on FIGURE 90 we can see that most customer cones are small, on average 80% of all cones are made of less that 10 prefixes. Only transit transit AS exhibits large cones, but very few of them (10%) have a cone containing more than 200000 prefixes, on the contrary in 60% per cent of the cases their cones contain less than 1000 prefixes.

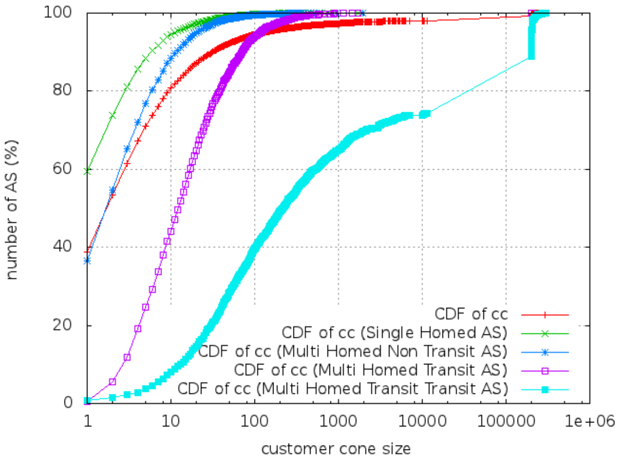


FIGURE 90 : SIZE OF CUSTOMER CONES (IN NUMBER OF PREFIXES)

11.9. GLOSSARY

11.9.1. BASIC DEFINITIONS

Name	Definition	Description
Network Service Provider (NSP)	A business or organization that sells bandwidth or network access by providing direct backbone access to the Internet and usually access to its network access points (NAPs). For such a reason, network service providers are sometimes referred to	In the ETICS Ecosystem, the NSP is responsible for assured quality traffic delivery.

	as backbone providers or internet providers (Source: Wikipedia)	
ETICS community	The set of ETICS NSPs as suppliers and buyers of ETICS network services (see [ETICS-D3.5] as refinement of the concepts introduced in [ETICS-D4.3])	The notion of ETICS community is flexible in terms of how the NSPs interact and do business. […]An ETICS community is described both in technological and in economics-business terms by a set of network capabilities, lifecycle management and business processes, as well as specific rules for conducting trade and for sharing revenues.
ETICS portal	The *ETICS portal* is the interface between the ETICS community and the ETICS customers. In this portal the customer can find information about all the regions that are served by the ETICS community (see Section 4.2.1).	ETICS customers, willing to buy an ETICS inter-carrier ASQ path, first connect the ETICS portal. The ETICS portal will help them finding which destination regions are reachable through the NSP community.
ETICS solution	The overall solution proposed by ETICS considering the inter-networking solution and network services as well as the ETICS core-system solution providing, managing and supporting inter-carrier ASQ paths and connectivity.	
ETICS Architecture	Architectural concepts enabling the realisation of the ETICS solution and its deployment in practice. Resulting deployments may be referred to as *ETICS System*. The interaction between the ETICS System and customers, e.g. NSPs, are handled by the help of the *ETICS Portal*.	
ETICS Core-system solution	Part of the overall solution focusing on ETICS Services and SLA management as well as providing links to underlying interconnected networks (and their technologies) and applications realised with the help of ETICS, e.g. session services.	
Facilitator	An entity which is not an NSP, its role is to perform the service composition.	Performs service composition in the fully centralized deployment scenarios
Edge or Access NSP	NSP which directly connects end users.	
Transit NSP	NSP which does not connect end users and is mainly responsible for carrying traffic between the different edge NSPs.	
Information	Content and applications provider, e.g.	

Service Provider (InfSP)	Google, Facebook, Twitter.	
Service Level Agreement (SLA)	A part of a service contract where the level of service is formally defined (Source: Wikipedia).	
Autonomous System (AS)	Within the Internet, an Autonomous System (AS) is a collection of connected Internet Protocol (IP) routing prefixes under the control of one or more network operators that presents a common, clearly defined routing policy to the Internet (Source: Wikipedia).	Associated IETF memo: http://tools.ietf.org/html/rfc1930
Aggregated Traffic	The part of the ASQ path where the traffic of multiple end hosts of a region goes through the same path.	The aggregated traffic is typically traffic from a PoI to a PoI.
Individual Session	Traffic that is related to an individual host or to an individual application within an individual host.	
Assured Quality (AQ) Service	An ETICS network service is more precisely termed Assured Quality (AQ) network Service, or for short "AQ Service". AQ services refer to network services with an ASQ nature (see ASQ).	
Assured Service Quality (ASQ)	Assured Service Quality (ASQ) is used as adjective referring to quality aspects, i.e. QoS guarantees, availability, etc., for provisioned network services. Please, refer to subsequent notions for more detailed usages.	
Assured Service Quality (ASQ) Traffic	Traffic being attached to ASQ paths and sold by the help of ASQ goods.	
Assured quality paths / Assured Service Quality Paths (ASQ paths)	*An Inter-carrier ASQ path* is an ASQ path that crosses multiple NSP domains. It results from the concatenation of two or more *single-NSP ASQ paths*.	Quantifiable network connectivity services that provide guarantees with respect to how data is carried (e.g. bandwidth, delay etc.). An Inter-carrier ASQ path in general is an ASQ path that crosses multiple NSP domains. It results from the concatenation of two or more single-

		NSP ASQ paths. An example for such services are those offered by some ISPs to their business customers in order to interconnect their distant sites and data centres spread in different locations.
End-user ASQ connectivity	The **End-user ASQ connectivity** is realized "on-top-of" already existing infrastructure level ASQ paths specifically addressing end-user / end-customer ASQ connectivity needs. From an NSP-to-NSP point of view this service is the so-called **PoI*2End-Point** network service.	
Single-NSP ASQ path	An ASQ path provided by a single NSP.	Thus, per definition, an Inter-carrier ASQ path results therefore from the concatenation of single-NSP ASQ paths.
Inter-carrier ASQ path	An Inter-carrier ASQ path is an ASQ path that crosses multiple NSP domains. It results from the concatenation of two or more single-NSP ASQ paths.	
ASQ good	Business products being used for selling ASQ paths to customers.	

Note on "AQ Service" vs. "SLA": AQ Service is the service to be managed by the ETICS service management (B2B) system (the system that realizes the "Network Service and Business Plane"). The AQ service has its reality in the network itself. The SLA on the other hand is a "tool" that is used between the trading parties and in the ETICS service management (B2B) system that help govern, assure and manage the AQ service. Hence, an SLA has its realization primarily in the ETICS service management (B2B) system and only indirectly it has implications in the network itself.

11.9.2. Traffic exchange and network services

Name	Definition & Description
Point of Interconnect (PoI)	Physical region where NSPs are interconnected
Interconnect Interface (ICI) **Inter-Carrier Interface (ICI)**	Physical interface by which NSPs interconnect. *A PoI may thus be defined by one or several ICIs.*
Traffic Delivery	Attached interface for a given network service on an ICI. *An ICI will therefore*

Point (TDP)	*generally include multiple TDPs*.
Point of Enterprise Interconnect (PoEI):	Identifies a point of interconnection between one NSP and an enterprise network. A PoEI also be explained by the notion “Enterprise end-point” (EEP). *Will be used as the basis for defining where the traffic is delivered (from/to) regardless of who would actually “profit” from the traffic delivery service*
End-user end-point (UEP)	The end-user end-point or consumer customer end-point typically corresponds to an end-user micro-flow and may be associated with a session connectivity service (cf. Section 4.4).
PoI*2End-Point / PoI*2EP	PoI*2EP services are used in order to enable End-user ASQ connectivity for NSP-to-NSP services. This service is used where the end-customer specific demand for bandwidth is not feasible for establishing an end-customer dedicated ASQ path. The handling (management and control) and support of such “on-top-services” are enabled by the so-called service enhancement functions (cf. Section 6.1). The exact PoI might be unknown by the supplier at a given point in time, i.e. being reflected in the the PoI* notation (see Section 4.4)
Host end-point	Refers to a single host, and can be that of a residential or consumer end-user, a business end-user, an enterprise server host, or a data centre host
Multipoint	Multipoint is the general term by which we refer to a specific set of given points in terms of either a set of PoIs, a set of PoEIs, or a set of host end-points.
Destination or Source Region	Set of host end-points, that is, end-user end-points typically designated by public IP addresses, given in terms of a set of IP prefixes.
Single-NSP network service	Network service where the NSP supplier has responsibility only across his NSP domain.
Multi-NSP network service	Network service where the NSP supplier has responsibility across his NSP domain and one or more other NSP domains.

11.9.3. Traffic Charging

Name	Definition
Traffic termination (TT)	The buyer NSP is sending ASQ traffic to the supplier NSP which is responsible for sending (terminating) the traffic according to the SLA. This is according to the sending party network pays principle.

11.9.4. NSP-to-NSP (Supplier-to- Buyer/Requestor) Network Services

Name	Definition	Description
PoI-to-Region	Buyer NSP is paying for ASQ traffic	**Traffic Termination:** Additional constraints will

	transported to a given region.	typically apply in order to assure end-to-end ASQ for host-to-host session services.
Region-to-PoI	Buyer is paying for receiving ASQ traffic from a given region.	**Traffic Origination**: Additional constraints will typically apply in order to assure end-to-end ASQ for host-to-host session services.
PoI-to-PoEI	For TT: Buyer NSP is paying for ASQ traffic transported to a given PoEI. For bidirectional: The buyer is paying for two-way traffic across the given segment.	**Traffic Termination, bidirectional**: Note that several variants of this service apply according to exactly where and how the traffic is terminated.
PoEI-to-PoI	Buyer NSP is paying for ASQ traffic transported from a given PoEI.	**Traffic Origination**: Note that several variants of this service apply according to exactly where and how the traffic is terminated/originated.
PoI-to-PoI	For TT: Buyer NSP is paying for ASQ traffic transported to a given PoI. For TO: Buyer NSP is paying for ASQ traffic transported from a given PoI.	**Traffic Termination, bidirectional**: Note that several variants of this service apply according to exactly where and how the traffic is terminated. For bidirectional: The buyer is paying for two-way traffic across the given segment. **Traffic Origination**: Note that several variants of this service apply according to exactly where and how the traffic is terminated/originated.

11.9.5. ABBREVIATIONS

Abbreviation	Meaning
BRPC	Backward-Recursive PCE-based Computation Procedure
GRX	GPRS Roaming Exchange
IGP	Interior Gateway Protocol
IPX	IP eXchange
IMS	IP Multimedia Subsystem
MNO	Mobile Network Operator
MPLS	Multi-Protocol Label Switching
PCE	Path Computation Element
PCRF	Policy Charging and Rules Function