AGAVE

A liGhtweight Approach for
Viable End-to-end IP-based QoS Services

IST-027609

D2.2: Specification of the Connectivity Service Provisioning Interface Components

Document Identifier: AGAVE/WP2/ALGO/D2.2	
Deliverable Type: Report **Deliverable Nature:** External	**Contractual Date:** 29 February 2008 **Actual Date:** 30 April 2008

Editor:	Panos Georgatsos, Algo
Authors:	*TID:* A. J. Elizondo, O. Gonzalez de Dios *FTR&D:* M. Boucadair, B. Decraene, B. Lemoine, J.L. Le Roux, F. Bersani *Algo:* E. Mykoniati, P. Georgatsos *UCL.uk:* D. Griffin, J. Spencer, J. Griem *UniS:* N. Wang, M. Amin, K. H. Ho, M. Howarth, G. Pavlou *UCL.be:* B. Quoitin, O. Bonaventure, L. Iannone
Abstract:	This document provides the final specification of the AGAVE approach to the issue of provisioning and delivering services in the Internet. The specification includes the connectivity provisioning interface between service providers and IP network providers and the architecture within network provider domains, which is built around the notions of Network Planes and Parallel Internets. The problem of defining Network Planes and Parallel Internets is elaborated and a greedy solution approach is outlined. The design of Network Plane Emulation Platform, which 'puts these concepts in motion', is also presented. Finally, the deployability of the proposed AGAVE solutions is assessed. A draft white paper positioning AGAVE with respect to known NGN architectures is included as an appendix.
Keywords:	Quality of Service, Network Planes, Parallel Internets, Connectivity Provisioning Agreement

Telefónica Investigación y Desarrollo	TID	Co-ordinator	Spain
France Telecom Research and Development	FTR&D	Partner	France
Algonet SA	Algo	Partner	Greece
University College London	UCL.uk	Partner	UK
The University of Surrey	UniS	Partner	UK
Université catholique de Louvain	UCL.be	Partner	Belgium

Project funded by the European Community under the
"Information Society Technology" Programme (2002-2006)

Executive Summary

To the problem of service provisioning and delivery in the Internet, AGAVE advocates:

- A 'clear-cut' interface between Service Providers (SPs) and IP Network Provider (INPs), which is based on the notion of Connectivity Provisioning Agreement (CPA).
- The concept of Network Planes (NPs) and Parallel Internets (PIs), allowing INPs to build and provide multiple Internet connectivity levels, as required by the end-to-end requirements of the services they offer.

This deliverable presents the final specification of the AGAVE architecture; the open connectivity interface between SPs and INPs and the required NP/PI-based functionality for supporting this interface. Specifically:

Service Providers interact with IP Network Providers on the basis of *Connectivity Provisioning Agreements* (CPAs). A CPA allows the IP connectivity requirements of a Service Provider to be defined in terms of QoS, resilience and availability guarantees with a specified scope. Further, a CPA allows for defining access control, shaping, flow forwarding and routing rules to be enforced at particular edges or across the defined scope. Through CPAs, Service Providers may also specify performance reports and notifications that wish to receive either for the assurance of their CPA or as feedback for driving their own dynamic service engineering functions. Clearly, the CPA notion substantiates in an open manner the interface between service and network providers.

A technology agnostic layer operating at the level of *abstract network-wide capabilities* is introduced in INP domains. These capabilities represent the dimensions along which the treatment of traffic flows can be differentiated e.g. QoS, availability and resilience dimensions. Depending on whether they refer to intra- or inter-domain scope, abstract network capabilities are encapsulated in the notions of Network Planes and Parallel Internets respectively. PIs are built from the perspectives of an INP, as appropriate to the requirements of the services it offers, by combining its NPs with those in other provider domains. The notions of NPs and PIs are specified and the problem of determining NPs and PIs is elaborated for gaining insight to the solution space, procedure and complexity. A greedy solution approach is specified. The information flow across the functional blocks within an INP for supporting connectivity provisioning requests is outlined.

The design of the Network Plane Emulation Platform (NPEP), prescribed by the project as a means for validating and exhibiting the proposed approach, is presented. NPEP provides a snapshot of an INP domain, embodying the essential aspects of the AGAVE approach. Emphasis is put on the required operations for defining NPs and PIs. Network configurations corresponding to alternative sets of NP/PIs can then be evaluated.

The deployment of AGAVE solutions in the today's best effort Internet can be made in an incremental fashion. Universal participation of INPs is not pre-requisite, neither is not required that all INPs have a common deployment or configuration of a particular set of traffic engineering (TE) techniques, mechanisms and protocols. AGAVE does not advocate a clean-slate approach, instead it relies and builds upon the existing IP-based Internet architecture and related protocols. Moreover, the set of mechanisms specified in AGAVE either co-exist with currently deployed protocols or make minor incremental extensions to existing protocols.

Finally, a draft white paper positioning AGAVE with respect to known architectures for NGNs, proposed by world-wide initiatives and standardisation bodies, is appended. The final white paper will be published on the project web-site by the end of the project.

This deliverable concludes successfully WP2 work.

Table of Contents

Detailed Table of Contents

List of Figures

1 INTRODUCTION

1.1 WP2 Objectives

This deliverable is produced by AGAVE Work Package 2, which focuses on *Connectivity Service Provisioning*. WP2 has been setup with the following objectives:

- To specify a unified interface supporting the common provisioning and control requirements for the connectivity aspects of end-to-end IP-based services within the Parallel Internets framework, with the ultimate objective to facilitate the rapid deployment of services.
- To identify the generic networking capabilities of Network Planes and design the Network Planes management interface to support the operations of the service provisioning interface.
- To specify an overall engineering approach and select appropriate implementation technologies.
- To design and implement the components realising the operations supported by the connectivity service provisioning interface and their interactions with the underlying network through the Network Planes management interface.
- To specify test requirements for evaluating the validity of the specifications and development focusing on specific service type and business model use cases.

The work presented in the deliverable addresses all above objectives. The documented results successfully conclude WP2 work.

1.2 Content Outline

This document provides the final specification of the AGAVE approach to the issue of provisioning and delivering of services in the Internet. The specification includes the connectivity provisioning interface between service providers and IP network providers and the architecture within network provider domains, which is built around the notions of Network Planes (NPs) and Parallel Internets (PIs). The problem of defining Network Planes and Parallel Internets is elaborated and a greedy solution approach is outlined. The design of Network Plane Emulation Platform which, 'puts these concepts in motion', is also presented. Finally, the deployability of the proposed AGAVE solutions is assessed.

The document is structured as follows:

- Chapter 2 presents the specification of the Connectivity Provisioning Agreement (CPA) on the basis of which service provider interact with underlying IP network provider domains for end-to-end service provisioning. CPA substantiates in an open manner the interface between service and network providers.
- Chapter 3 presents the specification of the required INP functionality for supporting CPAs. The concepts of NPs and PIs are specified. Emphasis is put on the problem of determining the PIs and NPs to realise in a network domain so that to match service requirements.
- Chapter 4 presents the design of the Network Emulation Platform (NPEP), prescribed by the project as a means for validating and exhibiting the proposed approach.
- Chapter 5 discusses the deployment of the proposed AGAVE solutions in today's Internet.
- Chapter 6 summarises the work included in the document.
- Appendix A -in the form of a draft white paper- positions the AGAVE approach with respect to known architectures for NGN, proposed by world-wide initiatives and standardisation bodies.

1.3 Major Updates

The specification of CPAs, NPs and PIs remains as in [D2.1]. The new contributions compared to [D2.1] are summarised below:

- NP/PI Problem -section 3.6: The problem of defining Network Planes (NPs) and Parallel Internet (PIs) has been elaborated for gaining insight into its solution space, procedure and complexity. A number of problems are specified and analysed. A greedy solution approach for the problem of determining the NPs to realise has been specified.
- NP Design and Creation feedback –section 3.7.4: Specification of the information required for evaluating the performance of the network given that a certain set of NPs and PIs has been instantiated. This specification complements the NP and PI definition specification in [D2.1], providing a 'reporting view' of their actual performance.
- NPEP –chapter 4: The design of the Network Emulation Platform (NPEP) is presented. NPEP provides a snap-shot of the proposed INP architecture, embodying the essential aspects of the AGAVE framework –Network Services, NIAs, NPs and PIs. The platform is built with the purpose of validating the concepts and notions developed by the project. Emphasis is put on the required operations for defining NPs and PIs.
- AGAVE deployability –chapter 5: The deployment of AGAVE solutions in the today's best effort Internet is analysed. A number of arguments are presented, which justify that deployment can be made in an incremental fashion.
- AGAVE positioning –appendix A: The AGAVE approach is positioned with respect to known architectures for NGN, proposed by world-wide initiatives and standardisation bodies. It is a draft of a white paper to be published at the project website by the end of the project.

2 INTERFACE TO SERVICE PROVIDER

2.1 Introduction

Defining the IP Network Provider (INP) as an autonomous role interacting directly with Service Providers (SPs) – from network layer SPs to higher layer Application SPs (ASPs) – introduces an interface between INPs and SPs, exposing the IP connectivity capabilities of the INPs in a generic service-provisioning-aware but not service-specific way (see [D1.1]). This interface allows for multiple services operated by different SP administrations to run over a common IP network infrastructure transparently, with the INP optimising the network performance overall and under the constraints of each service running over it.

The idea of defining an interface to clearly distinguish the operational concerns in the IP and service layer has always been thought of as a 'good practice', mainly on grounds of hierarchical system design, bringing amongst others the merits of encapsulation of lower level functions and separation of concerns. In addition to these engineering merits, the introduction of INP-SP interface advocates new business roles and therefore bears new business opportunities in the chain of service delivery in the Internet. In line with this view are emerging studies, which also 'break' the traditional role of an ISP along the lines of 'networks and services' proposing 'clear-cut' interfaces [FEAM06].

Beyond the forwarding and the QoS treatment of the traffic entering the INP's network from the SP's sites, the INP offers to the SP means to control the connectivity provisioning. Particular connectivity provisioning requirements are captured for the IP connectivity, VoIP and VPN service business cases studied in [D1.1].

The main ingredient of the INP-SP Interface is what we call the *Connectivity Provisioning Agreement* (CPA). SPs interact with INPs on the basis of such agreements. The interface provides for the necessary means to negotiate CPAs and execute the operational actions agreed in the CPA.

In a snapshot, the CPA allows for defining the connectivity requirements of the SP in terms of QoS and service availability guarantees in a specified scope. Further, the CPA allows for defining access control, shaping, flow forwarding and routing rules to be enforced at the edges or across the defined CPA scope. The SP may also specify performance reports and notifications that it wishes to receive either for the assurance of the CPA or as feedback for driving its own dynamic service engineering functions.

Our work draws from the SLS template [TEQUILA] and the provider SLS (pSLS) template [MESCAL] specifications of the IST TEQUILA and IST MESCAL projects. TEQUILA specified a service management and traffic engineering framework for *intra-domain QoS provisioning*, which prompts for standardisation of the notion of SLS, proposing a standard template to capture the IP QoS-aware connectivity services offered to end customers. MESCAL advanced the SLS template specification to incorporate *inter-domain aspects*. The provider SLS (pSLS) template models the interactions between peer providers for the provision of end-to-end IP QoS-aware connectivity services to end customers, assuming that peer providers co-operate for providing QoS guarantees in a cascaded way.

Although AGAVE CPA specification builds on the TEQUILA SLS and MESCAL pSLS, the CPA model is different from these in the following aspects: a) AGAVE focuses on the interactions between the -now distinct- IP Network Provider and Service Provider roles, while in TEQUILA and MESCAL these two roles were encompassed in the role of an ISP, b) CPA captures the connectivity provisioning requirements of SPs rather than the connectivity requirements of end customers and c) the AGAVE CPA does not depend on the strict cascaded model between peer providers as assumed by the MESCAL pSLS model.

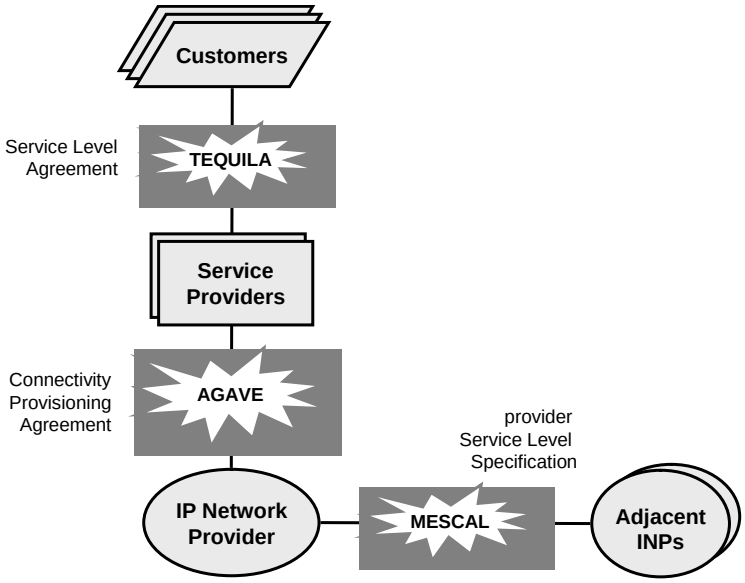

Figure 1: AGAVE CPA with respect to previous work.

2.2 CPA template specifications

The CPA is specified against the information elements (clauses) shown in Figure 2[1], see details in the following sections.

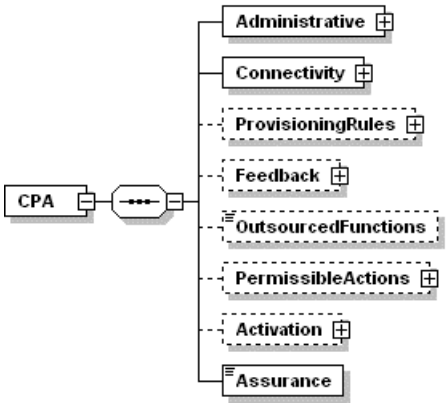

Figure 2: Connectivity Provisioning Agreement.

2.2.1 Administrative

The administrative information clause may include information on the Service Provider, the formula to be used for charging, etc.

2.2.2 Connectivity

The connectivity clause captures the IP connectivity requirements of the SP. The IP connectivity is specified in terms of *connectivity modules*. A connectivity module specifies capacity and guarantees within a defined scope. The desired guarantees are modelled as *connectivity classes*. The scope is defined as connections between specific *edges*, local or remote to the INP domain. Unlimited scope is allowed and is specified as a remote edge encompassing all possible destinations.

[1] The figures are drawn with the Altova XML editor. For an explanation of the diagram model, see section 5.1.2 Content Model View (pp 135-144) of *Altova XMLSpy 2007 Enterprise Edition User & Reference Manual*, available at http://www.altova.com/documents/XMLSpyEnt.pdf.

2.2.2.1 Edges

An edge (see Figure 3) denotes the ingress or egress border link or border router where the responsibility of the INP for delivering the traffic according to the terms of the CPA begins or terminates.

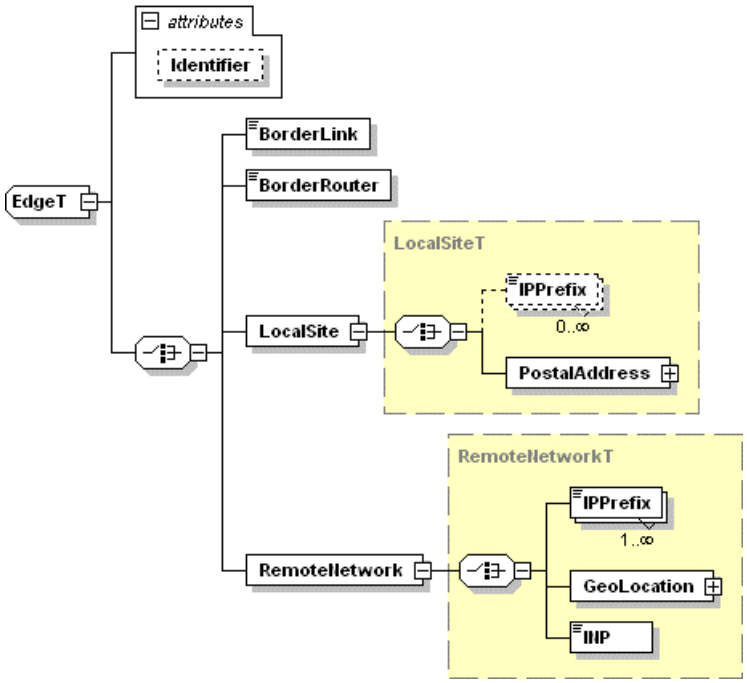

Figure 3: CPA - Edge.

Edges are used to determine the scope of connectivity modules (see section 2.2.2.3). An INP may provide guarantees for reaching directly attached sites only or for remote sites too, the latter is called a multi-hop CPA (see [D1.1], section 6.1.2). An edge can thus be specified as the border link or border router either local to the INP, or of a remote INP in case of a multi-hop CPA. Instead of the INP link or router, an edge can be specified as a SP or customer site directly attached to the INP, or as a remote network represented by a set of IP address prefixes. Note that, when the CPA edge is the final destination of the SP traffic, it is transparent to the SP whether it is local or remote to the INP. However, when the CPA edge is only an intermediate node for the SP traffic, this implies that the SP has another agreement with the provider (INP or SP) connected to the remote end, hence the SP is aware of the CPA edge location.

In case of a *local site*, the INP matches the site information with some registered information to derive the corresponding border link where the site is connected. In case of multi-homed sites, a local site edge will be mapped internally by the INP to several edges (border links), as many as the connections between the site and the INP. In case of a *remote network*, the INP translates the provided geographical areas or peer INP identifiers to IP prefixes and derives the downstream NIAs and associated inter-domain links towards these remote destinations to be used to carry the SP traffic, based on the downstream NIAs and routing configuration in effect. In this case also, one remote network edge may be translated to several local border link edges, or the opposite, many remote network edges may be translated to just one local border link edge.

2.2.2.2 Connectivity Class

The connectivity class (see Figure 4) captures the guarantees on IP packet transfer performance metrics that the INP agrees to offer to the SP traffic in the context of a connectivity module.

A connectivity class includes the following attributes, corresponding to the IP packet transfer performance and IP connectivity availability metrics against which guarantees are given:

- Delay guarantees, specifying the guarantees for the one-way packet delay as measured between specific ingress and egress points crossed by the SP traffic.
- Jitter guarantees, similar to the above.
- Loss guarantees, specifying the guarantees for the packet loss probability; this is defined as the ratio of the lost packets between specific ingress and egress points and the injected packets at ingress.
- Availability guarantees, specifying the percentage of the time over an agreed measurement interval, where the above QoS guarantees are provided as agreed; availability guarantees thus capture the frequency and persistence of physical failures and/or QoS degradation caused by congestion.

It is not necessary for all above attributes to be specified. Relevant metrics have been standardised (see [RFC2679, RFC2680]). However, the metrics supported by each INP may vary depending on its policies and capabilities.

Figure 4: CPA - Connectivity class.

The following aspects underlying the semantics of the above attributes are worth noting:

The following types of guarantees are distinguished: quantitative and qualitative. The guarantees to a particular metric are said to be quantitative, if they can be expressed in quantitative, numerical, values. Otherwise, they are said to be qualitative; possible qualitative values, as appropriate as per performance parameter, may include: high, medium, low or red, yellow, green. The quantification of the relative difference between the qualitative values is a matter of provider's policy e.g. 'high' could be twice good as 'medium', which in turn is twice as good as 'low'.

Quantitative performance guarantees are expressed as maximum (worst-case) bounds or as (sets of) percentiles or inverse percentiles, indicating also the granularity period of the associated measurements. The meaning of the values of qualitative performance guarantees and/or their relative difference should be clear to the SPs, while it should be backed-up with relevant historical performance data.

Each connectivity class specified in a CPA will be mapped to appropriate Network Planes and Parallel Internets internally in the INP (cf. section 3.7.7).

2.2.2.3 Connectivity Module

The connectivity requirements of the SP are specified in terms of connectivity modules. The connectivity module associates a connectivity class (guarantees) to a specified scope and capacity (see Figure 5).

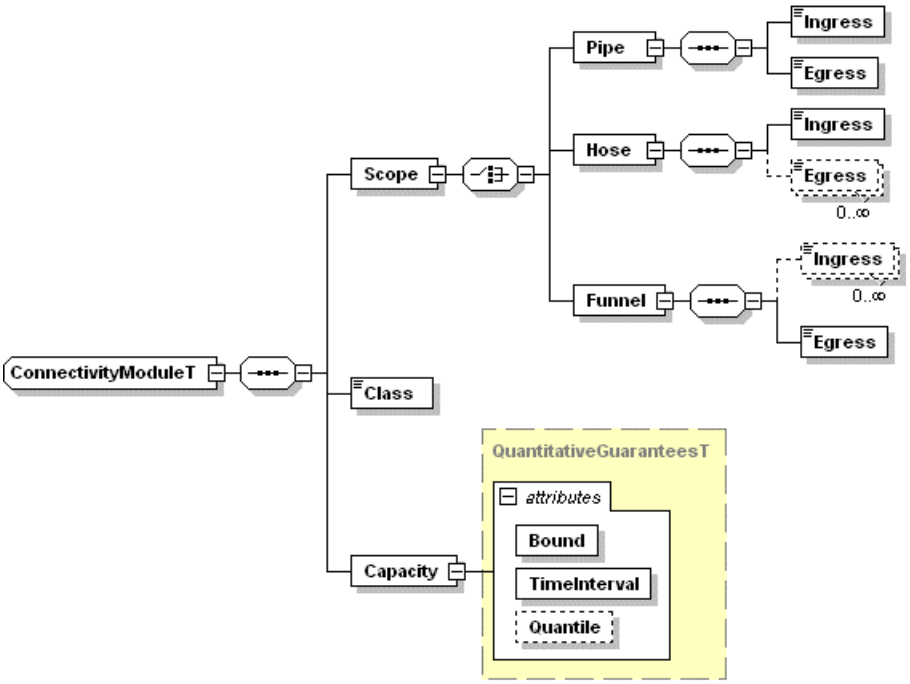

Figure 5: CPA - Connectivity module.

2.2.2.3.1 Scope

Scope explicitly identifies the geographical/topological region over which liability for the connectivity class guarantees ends, by indicating the boundaries of that region in terms of edges. It includes the following attributes:

- Ingress edge, indicating the entry point of the region over which the connectivity module is to hold
- Egress edge, indicating the exit point of the region over which the connectivity module if to hold

The following combinations of Ingress, Egress values are allowed:

- (1,1) implying an one-to-one communication; we call the connectivity module a pipe
- (1,N) one-to-many communication (N>1); we call the connectivity module a hose
- (1,any) one-to-any communication; we call the connectivity module an unspecified hose
- (N,1) many-to-one communication (N>1); we call the connectivity module a funnel
- (any,1) any-to-one communication; we call the connectivity module an unspecified funnel

Because connectivity modules are assumed unidirectional, the above taxonomy excludes the many-to-many communication (M, N); either ingress or egress attributes must be specified to exactly one interface identifier. Many-to-many communication can be achieved at the level of CPA, where a number of connectivity modules are combined.

In case where the specified edges are remote, their mapping to the provider's domain boundary links is subject to the NIAs and routing decisions in place. Hence, internally in the provider and transparently to the agreed CPA, traffic to remote sites may be merged over one boundary link or split to many boundary links turning a hose to a pipe, or a pipe to a hose, etc.

Usually, one of these attributes corresponds to an interconnection link of an SP site to the reference INP, while the other attribute is left unspecified or set to a set of destinations (remote network), the interconnection link of another SP site to another INP, or the boundaries of another INP. As an example, in the case of an Internet SP the value of the ingress would be the SP's interconnection points and the value of the egress would be left unspecified denoting "any"; the latter could be refined to denote the interface of a particular inter-domain link by the traffic engineering functions (cf. section 3.7.5), however, this is an internal matter, being not subject of agreement. In the case of a VPN SP, both ingress and egress would be clearly specified.

2.2.2.3.2 Capacity

The capacity clause determines the SP traffic volume that can be supported at the guarantees of the specified connectivity class in the specified scope.

When the scope of the connectivity module is a pipe (see section 2.2.2.3.1) the INP must verify that the required resources are available from the ingress to the egress. When the scope is a hose, the INP must verify that the required resources are available from the ingress to any of the egress points. INP is forced to allocate resources equal to the overall capacity across all ingress-egress pairs, resulting in a significant resource under-utilisation. This can be smoothed out, by constructing tree paths from the ingress to the egresses.

2.2.3 Provisioning Rules

Provisioning rules include rules on access to the specified connectivity, on the forwarding of flows across the edges of the CPA, routing rules for constructing the paths between the CPA edges, and finally rules for shaping the traffic at the egresses. Other types of rules may be included in the future.

2.2.3.1 Access Rules

Access rules (see Figure 6) specify how data flows are treated within the CPA, including policing at the CPA ingresses, assignment to the connectivity class for enjoying the associated guarantees, and marking policies at the CPA egresses. An access rule is defined for a particular micro-/macro-flow entering from one or a set of the CPA ingress points.

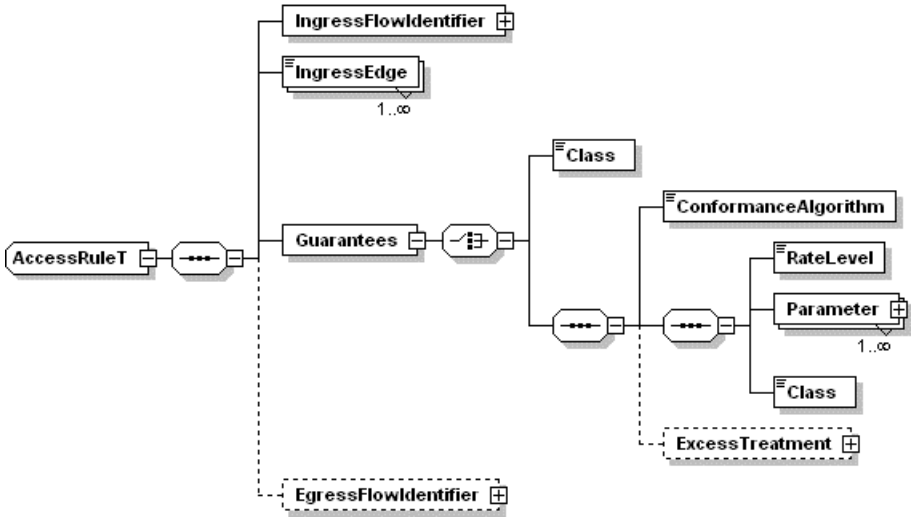

Figure 6: CPA - Access rule.

2.2.3.1.1 Ingress flow identifier

The ingress flow identifier sets the classification rules identifying the stream of IP datagrams constituting the flow to which the access rule is to apply. Classification is performed based on the IP

header fields (e.g., source and/or destination IP address, protocol, ToS/DSCP, etc.), and/or based on tunnel end identifier if tunnelling is used for interconnection.

2.2.3.1.2 Guarantees

The access rule determines the guarantees the identified flow is entitled to and the restrictions the flow must adhere to for getting the guarantees. Note that the access rule merely controls how much volume of which flows will gain access to the capacity associated to the connectivity modules; the access rule does not per se implies that all packets admitted following the defined access rules will get the associated guarantees, not if the rate of the overall injected traffic exceeds the capacity of the corresponding connectivity modules (see section 2.2.2.3).

The identified flow can be associated with a connectivity class (see section 2.2.2.2) unconditionally for all received packets, or subject to a conformance traffic profile. In the latter case, the following attributes are specified:

- Traffic conformance algorithm, specifying the mechanism used to unambiguously identify the packets complying with the traffic conformance criteria and those which do not, called the "in" and "out" of profile packets, respectively. Examples of traffic conformance algorithms are: leaky bucket, token bucket, combined token bucket with peak, a two-rate three-colour marker scheme [RFC2698] and an MTU-based scheme.
- Traffic conformance algorithms may allow for setting multiple levels of traffic rate conformance. Each traffic conformance level is characterised by associated conformance criteria in terms of rate (bandwidth) thresholds, captured in traffic conformance parameters like peak rate, token bucket rate, bucket depth and maximum transfer unit (MTU).
- The traffic that conforms to a particular rate level is assigned to a connectivity class.
- The treatment of the traffic in excess of the highest rate level is specified. Excess treatment may be dropping (default), shaping, or gaining access to a qualitative connectivity class. Note that the rate of excess traffic is unlimited, hence it is senseless to assign it to a connectivity class designed to deliver quantitative guarantees.

2.2.3.1.3 Ingress edges

Each access rule applies to one or more ingress CPA edges. If the CPA edge is a border link then this implies that policing rules corresponding to the traffic conformance clause will be configured at the corresponding border interface. If the ingress CPA edge is a border router, then the same policing rules will be configured at every external input interface. As a result, if a flow is distributed to more than one path from the upstream domain, it would gain access to a rate higher than the highest conformance level, as many times as the number of CPA edge border interfaces it is mapped to.

At every border router associated with an ingress CPA edge, forwarding will be configured so as to forward the packets of the packets conformant at each level to the Network Plane and Parallel Internet corresponding to the connectivity class of this conformance rate level. A validity check on the CPA specification will require that the ingress edge and connectivity class associated with a flow have a match to a specified connectivity module.

2.2.3.1.4 Egress flow identifier

Egress flow identifier captures the requirements for the marking and/or tunnelling identifiers to be applied to a certain flow at the CPA egresses. VPN traffic for example is expected to require DSCP transparency (see [D1.1], section 5.3.1.1), which forces the INP to maintain the ToS/DSCP values as in the original packets received at the CPA ingresses. Other flows may require re-marking with a particular ToS/DSCP value, because so it is expected by the other end following the CPA egress. Note that, delegating remarking at the CPA egresses raises interoperability and scalability issues for multi-hop CPAs (see [D1.1], section 6.1.1.5.1). By default when no egress flow identifier is specified, all INPs in the path between the CPA ingress and egress are free to remark SP's traffic.

2.2.3.2 Forwarding Rules

The clause of forwarding rules includes per flow route selection rules, specifying the egress CPA edge where the defined flow should be directed to. Such rules allow for overriding the routing of the INP, enabling the SP to implement its own routing logic over a logical topology where CPA egress edges are not the final destinations but intermediate nodes to paths controlled by the SP. Forwarding rules could be specified by overlay SPs, or by VPN SPs, etc.

2.2.3.3 Routing Rules

Routing rules determine constraints and preferences for constructing the path between the CPA edges, e.g. exclude a particular AS from the inter-domain path. A routing rule differs from a forwarding rule in that a forwarding rule indicates the egress CPA edge for a flow at a particular ingress edge, while a routing rule determines how the logical link between the CPA ingress and egress edges must be constructed.

2.2.3.4 Shaping Rules

Shaping rules determine profiles for shaping the SP traffic at a particular egress CPA edge. When the CPA edge is a remote location outside the domain of the INP, the INP must delegate its enforcement to the final INP in the downstream path where the egress CPA edge belongs. Identification delegation issues may arise in this case (see [D1.1], section 6.1.1.5.1).

2.2.4 Feedback

To perform fault and performance management each SP requires feedback from the network, especially in the case of qualitative guarantees. Feedback requirements are captured in the following clauses (see Figure 7):

- Monitoring tasks are specified in terms of the metrics (see [D3.1], section 4.2 for details) and data collection attributes like granularity, sampling frequency etc. The scope of a monitoring task may be limited to the INP domain or it may include inter-domain statistics extending as far as the final destinations or the CPA egress.
- Notifications and reports, determine the monitoring reports and/or alarms that the INP must produce and deliver to the SP periodically or on-demand. The notifications and reports are specified in terms of the metrics defined in the monitoring tasks clause, setting reporting frequency, alarm thresholds, etc. Applicable reports/alarms are related to network performance, failure incidents, security attacks, troubleshooting logs etc.

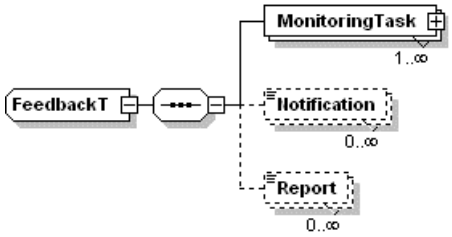

Figure 7: CPA – Feedback.

2.2.5 Outsourced Functions

An SP may choose to outsource the maintenance and the management of its IP equipment to the INP, making the latter responsible for firmware upgrades, performance monitoring, troubleshooting, etc.

2.2.6 Permissible Actions

The permissible actions clause (see Figure 8) determines the actions that the SP is entitled to invoke with respect to the particular CPA, for adding, deleting or modifying connectivity entries, provisioning rules or feedback requirements.

Each action is associated to one availability and one invocation profile. The action availability profile captures the guarantees and the conditions for authorising a particular action, e.g. the time of the day where this request can be performed, the probability for a request to be granted, restrictions on how frequently the action can be performed, restrictions on time to respond for the INP etc. The invocation profile captures information on how the action is to be communicated between the SP and the INP, such as the invocation protocol, and who is entitled to perform the action, allowing the SP to specify user authentication for securing and restricting the access to the CPA permissible actions.

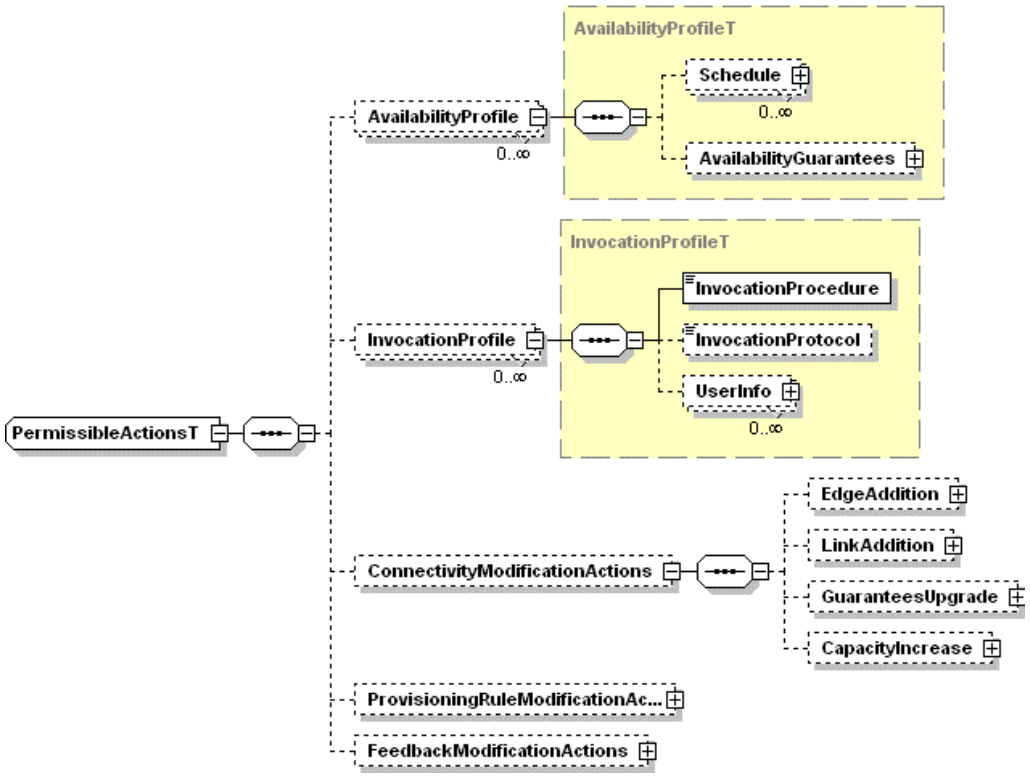

Figure 8: CPA - Permissible actions.

Typical examples of connectivity modification actions are the expansion of the CPA scope by adding new local or remote edges, the increase of the CPA connectivity density by adding new links between existing edges, the upgrade of guarantees associated to existing links and the increase of capacity.

2.2.7 Activation Info

Activation information specifies potential interactions that need to take place between the INP and the SP, beyond internal INP configuration, to complete the CPA activation. Such interactions may be establishing a peering connection between BGP speakers or between invocation protocol speakers for dynamic tunnel establishment and teardown, establishing of security associations between border routers at the CPA edges, configuring authorisation to invoking probing for CPA assurance, etc.

2.2.8 Assurance

The assurance information clause determines the verification methodology and sets the *Key Performance Indicators* (KPIs) and other applicable parameters for assessing the conformance of the SP and the INP to the agreed terms and conditions. Similarly to the specification of the feedback requirements (cf. section 2.2.4), notifications and reports may be associated to each KPI to notify the

SP for service degradation. Penalties may be associated to service degradation scenarios in terms of the specified KPIs.

Note that in general, assurance monitoring and reporting is not identical with the required feedback. Assurance specifies the requirements for assessing the conformance to the CPA, while feedback may involve providing additional information and in different time scales, for either a compliant or a non-compliant CPA.

As both the SP and the INP are expected to perform monitoring for verification and assurance of the CPA, for the results to be comparable, clock synchronisation may be required (see [D3.1], section 3.4). The SP also specifies the probing facilities to which it requires access for performing its own verification measurements in terms of protocols and probing scope, within the scope of the specified connectivity modules.

3 INP ARCHITECTURE

3.1 Key Concepts

For managing the provisioning and delivery of different 'types of traffic', AGAVE proposes a *technology agnostic layer* in INP domains, which is built around the concepts of *Network Planes (NPs)* and *Parallel Internet (PIs)*.

The technology agnostic layer operates on the basis of *abstract network-wide capabilities*, expressed in commonly understood technical terms rather than in the jargon of a particular technology. These capabilities represent the dimensions along which the treatment of traffic flows can be differentiated. Depending on whether they refer to intra- or inter-domain scope, the different abstract network capabilities are encapsulated in the notions of NPs and PIs, respectively.

PIs represent the Internet as seen from the perspectives of an INP. An INP builds PIs tailored to the end-to-end requirements of the services it offers, by combining its NPs that is, its local capabilities, with appropriate capabilities offered by other INP domains on the basis of respective interconnection agreements, called *INP Interconnection Agreements* (NIAs).

In essence, an AGAVE INP domain is hierarchically decomposed into three layers: the business layer, the technology agnostic NP/PI layer and the technology-specific layer. These layers can be directly mapped to known roles in existing network providers organizations; business development, network planning/design and network engineering/operations roles.

Interactions between the business and the NP/PI layers are based on the notions of *Network Services (NSs)* and *Engineering Guidelines*. Network Services represent the different 'types of traffic' that can be provisioned and delivered by the INP in terms of availability, resilience and QoS guarantees in certain topological scope. Building on them, the INP offers its products -CPAs to SPs and NIAs to other INPs. Engineering Guidelines provide rules for handling the demand for the supported services, including CPA requests.

Based on the defined Network Services and the set of Engineering Guidelines, the NP/PI layer determines the PIs and the NPs that need to be in place so as to best match service requirements and the underlying guidelines. Subsequently, based on the defined NPs and PIs, the technology-specific layer engineers the network suitably.

Figure 9 summarizes the above by depicting the key information entities introduced and their relationships.

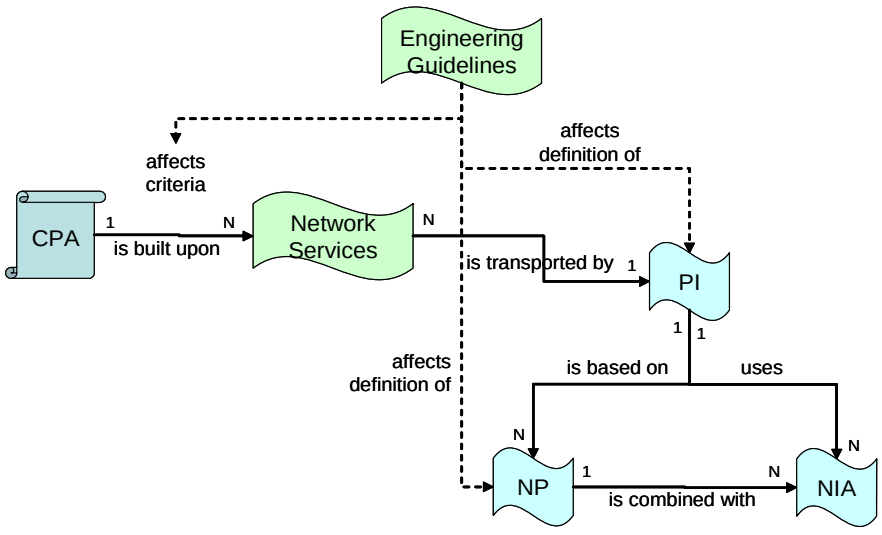

Figure 9: Key AGAVE notions in INP domains.

The value of introducing the NP/PI layer is that it offers a *well-defined communication bridge between business and network engineering/operations levels*. The NP/PI-based communication bridge is independent from specific network technologies, yet is powerful enough to accommodate both intra-domain and Internet-wide concerns. Specific benefits include:

- Increased levels of manageability, especially for service provisioning and reporting.
- Reduced time for putting new technologies in effect of business, thus accelerating RoI.
- Facilitates enforcement and evaluation of business strategies, avoiding monolithic approaches.
- Smooth interactions between development and operations within and across business and network levels.

Overall, we believe that the proposed 3-layer architecture contributes to effective and efficient network management and operations.

3.2 Network Services

Network Services (see Figure 10) are the QoS, availability and resilience performance guarantees that the business layer desires to offer as services to SPs and upstream INPs. The realisation of some of these Network Services may require the establishment of NIAs with other downstream INPs.

QoS and availability metrics can be expressed either as quantitative range values or qualitative values. Quantitative range values are set as range percentile values, e.g. delay from 100ms for 95% of the traffic up to 200ms for the same 95% of the traffic; qualitative values are set as weights that determine the relative difference of the specified network services.

The scope of a Network Service is specified as a collection of sets of ingress and egress edge domain. Egress domains can be adjacent, or remote, i.e. reachable through other INPs, limited to particular geographical areas, adjacent or remote peers or unlimited to any destination. The Network Service scope may be determined provisionally, to denote the desire of the business layer to serve traffic in a particular network service going to or through the specified egresses. If no scope is defined, the scope is considered unlimited. Business layer may desire to boost a certain Network Service only from some ingress areas of its network, e.g. in case these areas happen to be under-utilised.

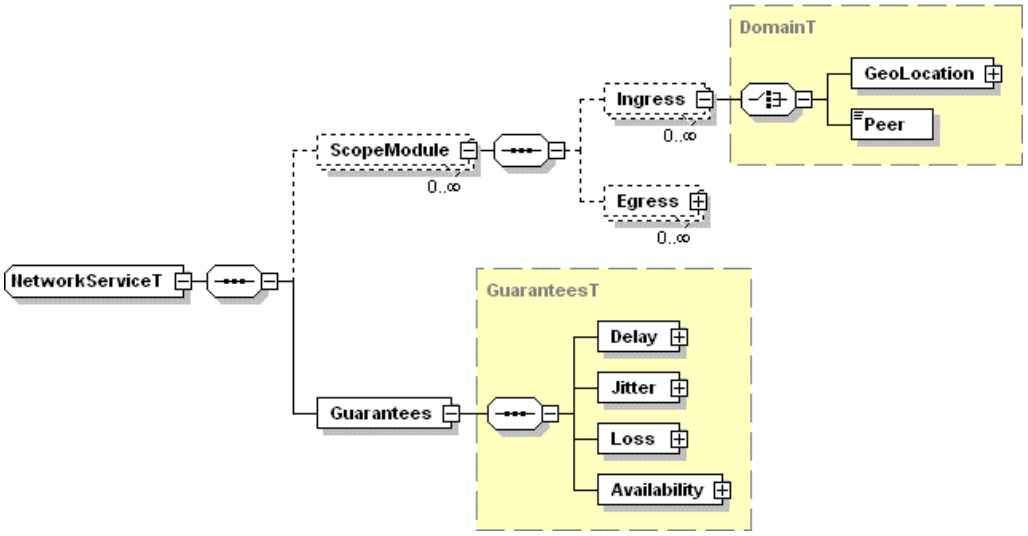

Figure 10: Network Service.

A network service is defined to inform the NP/PI layer on the permissible combinations of guarantees and scope and to allow for setting different Engineering Guidelines (cf. section 3.3).

3.3 NP Engineering Guidelines

Business layer provides guidelines on admitting or denying CPA requests and associated cost conditions, guidelines on resource allocation, routing constraints and preferences, etc (see Figure 11). These guidelines may be associated to either Network Service or CPA/NIA request aspects. The scope that a particular guideline is to be applied is called the *guideline scope*. The guideline scope can be:

- the entire network,
- a Network Service,
- a limited intra-/inter-domain topological scope,
- a CPA/NIA service request type or
- any combination of the above.

The target CPA/NIAs service request types are captured based on CPA/NIA attribute values. A CPA/NIA type may include a broad set of CPA/NIAs or be reduced to a specific CPA/NIA. Attributes may refer to the SP or upstream INP, to the volume of the anticipated demand, etc.

Guidelines for admitting or denying a CPA/NIA request are provided.

Cost constraints may be specified for the entire network, per network service or CPA/NIA type. The maximum cost may refer to intra-domain, inter-domain or total cost. The formula to calculate the cost is determined by business layer and fed to the NP/PI layer functions.

Resources are assigned to a network service overall or in defined scope, or per CPA/NIA types. Specifically, the resource allocation guidelines are expressed in terms of maximum, excess, and provisional resources. Maximum is specified as a weight or absolute value denoting the capacity to be allocated to the network service or CPA/NIA type without further consulting the business layer. The weight determines the share of overall capacity comparing to the other network services and CPA/NIA types. Excess is specified as a weight denoting the portion of the free capacity to be allocated to the guideline scope. Free capacity is the capacity that does not correspond to contractual demand from established CPA/NIAs. At bootstrapping all capacity is free capacity. Business layer may desire to reserve specific capacity provisionally for trial CPA/NIAs, or for CPA/NIA requests that are anticipated in the short-term with significant certainty. For allocation of resources, the guideline scope cannot be the entire network. Similarly, provisional capacity can be specified for Network Services or CPA/NIA types only within specific scope.

Routing preferences with respect to either adjacent or remote INPs may be specified for the entire network, per Network Service or CPA/NIA type. Note that individual CPA/NIAs may impose additional routing preferences and constraints specified by the SPs. Differentiation of routing constraints among CPA/NIAs within the same network service is allowed in general if the Network Plane realisation technique allows it.

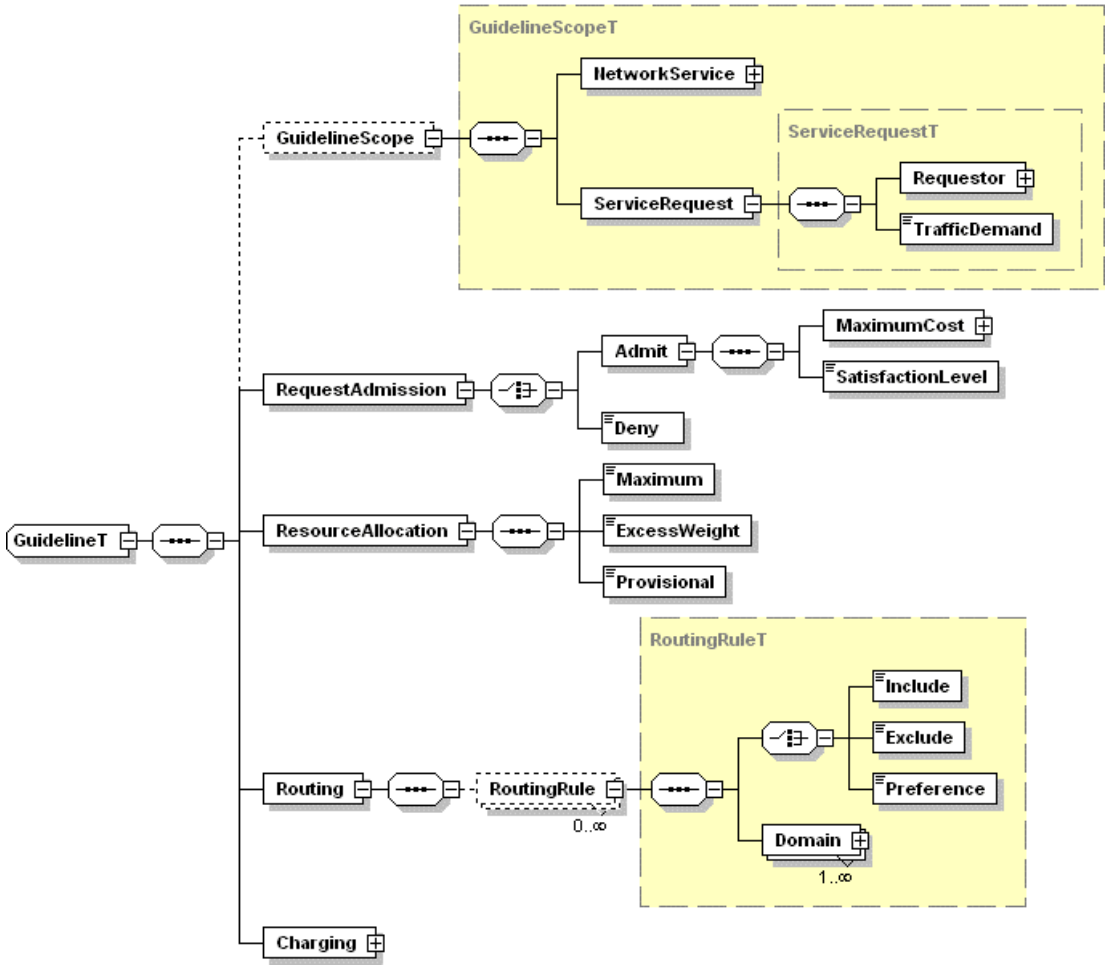

Figure 11: *NP Engineering* Guideline.

3.4 NP Definition

In an INP domain, a NP represents, in an abstract form, a certain set of capabilities for transporting traffic flows across the domain.

A distinct Network Plane is created to allow for setting different intra-domain traffic engineering targets. Differentiation in traffic treatment may exist even within Network Planes, however in more manageable forms.

Network Planes are defined in terms of their QoS and availability capabilities (see Figure 12). A Network Plane is therefore defined by the treatment granted to the traffic it carries. How to inject traffic to a Network Plane in general is not part of the Network Plane definition.

Each Network Plane may be designed to carry traffic of CPAs with diverse routing constraints and preferences, or with the requirement to maintain a unique identifier of their traffic across the Network Plane, with particular monitoring requirements, etc. As a result, each Network Plane may or may not need to support respectively differentiated routing within the Network Plane itself, flow identification, and particular monitoring capabilities.

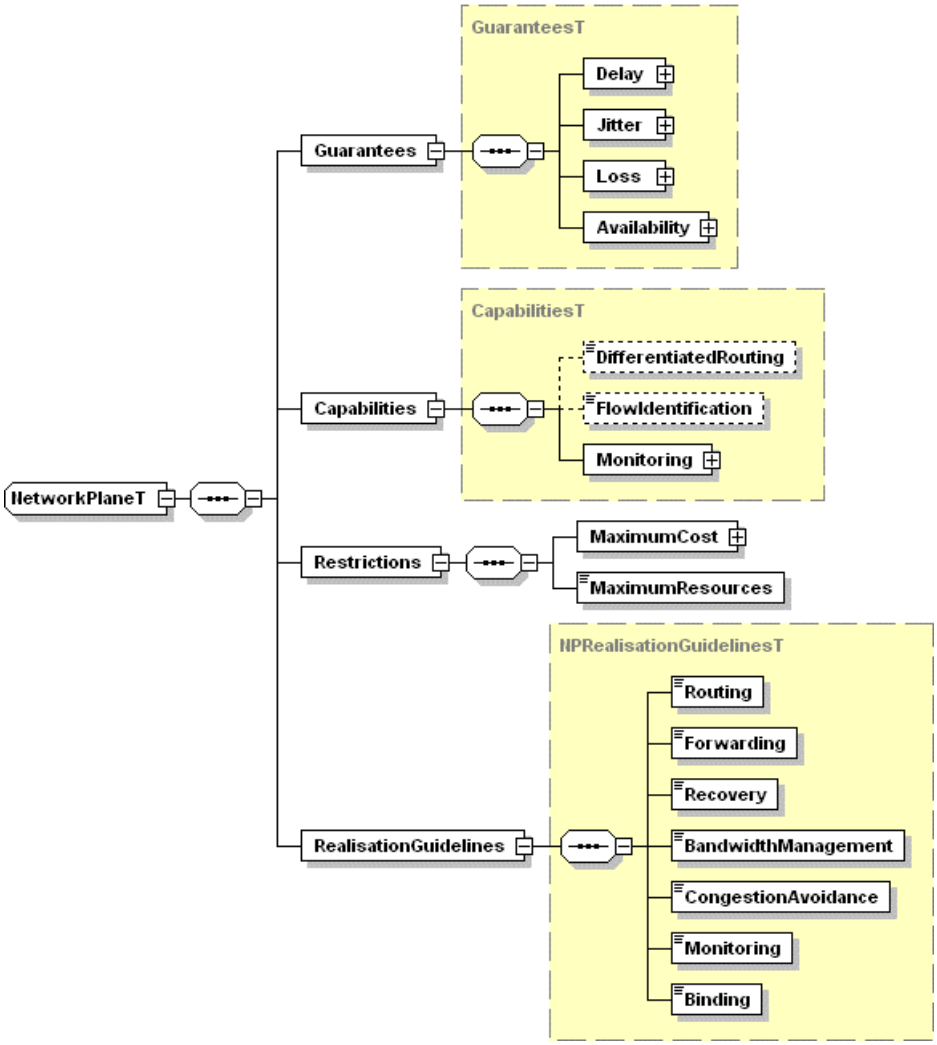

Figure 12: Network Plane Definition.

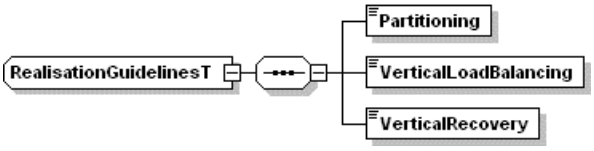

Figure 13: Network level realisation guidelines.

A Network Plane may be designed to accommodate traffic using internal resources in a way that the internal cost, calculated based on the formula provided by the business layer, is bound to a given value. This is considered a restriction for all traffic to be carried by the Network Plane.

Network Planes are designed for granting resources to existing CPA/NIAs in a way optimum from network administration perspectives and consistent to the Engineering Guidelines (cf. section 3.3). If new CPA/NIAs arrive, or if the anticipated demand from the existing CPA/NIAs is considerably different from the actual demand, then NP/PI layer may need to re-define Network Planes; the scale of the traffic demand plays an important role in deciding how to merge or split Network Planes for accommodating different CPA/NIA requirements. Hence, NP/PI layer may need to impose restrictions on the maximum amount of network resources that a Network Plane is entitled to without violating the assumptions that led to its definition.

Note that the restrictions on resource allocation per Network Plane set by the NP/PI layer should not be confused with the guidelines on resource allocation set by the business layer. The former are set as

a trigger to re-evaluate the defined Network Planes, potentially leading to their re-definition, splitting or merging. The latter are set a) to control the admission of new CPA/NIA requests, b) to favour certain network services and/or CPA/NIA request types among the existing ones by determining weights on distribution of the free capacity, and c) to invoke network planning when physical resources become scarce.

NP/PI layer selects the appropriate, among those available, techniques for the realisation of Network Planes. Such techniques may be applicable overall, e.g. techniques for partitioning the network to Network Planes or for vertical load-balancing across Network Planes (see Figure 13), or they may apply per Network Plane. Specifically, the following aspects are considered in Network Plane realisation:

- *partitioning* is the technique for differentiating traffic across Network Planes. The field(s) for unambiguously identifying traffic belonging to different Network Planes are specified after ensuring that the range of unique values is sufficient, e.g. 64 values for DSCP classifiers. A selection of techniques may be available based on differentiated forwarding and/or differentiated routing (see [D1.1], section 6.2.3),
- *vertical load-balancing* is the traffic engineering technique for balancing the overall load among Network Planes (see [D1.1], section 6.2.4.2.2),
- *vertical recovery* is the technique that relies on re-directing traffic across compatible Network Planes to recover from failures (see [D1.1], section 6.2.4.2.3), note that vertical load-balancing and recovery are traffic engineering enhancements, and as such optional,
- *routing* refers to route construction and route selection processes, routing can be common across Network Planes if routing is not used for service differentiation within this NP or specified per Network Plane; in the latter case, completely different techniques may be chosen per Network Plane resulting in a hybrid network provided that the partitioning technique allows it, or different values for the parameters of a common technique that allows differentiation internally, e.g. setting different route selection optimisation targets such as minimise delay versus load balancing (see [D3.1] section 4.1 for a selection of envisaged techniques),
- *forwarding* determines the scheduling technique, strict priority or Fair Queuing scheduling techniques, possibly specifying forwarding classes within a Network Plane, or sharing forwarding classes across Network Planes, or no forwarding differentiation at all,
- *robustness* encompasses techniques for recovering from physical failures, for ensuring protection against security threats, for ensuring resilience to environment changes, etc.; similarly to the routing technique, robustness techniques may be specific to a Network Plane or common across the network; in the former case completely different techniques may be chosen per Network Plane, e.g. recovery at the physical layer versus MPLS Fast Reroute, or different values for the target MTTR or operational parameters for a common technique,
- *bandwidth management* determines the approach for allocating bandwidth to competing flows, hop-by-hop explicit bandwidth reservation using RSVP can be one alternative, or admission control at the edges relying on model-based or measurement-based techniques to admit traffic, employing peak allocation or statistical multiplexing techniques with varying aggregation weights, etc.; differentiation on bandwidth reservation and policing techniques may apply at the CPA/NIA level within Network Planes when Network Planes are not created with the criterion to differentiate on bandwidth management,
- *congestion avoidance* determines whether ECN, RED etc. techniques are used to signal congestion building up to reactive applications in case bandwidth management uses statistical multiplexing,
- *monitoring*, specifies the use of active or passive monitoring techniques, the sampling frequency, the metrics, the granularity, the aggregation methods, applying to the entire network or to particular Network Planes,
- *binding* refers to the technology for interconnecting a Network Plane to a Network Plane downstream; binding includes a set of techniques, routing, forwarding, robustness, etc. to be used

at the domain's boundary; based on the corresponding intra-domain techniques only a set of compatible inter-domain techniques will be applicable, a Network Plane may be dedicated to use only one of these compatible techniques comparing to other Network Planes or all Network Planes may potentially use any compatible technique.

3.5 PI definition

From INP perspectives, a Parallel Internet represents, in an abstract form, a certain set of capabilities for transporting traffic flows in the Internet.

The NP/PI layer defines Network Planes matching the requirements of existing CPA/NIAs based on the intra-domain network capabilities and the capabilities provided by downstream INPs. As a result, each defined Network Plane is either designed to serve traffic to local destinations, or designed to be bound to a Network Plane of the downstream INPs to form a Parallel Internet. Note that Parallel Internets may span the end-to-end path but in general are defined from the perspective of the INP that binds its Network Planes with downstream Network Planes to form its own Parallel Internets.

A Network Plane may participate in more than one Parallel Internets, differentiated only at the inter-domain treatment. Parallel Internets are built to accommodate traffic from CPA/NIAs following the guidelines set by business layer in terms of Network Services. A Parallel Internet is defined by its capabilities to reach the final destinations and uses as a base a Network Plane (see Figure 14).

The Parallel Internet capabilities will be the basis for accommodating CPA/NIA requests to remote destinations. Similarly to a Network Plane, a Parallel Internet may support differentiated routing, flow identification and particular monitoring capabilities, as appropriate for accommodating CPA/NIA requests with such requirements. Additionally to the cost restriction, a Parallel Internet may, by design, need to adhere to restrictions on scope and inter-domain routing.

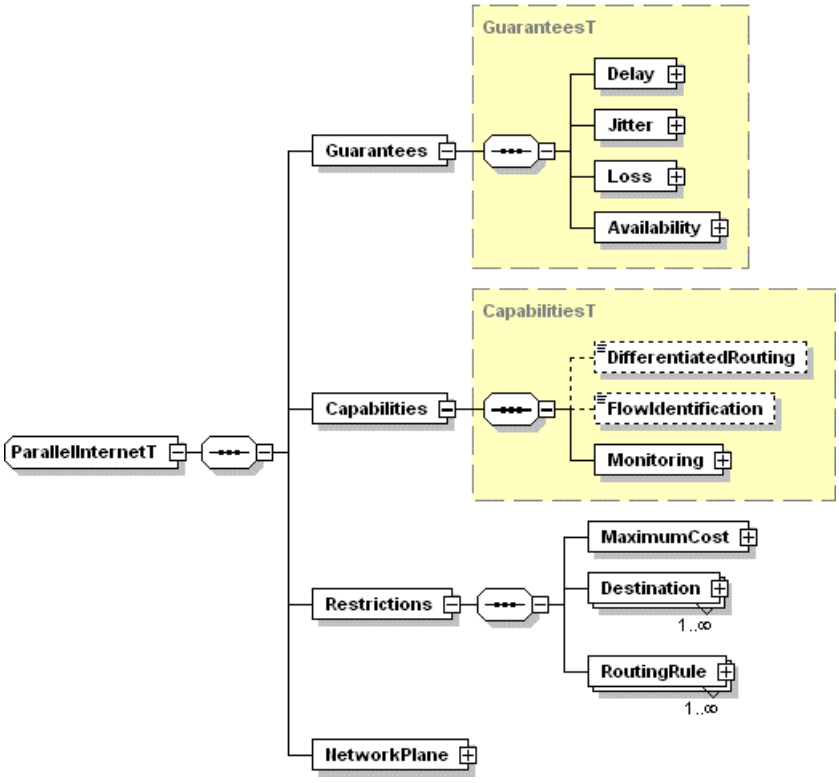

Figure 14: Parallel Internet definition.

A Parallel Internet can be formed by binding a Network Plane to a set of alternative inter-domain guarantees. The inter-domain guarantees are to be provided by the network capabilities made available by the downstream INPs. Alternative inter-domain guarantees may be provided. The exact NIAs to be

established will be formulated and negotiated by the traffic engineering and negotiation functions of the technology-specific layer.

3.6 The NP/PI Problem

3.6.1 Overall set-up

This section offers a theoretical treatment to the problem of NP/PI. Broadly speaking, the technology agnostic NP/PI layer in INP domains is concerned with the determination of solutions for the following equation:

$$NP \oplus \{NIA\} = PI \quad (1)$$

such that:

$$\{NS\} > \{PI\} \quad (1a)$$

$$\{NP\} > \{TC\} \quad (1b)$$

The variables in the above system are the set of Parallel Internets $\{PI\}$ that the INP needs to provide for accommodating the traffic flows of the services it offers (CPAs/NIAs), the set of Network Planes $\{NP\}$ to create locally and the set of Network Interconnection Agreements with downstream providers $\{NIA\}$ to establish for instantiating the Parallel Internets. It should be noted that these variables are mutually independent, in that each one of them needs to be determined individually in other words, cannot be derived from any combination of the others.

The set of the network services to offer $\{NS\}$ and the set of technology-specific capabilities $\{TC\}$ are assumed to be known; they are provided as input by the business layer and the technology-specific layers, respectively. For the sake of problem formulation, we have assumed that the technology-specific capabilities can be modelled in information entities (attribute value pairs), the attributes of which can be mapped to attributes of the information models representing NPs, PIs and NIAs.

The convolution symbol $\oplus$ denotes a generalized operation, of additive nature, which when applied to the values of compatible parameters (attributes) of NPs and NIAs yields a result value for the parameter. Note, that by their definition, the entities NP, NIA and PI have compatible attributes e.g. cost, performance guarantees. The generalized operation resolves to usual mathematical operations or well-defined algebraic expressions depending on the nature of the parameter under operation. For example, in the case of a cost parameter it resolves to the sum and in the case of a performance bound to the maximum.

The symbol $>$ denotes a generalized comparison operand, of less than or equal nature, which when applied to two sets of elements means that for every element in the left set there is an element in the right set which can 'accommodate' the element of the left set, in that the values of all parameters of the left element are less than or equal than the values of the corresponding (compatible) parameters of the right element.

Equation (1) intuitively says 'combine local and external abstract network capabilities and restrictions to yield the end-to-end characteristics of the different 'types of traffic' that can be provisioned and delivered by the INP'. Constraint (1a) says 'the different end-to-end provisioning characteristics should be such that they can accommodate the required characteristics of the services to be offered by the INP'. Constraint (1c) says 'the local abstract network capabilities should indeed be realised based on the specific capabilities of the particular technology(ies) employed in the INP domain'. Collectively, the system (1), (1a) and (1b) says: 'determine local and external abstract network capabilities that can be implemented, which combined together yield the required end-to-end provisioning characteristics of the services to offer'.

Clearly, the first question to be answered is: '*is the NP/PI problem feasible?' i.e.' are there any solutions to the system of (1), (1a) and (1b)?*'

We conjecture that indeed there are feasible and non-trivial -other than based on a single plane- solutions since (a) the set of network services to offer and the NPs to create are bound to a common set of technology-specific capabilities, (b) the external network capabilities are compatible with local ones, as they both refer to IP transportation characteristics and (c) there are technologies for differentiated routing e.g. MTR or differentiated forwarding e.g. DiffServ or differentiated inter-domain routing e.g. tunnelling approaches. Because of (a) and (b), equation (1) can be broken down to a set of independent algebraic equations, for which, obviously, there are solutions. Because of (c), non-triviality is justified.

As there is not a unique solution to the system (1), (1a), (1b), the next question that naturally follows is: '*out of the many solutions, which one to select to implement?*' Note that an INP domain can only have one set of PIs, therefore NPs and bindings with NIAs, configured in the network, based on which the offered network services will be provisioned and delivered.

Obviously, the selection of one out of the many feasible solutions requires that certain criteria are in place. These criteria depend on the specific business-driven policies for service provisioning and delivery and operational practices of the particular INP domain. As such, they should reflect INP business, network optimisation and operation targets.

As the entities/variables pertinent to the solution are mutually independent, the determination of the solution to select for implementation depends on the particular context where the solution is sought for, with the context being set according to which of the variables are considered to be known.

3.6.2 NP/PI Problem Space

The variables pertinent to the NP/PI problem assume discrete values and they are finite in number (the cardinality of their sets is finite). This is justified below.

The NPs to realise can be regarded as vectors in a multi-dimensional space, where each axis corresponds to a dimension along which service provisioning can be differentiated (cf. attributes 'guarantees', 'capabilities', 'restrictions' NP model clauses, section 3.4). In each axis there is an ordered set of finite values. These values reflect the level, the grade, of differentiation that can be provided along this 'service provisioning differentiation dimension', by means of the technology-specific capabilities of the INP domain.

The axes/dimensions of the NP space are determined according to the provisioning requirements of the Network Services and the requirements posed by the Engineering Guidelines. For instance, it may be necessary to incorporate a 'Differentiated Routing' axis, with values the number of routing planes that can be supported in the network, should this is deemed as the means to provide enhanced levels of availability. Or, a 'Maximum Resource' axis, with discrete percentage values, may need to be incorporated for enabling the enforcement of resource allocation Guidelines. Conclusively, the axes/dimensions of the NP space represent abstract network capabilities and as such, there should be ensured that their instantiation is indeed feasible by the technology-specific capabilities of the INP domain.

Last, it should be noted that NPs may not necessarily correspond to all possible combinations of values in the axes/dimensions. This is because there may be incompatibilities or interoperation problems between the technology-specific employed mechanisms.

Similarly, NIAs can be regarded as vectors in a multi-dimensional space, where the axes correspond to the traffic transport capabilities offered by INPs such as guarantees, bandwidth and cost. The NIAs are discrete and finite as the offered transport capabilities assume discrete values and the number of INPs is finite.

Finally, the space of PIs can be regarded as the Cartesian product of the NP and NIA spaces. As these spaces are discrete and finite, so is the PI space. It should be noted that PIs may not necessarily

correspond to all possible pairs of NPs and NIAs, as there may be technological incompatibilities between the underlying technology-specific intra- and inter-domain mechanisms.

3.6.3 NP/PI Problems

Depending on which of the entities/variables of equation (1) are known, a number of problems can be stated, as shown in Table 1. First, the following terminology is put forward:

$\{PI_{req}\}$: The set of the required PIs to provide. The PIs in this set are specified as far as their attributes that represent their provisioning characteristics (cf. 'guarantees', 'capabilities', 'restriction' clauses, section 3.5) are concerned; the attributes specifying their realisation in terms of the NP and NIAs to use, are not specified.

The set of the required PIs can be directly deduced by the Network Services and Engineering Guidelines determined by the business layer. The attributes representing the provisioning characteristics of the Network Services (cf. section 3.2) can directly map to corresponding attributes of the PI model (cf. section 3.5). The rest of the PI attributes can be specified by the Engineering Guidelines.

Note that, in the general case, there may not be a one-to-one correspondence between the required PIs and the Network Services. For example, Guidelines on 'routing' e.g. to favour a particular inter-domain route, may necessitate the provision of multiple PIs for the same Network Service.

The process of defining the set of required PIs is considered to be a process that requires human intervention in order to ensure that the set of the required PIs can indeed accommodate the provisioning characteristics of the Network Services, while it allows the enforcement of Engineering Guidelines.

$\{NIA_o\}$: The set of the NIAs that can be provided by downstream INPs. It is assumed that INPs have means to advertise the capabilities that can offer to other INPs and conversely, to discover the offerings of other INPs.

$\{NIA_{oest}\}$: The set of NIAs established with downstream INPs. This set is a subset of the previous set. An established NIA is an instance of an element of the set of offered NIAs, agreed with a particular provider for specific capabilities at certain bandwidth and cost.

$\{NP_s\}$: The set of Network Planes determined for realisation in the network –for instantiating the determined PIs.

$\{PI_s\}$: The set of Parallel Internets determined to be provided by the INP for accommodating the required Network Services. Compared to the set of required PIs, $\{PI_{req}\}$, the PIs in this set are fully specified i.e. their realisation in terms of the NP and NIAs to use, are specified.

Problem	Given (*)	Determine (**)
NP/PI Definition	$\{NS\}$ -set of Network Services. $\{NIA_o\}$ -set of offered NIAs.	$\{NP_s\}$ -set of Network Planes to realise. $\{PI_s\}$ -set of Parallel Internets to provide.
NP Definition	$\{PI_{req}\}$ -set of required PIs. $\{NIA_o\}$ -set of offered NIAs.	$\{NP_s\}$ -set of Network Planes to realise. $\{PI_s\}$ -set of Parallel Internets to provide.
PI Definition	$\{NP_s\}$ -set of Network Planes to realise. $\{NIA_o\}$ -set of offered NIAs.	$\{PI_s'\}$ -set of Parallel Internets that could be potentially provided.
NIAs-to-get Definition	$\{PI_{req}\}$ -set of required PIs. $\{NP_s\}$ -set of Network Planes to realise.	$\{NIA_o'\}$ -set of NIAs that ideally should be offered.
NP Realisation	$\{NP_s\}$ -set of Network Planes to realise. $\{TC\}$ -set of possible technology-specific capabilities.	$\{TC_s\}$ -set of technology-specific capabilities selected to realize the determined NPs.

Table 1: NP/PI problems.

(*) For clarity, the table depicts per problem only the input, which is related to the variables under consideration, omitting additionally required input such as network topology, traffic matrix, set of established CPAs, NIAs, bought or sold.

(**) For each problem, the determination of the solution variables is such that certain criteria are optimised, subject to the constraints of system (1).

The problem ‘NP/PI Definition’ boils down to the ‘NP Definition’ problem, if the set of required PIs is deduced by the set of Network Services, or an iterative procedure is followed whereby at each step a possible set of PIs to provide is specified.

In the above problems the set of offered NIAs, could be replaced by the set of established NIAs bought from downstream INPs. This is particularly the case for INPs, which are bound to certain types of NIAs or inter-domain technology.

The NP/PI problems may not necessarily be concerned only with the definition of NPs and PIs. The problems ‘PI Definition’ and ‘NIA-to-get Definition’ assume that the INP has decided on the NPs to realise. However, still these are valid problems, as their outcome can be of value to the business layer for decision making related to network development.

Finally, it is worth noting that the above NP/PI problems generalise the traditional traffic engineering (TE) problem. The TE problem relates to the mapping of user traffic requirements to the configuration of the particular intra/inter-domain TE mechanisms employed e.g. setting up of IGP weights, forwarding parameters. The NP/PI problem refers to a higher abstraction level, which sees abstract network capabilities, commonly understood in the IP world such as routing and forwarding, rather than individual TE mechanisms of a specific technology. The NP/PI problem is then concerned with the composition, putting together, of these abstract network capabilities to create *virtual network segments* as appropriate to accommodate the different requirements of the offered services.

In fact, the outcome of the NP/PI problems constitutes the required input to the traditional TE problem. Table 2 expresses the TE problem in terms of the variables pertinent to the NP/PI problem. The set of NPs and PIs determined to be realised in the network, specify the user traffic requirements, locally and end-to-end respectively, which are required for TE to apply.

Problem	Given (*)	Determine (**)
Intra/inter-domain TE	$\{NP_s\}$ -set of Network Planes to realise $\{PI_s\}$ -set of Parallel Internets to provide	$\{NIA_{oest}\}$ -set of established NIAs $\{NetConf\}$ -set of network configurations, specific to the particular technology employed

Table 2: The general intra/inter-domain TE problem.

(*), (**) See comments in Table 1.

3.6.4 The NP Definition Problem

This section elaborates on the 'NP Definition' problem in an attempt to gain insight into its complexity. Similar considerations apply to the other NP/PI problems.

3.6.4.1 Optimisation Criteria

As already discussed in section 3.6.1, the optimum solution, the set of NPs to realise, has to be sought for against certain optimisation reflecting business, network performance and operations targets. In particular, we see a set of optimisation criteria as follows:

- Maximise customer satisfaction i.e. integrity of INP in honouring the established CPAs/NIAs.
- Minimise network cost i.e. amount of resources required.
- Minimise operational cost and overhead.

Clearly the above set of criteria constitutes a triple trade-off, in that all three cannot be optimised, maintained at their desired levels, at the same time. Customer satisfaction is maximised with near-to-peak resource allocation schemes, which obviously increases network cost as well as operations for performance assurance. As the amount of network resources is tried to be kept at minimum, the operations complexity and therefore cost inevitably increases e.g. human intelligence and/or sophisticated mechanisms need to be in place.

3.6.4.2 Greedy Solution Approach

Since the problem space is finite (cf. section 3.6.2), a solution the problem can be found following a greedy approach, relying on exhaustive evaluation of all possible combinations of the variables pertained. The greedy approach is outlined below:

Step 0 - Init: Construct the NP solution space i.e. lay down the abstract network capabilities for accommodating the differentiated traffic flows characteristics.

As outlined in section 3.6.2, the NP solution space is constructed taking into account the provisioning requirements of the Network Services and the requirements posed by the Engineering Guidelines, having in mind the technology-specific capabilities employed in the INP domain.

This step is considered as a preliminary, initialisation step, requiring human intervention.

Step 1: Construct the set of feasible NPs, $\{NP_f\}$.

A feasible NP is a NP in the solution space determined in the previous step, for which there can be found NIAs in the set of offered NIAs so that if combined together, one of the required PIs is yielded, that is, it satisfies the following equation:

$$NP_f \oplus \{NIA_o\} = \{PI_{req}\}$$

By definition of the problem the latter two sets in the above equation are known. So, the above equation has one unknown and thus feasible NPs can indeed be determined.

Note that for a given required PI, a number of NP_f 's can be found and therefore, the set of the required PIs can be instantiated via a number of *alternative configurations* - combinations of NPs and NIAs. Say that there are Φ such alternatives and let $\{NP_f^{(i)}\}$ denote the set of feasible NPs in the ith alternative; the NPs contained in each of these alternatives, combined with appropriate NIAs, yield all required PIs.

The set $NPF \equiv \{\{NP_f^{(i)}\}, i = 1..\Phi\}$ constitutes the set of feasible solutions for the optimisation problem at hand.

Step 2: Find the optimum solution, set of NPs, $\{NP_s\}$ to realise the required PIs.

A) Evaluate each feasible solution determined in the previous step with respect to the optimisation criteria set for the problem. It is assumed that there exists an *evaluation function*, which for a particular alternative PI configuration, that is, a set of NPs and associated NIAs, $\{NP_f^{(i)}\}$, computes appropriate *metrics*, which substantiate the considered optimisation criteria. For example, such metrics could be goodput for customer satisfaction, average allocated link capacity for network cost and number of configuration complexity –weighted sum of configuration commands- for operational cost.

B) Select the 'best' solution, $\{NP_s\}$, by qualifying the feasible solutions on the basis of the metrics they yield e.g. by sorting, if the metric is numeric.

It should be noted that the NPs determined by the above procedure, may not necessarily correspond to the required PIs on a one-to-one basis. In general, the set $\{NP_s\}$ is smaller in cardinality than the set $\{PI_{req}\}$. In other words, there may the case that the same NP may be used in instantiating two or more required PIs. In such a case, the network should be able to classify the PI flows within the same NP, as these flows will receive different inter-domain treatment; such capabilities exist for instance in MPLS/DiffServ networks multiple DSCPs can be assigned for the same OA. Should the network cannot provide for such capabilities, the optimal solution should be searched with the constraint that the resulting NPs should be one-to-one with the required PIs.

The above analysis shows that the 'PI Definition' problem is hard even to be stated as a mathematical optimisation problem; let alone to be solved analytically. It is hard to find a closed-form expression for the optimisation function in terms of the variables pertained to the problem, NPs, NIAs, PIs.

Inevitably one has to rely on manual and/or computational methods for solving the problem. Although its solution has been described in a procedural way, a fully-automated procedure is hard to be developed. In any case, we believe that the 'NP Definition' problem better be solved in a manual way, as human intervention is required for customising, guiding, checking and controlling the execution of the solution procedure throughout its steps. The value of the above analysis lies in that it provides a systematic way to solving the problem, no matter whether it can lead to a fully automated procedure or not.

A key element in the above solution procedure is the existence of a function for evaluating the optimality of the various alternative configurations for instantiating the required PIs. For computing the required metrics, the function should incorporate the TE algorithms and mechanisms employed in the domain as well as it should provide for a (simulation-based) model for inferring the performance of the engineered network. The Network Plane Emulation Platform (NPEP), specified and developed by the project (cf. chapter 4) is an example of such an evaluation function. Clearly, the complexity of such a function adds to the overall complexity of the solution procedure and the optimality of the solution NPs is subject to the errors and assumptions inherent to the function.

3.6.4.3 A Differential View

In the following, the 'NP Definition' problem is looked from the standpoint of its solution space and the traversals therein towards the optimum solution. The idea is to gain insight on the how to move across the points, feasible solutions, in the solution space, from an initial point to another point and so on, until we reach the point where optimality of the objective function is attained. For instance, in optimisation problems of a single real variable, one should move along the direction, left or right, where the derivative of the objective function tends to zero. Evidently, should such 'direction indicators' could be found for the problem at hand, a more efficient than the greedy solution method could be specified.

Following the terminology introduced earlier in the chapter, in the set of feasible solutions NPF (cf. step 2 of the solution procedure in section 3.6.4.2) – alternative configurations for instantiating the required PIs- we define an ordering relationship, called *outclassing* based on the comparison operand $>$ (cf. section 3.6.1). That is, the jth PI configuration is said to outclass the ith, similarly the jth is downclassed to the ith or the ith is outclassed to the jth if and only if the following holds:

$$\left\{NP_f^{(i)}\right\} < \left\{NP_f^{(j)}\right\}$$

Effectively the above means that flows of certain required PIs will be transported across the domain through 'better' NPs.

Clearly, outclassing is a partial ordering relationship; for instance a NP with <delay=low, availability=high> in one PI configuration cannot be compared with an NP with <delay=high, availability=low> in another PI configuration.

The 'NP Provisioning' problem is an optimisation problem in a partially ordered finite set. A lexicographic type of ordering, yielding a complete ordering, could be defined, should the INP was able to prioritise the dimensions along which the offered services are differentiated.

We conjecture that there are maximal and minimal PI configurations in NPF under outclassing ordering as defined above. Maximal PI configurations contain the maximum possible NPs –intuitively, as many as the required PIs- and minimal PI configurations contain the minimum possible NPs – intuitively, just one- for instantiating the required PIs. Hence, maximal PI configurations compared to minimal have sets of NPs of smaller cardinality. In the general case, there may be multiple maximal or minimal PI configurations.

We call the PI configurations other than the maximal or the minimal ones as *intermediate*. Intuitively, the intermediate PI configurations lay between maximal and minimal configurations. From a maximal PI configuration we can reach an intermediate one by outclassing along certain (a subset of the) provisioning dimensions and so on until a maximal configuration is reached. We call this popping *NP-merging*. Similarly, through *NP-splitting* i.e. by downclassing along certain provisioning dimensions, from a maximal PI configuration we can reach a minimal through intermediate ones.

With the NP-merging and NP-splitting operations the set of set of feasible solutions NPF can be regarded as a fully connected graph, with nodes being the alternative PI configurations, in the sense that one can pop from one any other point. Intuitively, the maximal and minimal PI configurations form the perimeter of this fully connected graph.

Based on the above, the 'NP Definition' problem can then be stated as: starting from a maximal/ minimal PI configuration, how should I go NP-merging/NP-splitting to the end of reaching the configuration attaining the optimum criteria?

The optimum solution to the above formulated problem could be determined as a shortest path solution, provided there were means to substantiate the effect of NP-merging/splitting as the weights of the links in the fully mesh graph of the feasible solutions. This effect is the delta of the evaluation function used in evaluating configuration alternatives in the greedy approach (see previous section), with respect to changes in PI configurations i.e. sets of NPs to realize. The delta to NP changes is hard

to calculate, as the evaluation function depends, besides the set of NPs to realise, on multiple variables -input parameters- such as traffic demand estimates per required PI.

Intuitively, by NP-merging:

- Operational cost may be reduced as the number of NPs is reduced.

However, as traffic from different PIs is mixed in the same NP:

- Customer satisfaction may deteriorate, given the aggregate morality of the IP traffic engineering schemes, which usually avoid of relying on per flow reservation schemes for scalability reasons.
- There is the 'paradox' of provisioning different services through the same means, thus practically having the same cost intra-domain.

On the other hand, NP-merging may be justified when:

- The traffic volumes of the required PI flows are not such that to justify a different NP.
- Intra-domain differentiation for certain PIs (is proved by experience that) it does not play a significant role in end-to-end performance.

The above arguments indicate that even if there are means to compute the delta of the evaluation function for computing the effect of NP-merging/splitting, still there would be need for human intervention in order to guide and control the move from one feasible PI configuration to another.

The presented analysis raises mathematical interest. Further investigation along the outlined track, also possibly viewing the 'NP Definition' problem from alternative perspectives such as from mathematical topology or finite fields perspectives, is left for future work.

3.6.4.4 Dynamicity - 'On-line' Version

So far, the 'NP Definition' problem has been analysed in a static, so as to say 'one-off', form. An 'on-line' version of the problem can be considered. This problem version entails the determination of the optimum set of NPs to realise over a time period during which there are time epochs where specific conditions warranting the re-determination of NPs emerge; different sets of NPs may need to be determined at each time epoch. Examples of such conditions include:

- Changes in technology-specific capabilities.
- Introduction of new services i.e. types of traffic flows to handle.
- Emergence of new players, enhancing the options for NIAs and the potential for CPAs.
- Significant changes in PI traffic volumes e.g. caused by admitting CPAs.
- Deterioration of expected performance, intra/inter-domain.
- Changes of marginal effect of intra/inter-domain performance to end-to-end performance.

The 'on-line' version is formulated similarly to the static problem, with additional input a list of conditions to emerge expressed in probabilistic terms.

Compared to the static version, the 'on-line' 'NP Definition' problem is more practical and useful. Scenarios regarding network evolution -from business, traffic and infrastructure perspectives- may be executed and evaluated. However, it is of increasing complexity.

The increased complexity of the 'on-line' 'NP Definition' problem is because it applies to a time period –a series of epochs with conditions changing from one epoch to another. The set of NPs to realise should be determined against overall that is, over the period, optimisation criteria. A kind of 'best positioning' optimisation criteria should be specified.

Broadly speaking, as far as solving the 'on-line' 'NP Definition' problem is concerned, the greedy solution approach and the differential view of the static problem can still apply. A sort of 'look ahead' intelligence needs to be incorporated. The specification of a solution approach, even a greedy one, is of staggering complexity as the length of the look-ahead window increases. Even if a solution

procedure is feasible to specify, for small length windows, its validity would be questionable given the underlying assumptions and the errors inherent to the model used. A step-by-step, trial-based approach may after all be the best way to go around.

3.6.5 NP-based Performance – NP Realisation

As it became apparent from the previous analysis, the ability of measuring network performance under various PI configurations i.e. sets of NPs for instantiating the required PIs, is crucial in determining the set of NPs to realise. To that end, inevitably one resides on analytical and/or simulation-based methods, like NPEP.

For being able to safely, within reasonable statistical errors, predict network performance, *robust NP realisation* becomes a critical issue. Ideally, NP's should be realised so that to yield an almost-deterministic behaviour with respect to:

- the volume of traffic they can deliver according to the specified provisioning characteristics, and
- (the pattern of) the resources they consume.

Then, valid models could be derived for predicting the performance of NPs and the network as whole and answering hypothetical questions such as:

- What is the impact of a resource failure?
- Where and by how much, resources need to be upgraded?
- What is the effect of merging or splitting NP?

Robust NP realisation should be set as a criterion for selecting the most suitable technology-specific mechanisms for realising NPs with given provisioning characteristics, should alternative ones are available.

Finally, robust NP realisation allows the INP to maintain a complete and valid view of its network with respect to the services it offers.

3.7 INP Internal Interfaces

3.7.1 Overview

Internal interfaces within the IP Network Provider are defined based on the functional architecture specified in [D1.1].

Business-based Network Development sets the targets for the *NP Engineering* components to fulfil, specifically, the Network Services to be supported and the guidelines for handling the demand for these services. Target network services are expressed in terms of QoS and availability performance metrics and are optionally restricted to a defined set of local or remote destinations.

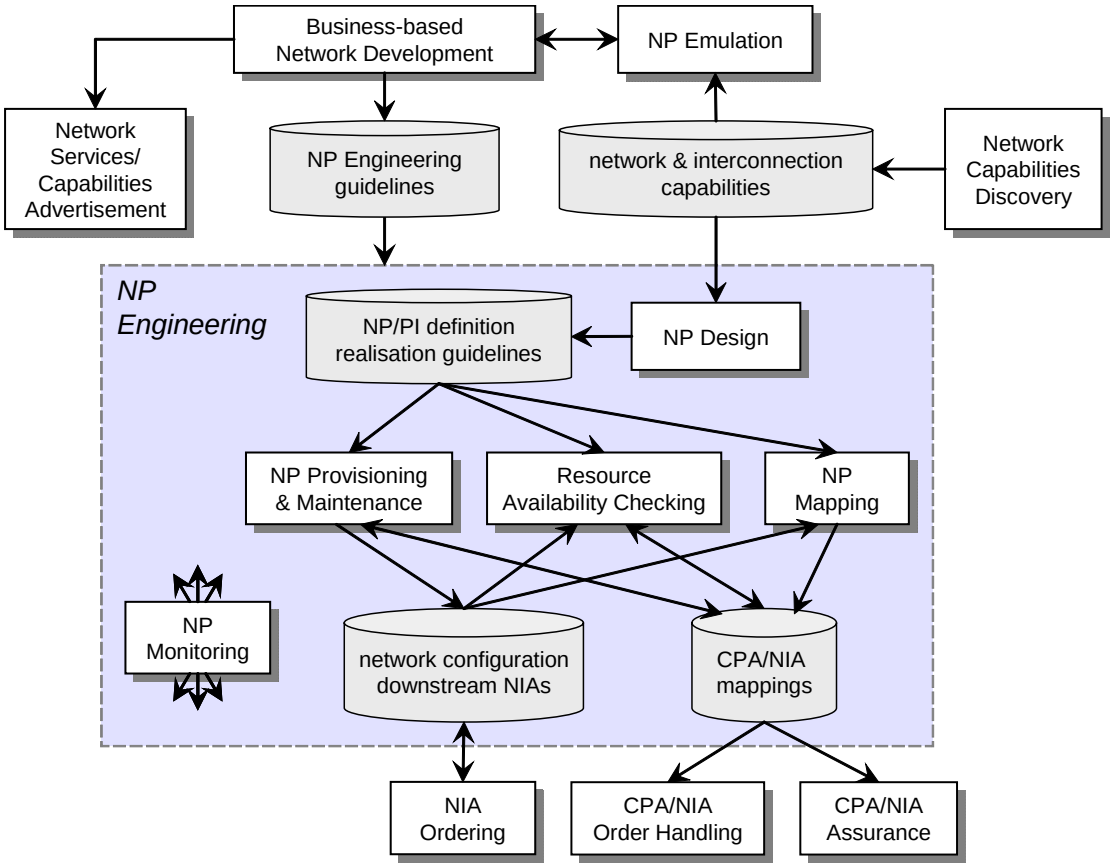

Figure 15: INP architecture and information flow.

NP Design & Creation defines the Network Planes and Parallel Internets, required to fulfil the *Business-based Network Development* targets. Network Planes and Parallel Internets are defined in terms of abstract networking capabilities. On the realisation of Network Planes and Parallel Internets, appropriate technologies are selected and directives are produced and fed to *NP Provisioning & Maintenance* which undertakes the actual implementation, producing the appropriate concrete network configuration and NIA orders, which will be negotiated and established by *NIA Ordering*.

NP Mapping produces candidate CPA/NIA mappings to Network Planes and Parallel Internets on the basis of compatibility of the CPA/NIA requirements to the capabilities of the Network Planes and Parallel Internets. The produced CPA/NIA mappings are used by *Resource Availability Checking* to deduce the admission or rejection of the CPA/NIA request by comparing the capacity in the engineered Network Planes with the demand of the CPA/NIAs. *NP Provisioning & Maintenance* also uses the CPA/NIA mappings to actually accommodate the CPA/NIA traffic demand.

Data gathered by *NP Monitoring* are used to generate notifications and reports for the *CPA/NIA Order Handling* and *CPA/NIA Assurance* to forward to SPs and upstream INPs, for the online traffic engineering functions in *NP Provisioning & Maintenance*, for *Resource Availability Checking* to derive appropriate multiplexing factors, for the *NP Design & Creation* and *NP Emulation* and *Business-based Network Development* functions to formulate a high level view of the network performance.

3.7.2 Business-based Network Development

The *Business-based Network Development* function takes into account growth trends for established NIA/CPAs and trends for new NIA/CPAs. It evaluates current and opportunities for new interconnections, options for investing in new equipment internally and/or for expanding the geographical scope, analysing associated costs and anticipated profits with the support of the *NP Emulation* function. Based on the above, *Business-based Network Development* decides which Service

Providers and IP Network Providers to target and proceeds to procurement of the networking capabilities required to support potential requests from SPs and INPs (network planning). Such capabilities include capabilities of the network elements (e.g. DiffServ, MTR, to enable QoS differentiation, IGP protocols, memory capabilities, etc.), network management capabilities (e.g. traffic engineering tools, monitoring and reporting tools, etc.), network topology coverage and capacity, interconnection to INPs which can support inter-domain requirements of target SPs and upstream INPs etc. Given the procurement of the required capabilities, *Business-based Network Development* sets the targets for the *NP Engineering* components to dimension the networking capabilities for accommodating CPA/NIA requests and associated traffic.

Specifically, the *Business-based Network Development* function produces the target Network Services (cf. section 3.2), and Engineering Guidelines (cf.. section 3.3) for *NP Engineering* functions (see Figure 16). Target Network Services are the performance guarantees that *Business-based Network Development* desires to offer to SPs and upstream INPs in defined scope, as a result of market trend analysis and based on the procured capabilities. These service targets are also provided as input to *Network Service and Network Capabilities Advertisement* functions. Guidelines for the *NP Engineering* functions are provided in terms of Network Services and CPA/NIA request types.

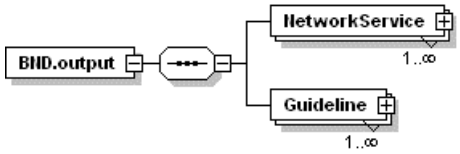

Figure 16: *Business-based Network Development* output.

3.7.3 NP Design & Creation

The *NP Design & Creation* function is responsible for designing the Network Planes (cf. section 3.4) and Parallel Internets (cf. section 3.5), matching the requirements of existing CPA/NIAs and the guidelines of *Business-based Network Development* with the intra-domain network capabilities and the capabilities provided by downstream INPs as discovered by *Network Capabilities Discovery*. *NP Design & Creation* operates with some additional non-functional objectives. The priority of these objectives can be determined by the *Business-based Network Development* or by the *NP Design & Creation* based on best practices. Such objectives include:

- optimisation of resource utilisation for existing CPA/NIAs and *Business-based Network Development* provisional resource requirements,
- minimisation of the cost paid for downstream NIAs,
- minimisation of the complexity of network administration,
- minimisation of network load caused by control messages and heavy processing at the network elements,
- maximisation of stability against modifications of *Business-based Network Development* guidelines, CPA/NIA requirements or traffic volume.

The local capabilities of the INP that are considered by *NP Design & Creation* include:

- *recovery techniques*: backup links, preconfigured backup paths with MPLS Fast Reroute or other techniques, option to allocate capacity on preconfigured backup paths, plain IGP rerouting, type of recoverable failures (single link, single node, etc.), etc.
- *intra-domain routing mechanisms*: IGP protocols and their extensions (e.g., M-ISIS [PRZY05], MT-OSPF [PSEN06]), MPLS explicit paths, overlay routing, IP tunnelling technologies, MRDV, etc.
- *inter-domain routing mechanisms*: the BGP protocol and its QoS extension (qBGP [MESCAL]), Tunnelling Service [QUOI05, QUOI06], Path Computation Elements (PCEs) and associated communication protocols, etc.

- *forwarding and other mechanisms*: priority and CB-WFQ scheduling, RSVP, RED, ECN, DSCP/ToS traffic differentiation, policing algorithms, etc.
- *QoS performance*: minimum edge-to-edge delay and sustainable throughput
- *off-line / on-line traffic engineering mechanisms*: measurement-based admission control, etc.
- *monitoring capabilities*: supported metrics and granularity, sampling frequency, aggregation techniques, reporting and notification capabilities, etc.

NP Design & Creation is also aware of the inter-domain capabilities of adjacent and remote INPs, discovered by the *Network Capabilities Discovery*. Inter-domain capabilities provide information on:

- QoS and availability guarantees within a certain inter-domain scope,
- indication of supported throughput,
- supported inter-connection mechanisms (e.g. Tunnelling Service),
- interconnection activation requirements (e.g. DSCP marking or tunnelling protocol for differentiation at the border routers, invocation protocols, etc.),
- cost formula, etc.

The problem that *NP Design & Creation* function needs to solve greatly varies depending on the above limitations. It is very unlikely that any INP will ever face a situation where *NP Design & Creation* has all options open. Hence, we consider that devising an algorithm for a universal solution is overkill. However, some principles apply to any *NP Design & Creation* problem setup, such as:

- the more gross differentiation in traffic treatment resulting in less NPs the better, merging should be pursued whenever possible to minimise complexity of network administration and to allow for higher aggregation leading in turn to improved resource utilisation,
- the more lightweight the partitioning of the physical network topology to Network Planes the better, choose NPs with the less impact on control traffic and processing at the network elements.

The output of the *NP Design & Creation* (see Figure 17) includes the definition of Network Planes in terms of their targeted QoS and availability capabilities, realisation guidelines specific to the employed technologies, and the definition of the Parallel Internets.

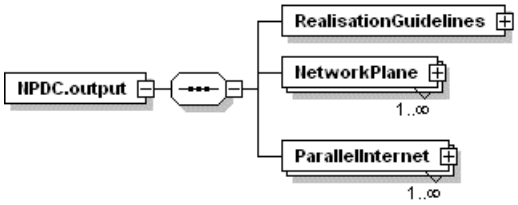

Figure 17: *NP Design & Creation* output.

NP Design & Creation defines Network Planes and Parallel Internets for accommodating existing CPA/NIAs and for allocating resources provisionally. As long as the requirements of future CPA/NIA requests fall into the capabilities and the restrictions of the existing Network Planes and Parallel Internets, *NP Provisioning & Maintenance* undertakes the re-shuffling of resources to the existing Network Planes and Parallel Internets without *NP Design & Creation* intervention. The point up to which *NP Provisioning & Maintenance* can expand Network Planes and Parallel Internets in terms of scope and capacity is determined by the corresponding *NP Engineering* guidelines set by the *Business-based Network Development* and the Network Plane resource allocation restrictions set by the *NP Design & Creation*.

3.7.4 NP Design and Creation Feedback – NP/PI Reporting

In order to evaluate the performance of the particular Network Planes and Parallel Internets design enforced in the network, *NP Design & Creation* receives feedback from the underlying *NP*

Engineering components, concerning the utilisation of the network, the satisfaction of the traffic guarantees, the operational overhead and the load of the physical infrastructure (see Figure 18).

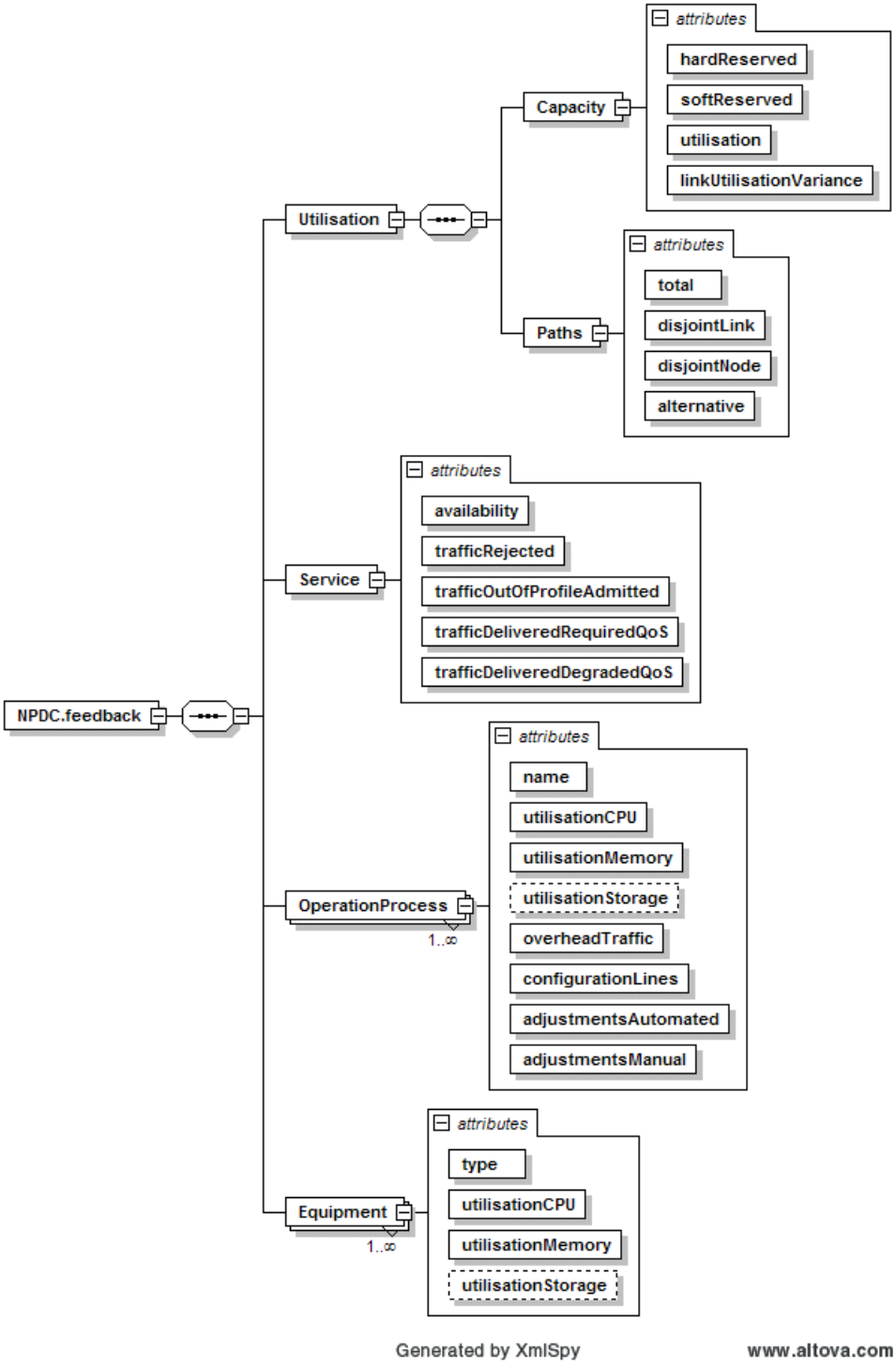

Figure 18: *NP Design & Creation* feedback

The scope of the received feedback can be the entire network, or a particular portion of the network and the traffic, over a particular time period. The scope can be specified in any combination of the following:

- Network Planes, Parallel Internets, Parallel Internets realisation over particular Network Planes,
- CPAs and upstream NIAs,
- downstream NIAs,

- topological scope over particular ingress and egress nodes and/or remote destination networks,
- macro-flows.

The reserved capacity and its utilisation are reported for the traffic that falls within the specified scope. Some scheduling algorithms (e.g. fair queuing) and traffic engineering techniques allow for a distinction between capacity that is wasted if not used for the specified traffic (hard-reserved), and the capacity that is made available to other traffic when not used by the specified traffic. An indication of the distribution of the traffic over the network resources is the variance of the utilisation across the links within the specified scope. The number of paths may be of interest for evaluating the granularity of the routing and forwarding configuration and the resilience of the dimensioned network.

The service experienced is reported in terms of service availability, of traffic that has been rejected at the ingress network node, of the excess traffic that, although out of the agreed traffic profiles, has been admitted into the network. Out of the admitted traffic, it is reported the proportion of the traffic that has experienced the agreed QoS across the INP's network, and its counterpart that has experienced degraded QoS.

Each network operation process utilises bandwidth, CPU, memory and possibly storage resources of the infrastructure. Network operation processes are considered the routing, monitoring, network management, security, etc. Moreover, increased network operation complexity requires acquiring expensive software and highly qualified personnel for the troubleshooting and performance analysis of the network. Therefore, it is important for the *NP Design & Creation* to evaluate the complexity of the *NP Provisioning* processes. Candidate indicators are the number of configuration commands, the frequency of automated and manual configuration adjustments per process.

Finally, a summary of the load of the equipment itself is reported, in terms of CPU utilisation, memory and possibly storage if so is relevant, providing a view of the stress imposed to the infrastructure by any particular Network Planes and Parallel Internets design. Different equipment of interest might be routers, AAA servers, firewalls, etc.

3.7.5 NP Provisioning & Maintenance

NP Provisioning & Maintenance undertakes the actual realisation of Network Planes and Parallel Internets and produces the appropriate network configuration, based on the technologies selected and the realisation guidelines provided by the *NP Design & Creation* (see [D3.2] for details on *NP Provisioning & Maintenance* techniques).

The *NP Provisioning & Maintenance* function triggers and coordinates the appropriate offline and online mechanisms which undertake the actual network configuration. Additionally, *NP Provisioning & Maintenance* produces concrete NIA orders subsequently executed by *NIA Ordering*, taking as input the PI definition produced by *NP Design & Creation* and the NIA capabilities discovered by *Network Capabilities Discovery*.

NP Provisioning & Maintenance is invoked to re-allocate resources by *Resource Availability Checking* when there are no available resources for accommodating a new CPA/NIA request, or by *NP Monitoring* when the utilisation targets per Network Plane are not met. The NP utilisation targets may be set by the *NP Design & Creation* or the *NP Provisioning & Maintenance* itself. *NP Provisioning & Maintenance* is triggered also as a result of the *Business-based Network Development* function changing the resource allocation guidelines.

Resource re-allocation is restricted by the Network Plane maximum resources set by *NP Design & Creation* and by the resource allocation guidelines set by *Business-based Network Development* at the granularity of network services and CPA/NIA types, mapped to the Network Planes by *NP Mapping*.

3.7.6 CPA/NIA Order Handling

The *CPA/NIA Order Handling* function analyses the received CPA/NIA request, performs administrative admission control based on the *Business-based Network Development* CPA/NIA type-

based guidelines on request admission and feeds the other *NP Engineering* functions with the required CPA/NIA information.

3.7.7 NP Mapping

NP Mapping matches the CPA/NIA request requirements with the capabilities and the restrictions of existing Network Planes and Parallel Internets, following the Engineering Guidelines set by the *Business-based Network Development* function.

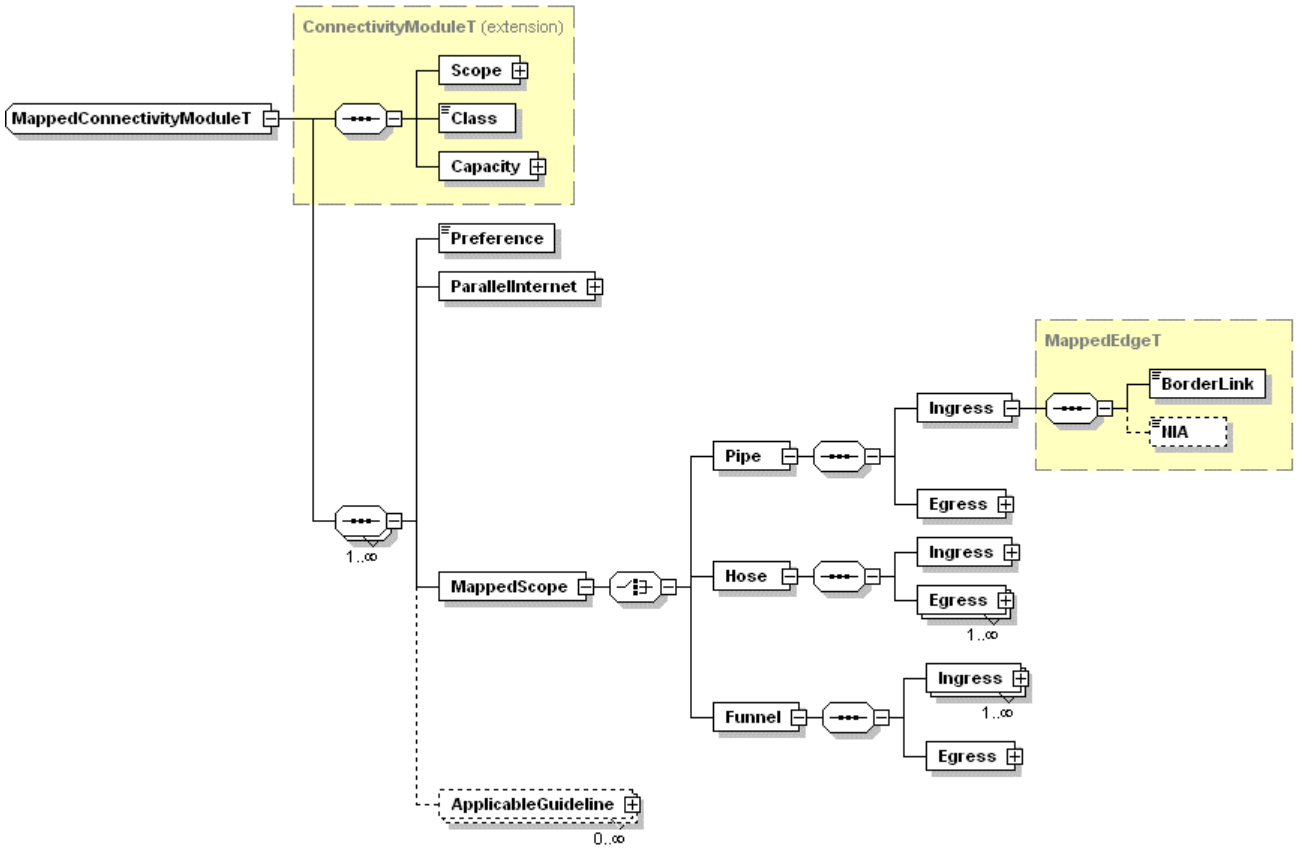

Figure 19: Mapped connectivity module.

NP Mapping maps the CPA/NIA connectivity modules (see Figure 19) to candidate Parallel Internets and downstream NIAs as follows:

- Matching the connectivity class to the guarantees of the Parallel Internet, this procedure potentially results in multiple candidate Parallel Internets.
- Restricting the candidate Parallel Internets on the basis of their capabilities for supporting the CPA/NIA requirements on provisioning rules, feedback, activation and assurance. *NP Mapping* finds and registers all the *NP Engineering* guidelines matching each CPA/NIA connectivity module. From the CPA/NIA request and the matching guidelines it deducts a set of combined requirements, such as cost restriction, route restrictions and preferences per connectivity module. *NP Mapping* then matches the connectivity module scope and guarantees, and associated set of requirements with the Parallel Internet and Network Plane restrictions. Note that the requirement to fulfil a CPA routing rule can be satisfied either when the same rule applies to the Parallel Internet, or when routing differentiation is supported inside the Parallel Internet.
- Further restricting the candidate Parallel Internets to those that reach the connectivity module edges. This procedure requires scope translation, i.e. mapping the scope edges to specific border links and, in case of remote edges, downstream NIAs. This step has been thoroughly studied in [MESCAL].

As a result, *NP Mapping* produces candidate mappings, for the potentially many Parallel Internets and Network Planes found compatible to accommodate the requirements of the CPA/NIA connectivity modules.

3.7.8 Resource Availability Checking

The *Resource Availability Checking* function uses the CPA/NIA mappings to deduce the admission or rejection of the CPA/NIA request by comparing the capacity in the engineered Network Planes and Parallel Internets with the demand of the CPA/NIAs.

Resource Availability Checking calculates the demand of a mapped CPA/NIA connectivity module provided by the *NP Mapping* function, applying a multiplexing factor to calculate the effective bandwidth. *Resource Availability Checking* attempts to fit the overall demand, anticipated by all CPA/NIAs, to the total capacity of the engineered Network Planes, downstream NIAs and Parallel Internets. The actual accommodation of the CPA/NIA request is undertaken by *NP Provisioning & Maintenance*, which may use any of the valid candidate mappings based on the actual network conditions.

The capacity is calculated by the *NP Provisioning & Maintenance* function and is specified in defined scope, in the form of a Resource Matrix, defining capacity per Network Plane between local ingress and egress border routers, and the capacity of the downstream NIAs (see [MESCAL]).

The objective of this function is to find at least one combination of CPA/NIA mappings such that:

- all connectivity modules are mapped to a Parallel Internet and Network Plane,
- the total derived demand from all connectivity modules mapped to a Parallel Internet and Network Plane does not exceed the corresponding capacity,
- the NP Engineering Guidelines for freely allocating capacity to new requests are not violated.

In case such combination is not possible, the new CPA/NIA request is forwarded to the *NP Provisioning & Maintenance* component, to potentially trigger a reshuffling of the network resources or the establishment of new downstream NIAs.

3.7.9 CPA Assurance

CPA Assurance is responsible for ensuring that the clauses of an offered (to a Service Provider) CPA have been met by the INP. This function interfaces with *CPA Verification* function implemented by a given Service Provider. The *CPA Verification* function is responsible for checking if the INP has honoured the clauses of a CPA. To do so, the CPA includes information about the verification methodology to be used and which parameters, also denoted as *Key Performance Indicators* (KPIs) will be reported by the INP. Note that an INP can adopt a formalism of two KPI views: the ones provided to its customers and ones used for internal purposes. The internal view is more fine-grained than the external one, unless the SP includes specific parameters in the agreed CPA. KPIs provide an exact view of the level of the provided IP service. Thresholds may be associated to each KPI so as to notify a service deterioration (i.e. the level of the experienced service is worse than expected). Furthermore, a CPA may also enclose the type, format and frequency of the reporting data.

CPA Assurance function interacts also with *NP Monitoring* in order to retrieve monitoring data related to a given CPA. Appropriate filters may be defined so as to select specific monitoring data or to configure and activate appropriate monitoring jobs. Note that beyond explicitly initiating monitoring jobs, *CPA Assurance* may need to ensure that the Service Provider has the agreed access to the monitoring facilities of the domain, e.g. by activating access for the ‘ping’ requests generated by the SP or access to probing facilities, etc.

An INP may also implement an interface between the *CPA Assurance* and *NP Provisioning & Maintenance* to issue notifications upon CPA violation (service deterioration). Notification actions may be configured and notification thresholds assigned. *NP Provisioning & Maintenance* may then undertake appropriate actions, for instance prioritising maintenance operations, to align the current NP

service level with the targeted one. Alarms may be correlated with billing tickets so as to take into account degradation of the delivered service. From the Service Provider perspective, service deterioration is assessed as a cumulative value, a CPA is considered to be honoured until the cumulative value of the service deterioration duration violates the service availability agreed in the CPA.

From the operational standpoint, an identifier should be assigned to each subscribed CPA so as to be unambiguously identified during NP-related operations.

Two complementary modes of CPA Assurance (implicitly of CPA Verification) are possible:

- *Active mode/on demand:* In this mode, the SP can issue real-time requests. These requests are handled by the INP which undertakes their execution like computing the round trip delay to reach a given network/service node, etc.
- *Passive mode/pe*riodic: The INP delivers non real-time reports to the SP so evaluate the level of the experienced service. For instance, a weekly report about the variation of the delay may be considered.

4 NP-EMULATION PLATFORM (NPEP)

4.1 Rationale

NPEP is a 'snapshot' of the INP functionality, embodying the essential aspects of the project work, definition and realisation of Network Planes (NPs) and Parallel Internets (PIs) according to business policies. NPEP also provides means for generating traffic and measuring the performance of the network in accommodating the generated traffic flows. The platform currently emulates IP networks with/MPLS. However, its design is modular and alternative IP network technologies/capabilities can be incorporated.

The platform is built with the purpose of validating the concepts and notions developed by the project, for exhibiting the business-driven engineering of INP domains and for running 'what-if' scenarios and comparison tests to assist decision-making in business policies on service provisioning, network upgrades and technology choices, including traffic engineering.

As it became apparent from the analysis of the NP/PI problem (cf. section 3.6), there needs to exist means for evaluating alternative set of NPs to the end of determining which set to realise for instantiating a required set of PIs. NPEP can provide such means.

4.2 Functional Overview

Figure 20 presents an overall view of the NP Emulation Platform. As it can be seen, it consists of (a) components pertinent to project work –interfaces for CPAs, NP engineering guidelines, NPs, PIs, NP provisioning algorithms and (b) generic components of an emulation system –traffic generation, emulation engine, reporting facilities.

Furthermore, it includes traffic engineering (TE) components, which based on the defined NPs/PIs produce the required network configuration for the emulation engine to execute; conversely, they mediate the emulation results to the NP/PI nomenclature. This part of NPEP can be replaced with alternative TE components as long as they adhere to the emulation system interface and to the schema representing the AGAVE entities, CPAs/NIAS, NPs/PIs. This way, different TE schemes can be incorporated in NPEP, providing also an idea of how AGAVE can be introduced in INP domains.

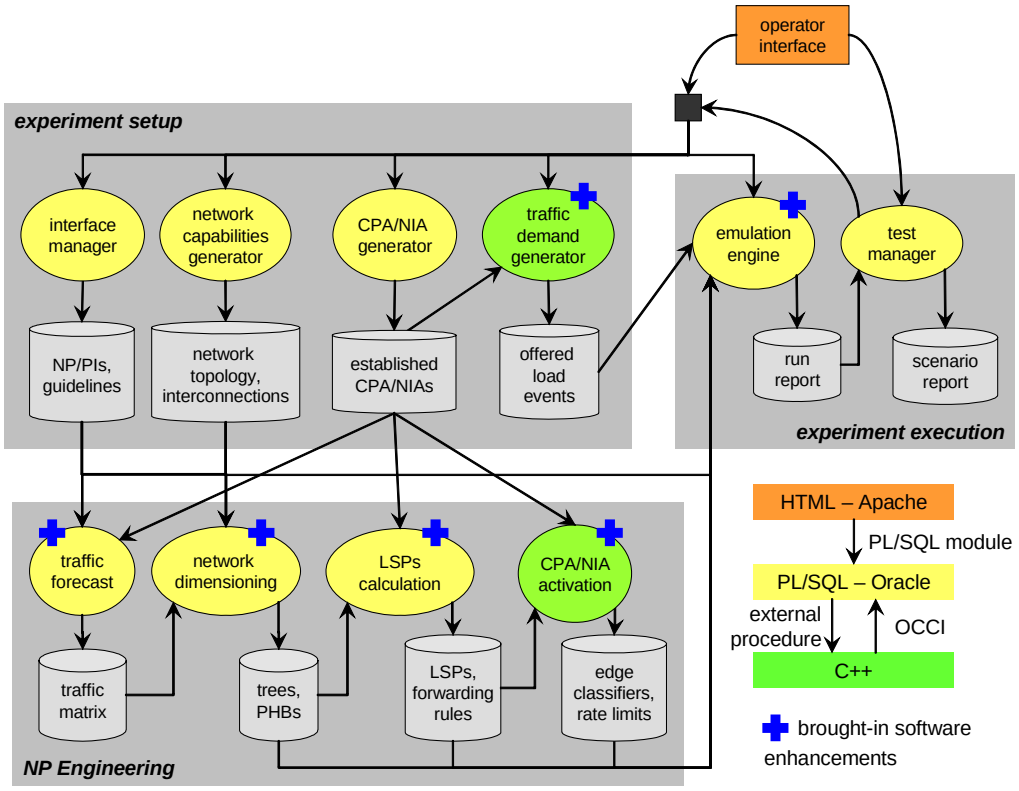

Figure 20: Overview of NP Emulation Platform.

The *operator interface* implements a Web interface provided to the operator of NPEP, allowing for setting up and running specific experiments. As part of setting up the experiment, the operator defines the business guidelines and the NP/PIs (cf. section 4.3). The *interface manager* component undertakes consistency checks and populates the corresponding repository. The operator also defines in high-level, the dimensions of the network topology and interconnections (how many and what capabilities). The *network capabilities generator* component generates the network nodes, core and inter-domain links, the NIAs in place etc. following the high-level settings of the operator. The *CPA/NIA generator* component generates established CPAs and downstream NIAs, based on which traffic will be generated for the experiment. The operator provides guidelines on the CPA/NIA types to generate, their number, contractual demand, etc. Finally, the *traffic demand generator* component (cf. section 4.4) generates the offered load events for the experiment, based on the established CPA/NIAs and operator settings with respect to the traffic profile and the demand to generate comparing to the contractual (distributions and parameter values for session inter-arrival and holding time, demand per session, etc.).

NP Engineering functions are invoked to configure the network. The *traffic forecast* component produces the anticipated traffic matrix over the local and remote edges for the defined NP/PIs, based on the established CPA/NIAs and the business guidelines. Note that in this prototype, a DiffServ QoS-class (similar to DiffServ PDB) is defined for each Network Plane. The *network dimensioning* component calculates the forwarding path trees and the PHB weights per link for accommodating the traffic matrix. The *LSPs calculation* component derives the LSPs that correspond to the calculated forwarding path trees and assigns the established CPA/NIAs flows to alternative LSPs. The *CPA/NIA activation* component establishes the traffic classification and policing rules at the ingress nodes for admitting and traffic conditioning the eligible traffic from the established CPA/NIAs.

The network configuration in terms of PHB weights, LSPs, ingress forwarding, classification and policing rules is fed to the *emulation engine* component (cf. section 4.5) which runs the experiment, taking also as input the offered load events generated during the experiment setup. The behaviour of the network is assessed during the experiment and at the end a number of suitable reports are

generated, both at the NP/PI abstraction layer and at the layer of the particular NP/PI realisation technologies.

The operator may choose to interrupt an experiment, modify the experiment setup and then resume the experiment execution at any time. The *test manager* component is responsible for automatically setting up and regulating the execution of one or multiple experiments, predefined for the project testing needs, or in the context of what-if scenarios. Depending on the logic of the particular scenario, a suite of tests may be executed, the results of which are correlated by the *test manager* to produce the scenario reports. Individual experiment and scenario reports are made available to the operator through the operator Web interface.

The following sections provide more information on the key NPEP components.

4.3 Operator Interface – NP/PI Definition Operations

This section is concerned with the part of the operator interface dealing with the definition of NPs and PIs in the context of an experiment.

First, it should be noted that NPEP views the information models for NIAs, NSs, NPs and NIs (cf. sections 3.2, 3.4, 3.5) as meta-models. As such, it allows for the operator to define the exact model for these entities. This capability is not offered as an operation of the Web interface. Rather, it is provided as a service. We consider that this approach enhances the applicability of NPEP and its porting to different INP domains. Note that the attributes -as meaning, syntax and default/possible values- representing the offered services and the dimensions along which services can be provisioned may change from INP to INP. A universal model, able to capture all possible cases, is hard to specify.

As it became apparent in section 3.6, the NP/PI definition problem requires human intervention. Inspired by the analysis therein, a number of appropriate operations have been designed with the purpose to facilitate the task of the operator. The following operations are provided in the context of an experiment:

- *NIA management*: Operations for entering, viewing, updating and deleting NIAs. A single set of NIAs can be defined per experiment.
- *NS management*: Operations for entering, viewing, updating and deleting NSs. A single set of NSs can be defined per experiment.
- *Required PIs management*: Operations for entering, viewing, updating and deleting the required PIs. A single set of required PIs can be defined per experiment.
- *Calculate required PIs:* This operation allows for the automated calculation of the required PIs based on the defined NSs, should such an automated procedure is available.
- *NP management*: Operations for entering, viewing, updating and deleting NPs. Multiple sets of NPs can be defined per experiment.
- *Calculate NPs:* This operation allows for the automated calculation of NPs based on the defined required PIs and NIAs, should such an automated procedure is available. With reference to the terminology of section 3.6.4.3, the maximal and minimal sets of PIs can be constructed.
- *NP-NIA association management:* Operations for defining, viewing and deleting associations between a selected set of NPs and the set of the defined NIAs.
- *PI-NIA association management:* Operations for defining, viewing and deleting associations between the set of the defined PIs and the set of the defined NIAs.
- *NP-NIA for PI associations management:* Operations for defining, viewing and deleting associations between a selected set of NPs and the set of the defined NIAs for instantiating a particular PI from the set of the required PIs.

Clearly, with the above operations, the operator can define multiple sets of NPs and respective associations with NIAs for instantiating the set of the required PIs. Then, NPEP can ‘run’ the corresponding network configurations for their evaluation.

4.4 Traffic Demand Generator Tool

The NP Emulation Platform Traffic Demand Generator Tool[2] (NPEP-TDG) (see Figure 21), integral part of the NP Emulation Platform, generates traffic load events based on a population of established CPA/NIAs, specified source profiles and parameters regarding the desired level of load to be injected in the network during an NPEP experiment execution.

The traffic load events are arranged into chronological order and present an aggregate of the traffic - over active traffic sources - to be injected in the network on behalf of a CPA/NIA from a particular access point.

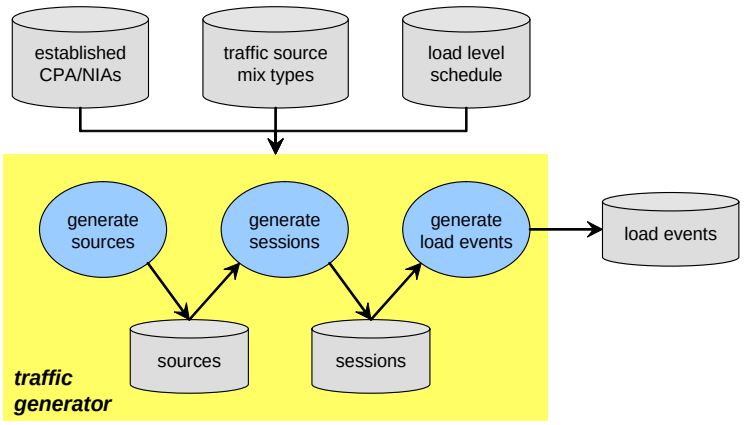

Figure 21: NP Emulation Platform Traffic Demand Generator Tool.

Specifically, NPEP-TDG takes as input:

- *Established CPA/NIAs*: the CPA/NIAs that have access to the network resources; each CPA/NIA comes with the access points, the connectivity requirements and the pools of source/destination IP addresses defined; established CPA/NIAs may be generated by the CPA/NIA generator tool or defined by the NPEP operator or a combination of the two;
- *Traffic source mix types*: specification of the types of traffic source mixes and association of each CPA/NIA with a type; a traffic source mix is composed by a population of sources with common traffic profiles, where a traffic profile (see Table 3) is defined in terms of session inter-arrival and holding times (distribution can be uniform, Pareto, exponential), bandwidth demand and source active/idle distribution;
- *Load level schedule parameters*: parameters that determine the load level evolution throughout an experiment. These parameters are set overall or per (type of) CPA/NIA. Load level is defined as the desired ratio of the total traffic to be generated over specific target values e.g. capacity or availability of network resources. Load levels may change in time. Based on these parameters and depending on the population of the established CPAs/NIAs, the number of traffic sources per CPA/NIA and their characteristics are determined; the more CPA/NIAs, the fewer the traffic sources to achieve a given load level.

The established CPAs/NIAs, in terms of their number and characteristics, and the load level schedule parameters are set manually or derived from a set of high-level parameters characterizing the particular traffic generation scenarios pertinent to a specific experiment.

[2] NPEP-TDG is based on the traffic generator function of the Traffic and Network Emulation Tool built in IST-TEQUILA project (see section 8.4.3 in TEQUILA deliverable http://www.ist-tequila.org/deliverables/D3-4a.pdf).

Session	Distribution
	Inter-arrival mean time
	Holding mean time
	Variance
Load	Distribution
	Peak rate
	On mean time
	Off mean time
	Variance

Table 3: NPEP-TDG source traffic profile parameters.

The generation of load events involves the following steps:

1. *Source generation*: For each CPA/NIA a number of traffic sources (customer applications) are generated following the load level schedule. The source type is determined following the settings of the associated traffic source mix, and the number of the sources is calculated so that the desired load level can be achieved. Each source is associated to an IP address from the source IP address pool associated to the CPA/NIA. The traffic source mix may change over time following the load level schedule parameters.
2. *Session generation*: For each existing source, sessions are generated following the source traffic profile (inter-arrival and holding times). Each session corresponds to an IP packet flow with source IP address the IP address assigned to its source and a destination IP address, randomly selected from the pool of destination IP address of the CPA/NIA.
3. *Load event generation*: For each generated session, load events (ON/OFF source model) are generated following the source traffic profile. Load events are aggregated over all active sources in a CPA/NIA to determine the total traffic load to be injected in the network from the particular edges that the CPA/NIA has been defined.

4.5 Emulation Engine

The *emulation engine* component[3] is further decomposed into the *workflow coordinator*, *traffic emulation* and *network performance evaluation* blocks (see Figure 22).

[3] The emulation engine component is based on the Emulation Engine, Injected Network Traffic Calculation and Network Performance Evaluation functions of the Traffic and Network Emulation Tool built in IST-TEQUILA project (see section 8.4.3 in TEQUILA deliverable http://www.ist-tequila.org/deliverables/D3-4a.pdf).

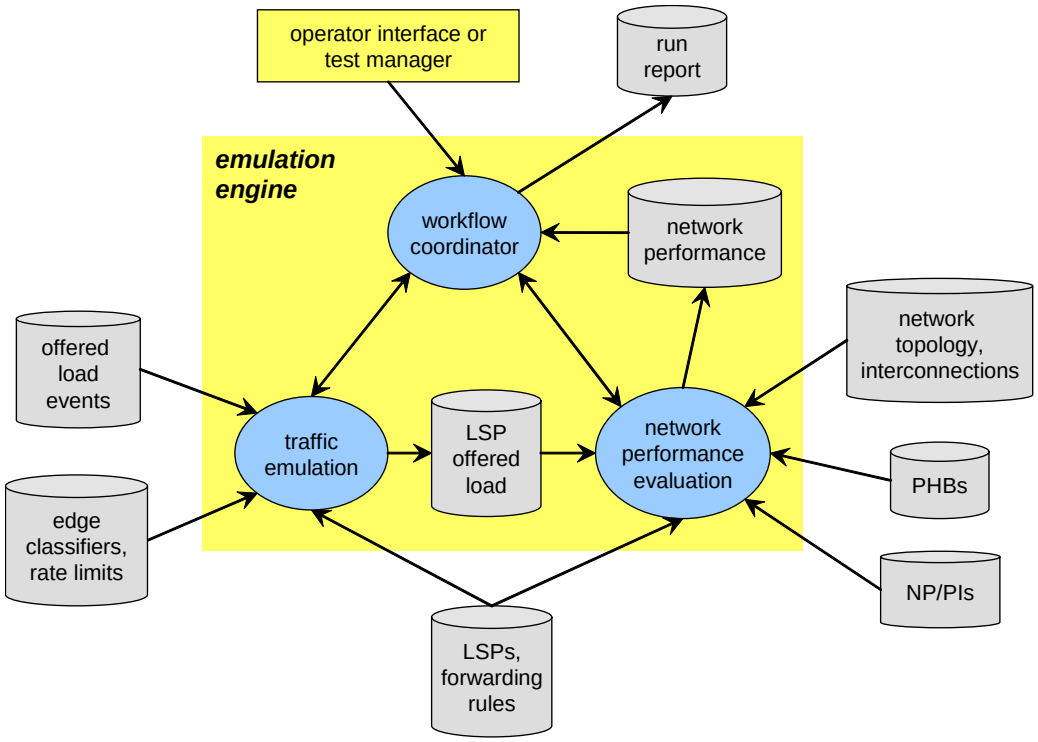

Figure 22: Emulation engine design.

4.5.1 Workflow coordinator

Workflow coordinator implements the following methods:

- *start_experiment (time units)*, called by the operator interface or the test manager, it initiates an experiment execution, running for the specified time units, generating a run report at the end;
- *pause_experiment*, called by the operator interface or the test manager, it pauses the ongoing experiment execution, allowing for experiment setup modification;
- *resume_experiment*, called by the operator interface or the test manager, it resumes a previously paused experiment;
- *abort_experiment*, called by the operator interface or the test manager, it aborts the ongoing experiment execution, generating a run report;
- *tick*, private method, called by start_experiment and resume_experiment, it calls traffic emulation and network performance evaluation to emulate the network behaviour for a single simulation time tick.

4.5.2 Traffic emulation

Traffic emulation implements the following methods:

- *tick*, called by the workflow coordinator, it calculates the offered load per established intra-domain LSP at the current simulation time tick;
- *calculate_offered_load*, private method, it calculates the traffic arriving at the ingress routers by aggregating the offered load events produced by the *traffic demand generator* component;
- *calculate_admitted_load*, private method, it calculates the load admitted to the network by applying the classification and policing rules produced by *CPA/NIA activation* component to the overall injected load calculated by calculate_injected_load;

- *calculate_offered_load_LSP*, private method, it calculates the load injected to each established LSP by applying the forwarding rules calculated by *LSPs calculation* component to the admitted traffic calculated by calculate_admitted_load.

4.5.3 Network performance evaluation

Network performance evaluation implements the following methods:

- *tick*, called by the workflow coordinator, it evaluates the network performance at the current simulation time tick;
- *calculate_link_load*, private method, it calculates the load down each PHB, given the PHB weights calculated by *network dimensioning*, the established LSPs calculated by *LSPs calculation* component and the LSP offered load calculated by traffic emulation;
- *calculate_delivered_traffic*, private method, it determines the delivered traffic per LSP, based on the link load calculated by calculate_link_load;
- *evaluate_network_performance*, private method, it evaluates the performance of the network by comparing the delivered traffic calculated by calculate_delivered_traffic against the NP/PI performance targets, set by the operator through the *interface manager* component.

5 DEPLOYING AGAVE

To the end of the problem of service provisioning and delivery across the Internet, AGAVE has specified a framework approach and a set of enabling algorithms and mechanisms. *The proposed solutions have been designed to be deployable in the existing best-effort Internet in an incremental fashion*, as explained in the following paragraphs.

Framework level

First of all AGAVE does not advocate a clean-slate approach; instead it relies and builds upon the existing IP-based Internet architecture and related protocols.

AGAVE clearly separates network and service concerns, by distinguishing IP Network Provider (INP) and Service Provider (SP) roles. The principle of network providers acting in a distinct role and offering *service-ready* connectivity to upstream SPs is by no means radical or new. This model is in-line with on-going initiatives and discussions regarding 'functional separation' and 'network neutrality'. However, the issue of whether INP and SP roles will be instantiated as distinct business organisational roles in their own right depends on market and business developments and also on appropriate regulatory frameworks. However, should the separation of roles be formalised, the AGAVE Connectivity Provisioning Agreements (CPA) specifications could be used as a basis of the interactions required between organisations acting in INP and SP roles. AGAVE's work has shown that CPAs can capture realistic service requirements, which, subsequently, can be used by INPs to appropriately engineer their network.

The following key aspects underlying the proposed framework for end-to-end service delivery contribute to the incrementally deployable nature of the AGAVE solutions.

- Universal participation of INPs is not a pre-requisite.
- It is not required that all INPs have a common deployment or configuration of a particular set of traffic engineering (TE) techniques, mechanisms and protocols.

The AGAVE inter-domain architecture is centred around the novel concept of Network Planes (NPs), which allows INPs to build and provide Parallel Internets (PIs) tailored to end-to-end service requirements by interconnecting NPs. NPs/PIs represent the service provisioning and delivery capabilities of INPs in an abstract technology-agnostic manner and can be incorporated as an additional functional layer above the TE / network management layer. PIs are not intended to be globally agreed or standardised; simply, they represent the view of the Internet from the perspectives of an individual INP. This approach leaves individual INP domains to decide which PIs and NPs to provide, how to engineer the network accordingly and with which technologies and mechanisms, and how to interconnect with NPs/PIs of other INPs.

Another aspect that positively contributes to the incremental deployment of the AGAVE framework is the process-oriented nature of the proposed INP functional architecture. The functionality is decomposed into layers which directly map into existing operational roles, such as: business development, network planning and design, and network operations. This clearly facilitates the adoption of the AGAVE approach as the functional separation mirrors the organisational decomposition of today's providers. Furthermore the AGAVE solutions provide a concrete set of models to facilitate communications between operational entities in an appropriate common language.

Individual mechanism level

AGAVE has specified and tested through implementation, a set of algorithms and mechanisms for realising NPs and PIs. In particular:

- Intra-domain mechanisms for enabling edge-to-edge service differentiation: Multi-Topology Routing (MTR), INP-layer overlay routing and Multi-path Routing with Dynamic Variance (MRDV).

- Schemes for achieving high service assurance as well traffic optimisation purposes in case of network failures: ASBR fast rerouting with RSVP-TE extensions, BGP Planned Maintenance and BGP/IGP based traffic engineering with resilience awareness.
- Inter-domain routing schemes: IP Tunnelling and QoS-enhanced BGP (q-BGP).

The above solutions are incrementally deployable, as they either:

- Operate above the IP layer.
  - This is the case for the specified overlay-routing scheme and the BGP/IGP based traffic engineering with resilience awareness algorithm.
- Co-exist with currently deployed protocols.
  - This is the case for the new proposed routing scheme MRDV, which has been designed to co-exist with widely deployed IGP schemes.
- Rely on 'smart' configuration of existing protocols or those currently under discussion in the IETF:
  - This is the case for the specified MTR-based intra-domain routing scheme and the IP Tunnelling mechanism for inter-domain routing which, is based on the emerging LISP (location identifier separation protocol) specifications by IETF.
- Require minor extensions to existing protocols, which are undertaken following the standard approach/mechanism prescribed by the protocol itself.
  - This is the case for the specified q-BGP protocol, the BGP Planned Maintenance procedure and the ASBR fast rerouting scheme with RSVP-TE extensions.

In conclusion, AGAVE solutions do not depend upon a clean-slate, wholesale redeployment of today's Internet. They are incrementally deployable from several perspectives: they build on current and emerging business models; they rely on existing protocols; mechanisms can be deployed using existing protocols in a smart way or with minor extensions; the solutions do not have to be universally deployed. It should be noted that a full techno-economic analysis of migration scenarios are out of scope of the project.

6 CONCLUSIONS

The essence of the AGAVE approach to the problem of service provisioning and delivery in the Internet can be summarised by the following statement '*Clear separation between service and network concerns*'. As such, AGAVE advocates:

- A 'clear-cut' interface between Service Providers (SPs) and IP Network Provider (INPs), which is based on the notion of Connectivity Provisioning Agreement (CPA).
- The concept of Network Planes (NPs) and Parallel Internets (PIs), allowing INPs to build and provide multiple Internet connectivity levels, as required by the end-to-end requirements of the services they offer.

This Deliverable presented the final specifications of the entities pertinent to the AGAVE Internet architecture, CPAs, NPs and PIs, and the required INP functionality. By elaborating on the NP/PI problem, a greedy solution approach for defining NPs and PIs was described. The design of the Network Plane Emulation Platform (NPEP) was presented, which 'puts in motion' the specified concepts for the sake of proving their validity; such a tool, is also required for evaluating alternative NP/PI configurations. Finally, the deployability of the AGAVE solutions was assessed. The following points are worth making:

- Service connectivity provisioning requirements lie in a number of dimensions such as QoS, availability, resilience, shaping, flow forwarding and routing.
- The proposed technology-agnostic NP/PI layer in INP domains bridges business and network levels, contributing to smooth and efficient network development and operations.
- The concepts of NPs and PIs are powerful enough to express a number of INP-related concerns; classical TE problems can be stated in terms of these concepts.
- The problem of defining NP/PIs based on a set of target services and related guidelines generalises the classical TE problem.
- The procedure for determining which NPs and PIs should be provided requires human intervention and requires a function for evaluating alternative network configurations.
- Robust NP realisation is essential and should constitute a criterion to determining suitable technology-specific mechanisms for realising NPs.
- The AGAVE approach and the proposed solutions can be deployed in today's best effort Internet in an incremental fashion.

7 REFERENCES

[D1.1] Boucadair M. et al., *Parallel Internets Framework*, AGAVE Deliverable D1.1, September 2006.

[D2.1] Mykoniati E. et al., *Initial Specification of the Connectivity Service Provisioning Interface Components*, AGAVE Deliverable D2.1, December 2006.

[D3.1] Wang N. et al., *Initial Specification of Mechanisms, Algorithms and Protocols for Engineering the Parallel Internets and Implementation Plan*, AGAVE Deliverable D3.1, December 2006.

[D3.2] Wang N. et al., *Specification of Mechanisms, Algorithms and Protocols for Engineering the Parallel Internets*, AGAVE Deliverable D3.2, April 2008..

[FEAM06] Feamster N., Gao L., Rexford J., *How to lease the Internet in your spare time*, Georgia Tech Technical Report GT-CSS-06-10, http://www-static.cc.gatech.edu/~feamster/papers/cabo-tr.pdf, August 2006.

[I4.1] Quoitin B. et al., *Initial Specification of Experimentation Platform and Tests*, AGAVE Internal Report I4.1, December 2006.

[MESCAL] Wang N. et al., *Final Specification of Protocols and Algorithms for Inter-domain SLS Management and Traffic Engineering for QoS-based IP Service Delivery*, http://www.ist-mescal.org/deliverables/d1.3_finalv2.pdf, MESCAL Deliverable D1.3, June 2005.

[PRZY05] Przygienda T. et al., *M-ISIS: Multi Topology (MT) Routing in IS-IS*, IETF Internet Draft, draft-ietf-isis-wg-multi-topology-11.txt, October 2005.

[PSEN06] Psenak P. et al, *Multi-Topology (MT) Routing in OSPF*, IETF Internet Draft, draft-ietf-ospf-mt-06.txt, February 2006.

[QUOI05] Quoitin B., Bonaventure O., *A Cooperative Approach to Interdomain Traffic Engineering*, in proceedings of EuroNGI, April 2005.

[QUOI06] Quoitin B., *BGP-based Interdomain Traffic Engineering*, PhD Thesis, August 2006.

[RFC2679] Almes G. et al., *A One-way Delay Metric for IPPM*, IETF RFC 2679, September 1999.

[RFC2680] Almes G. et al., *A One-way Packet Loss Metric for IPPM*, IETF RFC 2680, September 1999.

[RFC2698] Heinanen J., Guerin R., *A Two Rate Three Color Marker*, IETF RFC 2698, September 1999.

[TEQUILA] Damilatis T. et al., *Final Architecture, Protocol and Algorithm Specification*, http://www.ist-tequila.org/deliverables/D3-4b.pdf, TEQUILA Deliverable D3.4, part B, October 2002.

8 APPENDIX A: DEPLOYMENT OF IMS-BASED SERVICES UPON AGAVE-ENABLED IP ARCHITECTURES

8.1 Introduction

The emergence of new services such as video streaming and IP telephony requires IP networks to provide stringent guarantees not only in terms of traditional Quality of Service (QoS) metrics, but also in terms of availability (e.g. “five nines” for telephony) and robustness during emergency situations. Ever since the early stage of IP networking, proposals have aimed to capture and support the requirements of various services, especially in the realm of forwarding and routing. In 1994, the Nimrod initiative [1] was launched within the IETF with the ambition of providing service-specific routing in the presence of multiple constraints imposed by Operators and end users. RFC1992, one of the key documents produced by the Nimrod initiative, states that inter-network connectivity and services should be represented by maps at multiple levels of abstraction. Nevertheless, this recommendation has never been implemented. Additionally, QoS forwarding mechanisms like IntServ [2] and DiffServ [3] have been proposed, but have not been widely introduced into operational networks due to their complexity, the lack of clear views on the manageability of such mechanisms and the fact that Operators are not ready to get rid of their practices related to over-provisioning in favour of sophisticated Traffic Engineering techniques.

In order to handle the aforementioned challenges, mainly QoS and robustness ones, the AGAVE (*A liGhtweight Approach for Viable End-to-end IP-based QoS Services*) project introduces the concepts of *Network Plane (*NP*)* and *Parallel Internet* (PI) [7][8], a novel transport platform that offers end-to-end service differentiation across Internet. The proposed approach does neither require a single Internet-wide architecture nor even require universal deployment. This paper presents these concepts and associated functions. It also describes how IMS can make use of the AGAVE platform to offer QoS-enabled multimedia services.

8.2 Overview of AGAVE Architecture

8.2.1 Reference Business Model

The emergence of new players, such as Skype and Yahoo!, in the telephony service market as well as the trend for traditional telcos to migrate their services to run over all-IP networks is indicative of the separation of service and network planes. This is leading to a distinction between the *Service Provider* (SP) and *IP Network Provider* (INP) business roles (see Figure 1). It should be noted that business roles do not necessarily map one-to-one to distinct business entities; a business entity may implement more than one role.

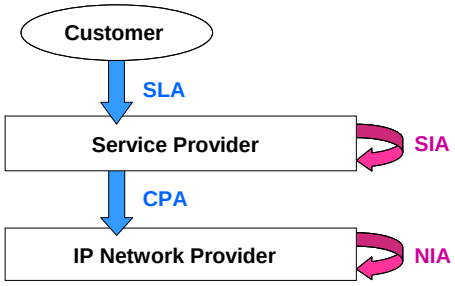

Figure 1 Business Model

INPs offer IP connectivity to SPs and do not offer their services directly to end customers. For expanding the scope of their IP connectivity, INPs interact with each other on a one-to-one relationship basis regulated by *INP Interconnection Agreements* (NIAs). A NIA specifies the QoS and availability performance of the traffic exchanged between the INPs, the scope and the profile of the

traffic entitled to the agreed performance and identifiers to capture distinct flows for providing them differentiated treatment.

The SPs offer IP-based services to end customers. SPs deploy the infrastructure required for the provisioning of the offered services, e.g. VoIP gateways or IP video-servers. To fulfil the IP connectivity aspects of their services, SPs establish *Connectivity Provisioning Agreements* (CPAs) with underlying INPs. Similarly to NIAs, CPAs specify the performance, constraints and identifiers of the service traffic entering the INP's network from the SP's sites. Beyond the connectivity specified therein, the INP offers to the SP means to control the connectivity provisioning, such as setting policing and routing rules and receiving feedback reports. The specific provisioning rules and required feedback are also agreed during the CPA negotiation.

To expand the scope of offered services, SPs interact with each other on the basis of *SP Interconnection Agreements* (SIAs). The content of an SIA is service-specific, e.g. a VoIP SIA may include telephony performance metrics like Average Success Rate or simultaneous calls capacity.

Customers are the target recipients of the services offered by the SPs. Services are offered on the basis of *Service Level Agreements* (SLAs), capturing the terms and conditions for the provision and use of the services.

8.2.2 Network Planes and Parallel Internets Concepts

AGAVE introduces *Network Planes* to differentiate the treatment experienced by IP flows when crossing an IP realm managed by a single INP. The NP notion is internal to INPs and its engineering can be undertaken before or after the formulation of service requirements as expressed by SPs. In addition to traditional QoS metrics, such as delay and packet loss, requirements such as availability are also considered. It is up to the INP to plan/select/(re-)engineer its NPs to meet these SP requirements. A given NP can be used to convey services' traffic managed by the same or distinct SPs in an aggregate fashion. In order to fulfil the service requirements specified in the CPA, INPs need to engineer corresponding NPs within their own network. Technically, a NP can be engineered through the combined tuning of several processes, which span one or multiple dimensions:

- *The Routing dimension*. To support heterogeneous service requirements, different paths can be implemented for individual NPs. Routing differentiation can be implemented at several levels, for example (1) assigning *dedicated topologies* to maintain several routing adjacencies towards the destination; (2) assigning *dedicated path selection configurations* so that multiple path selection configurations (e.g. routing metrics) can be installed, each being dedicated to one specific NP[4]; or (3) configuring *dedicated fast reroute procedures* for service resilience purposes such as pre-configuring backup paths/topologies inside high availability NPs.
- *The Forwarding dimension*. At the forwarding level, an INP can engineer its IP forwarding mechanisms so as to provide different packet scheduling behaviours by configuring different policies in a common scheduler, assigning dedicated scheduling resources, differentiating dropping policies, differentiating failure detection means, etc.
- *The Resource Management dimension*. The treatment experienced by IP packets can be differentiated by different shaping and policing, and the degree of traffic multiplexing, also denoted as over-provisioning factor.

INPs may select the most appropriate combination of mechanisms to implement specific NPs according to the service requirements. Furthermore, an INP will take into account its own operational objectives such as manageability, scalability and resource optimisation to provide cost-efficient NP realisation. In the forwarding dimension, DiffServ is a common platform for supporting service differentiation. As far as the routing dimension is concerned, multi-topology routing mechanisms [9][10][11][12] are regarded as suitable platforms for supporting service differentiation both within

[4] In this case, multiple diverse paths can be simultaneously maintained between individual ingress/egress pairs to support different service requirements from individual NPs.

and across NPs. Specifically, dedicated routing configuration, such as Multi-Topology-OSPF link weight setting, Multi-Protocol-BGP and QoS-Enhanced BGP tweaking, can be performed on top of different routing topologies, each serving a specific NP. Additionally, other mechanisms can be applied for implementing NPs: by using the functionalities of explicit routing and resource reservation of RSVP-TE, dedicated Label Switched Paths can be constructed to support hard QoS guarantees. Alternatively, QoS overlay routing and IP tunnelling [13] techniques can be used for realising NPs with less stringent requirements such as better-than-best-effort services. As far as service resilience is concerned, IP/MPLS fast rerouting techniques [14] can be used. A general overview of NP realisation is described in Figure 2.

The concept of *Parallel Internets* is introduced as an innovative way to enable end-to-end service differentiation across multiple INPs. Specifically, *PIs* are constructed through horizontal interconnection of *NPs* across individual INPs. In doing so, INPs need to negotiate and establish NIAs between each other to bind NPs with similar service characteristics. A salient novelty of the proposed approach is that each instance of PIs is not necessarily implemented with a homogeneous platform across multiple INPs. This aspect provides high flexibility for cooperating INPs to make local decisions in binding their own NPs to the PI.

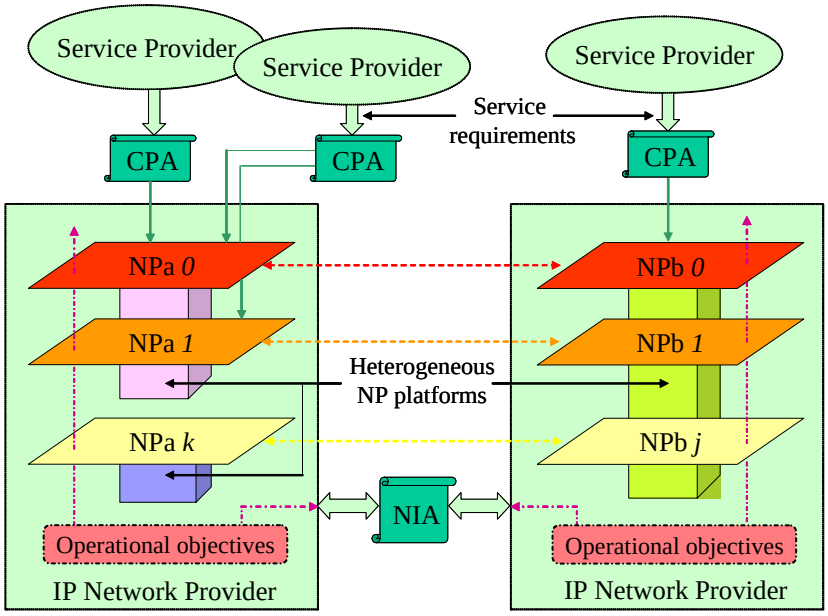

Figure 2 NP and PI Realisation

8.2.3 AGAVE Functional Architecture

8.2.3.1 Overview

This section analyses the interactions between the business roles of Customer, SP and INP and describes the functional blocks required to support these interactions focussing in particular on the internal functionality required to plan, engineer and operate Network Planes and Parallel Internets within an INP.

Building on the business model discussed previously and depicted in Figure 1, each agreement - CPA, NIA, SIA and SLA - is supported by three sets of functional entities corresponding to the three phases of the contractual relationships, see Figure 3. Each set is comprised of a pair of corresponding functional blocks in the business entities operating in the customer and provider roles pertinent to each agreement. Service advertisement and discovery blocks conduct pre-agreement interactions; agreements are subsequently negotiated via ordering and order-handling blocks; and post-agreement the performance of the service is monitored by verification and assurance functions.

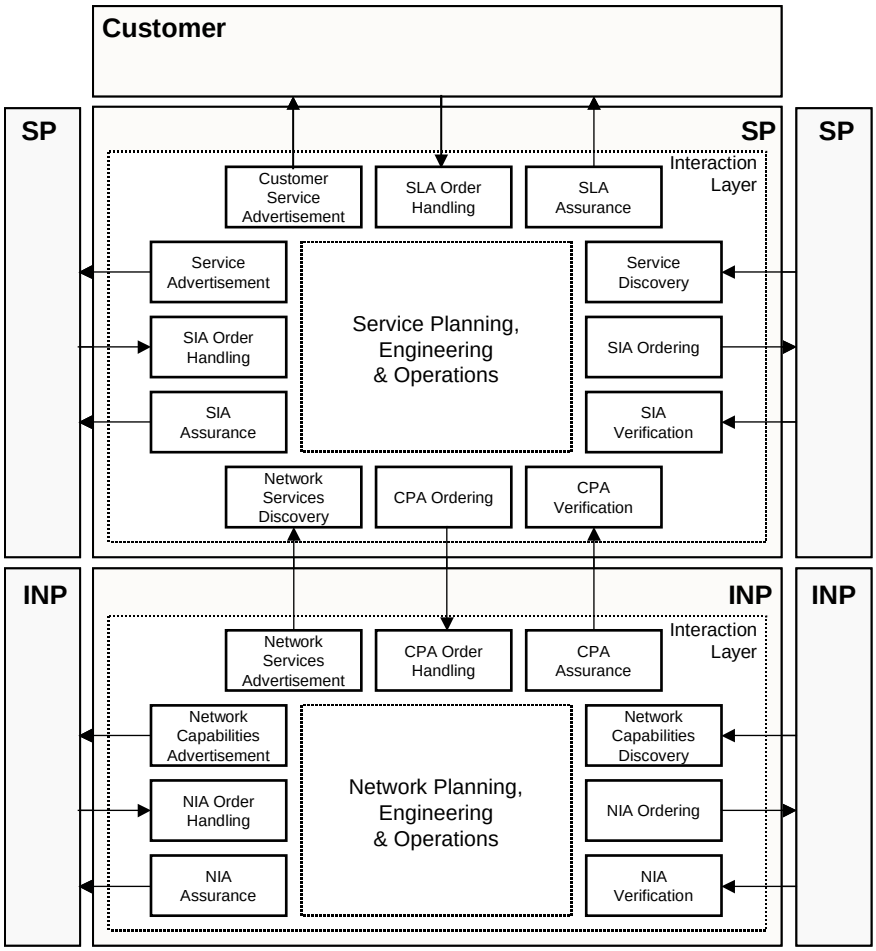

Figure 3 AGAVE Functional Architecture: Interactions Viewpoint

This interactions-focused view hides the complexity of internal SP and INP functional blocks contained in the *Service/Network Planning, Engineering & Operations* meta-blocks. The functional blocks of the INP are further decomposed as follows.

8.2.3.2 Rationale

The rationale behind the decomposition of functionality of the INP is to build a business-process view of the planning, management and operations tasks of the network. The goal is to mirror the internal organisational structure of a typical INP, the steps involved in building NPs and PIs to support CPAs and NIAs with SPs and customer INPs, and to model the interactions between the functional entities.

Three different perspectives of the INP's operational activities are identified. Firstly, the commercial view which is focussed on defining and ultimately selling connectivity services to SPs and customer INPs. Its main concern is to maximise the profit of the INP. The second perspective is concerned with network-wide optimisation of the INP's resources, given the services to be accommodated, their QoS and availability requirements and the anticipated demand. This is where NPs and PIs and their overall realisation objectives are defined. The third perspective is the one that focuses on network engineering and the implementation and configuration details of the NPs/PIs. The later view is heavily dependent of the technological aspects of the mechanisms selected for NP realisation.

While all three business processes are concerned with INP management and operations, they each have different perspectives and concerns as described above and they need to communicate with one another to achieve the network-level configurations that ultimately support the business objectives of the INP. The functional architecture facilitates this by defining clear boundaries between the entities implementing the decision-making processes and specifying the interactions between them based on issues of common concern. The interactions should be based on information models at an appropriate

level of abstraction but with sufficient detail to enable the delegation of tasks from higher to lower levels and the reporting of state and problems encountered in the reverse direction. The specification of the information entities is outside the scope of this paper but can be found in [8].

Figure 4 shows the functional blocks of the INP. The commercial perspective is handled primarily by the *Business-based Network Development* block, supported by *NP Emulation* and *Network Capabilities Discovery/Advertisement*. Network-wide optimisation concerns are dealt with by *NP Design & Creation*, while the detailed network engineering and configuration tasks are located in NP *Provisioning & Maintenance*. The functional blocks are described in more detail in the paragraphs below and their interactions are illustrated through two scenarios in the following section.

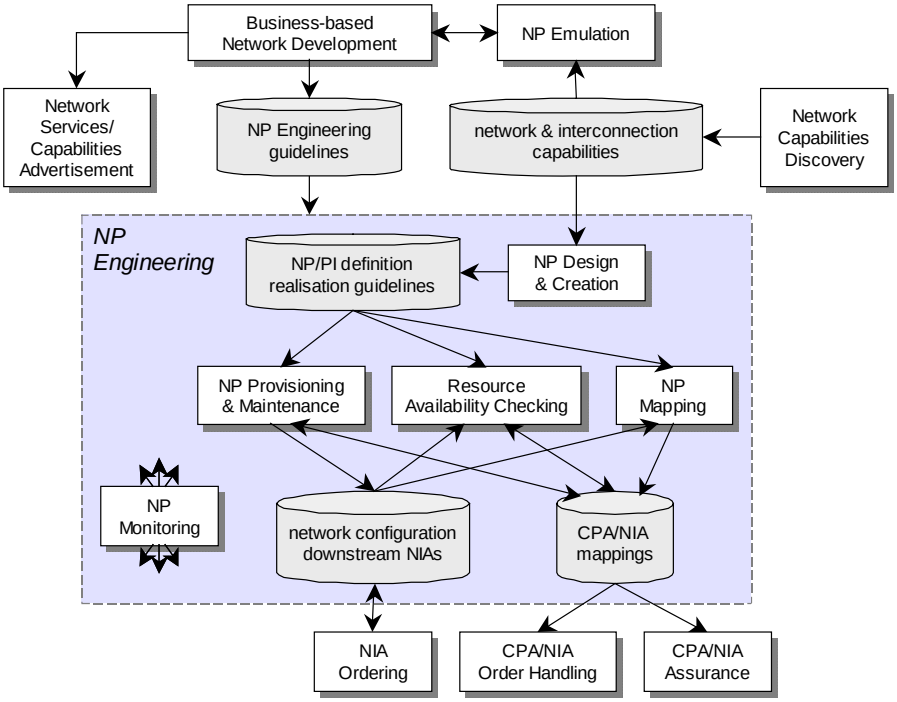

Figure 4 INP Functional Decomposition

8.2.3.3 Functional Blocks description

Business-based Network Development sets the targets for the *NP Engineering* components to fulfil, specifically, the network services to be supported and the guidelines for handling the demand for these services. Target network services are expressed in terms of QoS and availability performance metrics and are optionally restricted to a defined set of local or remote destinations.

NP Emulation provides the *Business-based Network Development* with data to support its decision-making process regarding the impact (financial, engineering, service capabilities, etc) of accepting new connectivity requests, introducing new connectivity capabilities, enhancing the infrastructure, establishing new interconnections, etc. A key purpose of *NP Emulation* is to allow *Business-based Network Development* to make deterministic decisions on the introduction of new services, increasing/reducing the traffic load of existing services and other *what-if* scenarios, by examining the impact of these changes on network performance and ultimately profitability without needing to be aware of the technical details of how the services are engineered/deployed.

NP Design & Creation defines the NPs and PIs required to fulfil the *Business-based Network Development* targets. NPs and PIs are defined in terms of abstract networking capabilities. For the realisation of NPs and PIs, appropriate technologies are selected and directives are produced and fed to *NP Provisioning & Maintenance* which undertakes the actual implementation, producing the appropriate concrete network configuration and NIA Orders negotiated and established by *NIA Ordering*.

NP Mapping produces candidate CPA/NIA mappings to NPs and PIs on the basis of compatibility of the CPA/NIA requirements to the capabilities of the NPs and PIs. The produced CPA/NIA mappings are used by *Resource Availability Checking* to deduce the admission or rejection of the CPA/NIA request by comparing the capacity in the engineered NPs with the demand of the CPA/NIAs. *NP Provisioning & Maintenance* also uses the CPA/NIA mappings to actually accommodate the CPA/NIA traffic demand.

Data gathered by *NP Monitoring* are used to generate notifications and reports for the *CPA/NIA Order Handling* and *CPA/NIA Assurance* to forward to SPs and upstream INPs, for the online Traffic Engineering functions in *NP Provisioning & Maintenance*, for *Resource Availability Checking* to derive appropriate multiplexing factors, for the *NP Design & Creation* and *NP Emulation* and *Business-based Network Development* functions to formulate a high level view of the network performance.

8.2.3.4 AGAVE Functional Architecture At Work

This section illustrates the invoked functional blocks and associated interactions to implement NPs and PIs, employing different routing and forwarding techniques. The first scenario – QoS-Inferred Parallel Internets is an ideal solution for a community of *adjacent* INPs who want to collaborate with each other in order to offer end-to-end QoS across their networks. The second scenario - Better-than-best-effort service, aims to provide less strict QoS between *non-adjacent* domains exchanging high traffic volume.

8.2.3.4.1 QoS-Inferred Parallel Internets

The hereafter QoS-Inferred Parallel Internet scenario relies on the use of DiffServ, the Meta-QoS-class [15] concept and the QoS-Enhanced BGP protocol [12].

Within this scenario, each INP domain is engineered to support a limited number of PDBs (Per Domain Behaviour) through NP Engineering functions (i.e. NPs are implemented as PDBs), one PDB to convey conversational traffic and one for best-effort traffic for example. The dimensioning of these PDBs, including individual PHB profiles, and associated DSCP (Differentiated Services Code Point) values are defined by *NP Design & Creation* and enforced within networks nodes by *NP Provisioning & Maintenances* functions. These PDBs are classified to well-known *Meta-QoS-classes* by the *NP Design & Creation*. Each INP advertises, through its *Network Capabilities Advertisement* function, the *Meta-QoS-classes* it supports. Other INPs can discover these capabilities through *Network Capabilities Discovery* and therefore request NIAs with the advertising INP to make use of offered *Meta-QoS-class* via the invocation of *NIA Ordering/NIA Order Handling* interface. When NIAs are agreed (results of *NIA Ordering* and *NIA Order Handling*), each peering INP activates QoS-enhanced BGP per *Meta-QoS-class* through *NP Provisioning & Maintenances* functions. The resulting QoS-enabled Internet can be viewed as a set of PIs, each offering QoS service levels associated with a specific *Meta-QoS-class* and running distinct instances of QoS-Enhanced BGP. The NIA agreement for a *Meta-QoS-class* makes it possible for INPs to benefit from their neighbour's inter-domain QoS capabilities, and enables them to reach anywhere in the QoS-Internet for that specific *Meta-QoS-class*. The *NP Mapping* function assigns identifiers (one for incoming traffic and another one for outgoing traffic) to be used in the inter-domain links. These identifiers will allow identifying the local PDB to be used to treat the traffic and therefore the PI in which flows will be routed.

8.2.3.4.2 Better-than-best-effort service

In this second scenario, we combine Multi-topology Routing and IP tunnelling techniques to improve inter-domain forwarding performance. We illustrate this combination based on the following example. A company is composed of geographically spread sites that use VoIP to place calls between its sites. Its objectives are to minimize the end-to-end delay between its sites and simultaneously to balance the traffic load in each site.

NP Design & Creation in each INP defines two NPs: NP1, dedicated to low latency service that accounts for a small proportion over the overall traffic, and NP2, used for best-effort traffic.

Furthermore, *NP Design & Creation* specifies how each NP is implemented. In this example, the MT-IGP protocol supports two sets of links weights, one that is optimised for providing the lowest latency paths and the other one that is designed to balance the *overall* traffic load. *NP Mapping* is responsible for assigning traffic flows to a specific NP according to the constraints handled by the *CPA Order Handling*. Traffic flows between SIP Proxy Servers and between VoIP customers and their outbound/inbound SIP Proxy Servers are attached to NP1; other traffic flows are attached to NP2. The assignment of traffic flows onto NPs can be based on packet fields (DSCP or source/destination ports). Based on these assignments, it is known what part of the traffic matrix is supported by each NP.

For what concerns the binding between the NPs in different sites, it is the responsibility of the Tunnelling Service Controller (TSC) to discover and select the paths with the best one-way delay between the local and remote sites, for the traffic assigned to NP1. The TSC discovers the possible ingress routers of the remote sites thanks to *Network Capabilities Discovery & Advertisement*. A communication is established with each remote site to discover its ingresses. Based on an exploration of the BGP routes received by the local AS border routers, the TSC identifies the egress routers that can reach the ingresses in the other sites. Then, the TSC performs a measurement of the latency between each pair of local egress and remote ingress requests (*NP Monitoring*). Consequently, it selects the lowest delay paths. Finally, the inter-site paths are configured in the network, using the *NP Provisioning & Maintenance* block. If tunnels are required, NIAs are established with the corresponding remote sites which result in tunnel establishment. Additionally, local routers must be configured to forward NP1 traffic destined to remote sites through the selected paths.

The off-line Traffic Engineering (TE) engine in *NP Provisioning & Maintenance* selects the IGP link weights that must be configured in the routing topology supporting each NP. The objective of the TE engine is to minimize the delay along the paths in NP1 while balancing the overall traffic load of NP1 and NP2 on the network resources. A typical realization of this optimisation is to assign link weights to the routing topology of NP1 so as to select minimum delay paths and to assign link weights to the topology supporting NP2 in order to spread the traffic load within the network.

8.3 Overview of NGN Architectures

According to ITU Recommendation Y.2001, a Next Generation Network (NGN) is a packet-based network able to provide telecommunication services and able to make use of multiple broadband, QoS-enabled transport technologies and in which **service-related functions are independent from underlying transport-related technologies**. It enables unfettered access for users to networks and to competing service providers and/or services of their choice. It supports generalized mobility which will allow consistent and ubiquitous provision of services to users.

8.3.1 Horizontal Architecture Segmentation

Traditionally, networks are horizontally divided into different segments that are engineered and operated by different entities/departments. The main reason is that although multiple technologies are used for reaching the end users, aggregated traffic is later on handled by a reduced group of technologies.

The access network is responsible for connecting the end users to the network. Traffic from users may come from single devices or terminals directly connected to the service provider network or from customer networks (residential or corporate ones). The access network is subdivided into the access segment (a.k.a. last mile), which stretches from the customer premises to the first network element (a.k.a. access node) and the aggregation segment (a.k.a. backhaul), which comprises the transport network elements that concentrate traffic from several access nodes and deliver it to the core network.

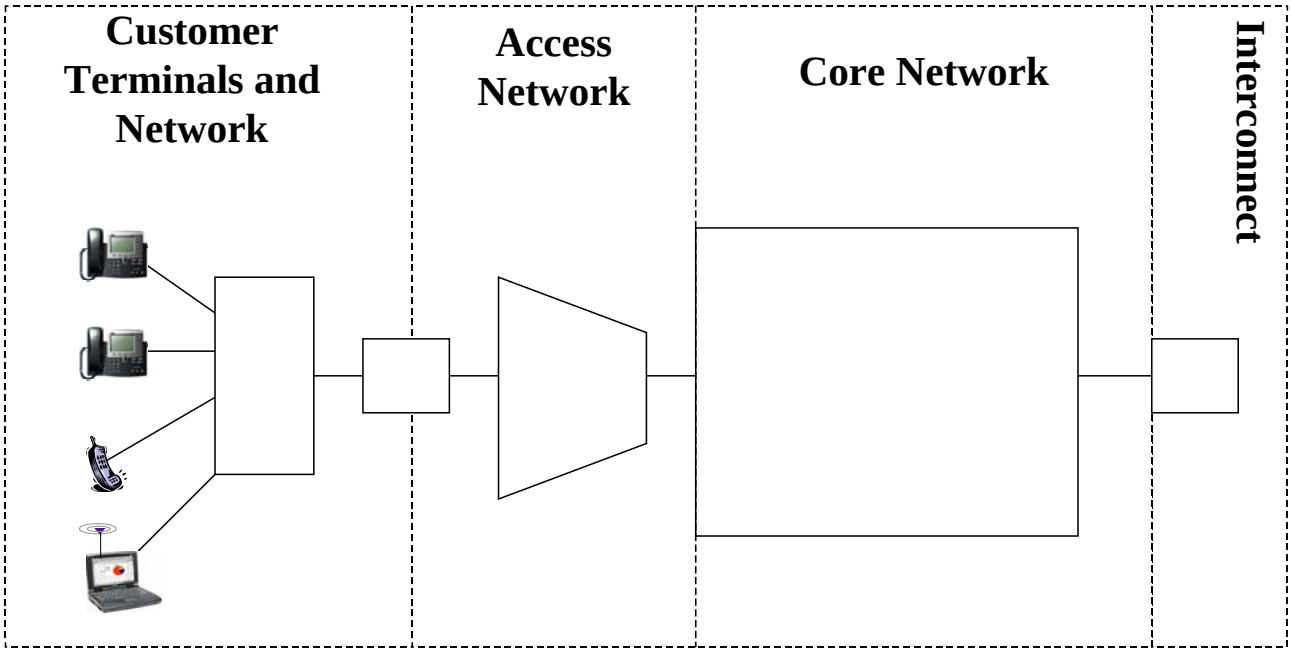

Figure 5: NGN System Components (ETSI TR 180 001)

NGNs assume that the core network is IP packet and that there are multiple access technologies. Hence, NGN architecture focuses on the core network and how this interconnects with other networks.

8.3.2 Architecture Vertical Segmentation

One of the main characteristics of the NGN is the functional separation of the services and transport layers. Orthogonally to these layers, there exist the user (a.k.a. data), control and management planes.

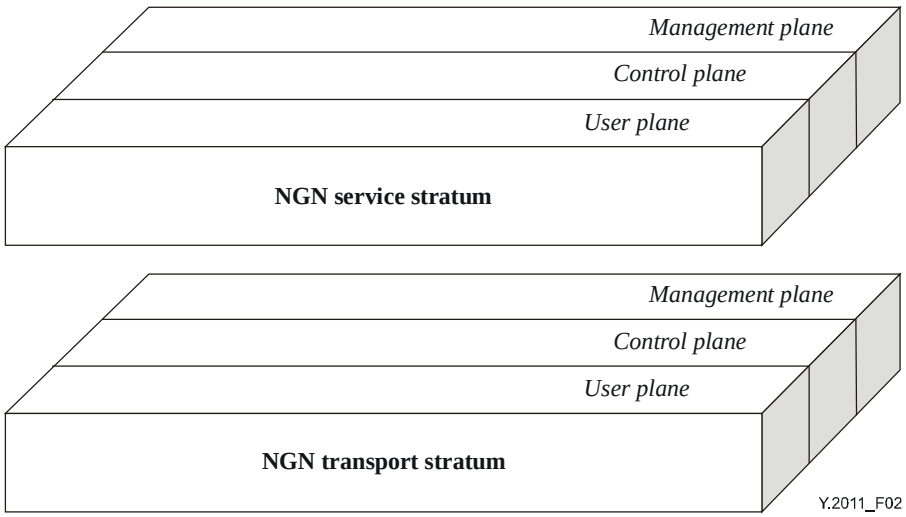

Figure 6 NGN Basic Reference Model (ITU-T Rec. Y.2011)

8.3.3 NGN Functional Architecture Overview

Many standardization bodies, such as 3GPP, ETSI and ITU, are working and collaborating on defining and refining the NGN architecture details.

Figure below shows an overview of the functional architecture of the NGN Release 1 of the ITU-T, where the functional separation of services and transport functions can be distinguished along with the associated control functions.

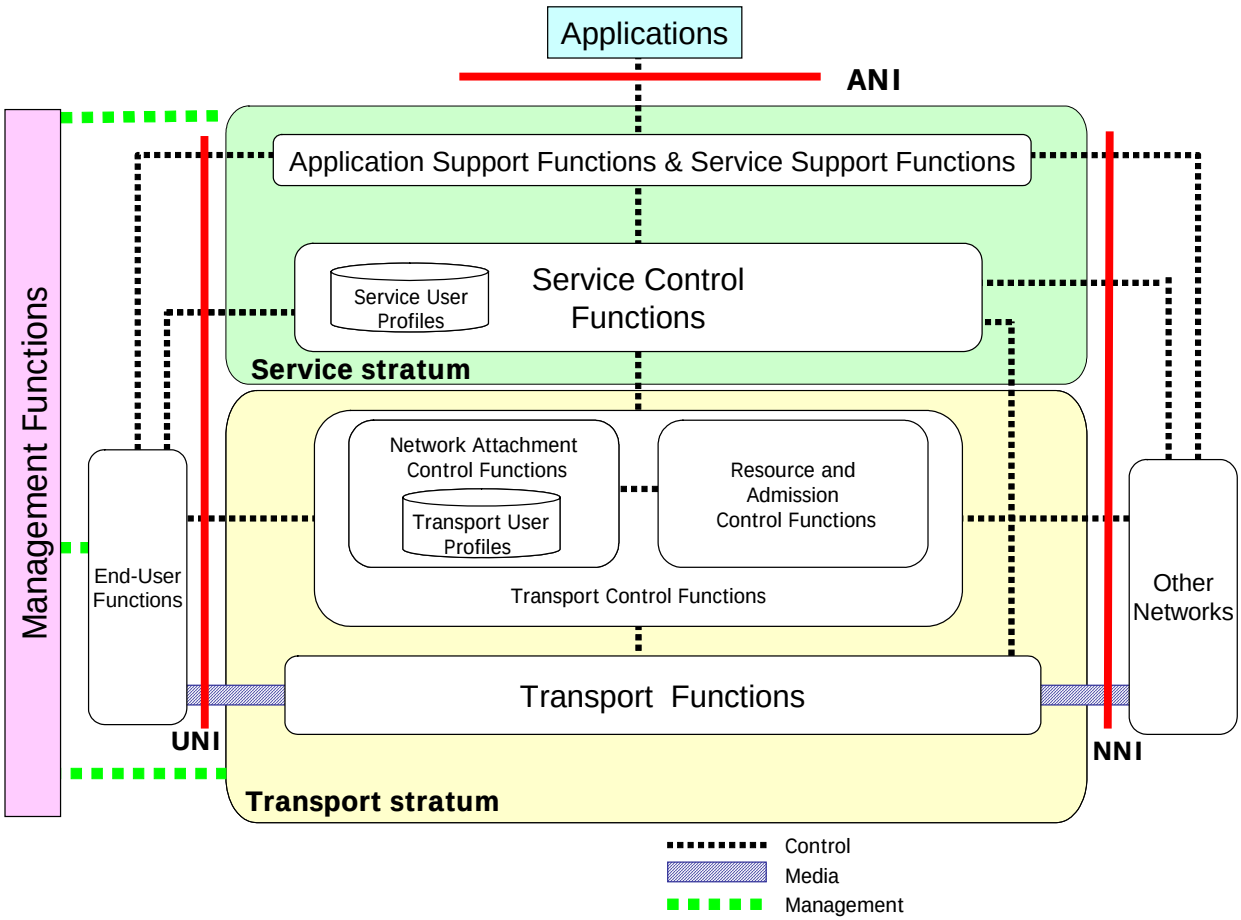

Figure 7 ITU NGN R1 Architecture overview (ITU-T Rec. Y.2012)

According to the ITU NGN R1, the delivery of services/applications to the end-user is provided by utilizing the Application Support Functions & Service Support Functions and related control functions.

The ITU NGN supports a reference point to the applications functional group called Application Network Interface (ANI), which provides a channel for interactions and exchanges between applications and NGN elements. The ANI offers capabilities and resources needed for the realization of applications.

The transport stratum provides IP connectivity services to NGN users under the control of transport control functions, including the Network Attachment Control Functions (NACFs) and Resource and Admission Control Functions (RACFs). Similar subsystems (RACS – Resource and Admission Control Subsystem, and NASS – Network Attachment Subsystem) are defined in ETSI-TISPAN.

Within the ITU NGN architecture [ITU-T Y.2001] and [ITU-T Y.2011], the RACF acts as the arbitrator between service control functions and transport functions for QoS related transport resource control within access and core networks. The RACF provides an abstract view of transport network infrastructure to Service Control Functions (SCFs) and makes service providers agnostic to the details of transport facilities. The RACF interacts with the SCF and transport functions for a variety of applications (e.g., SIP-based call, video streaming, etc.) that require the control of NGN transport resource, including QoS control, NAPT/firewall control and NAPT traversal. RACF decisions are

based on transport subscription information, SLAs, network policy rules, service priority, and transport resource status and utilization information.

The Network Attachment Control Functions (NACFs) provide registration at the access level and initialisation of end-user functions for accessing NGN services. These functions provide transport stratum level identification/authentication, manage the IP address space of the access network, and authenticate access sessions. They also announce the contact point of NGN functions in the service stratum to the end user.

With regards to service layer, IP Multimedia Subsystem (IMS) is seen as the core of the control plane of the NGN service layer, and it is in charge of supporting session-based services, and other services based on the session initiation protocol (SIP).

Additionally, one of the aims of NGN is to support PSTN/ISDN replacement. Therefore, the NGN provides support for PSTN/ISDN emulation as well as PSTN/ISDN simulation.

Other multimedia subsystems such as the streaming subsystem, content broadcasting subsystem, etc are being developed.

Finally, the service layer has also common components used by several subsystems, such as those required for accessing applications, data collecting and charging functions, user profile management, security management, and routing data bases.

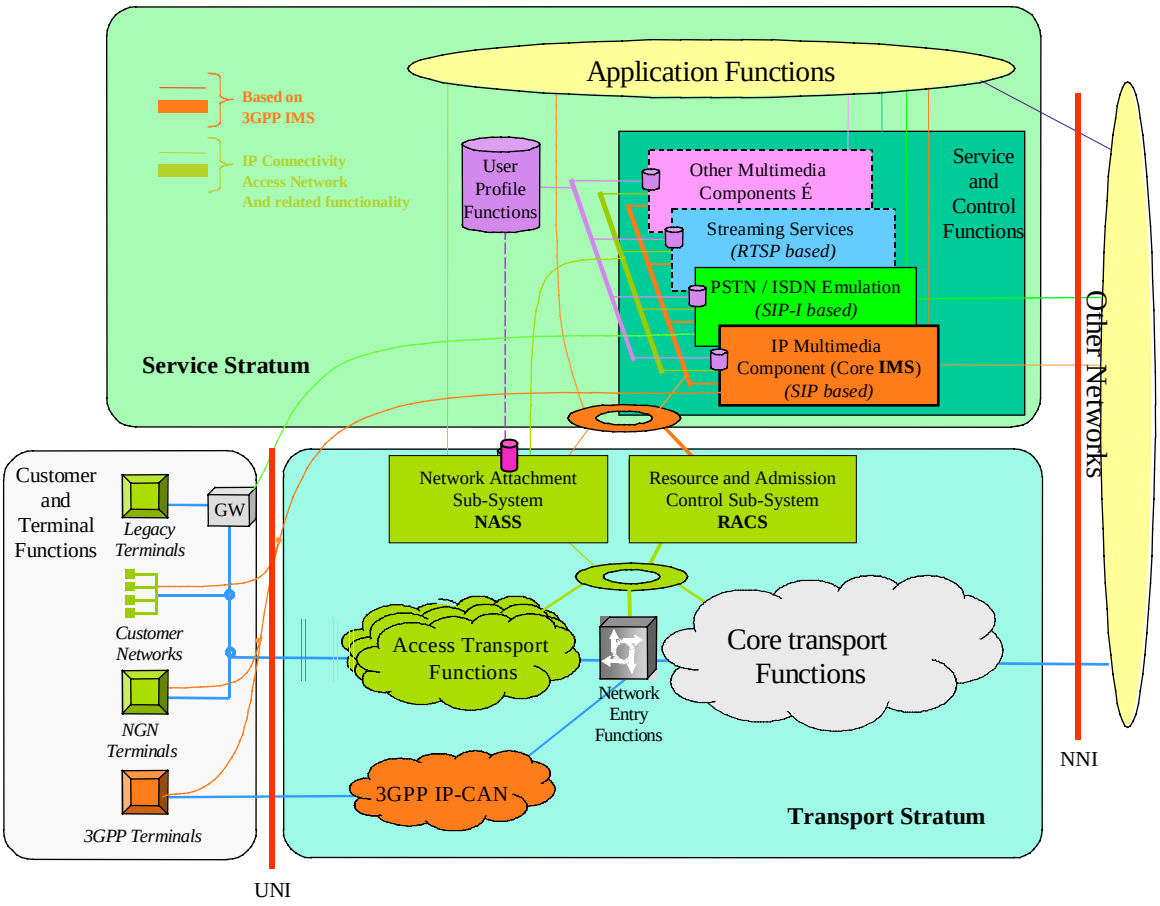

Figure 8 ETSI/TISPAN NGN R1 Architecture overview (ETSI TR 180 001)

8.3.4 Policy based systems

Policies are sets of rules to administer, manage, and control access to and usage of network resources, as defined in IETF RFC 3060.

According to the Common Open Policy Service (COPS) terminology, three functional elements are defined for using policies in a network: a 'Policy Repository', a 'Policy Decision Point' (PDP) and a 'Policy Enforcement Point' (PEP).

The policy repository contains the Policies Rules that are centrally defined in each of the domains. In order to define them, the operator must take into account the SLAs in force with the neighbouring networks, so that the end-to-end behaviour provided to end users are coherent.

The Policy Decision Points (PDPs) are situated within a network domain and they are in charge of deciding which Policy Action has to be applied upon the evaluation of Policy Conditions.

The Policy Enforcement Point (PEP) is the place in the network where Policy Actions are enforced.

8.3.5 RACS Functional Architecture

RACS is the TISPAN NGN subsystem responsible for the implementation of procedures and mechanisms handling policy-based resource reservation and admission control for both unicast and multicast traffic in access networks and core networks.

The RACS architecture is similar to the ITU RACF architecture. The ETSI/TISPAN Application Function (AF) is equivalent to ITU SCF and ETSI/TISPAN NASS is equivalent to ITU NACF.

Besides acting as a resource control framework, RACS also includes support for controlling Network Address Translation (NAT) at the edge of networks and assisting in remote NAT traversal. Furthermore, RACS also covers aspects related to the setting and modification of traffic policies, end to end quality of service and transport-level charging.

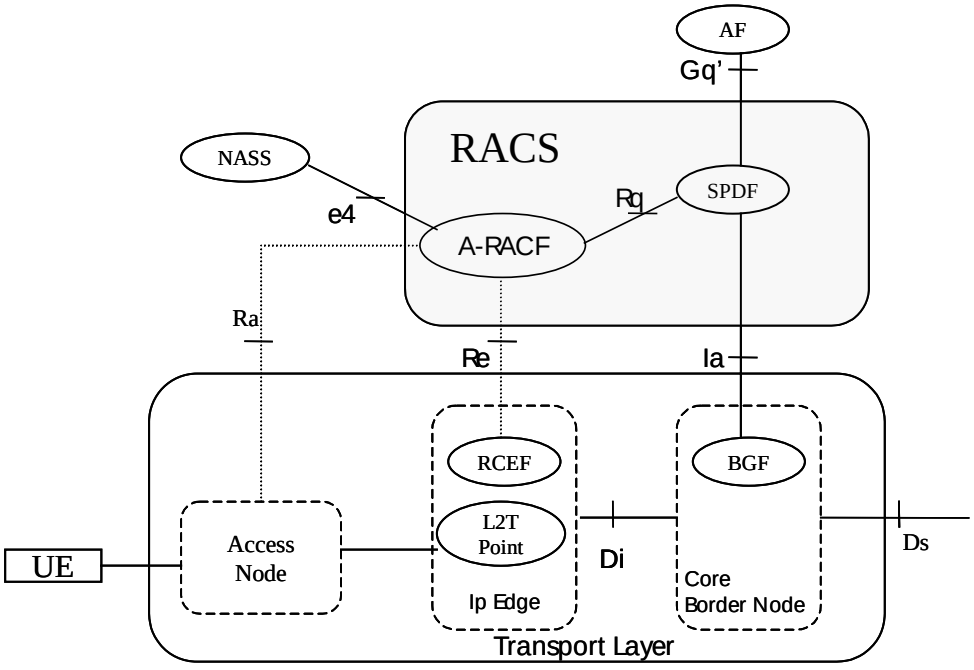

Figure 9 ETSI/TISPAN NGN R1 RACS Functional Architecture (ETSI ES 282 003)

The Service Policy Decision Function (SPDF) provides the AF (similar to the ITU SCF) with a single point of contact. The SPDF makes decisions purely based on service-related policies.

The A-RACF is always in the access network and supports resource reservation. The A-RACF receives requests from the SPDF. Based on these requests, information about the customer profiles obtained from the NASS, network policy information stored in the A-RACF and the status of the network resources within its control, the A-RACF may accept or reject these requests.

The RCEF and the L2TF are two different Functional Entities that are usually grouped into a physical entity called IP Edge Node. The Layer 2 Termination Function (L2TF) is the point where the L2

communication with the CPE is terminated. The Resource Control Enforcement Function (RCEF) is a logical element in the transport layer that enforces the traffic policies (gating, packet marking and traffic policing) by means of which RACS can assure the use of the resources.

The BGF is located anywhere in the transport network. It may be found between an Access Network and a Core Network (C-BGF) or between two Core Networks (I-BGF). BGF supports the same enforcement actions than RCEF but in addition performs NAT, NAT Traversal and usage metering.

8.3.6 RACF and RACS comparison

A comparison between the ETSI/TISPAN and ITU resource and admission control models shows that the RACS model is more specifically targeted at a specific application: support for triple-play services in a wireline access architecture. The combination of SPDF and BGF represents the functionality provided by current Session Border Controllers. Instead, ITU-T RACF model is more abstract and can be mapped to multiple network architectures.

Policy decisions are better separated in RACF. The RACF Policy Decision Functional Element (PD-FE) takes decisions based on service related policies and user profiles, and the RACF Transport Resource Control Functional Element (TRC-FE) takes admission control decisions based on network-related policies and bandwidth availability. The ITU Policy Enforcement Functional Element (PE-FE) is responsible for enforcement actions such as QoS marking, policing, gating, metering and NAT Traversal.

Finally, it should be noted that ITU Functional Elements allow for significant flexibility and variability. Multiple instances of PD-FEs may exist. TRC-FEs may be co-located with PD-FEs, but may also be integrated in element management systems or network elements in the data plane. Native admission control functionality in a specific network domain could also serve as TRC-FE. There may be zero, one or more PE-FEs acting on a specific application flow. For example, if an application flow comes from a trusted and well-known traffic source, policy enforcement in the network may not be required.

8.4 QoS hurdles in 3GPP architectures

Today, Voice over IP (VoIP) is one of the major fields of service innovation and most Service Providers plan (if they have not started yet) to migrate their PSTN infrastructures to IP. For this aim, IMS (*IP Multimedia Subsystem*, [4]) and TISPAN (*Telecoms & Internet converged Services & Protocols for Advanced Networks*, [5]) architectures have been respectively specified by the 3GPP and ETSI communities to meet Service Providers' requirements[5]. Nevertheless, as far as QoS requirements are concerned, 3GPP documentation introduces the notion of "*QoS Class*" but does not clearly define this notion. TS 23.107 identifies four QoS classes: Conversational Class, Streaming Class, Interactive Class and Background Best Effort. But TS 22.105 makes use of four "*groups of applications in terms of QoS requirements*" and points out that there is no strict one-to-one mapping between these groups and the classes as defined in TS 23.107. However TS 22.105 uses exactly the same names for its taxonomy as those of TS 23.107. A key issue resulting from this is: is a QoS class defined in terms of QoS parameter values or is it defined in terms of QoS requirements for a group of applications?

In addition, 3GPP relies on DiffServ to provide network services with the requested QoS parameters, but unfortunately the statement "*DiffServ is used to provide QoS*" says very little on how the network can be actually engineered to deliver the requested QoS. The engineering of QoS requirements, as well as robustness and availability requirements, do not appear to be addressed by 3GPP although the assumption is made that a QoS-enabled network is available, and that QoS can be requested on a per application flow basis especially during the session establishment (expressed as SDP – Session Description Protocol – offer). The success of such a session is a necessary condition for the reservation of appropriate resources in both directions of the call. An example of implementing this mode is the QoS preconditions as defined in [6]. This mode has several drawbacks, such as increasing

[5] Within this paper, we use the same terminology for both IMS and TISPAN.

the connection set up and release times, especially when crossing multiple telephony domains. Moreover, 3GPP specifications do not detail how to check the validity of the QoS requirements enclosed in SDP offers, what is the interface between the VoIP signalling protocols and the QoS enforcement mechanisms, how to validate the required QoS in both call directions, how requested QoS will be guaranteed, or how to ensure coherency of multimedia treatment when crossing several Autonomous Systems and IP Telephony Domains.

8.5 IMS Interaction with AGAVE

AGAVE offers an open interface for SPs to express their service requirements including QoS and availability. Thanks to this interface, SPs need not be aware of the IP engineering operations executed by the underlying INP. The IMS-based SPs can indicate their service requirements through CPAs and there is no need anymore to define the *Bearer Classes*[6] for IP clouds. Underlying INPs run their *NP Design & Creation* machinery to meet these requirements. The process is transparent for IMS-based SPs, decoupling VoIP signalling from the techniques that ensure QoS at the IP level. With this approach, IP resources are not reserved per call but per call aggregates and the IMS functions (e.g. PDF – Policy Decision Function – or RACF – Resource Admission Control Function) perform only a service-level access control and therefore will abandon the reservation per session mode. The IMS-based SP verifies if the VoIP service platform can accommodate the call only based on information like the number of active sessions and the amount of supported simultaneous sessions (especially within nodes embedding BGF – Border Gateway Function – such as SBCs – Session Border Controllers).

The figure below illustrates an IMS-based telephony Service Provider in which BGF nodes are interconnected through the infrastructure of an INP supporting several Network Planes. Connectivity provisioning aspects are negotiated between the IMS-based telephony SP and the underlying NP including code points (e.g. DSCP) for marking the SP generated traffic. In order to honour the IMS SP connectivity requirements (including QoS and robustness), the underlying INP engineers an NP suitable for transport of conversational services traffic. The creation of the NP is opaque to the IMS SP. Section 2 provides examples of engineering NPs.

In order to benefit from these connectivity guarantees, the BGF nodes must mark the outbound traffic with the DSCP code(s) as agreed during the CPA phase. When this traffic enters the INP domain, it is classified accordingly and is bound to the conversational services NP. This traffic is then delivered to the next BGF node and guarantees are met thanks to the NP technical realisations means.

[6] TS23.207 uses the notion of "*IP bearer services*" but never defines it. This notion largely predates the 3GPP work. In I210 ISDN Recommendation (1993), Bearer Services are introduced opposing to Teleservices. A "*bearer service*" is a network point to point relationship. In IP networks this notion has no real meaning: any host is ready to communicate with any other host in the world, and when a communication is established between two hosts, it normally brings no particular states in the network (according to the so-called *fate sharing* property). That would mean a host has always bearer services with all other stations in the Internet.

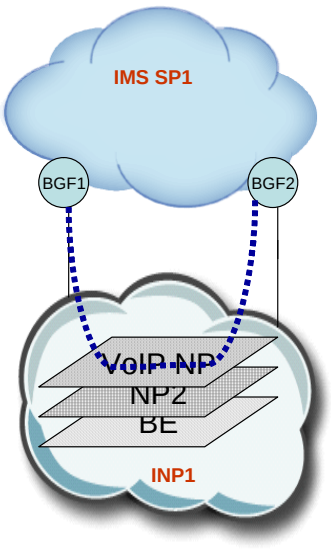

Figure 10 IMS and NPs

In order to offer QoS-enabled conversational services world-wide, IMS-based SPs should not only interconnect together but require also that the underlying IP infrastructure to be engineered in an appropriate manner. Thanks to the deployment of Parallel Internets, a coherent end-to-end QoS treatment is provided across several INPs. Concretely and for illustration purposes, IMS SP1 and IMS SP2 should agree CPAs with their respective underlying INPs. QoS, robustness and scope of the guarantees are part of these CPAs. These CPA are put in effect by engineering corresponding NPs and their binding with external ones to build a PI as illustrated in the following figure. Owing to the deployment of this PI, conversational traffic will benefit from an inter-domain QoS treatment. Note that two scenarios for building this PI are provided in Section 2. The first scenario is suitable for strict QoS guarantees and every intermediate domain needs to have a VoIP-friendly NP. While the second scenario provides enhanced QoS but without guarantees (intermediate domains are not assumed to implement VoIP-friendly NPs and only Best Effort treatment is sufficient.

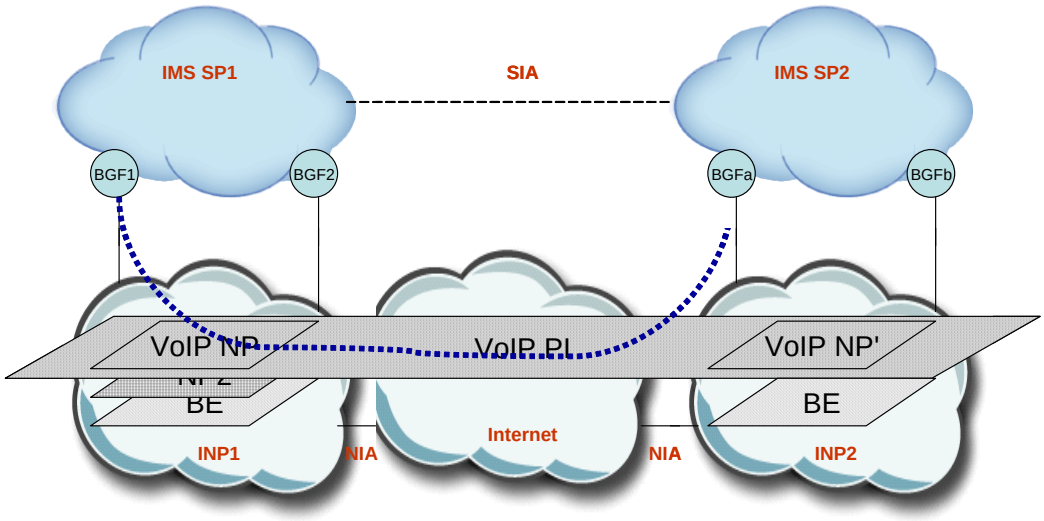

Figure 11 IMS and PIs

8.6 Conclusions

This article has proposed an approach to ease the introduction of differentiated services not only by acting at the forwarding level but also by tuning multi-dimensional techniques at the routing and resource management levels. The concepts of Network Plane and Parallel Internets are introduced. We presented a business model capturing the business actors and their relationship. The adopted business model assumes a decoupling between Service Provider and IP Network Provider roles. In addition, the AGAVE functional architecture is described, including the functions required to offer differentiated services. Functions for engineering Network Planes and Parallel Internets in order to satisfy heterogeneous QoS requirements set by SPs are presented in detail. Two scenarios to build Parallel Internets are provided. The first scenario is based on the use of PDB, Meta-QoS-Class and QoS-Enhanced BGP. The second scenario employs a combination of Multi-Topology Routing and IP Tunnelling techniques. This paper has also identified some QoS problems in 3GPP architecture and proposed a framework to ease the implementation of QoS-enabled multimedia services.

The merits of the presented approach are as follows. Firstly, it advocates decoupling *Service functions from Control functions* by specifying simplified interfaces between the two. Secondly, it is *lightweight* for the SPs since the complexity is pushed to the INP. Thirdly, the approach is *deterministic* thanks to the presence of *NP Emulation* function which assesses the status of the network and evaluates the impact of introducing new NPs and accepting new IP connectivity provisioning requests. Fourthly, it eases the *manageability* of the network resources mainly by optimising operational tasks. Fifthly, it abolishes the node-centric provisioning/configuration approach in favour of *network-based configuration* because the NP *Provisioning & Maintenance* generates the whole NP configuration ensuring configuration consistency. Sixthly, INP may easily evaluate the interference between service activation requests based on the analysis of service requirements. Seventhly, this approach abolishes service monolithic enforcement strategies and introduces a mediation layer to separate the service and network provisioning.

8.7 References

[1] N. Chiappa, “IPng Technical Requirements of the Nimrod Routing and Addressing Architecture”, RFC1753, December 1994

[2] B. Braden, D. Clark, S. Shenker, “Integrated Services in the Internet Architecture: an Overview”, RFC1633, June 1994

[3] S. Blake et al, “An Architecture for Differentiated Services”, RFC 2475, December 1998

[4] G. Camarillo and M. A. Garcia-Martin, “The 3G IP Multimedia Subsystem- merging the Internet and the cellular worlds”, John Wiley, 2005

[5] TISPAN, “Telecommunications and Internet converged Services and Protocols for Advanced Networking, NGN Release 1”, TR180001, 2006

[6] G., Camarillo et al, “Integration of Resource Management and Session Initiation Protocol (SIP)”, RFC3312, October 2002

[7] M. Boucadair et al., “Parallel Internets Framework”, AGAVE Deliverable D1.1, September 2006

[8] E. Mykoniati et al., “Initial Specification of the Connectivity Service Provisioning Interface Components”, AGAVE Deliverable D2.1, November 2006

[9] P. Psenak et al., “MT-OSPF: Multi Topology (MT) Routing in OSPF”, Internet-Draft, draft-ietf-ospf-mt, Work in progress

[10] T. Bates et al, “Multiprotocol Extensions for BGP-4”, RFC 2858, June 2000

[11] Callejo-Rodríguez et al, “A Decentralized Traffic Management Approach for Ambient Networks Environments", 16th IFIP/IEEE International Workshop on Distributed Systems: Operations and Management (DSOM) 2005

[12] M. Boucadair, “QoS-Enhanced Border Gateway Protocol”, Internet-Draft, draft-boucadair-qos-bgp-spec, Work in progress

[13] B. Quoitin and O. Bonaventure, “A Cooperative Approach to Interdomain Traffic Engineering”, Proc. NGI2005

[14] A. Raj, O. C. Ibe, “A Survey of IP and Multiprotocol label Switching Fast Reroute Schemes”, to appear in the Journal of Computer Networks (Elsevier)

[15] P. Levis, M. Boucadair, P. Morand, J. Spencer, D. Griffin, G. Pavlou, P. Trimintzios, “A New Perspective for a Global QoS-based Internet”, Journal of Communications Software and Systems (JCOMSS), November 2005