The AGAVE approach for network virtualization: differentiated services delivery

**M. Boucadair · P. Georgatsos · N. Wang · D. Griffin ·
G. Pavlou · M. Howarth · A. Elizondo**

Abstract This paper describes a new paradigm to realize network virtualization and defines two novel concepts, network planes and parallel Internets, to achieve service differentiation. These concepts are packaged in a technology-agnostic and a multi-dimensional approach for the delivery of Internet protocol (IP) service differentiation, both intra- and inter-domain. The definition of the aforementioned concepts covers several dimensions, mainly routing, forwarding, and traffic management ones. Unlike some radical "Post IP" proposals, this paper advocates an evolutionary approach for enhancing the level of experienced connectivity services (including quality of service and robustness) and therefore to enhance the Internet of the future. Both the rationale and the merits of our approach are explained. In addition, this paper focuses on the critical problem of determining the network planes and parallel Internets to be engineered by a given IP network provider to meet the service connectivity requirements of external service providers. Finally, in order to assess the validity of the proposed approach, a network plane Emulation Platform is described.

M. Boucadair (✉)
France Telecom R&D,
42 Rue des Coutures,
14066 Caen, France
e-mail: mohamed.boucadair@orange-ftgroup.com

P. Georgatsos
Algonet SA,
Athens, Greece
e-mail: pgeorgat@algonet.gr

N. Wang · M. Howarth
University of Surrey,
Guildford, UK

N. Wang
e-mail: n.wang@surrey.ac.uk

M. Howarth
e-mail: m.howarth@surrey.ac.uk

D. Griffin · G. Pavlou
University College London,
London, UK

D. Griffin
e-mail: dgriffin@ee.ucl.ac.uk

G. Pavlou
e-mail: g.pavlou@ee.ucl.ac.uk

A. Elizondo
Telefonica,
Madrid, Spain
e-mail: ajea@tid.es

1 Introduction

1.1 Context and challenges

Internet protocol (IP) has been adopted as the main transport protocol for a large variety of applications and services. New functionalities, features and capabilities have been progressively introduced to IP. By IP, we denote a constellation of layer 3 protocols covering both control and data plane functions. Moreover, and due to its usage to convey critical mass traffic, hard guarantees in term of quality of service (QoS), reliability, and availability are required to be natively supported by IP infrastructure. Additional requirements such as native support of mobility, management, traffic isolation, and security have been expressed by the community to which (some) solutions have been proposed. Furthermore, the end-to-end arguments are

not anymore valid mainly because of the proliferation of intermediate boxes and the needs of service providers to control and secure their service platforms. The introduction of the aforementioned features and capabilities did not take into account a "big picture" of IP leading therefore to the emergence of a complex environment for the value creation. Furthermore, Internet Engineering Task Force (IETF), the IP standardization body, excludes to investigate business issues. In this context, the value creators (i.e. service providers, IP network providers, etc.) are confronted with network engineering challenges, and no practices and guidelines are provided to them in order to ease, "orchestrate," and assess the compatibility of the individual solutions proposed to meet heterogeneous requirements (e.g., operational considerations such as the compatibility of security protocol and QoS ones has been never investigated by the IETF).

Several initiatives have been recently launched to promote innovative ideas in the field of networking as an answer to the current hurdles met by IP networks. Some of these ideas are not yet mature, and some "volatile" concepts have been introduced. An example of these concepts is the "Future Internet" [1]. This concept is used to denote alternative schemes and architectures that are candidate to replace the current deployed IP ones, but no concrete proposals have yet been produced. Moreover, the issue with this concept is that it groups heterogeneous proposals with no clear direction. Nevertheless, these proposals have been widely promoted in the USA under the FIND (http://www.nets-find.net/index.php) and GENI (http://www.geni.net/) programs and also in Europe under the FP7 program, which addresses "Future Internet" aspects among other, with several projects funded such as 4WARD [2].

In this context, both "Clean Slate" and incremental approaches have been proposed, with large-scale experimentation deemed important. Virtualization may be used as a means to achieve this goal. This paper focuses on an incremental approach that aims to solve some of the hurdles encountered by current Internet actors (mainly IP network providers and service providers). The concepts introduced in this paper do not advocate solving all technical issues met by Internet actors but argues that an "orchestration" function is realistic to ease the provisioning of QoS-enabled and robust connectivity services. These concepts represent a promising alternative to ease the delivery of differentiated connectivity services, including both QoS and robustness.

1.2 Network virtualization as a step forward toward the Internet of the future

IP networks are federated transport networks for various types of services. New services, such as real-time distribution of video streams, and the migration of traditional services, such as PSTN (public switched telephone net-

work), toward IP-based ones, demand hard guarantees especially in terms of the service robustness and the perceived QoS. Moreover, services deployed on IP networks are heterogeneous in terms of connectivity requirements, security support, sensitivity to delay and jitter, elasticity of traffic, demand matrix, etc. Taken together with the problems of convergence with mobile networks, which is also known as fixed-mobile convergence, these challenging requirements make IP the "hot" piece in the puzzle of service creation, deployment and operations. In addition to the new requirements, current service offerings encounter additional networking problems such as those caused by the proliferation of middle boxes, lack of deployment of security platforms, or the misuse of IP addresses as both service and locator/identifier.

Given the aforementioned analysis, some voices and initiatives are promoting the idea of "Post IP" or "Future Internet" [1] architectures and networking environments, which "will hopefully" be able to bypass current IP handicaps, provide better QoS and reliability features, and ensure native support to advanced network features such as security and multicast. The proposal is to replace the current IP network infrastructure with a new one designed from scratch, and this is also referred to as a "Clean Slate" approach.

AGAVE [3] was an FP6 European project that addressed the evolutionary approach presented in this study. From the AGAVE perspective, we fully agree with the analysis of the current situation and the problems faced by the networking community but believe that a revolutionary approach is not suitable in the mid-term for several reasons:

– Many problems are not due to the design of IP itself but due to the misuse of the model, e.g., the use of IP addresses/ports as service identifiers does not work anymore in real Internet environments, also with NAT and/or firewalls, application protocols such as the session initiation protocol should not carry layer 3 IP addresses in their message bodies
– Reliability of IP networks can be enhanced in the access segment by investigating techniques such as multiple in multiple out without IP architectures being abolished
– Privacy, security, and address space shortage can be solved through the use of IPv6 rather than persevering with IPv4 and the proliferation of intelligent service-aware border elements such as home gateways and corporate firewalls and today's simple NAT technology
– Routing may be enhanced by promoting overlay routing techniques without requiring for the underlying Internet architecture to be modified. Examples could include the use of inter-domain paths other than those selected by border gateway protocol (BGP)—using for instance explicit label switching path or other source-routing means

– Implementation of alternative solutions to IP are likely to be deployed in isolated "network islands" only—at least in the mid-term—because their introduction requires universal agreement. As evidenced by the delays in deploying IPv6, this can take a long time for several reasons: on the one hand, many telecom operators are currently migrating their services to IP, and large investments have been made in this technology, and on the other hand, billions of end-user devices are based on IP. Universal agreements can be difficult to reach because of the heterogeneity of involved actors and their interests.

– As far as IP networks are concerned, we believe that the current QoS approaches are incomplete, and another step forward needs to be investigated: The experience has shown that the proposed frameworks and architectures (e.g. ETSI Telecoms and Internet converged services and protocols for advanced network [4]) are heterogeneous and often deal with only one piece of the global QoS approach. Service synchronization with QoS benefits should be ensured.

From this standpoint, we believe that an evolutionary approach, and more precisely a virtualization-based approach, is the natural way to investigate how reliability, availability, and non-technical issues, such as usability, support for emergency services and acceleration of service innovation can be enhanced for the benefit of "service providers", "IP network providers," and "end users."

This paper presents our approach for virtualization as elaborated within the AGAVE IST project [3]. Our approach to achieve network virtualization is through optimized provisioning of network planes (NPs) and parallel internets (PIs). Within each autonomous IP network provider's domain, an NP can be described as a slice of network resources allocated for a specific set of services with common or similar requirements, including QoS and availability. The network resources used to implement NPs include the physical bandwidth and other "soft resources" such as routing/forwarding tables and dedicated packet treatment policies. By establishing a connectivity provisioning agreement with the underlying IP network provider, service providers may have their customer traffic treated in appropriate NPs that have dedicated network resources.

AGAVE "virtualizes" the network at a logical level by creating logical network segments through traffic engineering (TE) means, with the purpose of managing the complexity of offering services across the Internet. These AGAVE network segments do not, by themselves, constitute the end products offered by network providers to service providers. Instead, the AGAVE logical network segments are used in two ways: (a) internally by the network provider to serve the traffic from different services and service providers with similar requirements, and (b)

between network providers, on the basis of respective agreements, for extending the reach of a network provider's domain by "combining/interconnecting" its virtual network segments with similar segments of other network providers. In essence, the AGAVE logical network segments orchestrate, through TE, network resources in order to form a network tailored to best meet the requirements of the offered services as well as and the policies of the network provider.

1.3 Paper structure

This paper is structured as follows. First, the AGAVE approach to achieve network virtualization is presented. More specifically, this section presents (1) the business actors and relationships, which intervene in the delivery of end-to-end differentiated services, (2) the concepts of network planes and parallel internets, (3) the rationale behind the AGAVE approach and its merits and (4) a comparison between our approach and the one adopted by CABO. Second, a brief description of the AGAVE framework for implementing network planes and parallel internets is sketched. Then, emphasis is put on the problem of determining the network planes and parallel internets to be engineered in order to meet offered traffic requirements. Finally, a tool for assessing the validity of the AGAVE approach is described.

2 Network virtualization: the AGAVE approach

2.1 Actors and relationships

AGAVE assumes a clear separation between the "service provider" (SP) and "IP network provider" (INP) roles. INPs administer one or more IP domains composed of interconnected IP equipment, related resources, and functions (e.g. routing, switching, forwarding, etc.). They are responsible for ensuring service-ready connectivity at the IP layer. SPs administer a set of service-specific equipment, resources, and functions (such as user-billing means, authentication procedures, and customers' profiles), which are required for the delivery of the services they offer. INPs offer their IP connectivity services to SPs on the basis of respective agreements, which we call "connectivity provisioning agreements" (CPAs), made between them.

Horizontal interactions may occur between INPs and between SPs on the basis of respective agreements, "network interconnection agreements" (NIAs) and "service interconnection agreements," respectively. SPs offer their services to "end users" or "customers" through service level agreements (SLAs). SPs translate their SLAs to access control rules and policies enforced to appropriate nodes in their service domain, so as to allow their "end users/

customers" to access the subscribed services. SPs deliver the traffic flows of the services, as required by the SLAs, using the underlying connectivity services they have agreed with INPs; in essence, SPs map SLAs to CPAs on a many-to-one relationship.

2.2 Key concepts—network planes and parallel internets

Adopting the business relationships described in the previous section, INPs are confronted with the task of honoring the CPAs and NIAs established with customer SPs and peer INPs. CPAs and NIAs may present different connectivity service requirements in terms of a multitude of parameters such as packet transfer (QoS), resilience, and availability guarantees within specified topological scopes, access control, shaping flow forwarding and routing rules, and monitoring capabilities.

To the end of provisioning and delivering different "types of traffic" within and beyond their domain—each such type corresponds to a particular set of connectivity service requirements as outlined previously—AGAVE proposes a *network virtualization approach*, which is built around the concepts of *network planes* and *parallel Internets* [3].

The concept of network plane (NP) is introduced to denote the behavior that IP flows can experience when crossing an IP realm managed by a given INP. The concept of parallel Internet (PI) is introduced to extend the concept of network plane to inter-domain scope. A PI denotes the behavior that IP flows can experience up to the end of reaching a remote destination from a given originating INP domain.

NPs and PIs are defined in terms 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 can be viewed as coexisting parallel networks composed of interconnected NPs. PIs are constructed from the perspectives of each INP, by configuring for each NP different inter-domain routes to certain destinations, through the established NIAs with downstream INPs, based on local criteria.

NPs and PIs represent virtual network segments at a logical layer with specific performance characteristics. NPs can be regarded as local "virtual network segments" whereas PIs as end-to-end "virtual network segments", which are constructed by combining local "virtual network segments" with "virtual segments" of other INPs of *similar*

performance. The local virtual network segments, the NPs, are constructed by the specific TE means employed in the particular INP domain; although not necessary, inherent network resource virtualization techniques could also be considered.

It is evident from the above that the AGAVE network virtualization approach does not aim at creating virtual network segment as "slices for sale" to SPs or peer network providers. Instead, it aims at managing the complexity of honoring CPAs, that is, the provisioning and delivering of different 'types of traffic' within and across network domains. To the latter end, the AGAVE network virtualization approach presents a way that can be incrementally deployed in the today's best-effort Internet.

The NP and PI notions are internal to INPs, and their definition and realization, through TE, can be achieved before or after the formulation of service-specific requirements. SPs see only CPAs from an INP domain. The definition of NPs and PIs and their engineering are hidden from SPs.

The SP requirements for the transportation of the flows of its services are expressed, through CPAs, to an INP in terms of high-level connectivity service requirements in "human-readable" description; they are not formulated as technical implementation choices. It is up to the INP how to select and engineer its NPs and PIs in order to meet the SP requirements.

A particular NP and PI can be used to convey one or several services' traffic belonging to the same or different SPs. INPs and SPs agree on how traffic flows from an SP will be injected (especially, on IP packet marking and identification) and transported to a NP and subsequently to a PI. Therefore, the NP technical implementation is only meaningful to INPs, not to SPs. The correlation (i.e. the binding of a particular SP traffic flow to an engineered NP) between the SP connectivity service requirements and the network engineering (i.e. NP/PI selection and identification) is only of concern to INPs, not to SPs.

CPAs are built upon "Network Services" (NSs), which denote the distinct "types of traffic" that can be offered by a particular INP in terms of QoS, availability, resilience guarantees, and management capabilities within a certain topological scope.

The Network Services are defined by the INP business layer. In addition, business processes defines *Engineering Guidelines*, setting rules for handling the demand for the supported "types of traffic," including the admission of CPA requests. Based on the defined network services and the set of Engineering Guidelines, the INP determines the NPs and PIs to be enforced within a network.

Figure 1 summarizes the above by illustrating the key concepts pertinent to the AGAVE network virtualization approach and their relationships.

Fig. 1 Relationships between key concepts of AGAVE network virtualization approach
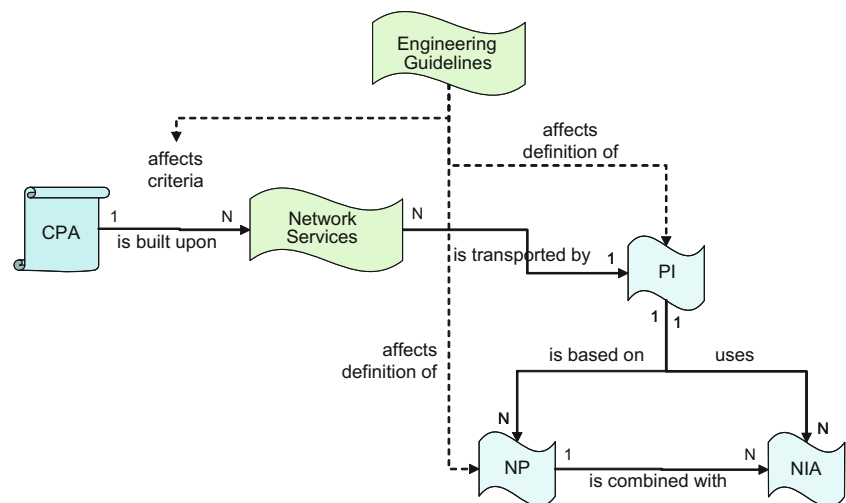
2.3 Network plane definition

A network plane is defined as the output of a combined tuning of several processes, which belong to one or multiple dimensions as listed hereafter:

1. The routing dimension: The treatment that will be experienced by IP packets can be differentiated, thanks to distinct routing policies and routing configurations within a particular NP. Examples of protocols related to this dimensions are multi-topology OSPF/ISIS (M-ISIS [5], MT-OSPF [6]), multi-protocol BGP (MP-BGP) [7] and QoS-enhanced BGP (Boucadair, work in progress). Several parameters can be tuned so as to implement differentiated routing as listed below:

    a. Dedicated network topology: This dedicated topology can be either a physical or logical topology. Therefore, several routing adjacencies can be maintained. These adjacencies are for instance the result of including/excluding nodes and links.
    b. Dedicated route selection process: Several route selection processes can be configured, each can be dedicated to one or multiple services. These multiple route selection processes can operate on the same topology, or for each topology, a route selection process can be dedicated. The behaviors of these route selection processes are not similar.
    c. Different fast reroute procedures: When errors or failures occur for a given topology, the routing process can be enhanced by means of fast-rerouting the IP traffic.
    d. Different policies and metrics: Another alternative to implement differentiated routing is to have dedicated metric settings for each NP. Therefore, the selected path can be different toward the same destination for different service traffic.

2. Forwarding dimension: At the forwarding level, an INP can engineer its IP resources and capabilities so as to have distinct forwarding behaviors by assigning distinct priority values for distinct traffic types, distinct scheduling mechanisms, distinct dropping policies, distinct failure detection means, etc.
3. Resource management dimension: The IP treatment experienced by IP packets can be differentiated by having different shaping and policing polices or the variation of the amount of granted traffic.

2.4 Overview of the AGAVE framework for implementing network planes and parallel internets

Figure 2 shows the functional blocks within an INP domain operating under the proposed framework. The commercial perspective is handled primarily by the *business-based network development* block, supported by *NP emulation* and *network capabilities discovery/advertisement*. Network-wide optimization concerns are dealt with by *NP design and creation*, while the detailed network engineering and configuration tasks are located in *NP provisioning and maintenance*.

Business-based network development and network emulation functional blocks are responsible for the planning of network operations, the production of evolution roadmaps, and network strategy, for expansion decisions of the network services and acceptance of service provisioning requests received from service providers.

The *network plane engineering* functional block is the place where the network planes are created, designed, implemented, and maintained within the network of a given INP. This macro-functional block is responsible for translating high-level requirements to network-specific ones. It is responsible to find the optimized network plane engineering parameters so as to implement the service

Fig. 2 AGAVE functional
architecture
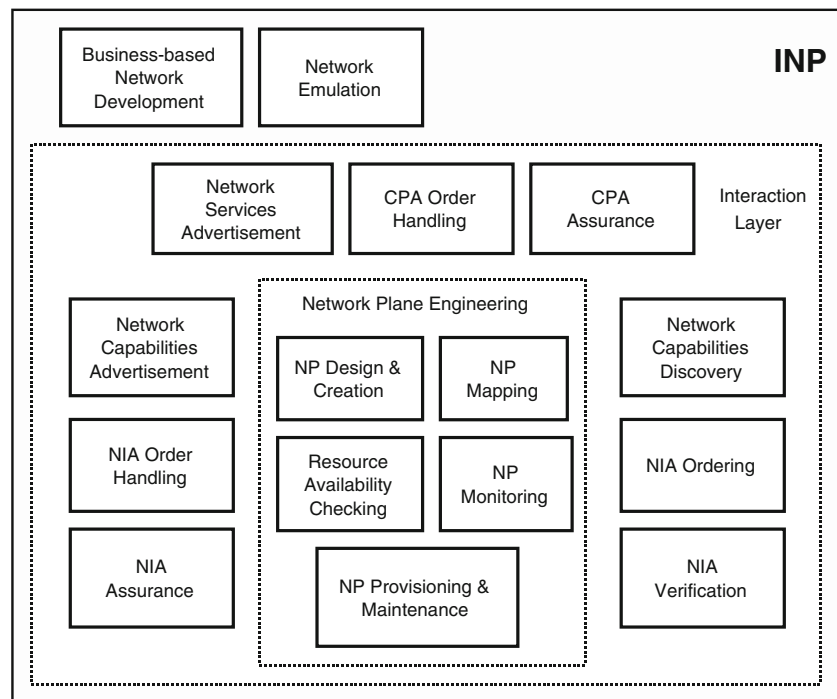
differentiation expressed in terms of network-specific requirements. This problem is denoted as "NP/PI definition problem" and is discussed later in this paper.

The NP design and creation function block is responsible for offline specification of network planes before actual enforcement within operational networks of a given INP. The design and creation phase aims to produce high-level specifications of the network planes in terms of qualitative and quantitative parameters associated with each dimension. This specification is translated into engineering configuration tasks by the NP provisioning and maintenance functional block.

This latter 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 and 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 TE functions in NP provisioning and maintenance, for resource availability checking to derive appropriate multiplexing factors, for the NP design and creation and NP emulation

and business-based network development functions to formulate a high-level view of network performance.

More details about the aforementioned functional blocks and implementation scenarios are provided in [8]. More detail in the informational model of NPs and PIs can be found in [9].

2.5 Discussion—merits and usefulness

As mentioned above, the AGAVE solution is built around the concept of parallel Internets that enable end-to-end service differentiation across multiple administrative domains. Parallel Internets are coexisting parallel networks composed of interconnected network planes. Network planes are established to transport traffic flows from services with common connectivity requirements. The traffic delivered within each network plane receives differentiated treatment both in terms of forwarding and routing, so that service differentiation across NPs is enabled in terms of edge-to-edge QoS, availability, and also resilience.

From an implementation standpoint, the adopted rationale for the design of the INP functional architecture is to build a business-process-oriented view for the planning and management activities of the operational network. From an INP perspective, this approach promotes an abstraction and technology-agnostic layer built around two concepts: network planes and parallel Internets. This abstraction layer is an answer to the need to take into account constraints related to internal organizational structure of an INP in the design process of the steps required in building NPs and PIs and therefore to offer a set of CPAs (respectively NIAs) to

SPs (respectively INPs). The proposed architecture offers a promising communication "bridge" between business and network engineering levels. The NP/PI-based communication "bridge" is independent of specific network technologies, yet is powerful enough to accommodate both intra and inter-domain issues. Taking into account such organizational considerations should facilitate and ease the introduction of the proposed architecture into real organizations and consequently into operational networks.

Several merits of the AGAVE approach can be highlighted, specifically:

- The approach advocates a decoupling of "Service"-related functions from "Control" ones by specifying simplified interfaces between the two and assuming a clear interface between Service Providers and IP Network Providers.
- It is lightweight for the SPs since the complexity is pushed to the INP and an abstraction layer is put at the disposal of SPs to express their connectivity requirements. As for INPs, the proposed framework introduces efficient procedures to manage and provision its IP resources. Operations are driven by NPs rather than specific services.
- The approach is deterministic owing 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.
- It eases the manageability of the network resources by optimizing operational tasks, especially for service provisioning and reporting.
- INPs may easily evaluate the interference between service activation requests based on the analysis of service requirements.
- This approach abolishes service monolithic enforcement strategies and introduces a mediation layer to separate the service and network provisioning. This approach facilitates evaluation and, subsequently, enforcement of various business strategies, avoiding monolithic approaches where the same policy is applied to the entire network for all services.
- When deployed, reduced time for putting new technologies in support of business, thus accelerating RoI (Return of Investment) should be experienced.

- It allows smooth interactions between development and operations within and across business and network levels.

2.6 The AGAVE approach compared to alternative virtualization architectures

An alternative to the proposed architecture is that proposed by concurrent architectures are better than one (CABO) [10]. A key difference between our proposed "network plane" and the concept of "network substrates" for network virtualization proposed in CABO is that an NP is completely managed by the underlying INP instead of being "leased" to external SPs who have the actual control over the "spliced" resources such as path selection decisions on each router. More specifically, the network resources allocated to each NP serve a set of SP's services in an aggregate fashion, rather than being dedicated to any single external SP who has the actual control over its own substrate. In this sense, our proposed approach exhibits a more scalable fashion, as the number of NPs does not increase linearly with the number of requesting SPs. As far as implementation is concerned, there exist two major strategies to realize NPs for service differentiation within individual domains. The first approach is to apply "multi-plane" aware protocols that naturally support differentiated traffic treatment, such as differentiated services [11] in the forwarding dimension, multi-topology IGPs (e.g., MT-OSPF [6], M-ISIS [5]) in the routing dimension. Alternatively, the INP may also deploy multiple co-existing protocols or mechanisms on top of the physical network infrastructure, each dedicated to the realization of a specific NP. It is worth mentioning that the realization of NPs is a completely local issue to be decided by each autonomous INP, and the relevant information on NP implementation is not necessarily exposed to external entities such as SPs or peering INPs.

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 compatible NPs across individual INPs. The aim is to allow individual SPs to geographically deploy their services across the Internet without the necessity to negotiate a dedicated CPA with each of the involved INPs (Fig. 3a). Instead, by establishing a CPA

Fig. 3 AGAVE NPs vs. CABO network substrates
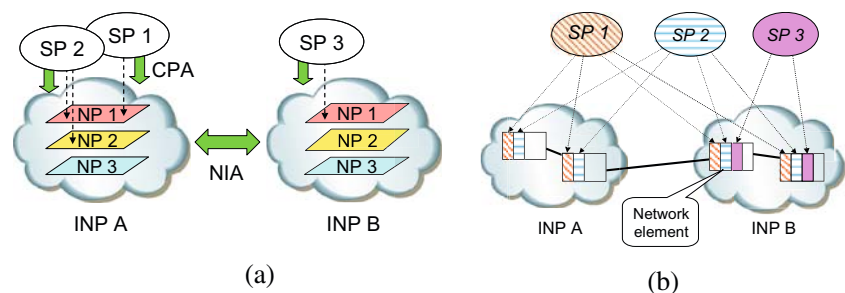
with one single INP, the inter-INP connectivity considerations are effectively outsourced to the horizontal NIAs between INPs. Toward this end, individual INPs need to negotiate and establish INP interconnection agreements between each other to bind NPs with similar service characteristics and requirements. In contrast, the CABO scheme requires the SP to interact with every underlying INP in order to have control over the corresponding network substrate allocated to it (Fig. 3b). Similar to the NP realization scenario, mechanisms used to implement PIs include "multi-plane" aware inter-domain protocols such as MP-BGP [7] as well as coexistence of multiple protocols, for instance plain IGP/BGP routing in conjunction with multi-protocol label switching (MPLS) based path computation services [12].

3 The "NP/PI definition" problem

3.1 Problem set-up

Broadly speaking, the PIs and the NPs are solutions of the following equation:

$$NP \oplus \{NIA\} = PI \tag{1}$$

such that:

$$\{NS\} \rhd \{PI\} \tag{1a}$$

$$\{NP\} \rhd \{TC\} \tag{1b}$$

The variables in the above system are the set of parallel Internets {PI} that the INP needs to provide for accommodating the different requirements of the traffic flows it transports, 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; each one of them 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.

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 $\rhd$ 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 that 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.

3.2 Problem space

The variables pertinent to the above problem assume discrete values, and they are finite in number. This is justified below.

The NPs can be regarded as vectors in a multi-dimensional space, where each axis corresponds to a dimension along which service provisioning can be differentiated. In each axis, there is an ordered set of finite values. These values reflect the level (or 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.

It should be noted that NPs may not necessarily correspond to all possible combinations of the 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.

3.3 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.

3.3.1 Optimization criteria

The optimal solution, the set of NPs to realize, has to be sought against certain optimization criteria reflecting business, network performance, and operations targets. In particular, we see a set of optimization criteria as follows:

- Maximize customer satisfaction, i.e., integrity of the INP in honoring established CPAs/NIAs
- Minimize network cost, i.e., amount of resources required
- Minimize operational cost and overhead.

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

3.3.2 Greedy solution approach

Since the problem space is finite, a solution to 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, initialize: Construct the NP solution space. As outlined above, the NP solution space is constructed by 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, initialization 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 the problem definition, the latter two terms in the above equation are known. Therefore, 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 $\left\{NP_f^{(i)}\right\}$ 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 $NPF = \left\{\left\{NP_f^{(i)}\right\} i = 1 \ldots \Phi\right\}$ constitutes the set of feasible solutions for the optimization problem in hand.
- Step 2: Find the optimal solution, set of NPs, $\{NP_s\}$ to realize the required PIs.
- Evaluate each feasible solution determined in the previous step with respect to the optimization 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, $\left\{NP_f^{(i)}\right\}$, computes appropriate *metrics*, which substantiate the considered optimization 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 operations cost.
- Select the 'best' solution, $\{NP_s\}$, by qualifying the feasible solutions on the basis of the metrics they yield.

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 $\{PI_{req}\}$. There may be the case that the same NP is used for 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/differential services (DiffServ) networks multiple differentiated services code points (DSCPs) can be assigned for the same ordered aggregate (OA). If the network cannot provide such capabilities, the optimal solution should be searched with the constraint that the resulting NPs should be mapped one-to-one to the required PIs.

A key element in the above 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 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. 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.

3.3.3 A differential view

In the following, the "NP definition" problem is looked from the standpoint of its solution space and the traversals

therein toward the optimal solution. In the set of feasible solutions NPF—alternative configurations for instantiating the required PIs—we define an ordering relationship, called *outclassing*, based on the comparison operand ▷ introduced earlier, as follows: the jth PI configuration is said to outclass the ith; similarly, the jth is down-classed to the ith or the ith is outclassed to the jth if and only if the following holds: $\left\{ NP_f^{(i)} \right\} \triangleleft \left\{ NP_f^{(j)} \right\}$

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; 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.

There are then *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 lie between maximal and minimal configurations. From a maximal PI configuration, we can reach an intermediate one by outclassing along certain provisioning dimensions and so on until a maximal configuration is reached. We call this popping *NP merging*. Similarly, through *NP splitting*, i.e., by down-classing 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 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 we go NP merging/splitting to the end of reaching the configuration attaining the optimal criteria?

The optimal 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 is effectively the delta of the evaluation function used in evaluating configuration alternatives in the previously outlined greedy approach, 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 realize, on multiple variables—input parameters—such as the 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 nature of the IP TE 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 sufficiently large to justify a separate NP.
- Intra-domain differentiation for certain PIs (is proved by experience that) 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 a need for human intervention in order to guide and control the move from one feasible PI configuration to another.

3.3.4 Dynamicity—"on-line" version

So far, the "NP definition" problem has been analyzed in a static, so as to say "one-off," form. An "online" version of the problem can be considered. This problem version entails the determination of the optimal set of NPs to realize 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 "online" version is formulated similarly to the static problem, with a list of conditions as additional input to emerge expressed in probabilistic terms.

Compared to the static version, the "online" "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 set of NPs to realize should be determined against the overall, that is, over the period, optimization criteria. A kind of 'best positioning' optimization criteria should be specified.

Broadly speaking, as far as solving the "online" "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, becomes 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 is subject to the underlying assumptions and the errors inherent to the model used. A step-by-step, trial-based approach seems the best way to go around.

3.3.5 Robust NP realization

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 realize.

For being able to safely, within reasonable statistical errors, evaluate network performance, *robust NP realization* becomes a critical issue. Ideally, NPs should be realized so that, to yield an almost-deterministic behavior 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 a 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 NPs?

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

4 NPEP: a network performance evaluation tool

The network plane emulation platform (NPEP) provides a "snapshot" of a network provider domain, operating based on the concepts and notions of the proposed virtualization framework. The platform allows for the definition and

Fig. 4 Overview of NP emulation platform
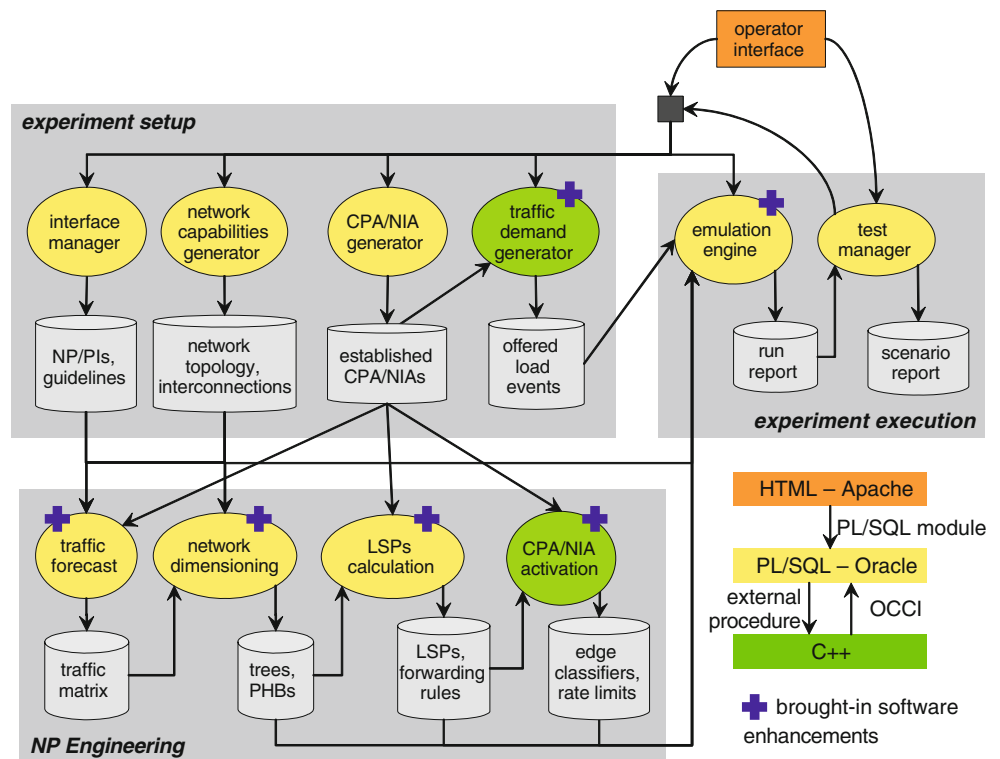
realization of NPs and PIs according to service provisioning requirements. In addition, for a defined set of NPs/PIs, it provides means for generating traffic and measuring the performance of the network in accommodating the generated traffic flows. The platform currently assumes IP networks with DiffServ/MPLS capabilities for realizing the defined NPs/PIs. However, its design is modular, and alternative IP network technologies/capabilities can be incorporated.

The platform is built with the purpose of validating and exhibiting the concepts and notions of the proposed framework. Furthermore, for running "what-if" scenarios and comparison tests to assist decision making on service provisioning, there should be network upgrades and technology choices. As it became apparent from the analysis of the NP/PI problem, there is a need to have a means to evaluate network performance against alternative sets of NPs to the end of determining, which set to realize for instantiating a required set of PIs. NPEP can provide such means.

Figure 4 presents an overall view of the NP emulation platform. As it can be seen, it consists of (a) components pertinent to the proposed framework, such as CPAs, network services, NP engineering guidelines, NIAs, NPs, and PIs, and (b) generic components of an emulation system, such as traffic generation, emulation engine, and reporting facilities.

Furthermore, it includes 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, and NPs/PIs. This way, different TE schemes can be incorporated in NPEP, providing also an idea of how AGAVE can be introduced in a given INP domain(s).

5 Summary

This paper introduced two new concepts denoted as network planes and parallel Internets. These concepts represent abstract network capabilities along which connectivity service provisioning can be differentiated. These concepts are packaged into the overall AGAVE Framework. This framework has been designed to ease the enforcement of differentiated connectivity services into an IP network provider domain and their delivery to service providers. This framework advocates a clear separation between IP network provider and service provider roles and a clear interface between them. Thanks to this NP/PI-based

virtualization approach, the complexity of operating connectivity services are hidden for service providers who can request a connectivity service, which is mapped internally by a given IP network provider to virtual instances of NPs and PIs.

This paper has presented the main benefits of the AGAVE virtualization approach, which are a smooth and efficient network operations taking into account both intra- and inter-domain concerns and a clearly defined incremental approach to service provisioning in the Internet, powerful enough to encompass any technical-level improvement. Compared to some other virtualization proposal, which rely on inherent virtualization techniques such as CABO, this paper has shown added value of the network planes and parallel Internets. Concretely, unlike CABO, our approach scales with evident economy because service providers buy connectivity as a "service" and not as a "network resource" for delivering their services, since AGAVE virtual segments are not visible to service providers. These segments are used for internal operations of IP network providers.

As part of our future work, we plan to undertake system and functional test campaign in order to fully evaluate the validity of our proposed approach.

References

1. Feldmann A (2007) Internet clean-slate design: what and why? Editorial note to ACM SIGCOMM. Computer Commun Rev 37 (3):59–64
2. Soellner M et al (2008) Technical requirements, 4WARD deliverable D2.1 (August 2008)
3. Boucadair M et al (2006) Parallel internets framework. AGAVE Deliverable D1.1 (September 2006)
4. TISPAN (2006) Telecommunications and internet converged services and protocols for advanced networking, NGN Release 1. TR180001
5. Przygienda T et al (2008) M-ISIS: Multi Topology (MT) routing in intermediate system to intermediate systems (IS-ISs). RFC 5120 (February 2008)
6. Psenak P et al (2007) MT-OSPF: Multi topology (MT) routing in OSPF. RFC 4915 (June 2007)
7. Bates T et al (2000) Multiprotocol extensions for BGP-4. RFC 2858 (June 2000)
8. Boucadair M et al (2007) A framework for end-to-end service differentiation: network planes and parallel internets. IEEE Commun Mag 45:134–143 (special issue on QoS Control in Next Generation Networks, IEEE, September 2007)
9. Mykoniati E et al (2006) Initial specification of the connectivity service provisioning interface components. AGAVE Deliverable D2.1 (November 2006)
10. Feamster N, Gao L, Rexford J (2007) How to lease the Internet in your spare time. Editorial note to ACM SIGCOMM. Comput Commun Rev 37:61–64
11. Blake S et al (1998) An architecture for differentiated services. RFC 2475 (December 1998)
12. Farrel A, Vasseur JP, Ash J (2006) A path computation element (PCE)-based architecture. RFC 4655 (August 2006)