UNIVERSITAT POLITÈCNICA DE CATALUNYA
Departament de Teoria del Senyal i Comunicacions

Grup de Comunicacions Òptiques

Towards a cloud enabler: from an Optical Network Resource Provisioning System to a Generalized Architecture for Dynamic Infrastructure Services Provisioning

Author: Sergi Figuerola Fernández

Advisor: Prof. Gabriel Junyent i Giralt

PhD Program: Teoria del Senyal i Comunicacions
Academic Year 2015-2016

Agraïments

Personalment vull agrair el support emocional que he rebut de molta gent per arribar fins aquí. No ha estat un camí fàcil, però una vegada s'ha fet, és quan retrospectivament es posa tota la feina feta en valor, i d'alguna manera o altre en sorgeix un orgull personal.

Així doncs, en primer lloc els meus agraïments a la meva família, qui dia si dia també han estat allà quan se'ls necessitava, a part, crec que també els hi fa particular il·lusió que hagi acabat aquest cicle ;).

Per altre banda, vull agrair el suport incondicional del Prof. Gabriel Junyent al llarg de tots aquests anys. Potser he acabat sent el més veterà dels seus estudiants de Doctorat, però el fet de tenir sempre el seu suport, i insistència en què acabes el PhD, ha estat un dels altres detonants d'aquest últim esforç. També a la Beni V., la secretaria de Doctorat, per la seva paciència i bon fer ;).

Per altre banda i ja més en un caire professional, vull agrair el suport que sempre he tingut del director d'i2cat, en Sebastià Sallent, qui sempre m'ha motivat a continuar endavant amb el meu doctorat, i m'ha donat la llibertat de deixar-me escollir l'estratègia de recerca tecnològica dins el grup que he liderat a i2CAT. Aprofitant que parlo d'i2cat, també vull agrair la llarga col·laboració durant tants anys amb l'Eduard Grasa, i amb qui vaig començar d'alguna manera l'aventura a i2cat. El seu punt de vista sincer i crític sempre m'ha ajudat a tenir perspectiva de moltes decisions, i recolzar-me amb ell per decisions en l'àmbit de la recerca. Per altre banda, també agrair a tots aquells amb qui he col·laborat, i a qui no enumero per no deixar-me a ningú (algun d'ells també presentarà la Tesi en breu, endavant Jordi F., has fet una molt bona feina en la recerca europea)

Last but not least, I want to acknowledge the collaboration I have had with some of my International colleagues. In particular I want to thanks Prof. Dimitra Simeonidou for her support since I began on the research community. Dimitra has acted as a mentor for me and has become a very good friend. Actually, many of my professional achievements are thanks to her support too. Moreover, I also want to thanks to Nicola Ciulli his unconditional support and long discussions about technology. You are an excellent engineer, and I feel proud of being your friend. I also do not want to forget about Artur Binczewski. He gave me the first chance to lead a research activity in an European research project: Phosphorus. Thanks Artur.

Not the fastest,

not the strongest,

but the most responsive to the change

- Charles Darwin -

The limit is our mind,

and our mind has no limits

- Sergi Figuerola ;) -

Table of contents

List of Figures

List of Tables

Part I

Chapter 1. Preface and Introduction

Chapter 2. Doctoral Thesis roadmap

Chapter 3. Sate of the Art

Part I

1. Preface & Introduction

This Thesis exposes and presents the work done on defining new architectures for Service Oriented Networks (mostly optical) and how it contributed to the current state of the art. The ideas and the outcomes developed and published along the Thesis show and validate that the research done during this Thesis is still a trend and is where the industry and research is focusing most if its efforts. Thus, the motivation and interests raised with regards to optical networks architectures at the beginning of the Thesis, were right and in line with the ICT sector evolution. The Thesis research challenge was based on how to define new architectures which could virtualize the optical infrastructure in order to provide coordinated, on-demand and dynamic services between the application and the network infrastructure layers, and so, create service oriented networks as enablers for the new generation of cloud network infrastructures that we have nowadays, and which are still a key topic for research and technology development. This Thesis started in 2005 and lasted until 2013, so the results obtained impacted the state of the art of that period.

A recent article, just presented on July 2015 at SIGCOMM 2015 [1], clearly shows the evolution that networks and technologies have had during the last ten years. Google Inc. has presented the research strategy they started ten years ago in networking. In fact, they already had a strategy towards building their own hardware and software-defined networking (SDN) solutions. When SDN and network virtualization were not yet even a research topic. Google Inc. (an internet search engine company at that time) exposes that most companies would not have considered to build their own data switches to integrate their network with their data-cloud infrastructures as they started to, with the goal to avoid the prohibitive costs and operational complexity on using the equipment's and technology available. Somehow, this Thesis' work came from the same concepts and ideas that motivated them, with the principle of bringing some of the advances of the IT world (e.g. virtualization, commodity hardware…) into the networking world. They already started to research on SDN in order to provide the solutions available nowadays, and which provide alternatives to conventional networking hardware. Solutions which are less expensive and requires less hands-on management, as changes can be done remotely by means of software. This trend is currently placing networking and optical networks as key elements to deploy a distributed and connected server's infrastructure capable to cope with the huge demands that Big Data has on analytics, among many other new service demands. Figure 1 is an example of the exponential growth of traffic in datacenters, and Internet itself, which shows the traffic evolution and so the need to develop architectures capable to coupe with this growing demand, besides the new QoS requirements that also emerged. Moreover, Google Inc. is making a movement that may change the business game. Google Inc. is opening up their infrastructures to third parties. This could be considered the latest trend on networking and IT/Cloud technologies research so, empowering the user and third parties on developing and deploying new services that facilitate the deployment of customized, dynamic and flexible network (and optical) infrastructures based on the demands that applications/users have.

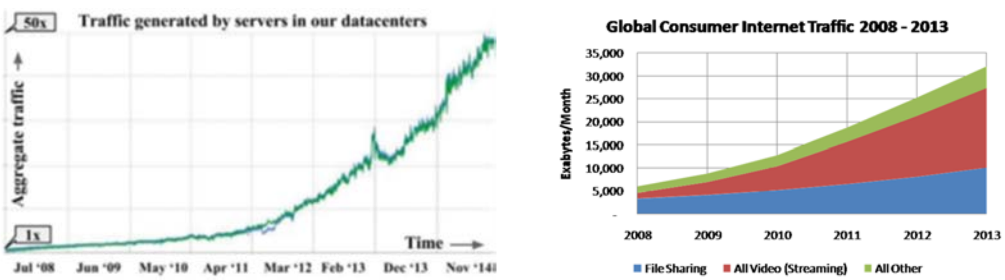


Figure 1 a) Aggregate server traffic in Google datacenter fleet, b) internet traffic growth

Back to the year 2005, when this Thesis research started, the technology and services offered by optical networks, and its management systems, were very different and less powerful compared to current standards. At that time optical networks and telecom infrastructures were mostly manually managed, fix and configured through proprietary networks management systems and interfaces. It was the moment were new optical systems (e.g. R/OADMs, WSS...) and GMPLS/ASON started to emerge as technologies to provide flexibility by means of dynamic configuration/provisioning and automatic recovery services for optical network infrastructures. These technologies provided new control planes to interact with the IP layer, and so provide traffic engineering functionalities at the optical layer too, among many other functionalities. However, these were complex technologies which also required large developments and changes at the operational level of the telecom industry, manufacturers and service providers. And their complexity did not facilitate the technology migration, which did not happen until a few years later.

The challenge was very clear, optical networks started to become key and crucial on the ICT evolution. Optical networks provided high-capacity infrastructure for serving the growing internet traffic demands. In that respect, it was already conceived that optical networks could have an increasing significant impact to our society and our quality of life. Many different organisations (governments, research agencies, universities and the telecommunications industry) started to invest in optical networking related topics and research. The end goal was to derive innovations that significantly would improve the capacity, performance and reliability of future networks.

Moreover, the telecom sector and the IT sector were also totally decoupled. Decoupled in terms of service provisioning, operation and management procedures and standards (without considering the fact of their different technical mentality approaches). It means that there were no tools nor mechanisms to coordinate services between optical networks and IT systems, compared to current open source frameworks capable to offer infrastructures (network and IT) as a service with the ability to integrate third party services on top too (e.g. OpenNaaS [2] , OpenStack [3], OpenDaylight [4]...). This simple fact shows us the fast evolution lived within this period. By the time of starting the research activities there were no open source frameworks which would allow the development of such a capabilities. However, part of the work behind the thesis was already on that direction. Basically, the services, architectures and technologies that we use today were not available nor thought ten years ago. During the last ten years there has been a vast change in terms of technology, user behavior and traffic usage patterns. Thus, not only the technology and the optical systems have changed, also its usage and the business models behind with the new telecom/vendors market demands. We truly live in a new era. The Internet era.

Actually, during the Thesis period, the research behind optical networks and the ICT society was facing a dual change. Technology needed to evolve, improve and change, but also the people barriers. Barriers among those dealing with different topics, technologies and departments needed to change too (i.e. NOC engineers and IT managers). More research exchange and interdisciplinary collaboration were needed too. Internet was starting to emerge and grow at an exponential speed. It was driving the evolution of the ICT sector and its technology. In fact, Internet impact was starting to change the industry and the users themselves, with new services and capabilities. Internet started to get into day by day lives of people and become an industry in itself. Up to the point that, nowadays, people cannot live without full internet mobility. The technology evolution and its usage have been so fast that none could foresee how important Internet would be for daily lives of people, for the business development and for the industry itself, and so, how critic would be the evolution of the internet and optical infrastructures. Said in other words, technology was forced to evolve in a way that network architectures became much more transparent, dynamic and flexible to the end users, whoever the end user would be (applications, user interfaces or simple APIs).

The evolution of optical networks, and internet globally, have been very promising during the last decade. And expectations towards the future are even higher. The impact of mobile technology, grid, cloud computing, HDTV, augmented reality and big data, among many others, have also driven the evolution of optical networks towards current technologies based on SDN architectures and NFV services. Moreover, the convergence of (IP/Optical) networks and IT services that started eight years ago, based on the new generation of service orchestrators and open source frameworks, is nowadays a reality.

As already exposed, this Thesis aimed at defining a new architecture for the management and provisioning of virtual optical infrastructures, and its convergence with IT infrastructures. Thus, the state of the art presented along chapter three reviews the evolution of optical networks and the surrounding technologies by the time the Thesis started (2005). Actually, this state of the art section will end when novel architectures based on Service Oriented Networks started to emerge. It was the period when the Thesis research activities started to produce valuable outputs (2006-20013).

Along this section there will be a presentation of the evolution of the research done, together with the evolution of technology with regards to optical networks. Moreover, a list of the research projects supporting this Thesis will also be presented. Thus, and since the Thesis covers a specific time period, the present/current state of the art is not directly explained. It is not key for the research work done under this Thesis, although it clearly states that the work done was aligned with current research trends.

Based on author's opinion, the work carried out along this thesis has impacted the way we understand technology nowadays. Thus, the research achievements presented (new architectures, new service layers, research publications and articles, conference papers and research projects outcomes) certify how this research has helped on the evolution of optical networks and the telecom sector somehow. Actually, it also opened the door to new research topics and ideas. This research work contributed to the aforementioned evolution of the state of the art of optical networks, with the development of new architectures and advanced virtual infrastructure services. In that way, the research produced and the definition of new architectures went beyond the established ideas of that time, and empowered the user and the applications (first Grid technology and later the Cloud) to become an active actor for the provisioning and management of new optical virtual infrastructure. Somehow, the research done brought back the key principle of internet, which based its growth on empowering its borders (the users).

Summing up, the work done during the research period was focused on the provisioning of virtual infrastructures from the architectural point of view of optical networks and IT infrastructures, together with the design and definition of novel service layers. It means, architectures that enabled the creation of virtual infrastructures composed of optical networks and IT resources, isolated and provisioned on-demand and in advance. With infrastructure re-planning functionalities, and a new set of interfaces to open up those services to applications or third parties.

Nevertheless, the research achievements presented have also been successful due to the effort of many researchers and engineers working on the projects where this research took part. Their contribution, and work on the implementation and validations of the results, was key for achieving great impact along the optical networks evolution.

2. Doctoral thesis roadmap

The PhD project research Roadmap's evolution (Figure 2) aims at presenting the technology evolution that took place along the research activities developed in this Thesis. This roadmap not only provides and understanding of the research activities done, it also established somehow the link between the different outcomes generated during the PhD duration and their alignment. This roadmap is very important in order to get an overview and to understand the work within a defined time line. The technology roadmap schema from Figure 2 measures the technology enhancements with respect to the state of the art by means of different attributes. These attributes are the level of convergence, dynamicity and virtualisation achieved by the technology itself along the period. Moreover, since the improvement of these attributes directly impacted the cost efficiency ratio of the technology, the Capex and Opex impact is also considered in the roadmap, although not assessed. Capex and Opex implications are important because what makes a technology successful, is not its complexity or robustness, but how good the innovation and business model behind are. This is one of the key conclusions (nontechnological) from the Thesis. On the other side, another attributed measured is based on the evolution of management architectures, which moved from being purely managed in a centralized way to a distributed one. Although some of nowadays trends are centralizing again the control decisions (i.e. openflow or some SDN approaches), they can also be deployed in distributed architectures under different types of approaches.

The research roadmap evolution picture (figure 2) is divided into two phases. These phases represent different time periods. Phase I represents the technology available when the Thesis started. The Thesis research started when network virtualisation was simply understood as the consumption of a VPN or VLAN services, without the complexity hidden when abstracting and partitioning a network resource to offer new advanced services like Network/Infrastructure as a Service. Actually, optical network infrastructures were not virtualised yet, and there were no tools nor architectures able to provide on demand virtual infrastructures, which is one of the main outcome of the Thesis. At that time, GMPLS/ASON architectures, together with the advances of new optical technology (WSS, colourless ROADMS...), were emerging as key technologies to provide intelligence to the optical network, and a liaison with the IP layer for Traffic Engineering (TE) functions and set-up of dynamic requests. During this phase, the study of the optical networks evolution and the technology supported was key to identify the gaps that needed to be covered in order to offer new solutions based on new service oriented architectures. It means architectures that could enhance the complexity behind multi-layer expensive approaches, or provide new approaches to multi-domain service provisioning, or simply empower the user by providing new services tight to new applications requirements (SOA was the new trend considered for that purpose).

Phase 2 of the roadmap corresponds to the period were the research activities of the Thesis started and results came. During this period the main focus was on how to provide virtualisation to optical networks with new interfaces for dynamic service provisioning. Research started with the study of Customer Empowered Networks (CEN) and so, the development of new architectures and services for optical network virtualisation, while dealing with new solutions for the so called Articulated Private Networks (User Controlled Lightpath Provisioning). CEN would allow the virtualisation and partition of optical networks in order to provide sliceable networks. And so, telecom operators, or service providers, would be able to offer optical infrastructures to third party operators as a service, something not yet technically conceived nor in terms of business model either. It evolved into Infrastructure as a Service (IaaS) down to the optical infrastructure. Actually, UCLP was one of the first, if not the first approach offering IaaS to optical network infrastructures. Later on it evolved into novel and emerging SON architectures and the integration with GMPLS and/or UCLP together providing new capabilities to deliver services over virtualized optical network infrastructures. During this second phase the research evolved from the proposed SON architecture to new architectures that consolidated the virtualisation work while enhancing the capabilities to upper layers, so fully integrating the optical network infrastructure into the cloud environment, and so providing an architecture that enabled cloud services by means of integrating the request of optical network and IT infrastructure services together at the same level. It set up a new trend into the research community. This technology was a game changer and so, from that point of view, most research started to consider this type of conception when providing new services. It evolved in many different ways until the technology that we use today based on Software Defined Networking (SDN) and Network Functions Virtualisation (NFV).

During the last period of the research work, the author became co-chair, together with the Prof. Dimitra Simeonidou from UK (University Of Bristol), of the Cluster of Optical Networks (CaON) of the FP7 from the European Commission. This cluster grouped all the projects of the FP7 dealing with optical networks, and coordinated their activities and research goals [5] in order to identify a common and aligned roadmap for the optical research community, with the aim to cover the needs and trends detected in the sector. The co-chair role was approved by the coordinators of the different research projects in optical networks and by the European Commission, which somehow contributed on certifying the work done so far within the community.

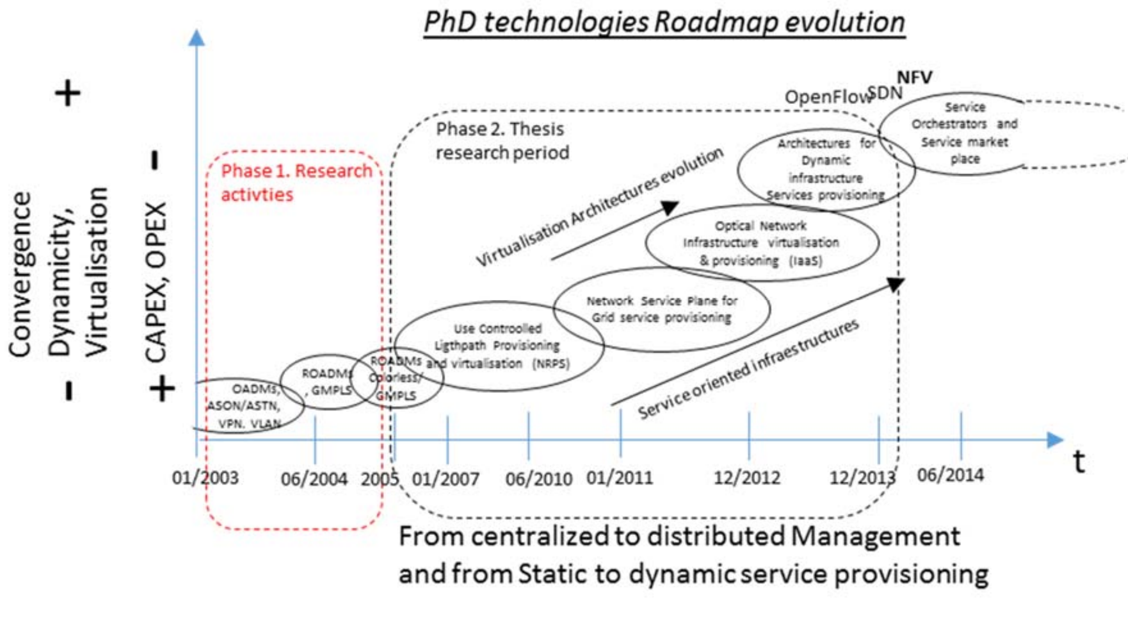


Figure 2. PhD research technology evolution

On the other side, Figure 3 enumerates, within the same timeframe of figure 2, the main research projects that sponsored the research activities of this Doctoral Thesis. Those research projects that also leveraged some of the Thesis outcomes are also presented (with an orange doted circle) in the roadmap picture. The three main projects presented were key for the evolution of the work, the ideas and the thoughts behind the Thesis. These projects also facilitated the testbed and ecosystem for the validation of these ideas. Actually, each research project took know-how, ideas and outcomes from the previous ones, so that, they were enhanced and improved according to new challenges.

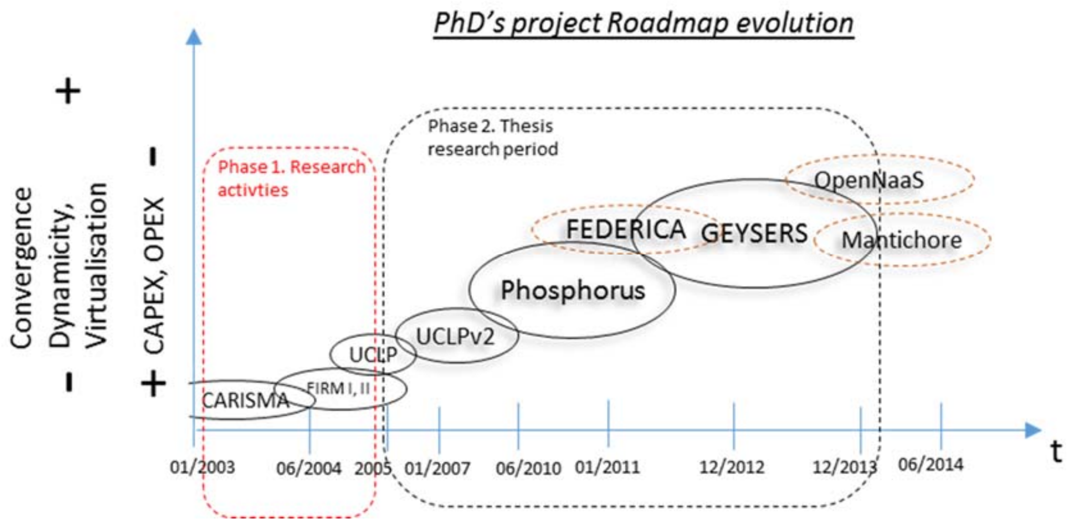


Figure 3. PhD projects Roadmap evolution

Another consideration for the understanding of the work done, and the impact of its outcomes, is the different roles the author had on these research projects from the phase 2 of the Roadmap. These roles are described below, together with a high level description and categorization of the project to identify whether the project was International (I) or European funded (EC FP6, FP7). It is important to mention that almost all the research work done under this Thesis has been achieved through the participation in research projects sponsored by the European Commission, mainly under the Framework Programmes 6 and 7.

The table below give a description of the projects that supported the research done along the Thesis, and the role the author had on these projects:

Project acron.	Website	Name	Type	Description	Role
UCLPv2	NA	User Controlled LighPaths V2	I	An Articulated Private Network Architecture that allowed a physical network to be partitioned in several independent management domains, and exposed the network resources belonging to each partition as software objects or services. These objects could be put under the control of different users so that they could create their own IP Network topologies. It allowed the end users (being humans or sophisticated applications) to create their own discipline or application specific IP network.	TL
PHOSPHORUS	http://www.ist-phosphorus.eu	Lambda User Controller Infrastructures for European Research	FP6	It addressed some of the key technical challenges to enable on-demand e2e network services across multiple domains. The Phosphorus architecture made applications aware of their complete Grid resources (computational and networking) environment and capabilities, and able to make dynamic, adaptive and optimized use of heterogeneous network infrastructures connecting various high-end resources.	WPL
GEYSERS	https://www.geysers.eu	Generalized Architecture for Dynamic Infrastructures Services	FP7	GEYSERS's vision is to qualify optical infrastructure providers and network operators with a new architecture, to enhance their traditional business operations. Infrastructure Providers will compose virtual infrastructures and rent them out to Virtual Infrastructure Operators, which will run cost-efficient, dynamic and mission-specific infrastructures by means of integrated control and virtualization management techniques.	TC

Table 1. Overview descriptions of the projects where research has been developed (TL = Team Leader, WPL = Work-Package Leader, TC = Technical Coordinator)

3. State of the art

This section gives an overview of the evolution of Optical Networks architectures, and the trends towards optical infrastructure virtualisation and IT convergence. It also highlights some of the thesis's topics related to: new architectures for virtualisation applied to optical networks; optical infrastructures as a service and its convergence with IT infrastructure, and how the evolution of technologies empowered the users to control their virtual optical infrastructures. So, understanding the user as the consumer of the service or the cloud application itself.

As said, this Thesis was developed during a period of eight years, so that, the state of the art is divided into two sections. This division will help on better understanding the impact of the thesis' outcomes on the evolution of optical networks.

Section I will present the state of the art of optical communications just before the research activities developed in the thesis started (corresponding to the research activities the author participated before starting the PhD research) will give an overview of the evolution of optical networks before the emergence of Service Oriented Architectures. While section II will highlight the state of the art during the first years of this Thesis, which corresponds to the Phase 2 of the Thesis roadmap. This is the period where virtualisation started to become much more consolidated and so the convergence of network and IT infrastructures by means of SON. The reason to provide this second phase overview is to give a perspective of the topics that emerged when the research started. Topics that became consolidated during the Thesis research period.

3.1. Section 1: Evolution of Optical networks architectures

3.1.1. Introduction

Back to the year 2005, Internet was already emerging as a powerful tool/service that would change the way we understand technology nowadays. It was the beginning of an internet revolution with many new open paradigms for the research community (from control planes to protocols extensions, virtualisation and new architectures). Optical networks were becoming the key technology for transporting the data traffic that was already envisaged due to the growth of Internet. Optical networks were based on fix optical transport systems and networks infrastructures with very limited flexibility and very high implementation costs, not only in terms of economics but on data process too.

Optical networks were mostly based on traditional transport SONET/SDH networks, and were designed and deployed thinking about delivering telephony services. They provided a framework that standardized: line rates, coding schemes, bitrate hierarchies, operations and maintenance functionality, as well as types of network elements required, network architectures that vendors could implement, and the functionalities that each node should carry out. These networks were optimized for voice, providing fixed bandwidth and circuit allocation, and they had a versatile layer on top of SONET/SDH to support a better flexible allocation of bandwidth as well as the integration of data and voice services. This technology was called ATM (Asynchronous Transfer Mode) and was promoted by standardisation bodies back to 1980.

During this period optical network architectures were evolving to adapt to new traffic patterns, and to the introduction of new types of IP-based services and broadband applications to cope with the growing demand of internet. The common architecture was based on a layer stack composed by IP over ATM over SONET/SDH over WDM. However, this network architecture was not optimal in terms of operational and capital expenditures nor enough flexible to provide the requirements to support the emergence of new services. The advent of new data intensive applications with strict requirement in terms of bandwidth and QoS forced the research community to find new architectures capable to map with these foreseen demands coming from upper/user layers (e.g. IP). Thus, the research on network architectures design became a crucial topic.

Although the research trend was clearly towards a full stack network architecture with a client layer (e.g IP) on top of the transport networks (IPoWDM), Telecom operators still used circuit-oriented infrastructure, based primarily in time division multiplexing for delivering connectivity services. Thus, in order to deliver these new emerging IP services and applications, enhancements at the optical network layer were required. The main requirements emerging from new applications and services were: dynamicity, on demand and end to end provisioning, protocol stack transparency and new service interfaces. Together with convergence with IT services.

3.1.2. IP over ATM over SONET/SDH over WDM

This was the most common architecture for the time being. An IP layer on top was the choice, due to its flexibility, computability and end user usage. In fact, most applications and services where already based on IP. ATM was the responsible for statistical multiplexing and multiservice integration (voice and data), and SONET/SDH the switching layer providing statistical multiplexing protection and TE [7]. Actually, ATM was also managing the lower layers, while SONET/SDH provided fix bandwidth allocation and added value mechanisms like protection and restoration [8]. WDM was basically considered as the physical transport medium with bandwidth capacity associated to wavelengths provisioning. Nevertheless, with this architecture, the resource usage was not efficient for the increasing volume of data coming due to the growth of Internet. Even though it offered a robust set of QoS mechanisms, it could not handle the requirements and needs of IP packets networks. IP packets length was variable (from 40 to 1500 bytes) [9] and had to be segmented into fix length ATM cells (53 bytes) in order to be assigned to different virtual connections which, at the same time, were packed into SONET/SDH frames and transported over SONET/SDH transport networks based on SDH Digital Cross-Connects (DXC) and add/Drop Multiplexers (ADMs). WDM systems were used to increase fibre capacity by means of using different wavelengths in the same fibre. In WDM systems the wavelength needed to be chosen among a set of wavelengths fixed by the International Telecommunications Union (ITU) frequency grid (i.e., with 50- or 100-GHz spacing). This architecture however, had two main drawbacks:

- Bandwidth inefficiency due to overhead. Different studies on Internet traffic identified that almost 50% of IP packets were between 40 and 44 bytes long. IP packets are encapsulated on ATM cells, which consist of fixed, small data structures of 53 bytes, where only 32 bytes were available for the payload. That meant that 50% of IP packets needed two ATM cells to be transmitted, while the second one used to be empty. This caused an overhead up to 30%, which was a big bandwidth waste.

- Scalability problems. The hardware needed to fragment IP packets into cells became very complex and expensive as the transmission rate increased. In fact, the maximum bit rate available in ATM networks was 622 Mbps (STM-4); although the standard specified interfaces up to 2.5 Gbps (STM-12).

These issues brought up to 10% of overhead due to the complexity of segmentation (one IP packet into many ATM cells). So, IP routers, connected to ATM switches to provide ATM virtual connections through an ATM network, caused large overhead due to the fact that the current multi-layer approach was not scalable. There was the need to use different type of equipment (IP router, ATM switches, SDH ADMs/DXCs) for each layer and different management systems for the different layer, so extra complexity was added.

First solution to overcome this lack of flexibility was the implementation of Virtual Concatenation (VCATS) and Link Capacity Adjustment Scheme (LCAS) for dynamic bandwidth allocation. VCAT enabled the transport of variable bit data streams by dividing the original payload into separate channels and recombining them at the destination endpoint. LCAS allowed to dynamically increase or decrease the bandwidth of the VCAT containers on demand and to remove failed members of a Virtual Concatenation Group (VCG) maintaining the connection active to avoid excessive protection bandwidth allocation. However, this solution was not scalable either.

3.1.3. IP over SONET/SDH over WDM

Another proposed architecture supported by a large number of data service providers, and which avoided the use of ATM technology, was based in a three-layer architecture. It was called Packet over SONET/SDH (PoS). Packet over Sonet technology was used for transporting IP data directly over SONET/SDH and to eliminate the inefficiencies of the ATM layer. Although this architecture provided reliability and a fast protection mechanism, it suffered from many deficiencies:

- Scalability problems: Due to the complexity of network management, and other technological problems like byte stuffing, building a big SONET/SDH network was complex and very expensive.
- Bandwidth granularity: In SONET/SDH networks bandwidth was offered in fixed increments. If an STM-1 (155Mbps) connection was not enough for an institution, the next step was an STM-4 (622 Mbps), which could be too much. This caused bandwidth inefficiency, which made SONET/SDH circuits more expensive.
- Dynamic bandwidth: If the bandwidth of a SONET/SDH circuit wanted to be changed, it could take up to weeks for the network operator to make the changes due to the complexity of SONET networks.

3.1.4. Next generation optical networks

This new concept came due to the need of evolution required by the traditional communications services revenue models, motly based on legacy technology without scalability features. Thus, the main goal was on finding new models with shorter equipment life-cycles as well as customer demand for newer and more customizable services. The main source of inefficiency of the IP/ATM/SONET/WDM and the IP/SONET/WDM architectures was due to the fact that the optical layer was too simple, in the sense that it only provides raw capacity and wavelength multiplexing. The challenge therefore was on transforming the optical layer, from being a group of point to point pipes to a resilient and manageable optical network which could offer more advanced transport network functions. These functions/mechanisms were:

- Cross-connection on the wavelength granularity.
- Transport network (layer 1) functions:
    - Protection and restoration, Monitoring, Management and supervision on the wavelength granularity and Circuit or wavelength dynamic provisioning.

To perform all these functions more intelligence had to be added to the optical transport network and to the upper layers, but without using the current well established technology of these times. The trend was clear, technology needed to evolve in order to eliminate the SONET/SDH and ATM layers and transform the backbone network into a two-layer network (IPoWDM)

In that sense, a data framing was still necessary to encapsulate the IP payload. So, Gigabit Ethernet or 10 Gigabit Ethernet was supposed to be the best option. This framing provided the following advantages:

- Bandwidth granularity (from 10 kbps up to 10 Gbps).
- Flexibility.
- Simple network management.
- Low cost.

During this period, the Multi-Protocol Label Switching (MPLS) [10] technology emerged as the standard way to integrate the IP and the ATM world. The reason was the need to migrate from current TDM-based networks into a more flexible, dynamic and more-cost effective network solutions. MPLS provided mechanisms to run IP and ATM protocols on the same layer. It supposed the integration of the IP client layer with the optical transport network, while some of the functionalities performed by ATM and SONET/SDH layer were moved to the IP layer. MPLS became in charge of the Traffic Engineering functionalities performed by the ATM technology, while the transport capabilities of SONET/SDH (e.g. protection and bandwidth management) were absorbed by the optical layer thanks to the evolution and technology advances of the optical systems technology.

With MPLS networks, which are also extensible deployed nowadays, traffic was forwarded using labels. The ingress router, also called edge Label Switch Router (LSR), marked IP packets with a label, which was used by the LSRs within the core network to forward the traffic along the desired path, while the label was removed at the egress site by an edge-LSR. The main advantage of MPLS, besides de QoS mechanisms that it provided, was the fact that the traffic going to the same IP network could be aggregated into the same label, so the growth of the IP forwarding table (the label forwarding table) could scale. Moreover, the advent of WDM, extended the transmission capacity of the fiber into the terabit range, with a spacing between lambdas from 10 to 20 nm (CWDM) to 0.1 nm (DWDM).

The advantages of label switching were based on speed, delay and jitter (faster than traditional forwarding), scalability (large number of IP addresses could be associated with few labels), resource consumption, route control, traffic engineering (support of Traffic Engineering to links and nodes according to the traffic considerations).

Besides these technological advances however, there was the clear need of enhancing the optical network capabilities and evolve towards a two layers architecture based on IPoWDM. These enhancements meant more intelligent functions to quickly respond to changes in the network topology and service distribution. Functions like wavelength configuration, automatic power balance, optical-layer performance monitoring that would help on reducing the operational expenditure costs. At that moment in time, the ITU-T supported the concept of Optical Transport Networks (OTN) to extend the capabilities of optical networks [11]. ITU-T/OTN defined the network as a set of optical network elements connected by optical fibres able to provide functionalities of transport, multiplexing, switching, management, supervision and survivability of optical channels carrying client signals. OTN was designed to provide support for optical networking using wavelength-division multiplexing (WDM) unlike its predecessor SONET/SDH, so combining the Operation, Administration, Maintenance and Provisioning (OAMP) functionalities of SONET/SDH with the bandwidth capabilities of DWDM. Some of the features included in OTN were the addition of Forward Error Correction (FEC), protocol transparency or multi-wavelength support.

The main problem however, was on deciding how to divide functions and tasks between the optical layer and the IP layer, and how to coordinate them in order to optimize network investments and maintenance costs while guaranteeing network reliability and satisfying QoS requirements. In order to perform these functions at the optical level, new technologies were needed. Actually, in terms of planning, the optical network could not stay anymore in the configuration defined during the network and resource planning phase. Optical network configurations changes were needed and should occur during the planning phase due to either traffic change (usually an increase) or link/node failures. It is during this period were the introduction of new optical systems: reconfigurable optical add/drop multiplexing (ROADM), wavelength selective switching (WSS) technologies and new tunable filters and tunable lasers among others, together with new network control planes, generated an increase of expectation towards offering 'lambdas on demand' within the scope of next generation networks. Together with the control plane they would enable an additional degree of freedom to the implementation of network traffic engineering, allowing the reconfiguration of connections when traffic demands increased or new connections were requested [12]. The trend was: a) flattening the layered architecture and b) adding more intelligence to the optical layer; that meant adding a control plane (GMPLS or OBGP based).

In this way, next generation of management and control planes for optical networks were clearly going to be based on direct (dynamic) provisioning of individual wavelengths and services on those wavelengths, through lightpaths that would allow applications to signal their requirements directly to the intelligent network devices, so that they could provision their own resources dynamically [13].

What was absolutely clear was that intelligence to the network would be controlled by an IP-based control plane. Research efforts were focused into models based on architectures capable to provide to the end user, and/or customer, with a much greater degree of freedom in selecting options, which was a scenario not yet possible at the time being.

3.1.4.1. *Automatically Switched Optical Networks (ASON)*

Arrived at this point, the evolution of optical network was towards Automatic Switched Optical Network (ASON) architectures (Figure 4). These architectures provided intelligence and functionalities to the optical transport network without changing the transport plane functionalities (the IETF defined ASON as an alternative/supplement to NMS based connection management) [6]. The optical element that facilitated this evolution was the optical cross connect (OADM) that evolved into Reconfigurable OADMs (ROADM). This element allowed to move from a traditional SONET/SDH ring to a full mesh network. This type of mesh network added extra complexity that required some level of intelligence to perform automatic functions such as provisioning and restauration, among others.

ASON[14] was aimed to automate the resource and connection management within the network in order to provide on demand bandwidth provisioning, together with the support of TE mechanism directly to the optical network. It offered dynamic signalling-based policy-driven control over OTN and SDH networks via a distributed (or partially distributed) control plane that provided auto-discovery and dynamic connection set-up. This enabled: a) improved support for current end-to-end provisioning, re-routing and restoration, b) New transport services such as bandwidth on demand, rapid service restoration for disaster recovery, switched connections within a Private Network, etc. and c) support for a wide range of narrowband and broadband clients signals such as: SDH/SONET, IP, Ethernet, ATM, Frame Relay, ESCON, FICON, Fibre Channel and Audio/Video.

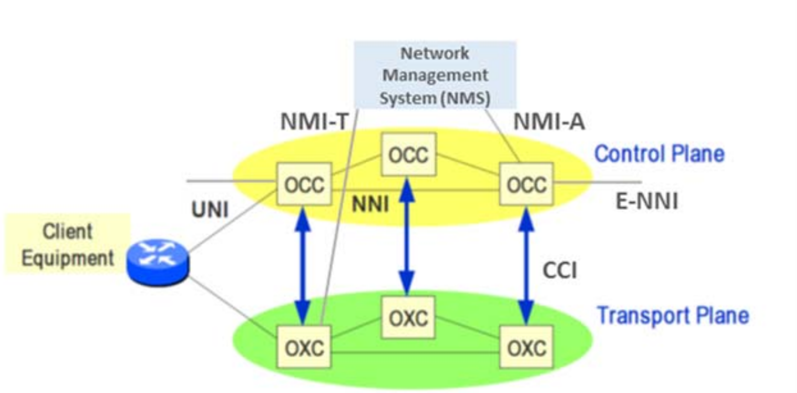


Figure 4. ASON architecture. A standard layered model for the automatic Switched Network management, standard interfaces between the management and control plane (NMI-A), and between management and Data Plane (NMI-T) are shown

The terminology used in the ASON [15] [16] architecture and the open interfaces capabilities of its domain boundaries were as follows:

- OCC: Optical Connection Controller
- OXC: Optical Cross Connect (also called Network Element –NE-)
- NMI-A/T: Network Management Interface (A: ASON control Plane, T: Transport plane)
- UNI: User-Network interface
  - Enabled: client driven end-to-end service activation, multi-vendor interworking, multi-client (IP, Ethernet, TDM,..), Multi-service (SONET/SDH, Ethernet…), Service monitoring interface for SLA management.
- E-NNI: External network-network interface
  - Enabled: End-to-end service allocation, multi-vendor interworking, Multi-Carrier inter-working, Independence of survivability schemas for each domain.
- I-NNI: Internal network-network interface
  - Enabled: Intra-domain connection establishment, Explicit connection operations on individual switches.

ASON architectures enhanced the capabilities provided by previous architectures, and besides supporting basic leased-line connections (permanent: provided manually by a management system), it supported two new types of transport services:

- *Soft-Permanent Connections (SPC): The set-up was triggered by the Network Management System (NMS) (located at the management plane) and the configuration was carried out through signalling and routing protocols generated by the network.*

- *Switched Connections (SC): The connections could be set up or released from a customer and on demand using signalling and routing protocols.*

The main functionality behind the ASON architecture was the Network Control Plane (NCP). ASON control planes are still running on many operators infrastructures of nowadays. These control planes use to reside in a distributed network intelligence constituted by the optical connection controllers (OCC). The OCC could run in separate workstations as the element managers but generally resided into de control unit of the equipment dedicated to switching at each network layer. The OCCs was interconnected via the interface called network to network interface (NNI) and ran the control plane protocol suite having the following functions: network topology discovery (resource discovery), address assignment to an equipment port when it is discovered and address advertisement to all the networks, signalling (connections setup, management, and tear down) and connections routing among others.

The main goal of the NCP was to reduce the complexity by adding intelligence to perform automatic functions. With the introduction of a NCP between the traditional management and transport plane control plane, the tasks delamination and responsibilities of the different layers were much clearer. The transport plane comprised the optical devices and physical links where the data traffic circulated. The control plane controlled the network resources to provide routing and signalling capabilities for the establishment of optical connections. And finally, the management plane was responsible for the management and supervision of the underlying layers.

The NCP became an important technology for the evolution of optical networks and so, it also become a major focus of work for the different standardisations bodies (i.e., ITU-T, Internet Engineering Task Force (IETF) and Optical Internetworking Forum (OIF)). The main functionalities of the NCP included routing information dissemination, path computation, signalling, connection establishment, and resource management, implemented in terms of control protocols executed between communicating entities.

The NCP could be based on three different models, depending on the type of information exchanged between the nodes from the optical transport networks (ASON) and the client networks (IP/MPLS). These models were:

- *Overlay Model:* Under the overlay model, the IP layer routing, topology distribution, and signalling protocols were independent of the routing, topology distribution, and signalling protocols within the optical domain, so that IP/MPLS and ASON networks had separate control planes. The IP network acted as one of the clients of the optical domain that provided point-to-point connections and the clients setup optical channel requests through the User Network Interface (UNI) [17].

- *Advantages: no visibility of the optical topology to the client layer, client-independent, optical layer keept the control of the optical transport plane.*
- *Disadvantages: two control planes, low interoperability.*

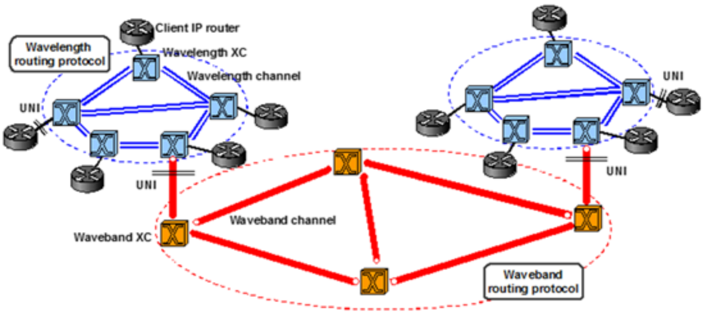


Figure 5. Overlay model

- *Peer Model:* Under the peer model, the IP control plane acted as a peer of the optical transport network control plane. This implied that a single instance of the control plane was deployed over the IP and optical domains avoiding unnecessary duplications of functionality and simplifying communications among layers. Thus, a router could compute an end-to-end path across an optical infrastructure, just because the IP routers could request an optical connection with other IP routers [19].
  - *Advantages: better interoperability and more efficient survivability.*
  - *Disadvantages: optical topology was visible to the client, weakly adapted to non-IP clients, optical layer as a slave and not a server of the client layer, optical layer lost the control of the optical transport plane.*

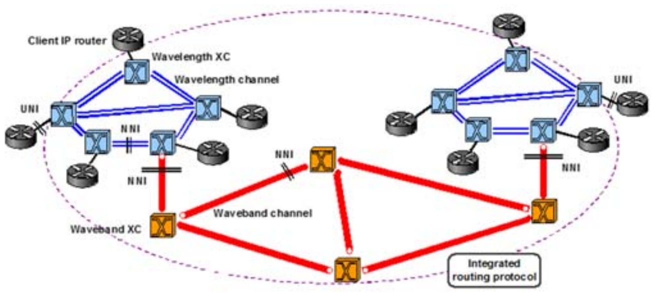


Figure 6. Peer model

- *Augmented Model:* Under the augmented model there were separate routing instances at the IP and optical domains, but certain types of information from one routing instance could be passed through to the other routing instance. This model provided a mechanism for limited information sharing.

The most widely acknowledged NCP for the ASON architectures was the Generalized Multiple Protocol Label Switching (GMPLS) which leveraged the MPLS-TE mechanisms. GMPLS was especially suitable for the envisioned paradigm of IP-over-WDM [18] [20]. Thus, an integrated GMPLS-based multi-layer and multi-domain control plane would allow a more efficient use of the network resources, so that it became the key topic for NCP research in the upcoming years.

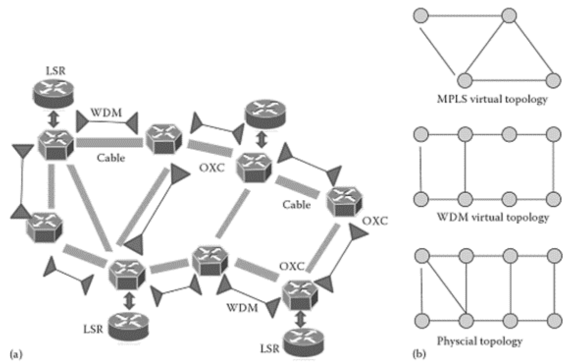


Figure 7. Schema of three layer network where the deployed equipment is shown in the physical topology plot (a) and the virtual topologies at each layer are evidence as graphs (b)

The network architecture presented (Figure 7 a) already indicated the type of complexity that new ASON architecture [21] would have to solve. As a consequence of this complex network architecture (but simple to address with the current state of the art of nowadays technology), three network topologies could be proposed: a physical network topology composed of fibre links, a WDM virtual network composed by WDM systems and OXCa, and a MPLS virtual network topology composed by LSRs with the connectivity offered to them by the underlying transport layer.

This example could be replayed in all multi-layer network architectures, identifying a specific topology (physical or virtual) at each layer of the network. Identifying the correct topology would be a key challenge since in a multi-layer network managed by a single control plan, every routing element would refer for its routing protocols to the topology that was specified at that layer.

The optical network essentially provided point-to-point connectivity between routers in the form of fixed bandwidth lightpaths. The topology of the virtual network interconnecting routers was defined by the collection of lightpaths, which were assumed to be bidirectional [28] and should support some advanced features. The following mechanisms were required to support automated provisioning of lightpaths:

- **Neighbour discovery**: It was the first step towards network wide link state determination. Each optical cross-connect (OXC) had to determine the status of each optical link (up/down) , the bandwidth and other parameters of the link, as well as the identity of the remote end of the link. The determination of these parameters was based on a combination of manual configuration and an automated protocol running between adjacent OXCs. The characteristics of such a protocol would depend on the type of OXCs that are adjacent.
- **Topology discovery**: was the procedure by which the topology and resource state of all the links in a network were determined. This function could be done as part of a link-state routing protocol, such as OSPF or IS-IS, or it could be done via the management plane. The implementation of a link-state protocol within a network meant that the same protocol ran in every OXCs.
- **Protection and restoration models**: There could be local and end-to-end mechanisms for restoration of lightpaths within a network. Local mechanisms used to select an alternate link between two OXCs when a failure occured. When local restoration was not possible, end-to-end restoration had to be performed, which meant the affected lightpath had to be rerouted over an alternate path to avoid the fallen link. Alternate paths could be pre-computed to expedite the recovery time. End-to-end protection could be based on two types of protection scheme.
    - Under “1+1” protection, a back-up path was established for the protected primary path along a physically diverse route. When a failure occurred there was an immediate switch-over to the back-up path.
    - Under shared protection, backpaths would share the same network resources. In case of a failure in the primary path, it was assumed that the same failure would not affect the other primary paths whose back-ups shared resources.
- **Route computation**: The computation of a primary route for a lightpath within an optical network was a constraint-based routing problem. The constraint was mainly the bandwidth required for the lightpath, sometimes along with administrative and policy constraints. The objective of path computation could be to minimize the total capacity required for routing lightpaths.

- **Path establishment**: The signalling protocols for provisioning Label Switched Paths (LSPs) in MPLS, such as Constraint-based Routing using Label Distribution Protocol (CR-LDP) or the ReSerVation Protocol (RSVP) could be adapted for provisioning paths in optical networks.

- **Optical Internetworking**: It had to provide the possibility to dynamically provision and restore lightpaths across optical networks. Therefore:
  - A standard scheme for uniquely identifying lightpath end-points in different networks was needed.
  - A protocol was required for determining reachability of end-points across networks.
  - A standard signalling protocol was required for provisioning lightpaths across networks.
  - A standard procedure was required for the restoration of lightpaths across networks.
  - Support for policies that affected the flow of control information across networks would be required.

3.1.4.2. GMPLS-based control plane

The GMPLS control plane was defined by the IETF and therefore strongly associated with IP-based data networks. In fact, GMPLS inherited IP protocols and concepts. It was the natural evolution of MPLS [22] technology within the optical domain, which in the same way, was designed to improve the efficiency of data networks. So that, with the GMPLS [23[24] technology, MPLS was generalized and extended to cover the circuit-oriented optical switching technologies.

With GMPLS, the introduction of 'intelligence' also in the optical layer opened the possibility to carry out more efficient multi-layer Traffic Engineering strategies and brought new fields of research, which ended up with a very robust, and complex set of implementations. Research around GMPLS started to become a key topic and lasted for a long period. A period in which many advances were produced. However, besides the large effort and investment put on GMPLS, the migration to this technology was never easy, basically because telecom operators wanted to leverage previous existent technologies, and because the cost of equipment and knowledge migration was quite significant. However, and after a few years, and when the technology become more mature and consolidated, many investments where done and different NCP GMPLS based made into operational environments. Then, interoperability between GMPLS implementations become a key topic too. Basically because different proprietary implementations were developed by different vendor's manufacturer, and service providers had heterogeneous vendors on their network that needed to be controlled from a central entity.

Generalized MPLS (GMPLS) differed from traditional MPLS in that it supported multiple types of switching, i.e., the addition of support for TDM, lambda, and fibre (port) switching. The support for the additional types of switching drove GMPLS to extend certain base functions of traditional MPLS and, in some cases, to add new functionalities. These changes and additions impacted basic LSP properties (e.g. how labels were requested and communicated, the unidirectional nature of LSPs, how errors were propagated, and information provided for synchronizing the ingress and egress LSRs). GMPLS not only provided new protocols, it also changed some of the already existing network protocols mostly used in MPLS networks. As an example, the Link Management Protocol [25][13] to handle the negotiations of signalling, routing, link management and Traffic Engineering between the adjacent nodes, was arisen in part as consequence of GMPLS evolution. LMP could separate data and control channels so that each could be protected and accounted separately. It verified physical connectivity on the transport plane, maintained the reliability and the integrity of the network by protecting signalling messages, correlated the link information on the adjacent nodes, helped with link-fault localization and reduced the probability of error in provisioning services. GMPLS also brought important and significant changes into Open Shortest Path First (OSPF) [27] for routing and Resource signalling Protocol (RSVP) [26] for signalling, both with Traffic Engineering extension. They were the most widely used GMPLS protocols. GMPLS key topics were focused on how to facilitate rapid fault detection, fault isolation, and switchover to alternate channels, minimizing network downtime.

GMPLS introduced new functionalities in the transport layer such as resource discovery, Constraint Based Routing (CBR) and connection management. The communication protocols were used to exchange information related to these functions.

GMPLS developments provided a new range of services for optical networks. Below there is a list of the main services that GMPLS delivered.

- Point-to-Point connection provisioning: Operators only needed to signal the ingress node with all the parameters required for establishing a connection, while this node would forward the request through the network using routing and signalling protocols. The whole procedure could be done within seconds instead of hours.
- Bandwidth on demand: It extended the concept of provisioning by allowing the client devices that connected to the optical network and request the connection setup dynamically in real time as needed. The request could be done through a UNI to the

transport network control plane (switched connections) or through an interface to the management system that activated a soft-permanent connection.

- Optical Virtual Private Networks (OVPN): It allowed users to have full network resource control of a defined partition of the carrier network.

Comparatively, in terms of standardisation, ASON was standardised by the ITU-T community which was (is) characterized by a traditional telecommunications networks background. ASON concept was based on a network view based on legacy transport networks, such as SONET/SDH and ATM. IETF defined the GMPLS protocol architecture for the peer, OIF defined the OIF UNI and the external network-network interface (E-NNI), and TMF started the activities on the use cases (it would evolve into the Service Delivery Framework TMF working group – SDF-) (Figure 8, 9)

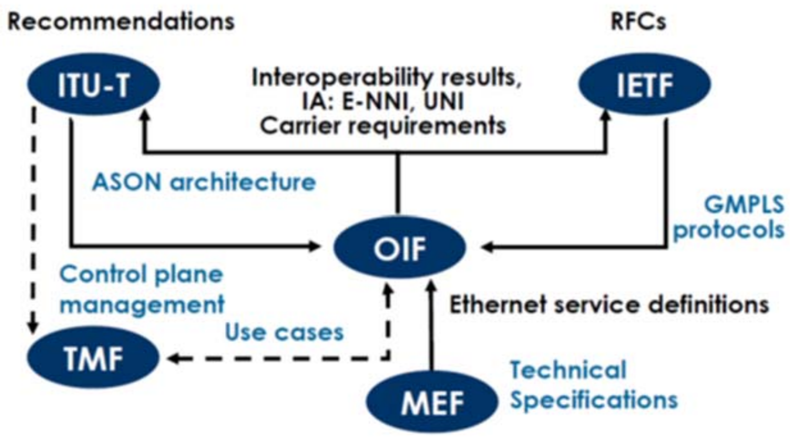


Figure 8. Relationship between standardisation bodies

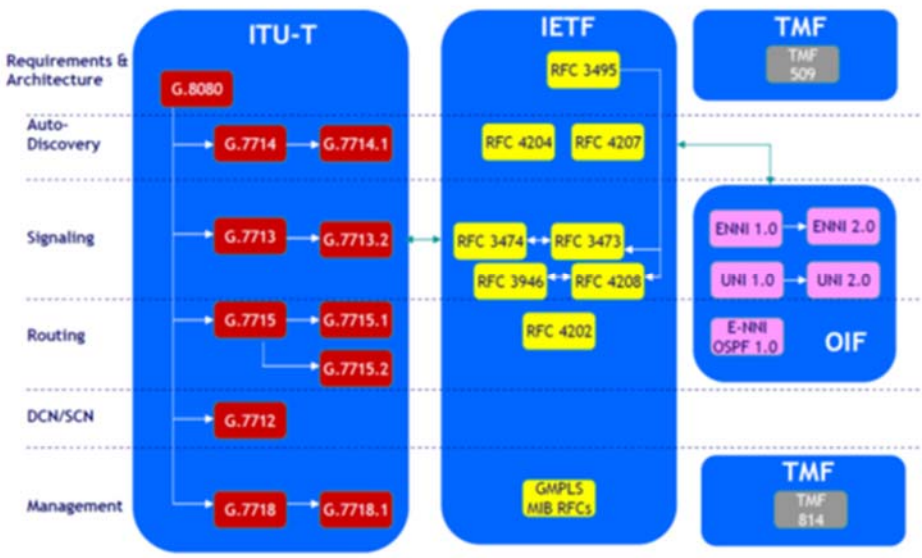


Figure 9. GMPLS initial standards activities

Some of the main advantages that GMPLS brought to the evolution of optical networks were as follows:

- Customers could use different types of traffic (ATM, SDH, IP, Ethernet).
- Accelerated service provisioning.
- Efficient usage of network resources.
- Flexible usage of network resources.
- Reduction of the optical signal distortion by using the concept of Label Set and Explicit Label control.
- Allowed bi-directional LSPs to be set up using a single message exchange.
- Able to transport a wide range of data streams and very large volume of traffic.
- Easier to manage and to scale.

On the other side, the main drawback of IP/GMPLS was its assumption that the same instance of the routing protocol was running in the whole optical network. Thus, there was no way to set up an end-to-end lightpath between independently managed optical networks. The multi-domain provisioning kept for a long time as a crucial research since the interconnection between different providers, while keeping the SLA, was a must [29] [30] [31]. Thus, the optical user network interface (O-UNI) [3], as interface for the user request of on-demand services, became a considerable promise as a client interface to request the setup of an optical circuit or VPN using either GMPLS or ASON/ASTN [32].

During many years, many researchers put lots of effort in the research of ASON/GMPLS, and even though it is largely deployed, it has not evolved as the unique key technology for the provisioning of the new and advanced services we have nowadays. ASON/GMPLS control plane followed the basic principles of transport networks and had large business impact expectations. It supported on-demand services over intelligent optical networks and enabled cost-effective transport of high-speed data. It was the preferred standard based solution for seamless interworking in complex environments, comprising multi-vendor network elements, multi-domain and multi-layer. Another concept within this type of ASON architectures was the administrative domain. The administrative domain was a part of the network that had its own autonomy. This concept was needed to allow the network to hide its own internal data to the overall network, and encapsulate all the internal addressing and topology so that they could not be seen by other administrative domains. Therefore, networks would have different administrative domains [33]. GMPLS enthusiasm started to decrease a bit its popularity around 2013, with the advent of virtualisation techniques, Openflow and SDN/NFV. As said, nowadays GMPL is deployed in many carriers' operators, however, new deployments are in stand by and waiting for the impact of implementing new transport SDN/NFV solutions. Nevertheless, GMPLS has kept as the main control plane for transport optical networks.

3.1.5. Peer models Vs. new peer-to-peer models

Economic models of optical networks were moving to more dynamic architectures, which would generate new research opportunities in the context of next generation networks. These models needed to be reflected into new infrastructure architectures, technologies and methods, enabling powerful new applications and services. As said, optical networks evolved in terms of flexibility [34], providing several types of services based on dynamic provisioned wavelengths or ligthpath (unidirectional path in an optical network with guaranteed resources). However, an analog wavelength was fix in bandwidth and could not be multiplexed, merged or otherwise modified between two points in the network [35]. In parallel to the emergence of GMPLS, and within the research academic network environment (NREN – National research Educational Networks), there were some discussions about the advantages of peer-to-peer models applied to optical networks architectures. The architecture of peer-to-peer optical networking would allow multiple optical network domains equally control the links among them without centralised control and mutually provide transit service to each other based on an open access policy.

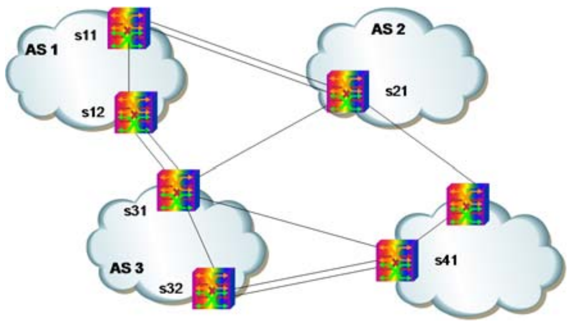


Figure 10. Peer-to-peer optical network architecture.

Thus, most important differences between the peer-to-peer architecture and the peer model based client-server network architecture needed to be clarified [36]. Compared to the client-server architecture, the peer-to-peer architecture had two key features. Firstly, each domain could not only receive transport services from other participants but also contribute with new transport services to other domains; thus, each domain could behave as both, a client and a server. Secondly, a link between two domains was equally controlled by both of them as opposed to being controlled as an access link.

Peer-to-peer optical networking did not aim at replacing the client-server models; it had different applications. Distributed protocols used in centralized models such as GMPLS and O-UNI would play significant roles when thousands of connections and strict service agreements were in place. However, for a small number of connections with large amounts of data, particularly for research and education networks, some other approaches not requiring a centrally managed network could be possible. At that point there were no clear winner since each model had its pros and cons.

The peer model needed a common routing protocol with optics-based extensions to allow routers to compute an end to end path across an optical network. This meant that the IP routers could see the optical network topology, but could only access it through the optical user-network interface (O-UNI). Therefore, there was a permanent client-server relationship between the IP and the optical worlds. This limited access to the optical infrastructure resulted in some lightpath management limitations:

- It was not possible to cross-connect virtual private networks (VPNs) inside a domain, which could make the VPN topology less optimal from the client viewpoint.
- It was not possible to change the topology or the bandwidth without re-signalling. This drawback would not allow centralized networks to support some of the grid computing applications, such as GridFTP.

Security was somehow easy to control, since clients could only access the network through the UNI, and restoration could be done in a quick way using pre-computed paths, both for 1+1 and shared protection. Moreover, an underlying assumption for the peer model was that all the optical networks belonged to the same autonomous system (AS), or that they ran the same instance of the routing protocol. A consequence of this principle was that a client could not signal nor cross-connect a lightpath across its enterprise (or campus) into a choice of multiple carrier optical domains. Although edge IP routers could take a look into the optical network topology, the network was closed to the end user. Big carriers and corporations who controlled the optical networks could take profit of this situation, being able to discriminate against those users and content providers with whom they did not have specific business agreement. Thus, the danger of a centrally managed approach [37] could result in an Internet dominated by a few firms, which could cause:

- Higher user prices, and a higher profitability of the major firms.
- A slowing of innovation and upgrade, as services would only be created by the network operator at the core of the network, and not by the users at the edges.

- Increased power of major Internet firms over:
    - Its governance, standards and protocols.
    - Access by content and application providers.
    - Hardware providers

On the other hand there was the peer-to-peer (p2p) model approach. In such architecture model multiple optical networks could equally control the links among them without centralized control and mutually provide transit service to each other, based in an open access policy. This feature meant an overcome of the main drawback of the peer model: how to do optical internetworking between multiple carrier optical networks. Moreover, this model provided an environment that enabled the end users to create new services and applications, and therefore promote innovation.

Each domain received transport services from other participants, and also contributed with new transport services to other domains. As opposite to the peer model, a temporal client-server relationships was established between the participant domains. Another important consideration in p2p networking was that the availability of network resources was dynamic; domains could dynamically join or withdraw from other domains. This feature added flexibility as well as complexity to the resource discovery and the inter-domain routing protocols which had to be aware of who, and when, was attached to the network. And so, change the routing information in consequence. That's why the inter-domain routing protocol needed to interact with the intra-domain routing protocol (they needed to exchange availability information).

Although there was no central authority, an autonomous blocking control had to be implemented in case there were not enough network resources to satisfy all the requests, and so, prevent "unfriendly" applications from stealing the resources reserved to other applications.

The main driver of p2p optical networking was its capability to create new business models based in open-access policies. Instead of paying the carrier for a transport service, some kind of optical transport services exchange could be expected; in which participants would make their network accessible to other domains to let them access third party domains. This feature was ideal for economic environments looking into foster knowledge transfer, such as universities and research institutions, although not purely for service providers.

The main problem behind peer-to-peer networking was the security, this issue was critical due to the p2p decentralized nature. If the network was open and accessible to other domains, there had to be a clear control on how these domains would access the network, and which usage of network resources would they make. A lot of work had to be done in this area in order to let p2p networking become a reality.

Peer-to-peer model offered a more flexible, scalable and open approach, which would lead to a more competitive optical Internet, with many optical networks service providers cooperating to offer to the client less expensive and better services. Nevertheless, peer-to-peer optical network architectures had a set of functional requirements that needed to be considered:

- **Resource discovery and inter-domain routing protocol:** In peer-to-peer optical networking, domains would dynamically join in or withdraw from a group of domains; a domain would choose to contribute part of its network resources for p2p optical networking while reserving the rest of its network resources for internal use or specific applications. Therefore, the availability of network resources should be dynamic. Peer-to-peer optical networking should provide availability information of channels to be disseminated dynamically.

  Another requirement of an inter-domain routing protocol was the capability of interacting with an intra-domain routing protocol. This interaction should be a two-way process:
  - Firstly, a mechanism would need to be introduced to make inter-domain routing information learnt from outside of a domain, and carried within the domain to related nodes.
  - Secondly, a mechanism would be required to inject intra-domain routing information into inter-domain routing protocols.
- **Symmetric inter-domain signalling mechanism:** A signalling mechanism that establishes, terminates or maintains connections would be required in peer-to-peer optical networking, basically because it operated in a circuit-switched mode. Since under p2p optical networking all the domains could operate both, as clients and as service providers, a new symmetric signalling mechanism would also be required. An interesting proposal for the inter-domain wavelength signalling [38] was the symmetric O-UNI signalling. The basic idea was to build a p2p relation of neighbouring domains based on mutually provided services to each other. When the first domain would initiate a request for a lightpath to the second domain, then they would temporarily form a client-server relationship in which the first domain would be the client. When the second domain would initiate (or carries from other domains) another lightpath request to the first domain, the first domain would act as a server to provide transport services.

- **Autonomous blocking control:** A distributed blocking control mechanism would be required to co-ordinate the requests on shared network resources. Because every domain had equal authority, the blocking control mechanism should operate in an autonomous mode. There were three possible applicable strategies to the design of the blocking control:
    - When there are sufficient network resources to fulfil all potential requests, no blocking control mechanism would be required.
    - When network resources could not satisfy all requests, blocking control should be introduced, and when blocking would happen in a link, all the connections that would go through this link should back off their usage on this link, after proper notification.
    - Every domain should audit and log network operations from its local point of view. If any significant misbehaviour would happen, a central committee could issue appropriate punishments.
- **Data plane or physical layer internetworking:** All peer domains should be considered equals, but they might not share the same capability in the data plane or physical layer. Therefore the control information carried in either inter-domain routing protocols or inter-domain signalling protocols or both should be enhanced.

3.1.6. Conclusions and next evolution

During the expansion moment of GMPLS, and after considering some of its advantages at the architectural and service level, new types of architectures started to be addressed. They were based on application-to-network interaction for the provisioning of automatically switched connectivity supported by Network Control Plane (NCP)-enabled transport networks (e.g. GMPLS networks). The Service Oriented Optical Network (SOON) become the first fully-distributed service architecture expressly conceived for NCP-enabled transport networks. SOON architectures had a service layer (Service Plane - SP-) that translated application’s service request, expressed in terms of QoS and resources addresses, in a set of CP directives at the boundary of the transport network. This SP facilitated the decoupling of network technologies from future evolution of the network services, so that, the NCP was freed from service oriented functionalities and could focus on the provisioning of connectivity services (e.g. LSP creation). The emergence of the so called Service Planes allowed a new generation of applications such as Grid and Storage on-Demand to move from ad-hoc scenarios to transport network where the technology details were usually hidden to third party application providers.

Moreover, infrastructure and resource virtualisation started to move forward and to extend its boundaries from the IT technology scenario to the optical networks domains. With it. A new approaches to virtualise the optical network emerged under the concept of Customer Empowered Networks (which considered the peer-to-peer models), or Managed and controlled Networks were sorted out with the implementation of the User Controlled Lighpath Provisioning (UCLP) architecture. Both of the novel concepts, SOON and CEN, definitely brought a new era of architectures and so the starting point of the research activities carried out along this Thesis.

At present, Infrastructure virtualisation techniques for SON and network orchestrators, together with the advent of SDN and NVF, are the main research trend of the world class research community.

3.2. Section 2: Service Oriented Networks and Customer Empowered Networks

As shown along section 1 of chapter 3, there were several approaches to implement different architectures and protocols for the management and control of optical network. However, and based on the work developed on this thesis, two main approached were considered as main interests. Which were the starting point of the research activities. Service Oriented Networks were a clear trend and an open topic for new types of research. SON architectures would allow applications direct access to IT resources with storage or computational capabilities that could be located locally or distributed throughout the network, in data centers and clusters of servers.

At the same time, and considering that ASON/ASTN and GMPLS technologies were well suited to traditional centrally managed hierarchical networks in telecom wide area network environments, a new type of wide area network architecture emerged following the SON trend. It was called Customer Empowered Networks (CEN) or Customer Controlled and Managed Networks. This approached, based on the SON premises, and were the network could be integrated with other web service applications, aimed at bringing network virtualization down to the optical layer. It started to become of interest among large enterprise networks, university research networks, and government departments. Customer-controlled and -managed networks concepts were radically different from the traditional centrally managed networks in that the enterprise not only would manage and control its own internal local area or campus network, but also control and manage its own wide area optical network. As a consequence, traditional management and hierarchical optical network technologies, which were premised on central provisioning of optical VPNs to customers, were largely unsuitable for customer management of their own optical network.

The challenge was on developing an architecture to implement a production level optical transit network to enable users to directly connect to each other by setting up LightPaths (LP) and developing software tools and reference models to allow users to provision and manage LPs themselves. The CEN approach was born in Canada, under the research activities of the CANARIE Inc. research organization (CANARIE is the Canadian Network and Research Educational Network).

The following sections will give an overview of the SON and CEN state of the art during the first years of the Thesis studies. These architectures, and its associated technologies and services, evolved together with the research activities developed along the Thesis. In fact, these technologies have been very important for the development of the research outcomes achieved. These outcomes supposed a step beyond in terms of architectures and technologies from those already available and discussed in section 1 of chapter 3 and represented in Phase I of Figure 2.

In fact, some of the main research contributions from the Thesis are based on architectural designs that are a merge of both concepts (SON and CEN), providing a new approach for the provisioning and management of optical lighpaths across different domains. In some way, it provided a solution towards the multi-domain approach, which was already discussed across many research papers and forums. This Thesis worked on a solution that decoupled the optical network (data plane) from the control plane. Thus, dealing with different administrative domains provided different capabilities to the end users. This concept has currently evolved towards current Software Defined Networks with Network Functions Virtualisation.

3.2.1. Sliceable research networks

At this point in time it was already clear that the original Internet architecture needed to improve in terms of security, management, monitoring and mobility, among many others. Internet evolution was very fast, almost without time to comprehensively think about architectural designs. Many open discussion were happening within the research community. These discussions had two different positions with regards to the evolution of future internet. In one side there was a strong position towards a revolutionary approach (clean slate), while on the other side the position was on an evolutionary approach (incremental). Both of them promoted different types of research while targeting the same goal, provide a more robust, flexible and dynamic architecture. In any case, future internet needed to expand its borders from the physical layer to the applications layer while improving its performance globally.

New research activities were developed around the world with different approaches for novel architectures and protocols aiming at innovation within the core of internet. Pioneering clean slate initiatives such as Future Internet Design (FIND) [39], Global Environment for Network Innovations (GENI) [40], and Stanford University's "Clean Slate Design for the Internet" in the U.S. brought new areas for innovations within the scope of Internet of the future. These areas included: addressing and identification, cross-layer design, network virtualization, routing and traffic engineering, dynamic switching of optical circuits, decoupling of control and data, service discovery and composition, as well as management. While in Europe, the European Commission launched the Future Internet Research and Experimentation (FIRE) [41] initiative in its 7th Framework Programme (FP7) based on early experimentation and testing in large-scale environments similar to the Internet.

An example of this initiative was the FEDERICA [42] European research Project which provided a physical infrastructure that could be sliced in independent parallel networks controlled by the network researchers. It allowed researchers to deploy, test and validate new Internet architectures and protocols in their slice without impacting the experiments carried out in other slices. Actually, some of this Thesis research activities leveraged this research project too, by enhancing previous work on optical virtualization. It will be explained along Part II.

In parallel, other initiatives dealt with techniques for virtualization in computer science. PlanetLab [43], was created as a global research network that supported the development of planetary-scale network services. It emerged from an IT community that set up a testbed in 2002 using virtualization as main principle. The GENI concept addressed these shortcomings of PlanetLab and extended the aforementioned virtualization principle to the whole infrastructure with the notions of substrate, slicing and federation. The research challenges behind these topics were based on slicing up the network in small networks controlled by different users, following the CEN approach, so that one or more network substrate providers could select some unused resources of their physical network, add them to a slice and assign the control of the slice to another organization. This is equivalent to the condominium model, where the administrative owner assigned the control of different wavelengths of the network to the users that have acquired them. Another feature was the federation of slices from independent providers; i.e.: create a bigger slice from two different network provider slices. Again, this was equivalent to the CEN scenario, where the user of the infrastructure could get resources (dark fiber, network equipment, wavelengths) from different providers and integrate all of them into a single management domain.

3.2.2. Service Oriented Architectures and Grid services

Together with the growth of Internet, a new generation of data intensive and scientific applications started to emerge too. These applications dealt with large amounts of data that could only be provisioned by means of high-speed optical networks. Many of these sever applications had requirements based on specific constrains: determinism (e.g. guaranteed QoS), shared data spaces, large transfer of data, and latency requirements that are often achievable only through dedicated optical bandwidth (lambdas). The advent of high capacity optical networking provided the raw capacity to carry vast amounts of data, but software tools and frameworks, addressing end-to-end user and application-level access, as well as provisioning-on-demand of such bandwidth needed to be developed in coordination with other resources, such as CPU and storage.

Thus, this rapid usage and deployment of Internet directly impacted network architectures. Service Oriented Networks [44] become a real need. As said, these were networks that offered applications direct access to IT resources with storage and computations capabilities located locally or distributed across the network. Being one of the main requirements bandwidth capacity and on demand provisioning of the infrastructure and its resources, together with a higher degree of intelligence and control. In order to efficiently manage the resources why these applications should be able to configure the network in the way it better suited their needs. Actually, this was the reason these architectures were called Service Oriented Architectures. The concept of Service Oriented Optical Networks brought closer Service and Optical layers in order to reduce drawbacks and facilitate the interoperation between them.

Highly-dynamic, data intensive and e-science applications required access to the following to maximize scientific discovery:

- Application-level middleware providing the execution environment of generic high demanding applications including all necessary service abstractions exposing advanced Grid-like functionalities combined with network services.
- A new generation of management and control planes with strong interaction with near-real-time resource (compute and network) scheduling, allocation and reservation.

The fact of provisioning on-demand network resources across multiple administrative and network technology domains became a key enabler for emerging Grid services of that time. With that, new complex service connections would be needed, since available point-to-point services were insufficient for covering the point-to-multipoint or multipoint-to-multipoint connection services needs of Grid applications. Actually, asymmetry or bi-directionality became key too.

During the Thesis evolution there were a tremendous research effort and new development in the Grid community in terms of Grid services infrastructure and Grid application development. However, there was very little work done in the area of using network as a first-class Grid resource, and so coordinating the provisioning of both resources. In fact, besides some of the work done behind UCLP, there were no existing implementations that could demonstrate the power of exploiting the optical network as a first-class Grid resource and the challenges that would arise in provisioning end-to-end light-paths across different management and control plane technologies spanning multiple administrative domains.

Therefore, the research community started to focus their efforts on service-centric infrastructures supporting the deployment of mission-critical applications on a global scale. Said in another way, architectures that could significantly enhance the capability of data intensive and e-science applications, providing a unified network/Grid infrastructure that could flexibly adapt to the strict demands of applications, combining requirements on CPU, memory and storage resources as well as on the communication network. Therefore, these architecture should provide the applications the capability to rely on a network infrastructure that could be adapted to the application, rather than having the application adapting to the network, as proposed in many research efforts of the time (i.e. EU-funded projects like MUPBED, EU-QOS) [45][46].

Grid applications were usually compute and/or data-intensive applications executed in a grid environment. Grid technology allowed the secure and reliable sharing of distributed resources (typically computing power, memory, and storage) in a heterogeneous environment divided into multiple administrative domains. Grid applications often requested a number of resources to the Grid, which used a metascheduler to find a suitable set of resources that would match the application's requirements, and so reserve them in advance. At the time that a requested set of resources were available, the application was executed in a distributed fashion (using all the computing nodes that the application would have previously requested).

Grid technologies allowed the division of huge tasks into smaller jobs to process them separately in remote locations and then reassemble them to obtain the resulting information. This idea was applied in general to any IT resource such as storage, computing, processing or any collaborative or distributed service. A grid application used services and functions defined by OGSA [47] and specified by OGSI [48] along with

Grid infrastructure to accomplish specific work-related tasks that solved business and technical problems.

Increasingly more and more grid applications had to move large amounts of data between storage and computing resources, therefore they required high bandwidth connectivity between the sites contributing resources to the Grid. To let the Grid efficiently manage the network resources (i.e. only setup the connections that were required by the applications that was currently being executed), network resource availability information and network resource reservation had to be part of the Grid management infrastructure. To achieve this goal, the network interface had to be able to provide availability information and also had to be able to reserve network resources in advance (such as dedicated bandwidth pipes between endpoints). Thus, a new service layer was needed at the architecture level.

3.2.3. Customer empowered networks

As previously commented, with customer-controlled and -managed networks the organisations not only were able to manage and control its own internal local area or campus network, but also control and manage its own wide area optical network, assuming responsibility for direct peering and interconnection with other like-minded networks. There were two main types of CEN, the customer controlled [49] and managed dark fibres networks and the wavelengths networks.

The first referred to those organisations which were acquiring their own metro dark fiber (schools, hopitals, and governments). These institution participated in what was called 'condominium' [50] dark-fiber networks, so they could better manage and control their connectivity and bandwidth requirements. The advantage of customer-owned metro dark-fiber networks was that traditional "dollars per megabit" business model for bandwidth was largely replaced by the much lower cost for the one-time capital cost for the dark fiber and initial equipment outlay. Thereafter, any increase in bandwidth only required a simple equipment upgrade. Customers could take advantage of the inexpensive metro Gigabit Ethernet and 10-Gb Ethernet equipment for lighting up the fiber.

The second, the customer owned managed wavelengths networks, came due to the fact of availability of long-haul dark fiber and the dramatic drop in the costs of long-haul optical equipment which allowed large corporations and a number of research networks to deploy their own long-haul optical network. Many carriers started to sell or lease point-to-point wavelength services to large enterprise and university research networks. A good example of this model was the Canadian national research network CANARIE's

CA*net 4, which purchased point-to-point wavelengths from three separate carriers. The wavelengths terminated on CANARIE- owned and –operated optical add–drop and cross-connect equipment at various nodes across Canada. Thus, members of condominiums wanted to independently manage their own optical add–drop multiplexing (OADM), optical cross-connect (OXC) to other clients, and offer optical VPN services to third parties. This was the main motivation behind the development of the UCLP [51]. Basically, CEN allowed a reduction of bandwidth costs by means of participating into condominiums [52], and an indirect cost saving thanks to the internet reduced costs due to remote peering and transit. Customer control of the cross connect allowed the user to change the peering relationship without having to contact a central management body or pay expensive Internet transit fees. In fact, customer-controlled and -managed networks also provided significant technical advantages, particularly in support of end-to-end (e2e) lightpaths and Quality of Service (QoS) for large file transfer, storage area networks (SANs), and the nascent grid services [53] and SON. These applications required substantial bandwidth links, in the order of gigabits, that needed to be provisioned rapidly across multiple independently managed optical networks. To that date, only few commercial carriers offered intra-network optical VPN services with such capacity, and even fewer, if any, offered this capability across multiple independently managed networks.

Some challenges behind this concept were as follows [54]:

- Only users (customers) and no single provider had full visibility of their network and could see all the network elements, besides GMPLS and ASTN/ASON, where the carrier within its management domain had full visibility of all network elements and a common interface to them.
- Current inter-domain service models assumed a multiple independent network models of carrier-to-carrier signalling serving as proxy for customer request, rather a customer at the edge negotiating directly with the separate independently managed networks
- These type of networks require common equipment for the optical links across the network, and it was not either practical nor cost effective to have independent optical repeaters, ROADM's and OXC for each separated customer owned wavelength, so new tools and services need to be provided to allow them to manage their own restoral and protections schemas and independently provide optical VPN services.

It is important to notice that the challenge of customer owning and managing their own networks was very similar to the challenge of independently managed High Performance Computers (HPC) [55] or network storage systems who wished to share their resources with a community of interest of users. The concept of 'grid' system, where these services could be advertised, discovered and consumed by users in a particular community, had a clear convergence with the management of next generation optical internet.

3.2.3.1. UCLP

The UCLP approach started its discussion based on the concept of administrative domains and the different implications, or understandings, it could have into optical network. The premises UCLP had behind the administrative domain capability were based on the fact of applying virtualisation and partitioning functionalities at the optical network. It proposed the idea to create an abstraction model of the optical nodes in order to create different administrative domains for the same infrastructure, and so, allow the users to manage their own virtual infrastructures as they would wish. It was clearly empowering the user capabilities at the optical domain level as no one approach did before. The major goal of this architecture was on providing to the customer with capacities to set-up its own lightpath along a network and through different administrative domains using a distributive control plane. This idea of a distributed control plane was based on the peer-to-peer networking model. In such a model, all the nodes (also called peers) equally controlled the links among them and provided transit service to each other.

UCLP (user ControllerLightPaths) provisioning system was the name given to the system implementing the CEN approach. The UCLPv1 Research Program was put in place through an RFP (Request For Proposals) by CANARIE (a non-for-profit organization founded by the Canadian government to promote the development of advanced networking to enable new internet based applications) in September 2002. Four implementation proposals from four different teams were accepted, one of them being led by the Communications Research Centre (CRC), with the University of Ottawa (UofO) as a partner. The Grup de Comunicacions Òptiques (GCO) of the Universitat Politècnica de Catalunya (UPC) and the Fundació i2cat joined this team in April 2004.

The UCLP Software allowed a physical network to be partitioned in several independent management domains, and expose the network resources belonging to each partition as software objects or services. These objects could be put under the control of different users so that they could create their own IP Network topologies.

UCLP allowed the end users (being humans or sophisticated applications) to create their own discipline or application specific IP network. For instance, a community of high energy physicists could take a subset of the resources of several optical networks and create their own network, whose topology and architecture would be optimized to better fit their specific applications or service needs. They could also reconfigure their network any time they wanted without the intervention of the physical network owners, hence the name of User Controlled LightPaths. Although the concept could be applied to any type of Information Technology (IT) infrastructure, the original focus was the user control of optical networks and specifically the setup of multi-domain end to end lightpaths.

The motivation behind this new approach came due to the fact that many researchers, carriers or even manufacturers, started to explore alternate network architectures that could obviate some of the short comings of Internet and provide a more robust and secure infrastructure. This architecture concept made more sense for the scientific community and academic research networks, which needed a huge volume of data to be transferred and certain degree of network control by the user (i.e. universities, schools, research centers and governmental or local institutions).

This included building virtual private IP networks, deploying centrally managed optical overlay networks, binding applications to the transport network and so forth. Therefore, most of these new proposed architectures eliminated or severely restricted many of the critical features such as the Internet end to end principles [56] that made the Internet we know today so innovative and successful.

The needs of the research and scientific community were very different compared to the Internet ones, were IP Networks were optimized for thousands of users with relatively small traffic flows. Thus, worldwide scientific distributed communities at large needed to exchange high volumes of data (grid applications, e-science community, sensor networks). These communities were not able to use the public Internet to conduct their experiments; they required dedicated IP Networks where the routers or the computers where the applications were executed had end to end dedicated links. These dedicated links needed to cross the networks of different providers without experimenting quality of service (QoS) degradation, which was possible using network protocols designed for a single administrative domain. An example of these communities using application with high bandwidth requirements was the US/Canada NEPTUNE [57] project (an undersea sensor network).

Thus, another key topic would deal with the QoS, by assuring a guaranteed bandwidth for large data file transfer across different administrative domains. And so, allow the customer to independently change VPN topology and bandwidth of their connections, cross connect VPNs from different autonomous customers within a carrier cloud, support the setup of inter-domain optical services and support peer to peer management and control as is done in the internet. Although client server with GMPLS and O-UNI protocols fitted well in models where there were thousands of connections and strict service agreement in place, the UCLP concept was focused in solutions for research and education networks, where there were a small number of connections without a centrally managed network solution.

By then, optical networks did not allow the implementation or even the evolution to a peer-to-peer optical network approach, since optical devices with more than one management interface or control plane were not feasible, so that that equipment could only be operated from one management entity. Therefore the goal consisted in dividing the rights of us of the available resources provided by an optical device through all the customers that had rights over the device, allowing them the possibility to control and manage their resources, and even more to share available resources among them. The real challenge was on the capability to provide dynamic partitioning of a resource and give rights of this resource to different customers. There were different possible ways to do so, even though more research would be needed. This research was based on the development of new drivers to the management and control software and implementing it in a real optical network. With the peer-to-peer architecture the telecoms operators' model was somehow transformed. They could also act as infrastructure providers. Thus, the user would decide to which service provider connect in order to establish a peering connection. In peer-to-peer optical networking, multiple optical network domains equally controlled the links among them, and mutually provided transit service to each other based on an open based policy. This was a relationship that widely existed in the Internet. Then, the collaboration among multiple independent customers without the need of a coordination centralized management would be somehow solved. Although mechanisms for policy enforcement, authorization and authentication had to be developed and applied. In particular, the UCLP architecture was based on the creation of articulated private optical networks[58], and so, it was the first initiative that put the emphasis of virtualisation at the optical transport network, as an alternative to NCP (e.g. GMPLS).

Another challenge of the UCLP model was the use of new extension for the BGP protocol. Inter-domain network architectures of internet were well represented by the applications of the inter-domain routing [59] protocol Border Gateway Protocol (BGP). So, noticing that its power lied on its attributes and its route filtering techniques, attributes were simply parameters that could be modified to affect a decision, and route filtering could be done on a prefix level or a path level. Thus, some implementations of

BGP evolving to an Optical BGP (OBGP) [60] served for the UCLP purposes. Actually, the OBGP extension was proposed for the manipulation of OXCs and to allow them to be automatically setup and configured as BGP speaking devices to support multiple direct optical lightpaths between many different autonomous domains. OBGP allowed customers at the edge to control a subset of lightpaths within another network's wavelength cloud so that they could manage their own lightpath routing within that cloud.

Therefore, UCLP allowed the creation of new BGP paths between two routers through the establishment of a dedicated optical direct links. This way large institutions and regional networks could directly connect to each other through a peering relationship instead of exchanging traffic through a hierarchical IP Network. Figure 11 depicts this scenario.

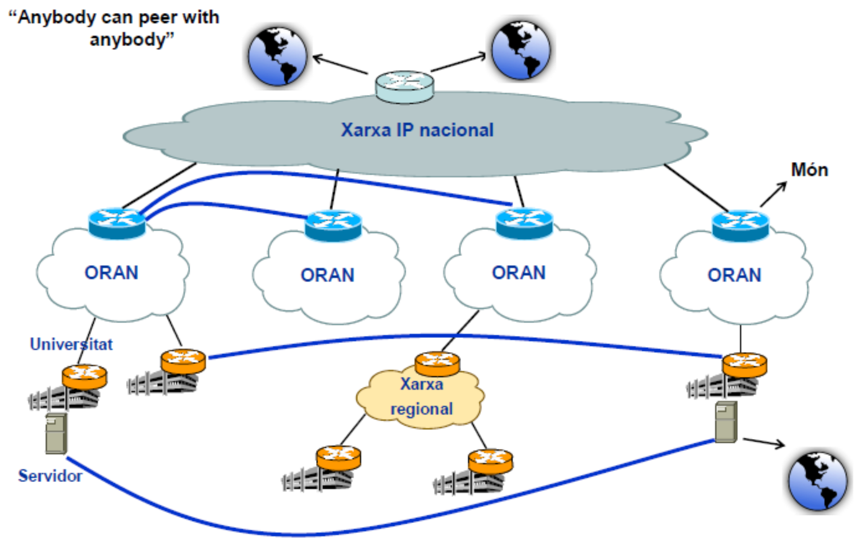


Figure 11. Universities and regional networks peer with each other directly instead of doing it through its transit provider (the national IP Network)

The first implementation of UCLP, called UCLPv1, was a software solution that allowed the users to control and manage their own optical network elements to establish end-to-end connections through optical networks belonging to different administrative domains. The software was based on a service oriented architecture (SOA) implemented with Grid (Globus Toolkit 3) and Jini technologies [61].

This approach provided a new set of optical network management tools and protocols. These tools and protocols, based on the Open Grid Service Architecture (OGSA) and other space based distributed protocols such as Jini and Javaspaces, allowed end users to independently manage their own portion of a condominium wide area optical network and control their own optical resources available in any autonomous system or cross-connect device. Thus, customers could manage their own restoral and protection schemes, optical add drop multiplexing or cross-connect to users on a peer-to-peer basis without signaling or requesting services from a centrally managed entity.

The UCLPv1 System was divided in federations. A federation was a logical partition composed by heterogeneous network resources of different technologies (WDM, SDH, Ethernet, GMPLS) that were under the control of a single organization. Each federation was an independent management domain with its own set of UCLP services. UCLP virtualized the optical network elements, in the sense that it abstracted all the optical network resources (interfaces and links) as software objects. This architecture faced a number of technical challenges in order to manage networks with resources from different sources and co-ordinate the protection and restoration involving multiple providers.

The use of grid technology, together with web services, allowed non-traditional telecommunications organizations, who had acquired their own dark fiber or wavelengths (i.e. universities, regional networks and large enterprises), to advertise, discover or interconnect lightpaths as a service between themselves. Those services could be signaled over a multi-domain network through the use of web services. Web Services were self-contained, self-describing, and could be published, located, and invoked across the Web. Web services performed several type of functions, from simple requests to complicated business processes [62]. In that sense, Web services were not only viewed as the new middleware for network management, they were also used as a signaling approach to deploy network services over multiple domains like VPN, Ligthpath signaling, VoIP services, VLAN services, SAN services, etc.

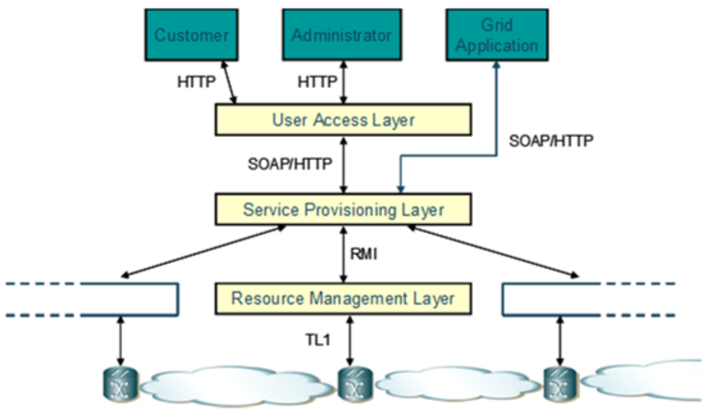


Figure 12. Control and management plane for a peer-to-peer architecture with a web services model

They aimed at providing a systematic and extensible framework for application-to-application interaction. Therefore, another main features of using Web Services for Customer controlled networks was the cross-connection associated with lightpaths, and partitioning the OXC (optical cross connect) or an OADM (optical add and drop) of a network into multiple separate control and signaling planes. A UCLP model for the peer-to-peer approach under a web service solution was already available (figure 12).

An important development for UCLP was building the model for controlling an OXC or OADM (ROADMs were on its initial phase). This model was based on the design proposed in the Figure 13. Control plane partition to manage an OXC (named by CA*net 4 a Mini-IX)Figure 13, where each partition would be managed and controlled by individual users willing to connect and link their lightpaths. So, it would allow the users to operate its own grooming, routing and discovery protocols independent of the routing and discovery protocols operated by the other network connected to the switch (OXC or OADM). In that model, all the services such as grooming, alarms signalling and control (different control planes architectures would fit within the model) were rendered as web services within the defined web resource.

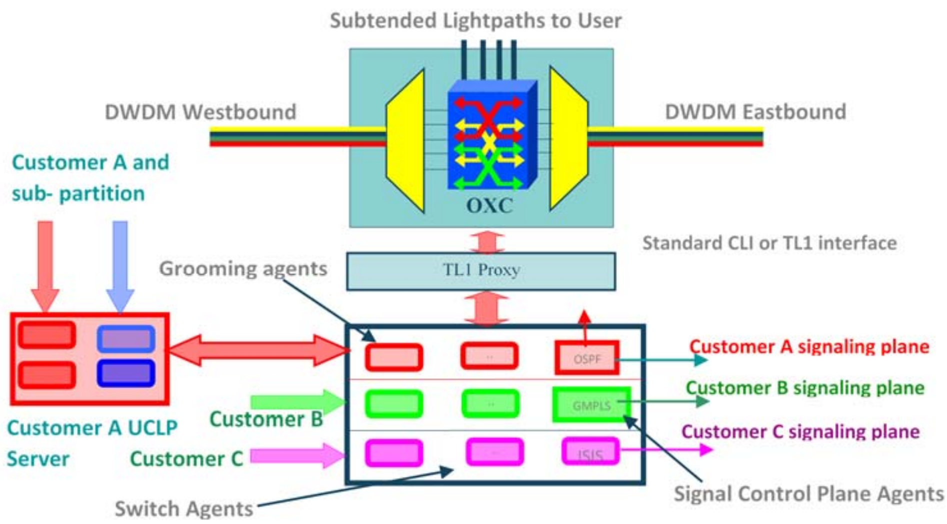


Figure 13. Control plane partition to manage an OXC (named by CA*net 4 a Mini-IX)

At that time, and considering the state of the art of grid technology and software tools to build SON, the software Jini/Javaspaced architecture to implement the proposed model for the control and management plane (figure 14) provided some key advantages. Jini ran on top java and used Remote Method Invocation (RMI) to access remote services. Jini also had the Jini Lookup Services (JLS), a distributed service registry that allowed users to find services without having to know anything about where they were located, thus through a federations of JLS’s a user could find any service in any domain. Then, Javaspaces provided a distributed data store for java. This solution provided more functionalities for storing Lightpath objects in a distributed fashion, and was more powerful and mature than XML (web services) while used RMI to pass java objects instead a XML schema definition (XML would evolved as key technology for management systems, as it is in today's state of the art). A drawback was the fact that it was limited to Java languages. XML was still an independent language. However, Jini internal service calls were transparent to the user and other applications that would use OGSI (Open Grid Service Interface) provided by the GSAP (grid Service Access Point). Figure 15 presents the Service Orietned architecture for UCLPv1.

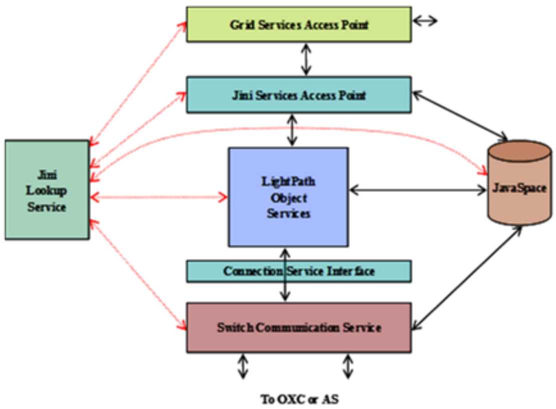


Figure 14. Jini and Javaspaces proposed model for the control and management plane

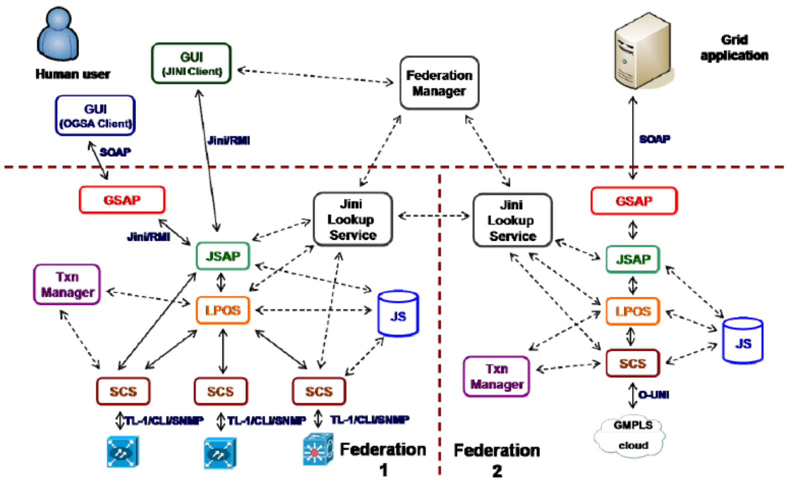


Figure 15. UCLPv1 Service Oriented Architecture

Part II

Chapter 4. Introduction

Chapter 5. Thesis challenges

Chapter 6. Articulated Private Networks

Chapter 7. Service Plane for resource virtualisation and service provisioning

Chapter 8. A Generalized Architecture for Dynamic Infrastructure Services

Part II

4. Introduction

This second part of the Thesis will cover and present the research activities developed along the whole research period, together with its associated research outcomes. These outcomes have already been published in different journals and conference papers, which validates the impact of the work done. The Thesis research activities started with the emergence of CEN and SON around 2005-2006. Thus, the following chapters will allow the reader to understand how the research activities performed impacted the continuous evolution of the state of the art presented in section Part I.

As said, Part II will expose and present the work done on defining new architectures and services for Optical Networks while empowering the user on its control, together with different solutions and architectures to cope with the emerging needs of Internet, Grid and Cloud applications. In fact, this section is divided into different chapters that will cover the three main topics of the research done. As commented on the evolution roadmap from Part I, the research developed shows an architectural evolution. It first starts with an architecture for Network Resource Provisioning System, based on the CEN principles, which allowed the virtualization of the optical networks devices and the creation of Articulated Private Networks. Then it evolves into the design of a SON architecture of a Network Service Plan for the provisioning of multi-domain services to support Grid applications. This service plane layer facilitated the virtualization of optical network infrastructures and so the provisioning of inter-domain services with specific requirements. And it ends up with the design of a whole architecture to make optical networks become cloud enablers, with the provisioning of virtual infrastructures managed by advanced control planes that together facilitated the capability to map down, to the virtual and physical infrastructure, the cloud applications SLAs. So, it meant the first architecture capable to provision optical networks and IT resources in one single step. This architecture brought many new advanced functionalities to the community since it was one of the first cloud enablers for optical networks. This architecture enabled the paradigm of SON in support of cloud applications.

5. Thesis challenges and research contributions

5.1. challenges

The objectives and challenges of this Thesis were to investigate and design new architectures and service layers for the delivery of advances services for Service Oriented Networks. These advanced services had to fulfil the following requirements:

a) Allow an infrastructure provider to virtualise, abstract and partition his physical infrastructures in several independent virtual infrastructures, segments or slices.
b) Allow several infrastructure providers to share their resources and create virtual infrastructures composed of resources from different infrastructure providers.
c) Allow the virtualisation and partitioning of an optical network.
d) Allow the provisioning of virtual resources under different advance reservation schemas and dynamic on-demand provisioning services to end user.
e) Allow third parties to provide e2e services across different domains
f) Create infrastructure applications aware with a well-defined set of API's so that users can customize the infrastructure as a service.
g) Allow third parties to deploy their network control planes over virtual infrastructures composed of physical infrastructures not belonging to them and on-demand and dynamic provisioning of virtual infrastructures (Net+IT)
h) Provide a cloud enabler infrastructure for optical networks.
i) Reduce the SON provisioning set-up time from days to minutes.

5.2. Research contributions

The work presented along the following chapters fulfilled the requirements and challenges described in section 5.1. Actually, the table below establishes a link between the research outcomes presented along the following chapters, and the requirements and needs they covered.

Research Outcome	**Activities done within this Thesis**	**Requirement covered**	**Chapter**
A Service Oriented architecture for Virtual Networks and an enhanced Network Resource Provisioning Service (NRPS)	Researcher in the architecture design of the NRPS-IaaS	a), b) c)	6

A virtual optical Network Service Plane in support of future internet and Grid services. A new architecture with a new Network Resource Provisioning Service (NRPS)	Proposer of the NSP idea and architecture and participation on the NSP design	c), d), e)	7
An architecture for Network and IT convergence	The initiator of the vision and one of the main designers of the whole architecture	f), g), h), i)	8

Table 2. Relationship between research outcomes, Thesis author involvement and requirements achieved

6. Articulated Private Networks

Due to the evolution of technology and the increasing demands of Internet and grid applications, the aforementioned UCLPv1 evolved towards a more robust and flexible architecture which produced a beta-quality product that could be deployed in a production network. Being this new architecture the UCLPv2. User Controlled LightPaths version2 (UCLPv2) provided an alternate future Internet architecture that enabled users to define their own packet- or switched-based network architecture including topology, routing, virtual routers, switches, virtual machines, and protocols, based on the concept of many separate, concurrent, and independently-managed articulated private networks (APNs) operating on top of one or more network physical substrates across different ownership domains (

Figure 16).

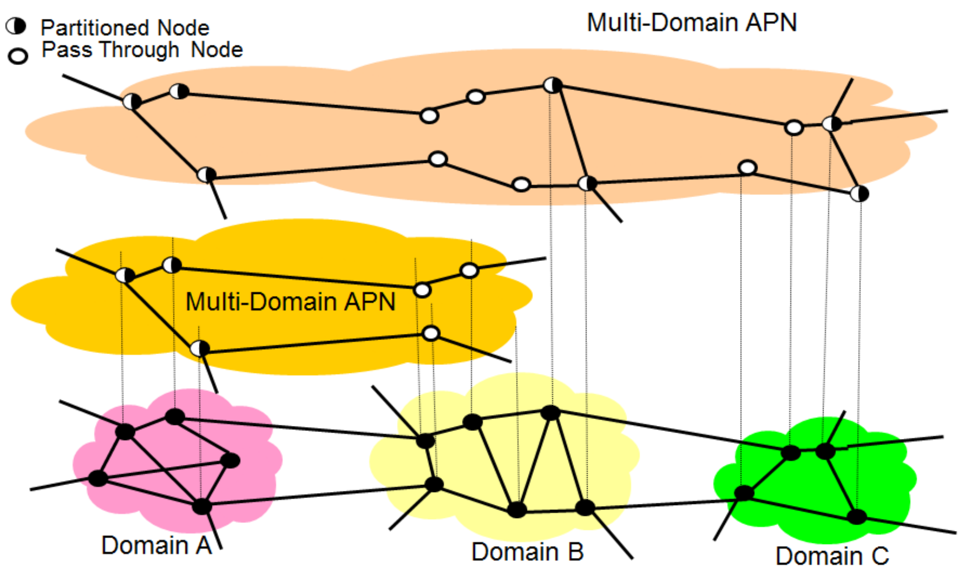


Figure 16. APN Multi-domain APN

APNs were considered as a next generation virtual private network (VPN), where a user could create a complex multi-domain network topology by binding together layer 1 through layer 3 network links, computers, time slices, and virtual or real routing and/or switching nodes. This UCLPv2 capability was realized by representing all such network elements, devices, and links as Web services and by using Web services workflow as the tool to enable users to bind together their various Web services to create a long-lived APN instantiation [51]. Thus a new way of manage networks started to emerge.

6.1. Strategies in network management

Network and service management started to become an important topic involving a whole set of evolutionary and novel approaches. Research activities provided scalable and federated platforms based on various management approaches while applying the infrastructure virtualization for network slicing as a key topic.

The functional requirements of traditional network management were already summarized by the well-known term fault, configuration, accounting, performance, and security (FCAPS). However, during the thesis period, the importance of configuration management as well as network monitoring and measurement (i.e., fault and performance management) increased. While accounting and authorization (i.e., security management) was also becoming more and more important in virtual environment.

Classical management frameworks such was the open systems interconnection (OSI) network management framework, telecommunication management network (TMN), or Internet network management framework were following the traditional agent-manager centralized paradigm. As networks grew, management complexity and service requirements made those paradigms no longer adequate and should be replaced with distributed management concepts.

That was clearly illustrated by the evolution of the Simple Network Management Protocol (SNMP) for Internet Protocol (IP) networks. As an early effort in 1992, SNMPv2 introduced the concept of an intermediary manager for distributing management functions. Although SNMPv3 with embedded security features was widely deployed for monitoring network devices, it had some shortcomings in configuration management and managing complex systems [63], including virtualization-capable network devices, servers, and PCs.

The new generation of routers and switches become highly programmable devices and could accommodate management functions directly on the device. This design trend led to new management approaches based on Extensible Markup Language (XML) and Web services, as well as some vendor-specific solutions. These trends were not called revolutionary because they had been around for several years and did not really change the network administrators’ way of thinking. Since XML could be efficiently used as a textual encoding mechanism for application protocols, and object and interface specifications, the XML-based management technologies seemed the obvious choice for the next evolutionary step. In parallel with some proprietary developments like Juniper’s

JUNOScript [64] the benefits of the combined XML and Web-based approaches became clear. Such low-cost infrastructure with proven scalability and security features was developed by the Distributed Management Task Force (DMTF) and called Web-Based Enterprise Management (WBEM). The ideas behind JUNOScript and WBEM formed the basis of the Internet Engineering Task Force's (IETF's) standardization effort on the Network Configuration (NetConf) protocol [63]. Like SNMP, the WBEM and NetConf also followed the agent-manager paradigm, albeit with important differences regarding the capabilities of the management protocol and the associated information model. NetConf was qualitatively different from all the other approaches, given that it followed a document-based concept as opposed to an individual managed object access approach.

6.2. *Challenges in virtualized network management*

It was clear, management strategies had to deal with virtualized resources. Some work was already done in UCLP and other research projects like MANTICORE [66]. An important aspect of virtualization were the design requirements of isolation and federation. Isolation meant that resources from one slice, which means a set of virtual resources logically interconnected, and physically connected (Figure 17 shows an example of the procedures to create an slice) and allocated to a specific user had to act independent of the resources allocated to another user when both were sharing the same physical infrastructure. Federation meant that the resources in a slice could cooperate with resources from outside the slice.

The design process was considering the virtualization technology within a service oriented architecture (SOA) model. As an example, virtualization could be considered in the context of infrastructure as a service (IaaS) –section 6.5.1-. IaaS was a technique that represented a physical network device or network substrate as a software entity (e.g., aiming for quick and easy deployment of management approaches using Web services). The idea was that the users did not need to purchase physical hardware, but could instead pay the hardware owner for the usage of the resources as a service. So, during the lifetime of the service, the user could control and manage the requested infrastructure like the owner does.

The real challenge behind network virtualization was on providing complete virtual infrastructures or slices as a service, which brought some advantages in respect to legacy management system architectures. In terms of scalability, the virtual infrastructure could follow business or project needs, since resources could scale efficiently, and when the requirement was over, network resources could be brought back to the provider.

The virtualization system services could be exposed as Web services, which could be exported or connected directly to external SOA applications. SOA helped to eliminate the network of layers, and allowed network components and services to be accessed horizontally, vertically, and externally.

A particularly important challenge was the capability to federate different virtual infrastructures. In that sense, infrastructure integrators could integrate resources from different management domains into a single management domain, thus federating disparate resources into a single infrastructure. Particularly, there were some ongoing efforts on this side to define a service interface in order to federate resources like FEDERICA, OnaLab2 or even GENI [67] [68][69].

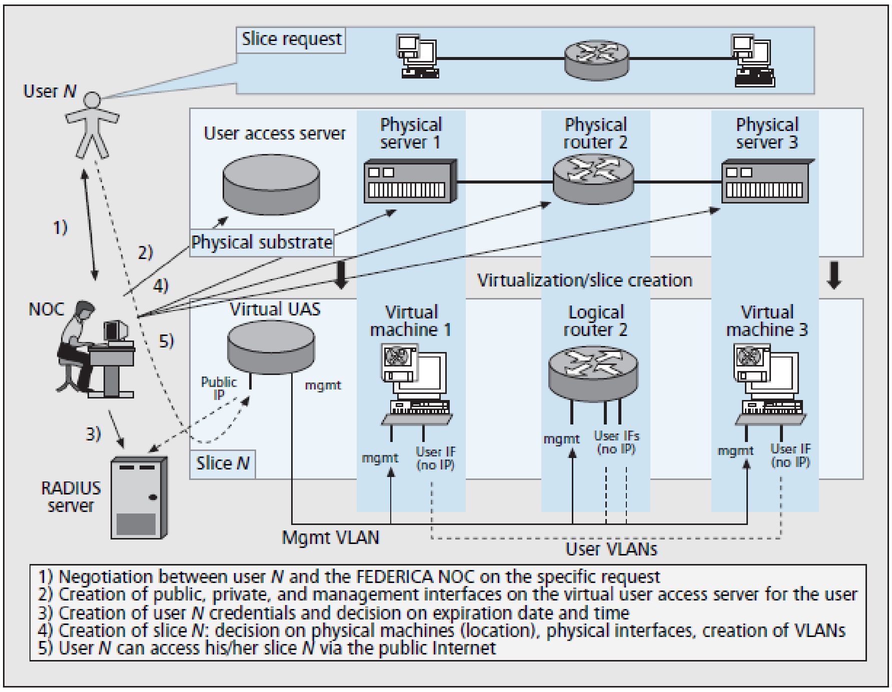


Figure 17. Basic slice provisioning procedure

From there and on Internet would develop according to two basic requirements: interconnecting a constantly growing number of elements and offering ubiquitous permanent connectivity to the majority of its users. The possibility for any user to connect to any other user was also considered fundamental to allow access to all kinds of information, as well as for sensors and objects. These requirements mandated a robust authorization and authentication infrastructure that should be available at all layers of the Internet. Thus, the Internet would continue to be implemented as a set of interconnected autonomous domains due to scaling requirements, each with different technologies and capabilities, and without centralized management or control. Mobility, reachability, and security requirements would impose strong interdomain communication of various types of information [70]. Monitoring capabilities should also extend to mobile users traversing administrative boundaries and multiple parallel virtual networks on the same physical infrastructure, while such advances would require ad hoc support in the hardware and development of new standards for virtual resource representations. In particular, the need was for a richer information system capable to track the relationships between the entities in each domain.

6.3. *Overview of UCLPv2*

The UCLPv2 SOA architecture is depicted in

Figure 18. Each box represented a different service. According to this new architecture, the resource management services, or network element Web services, were a group of services that managed and controlled the resources on a physical device; and each single Web service dealt with a different technology. Resource virtualization services provided a layer of virtualization, so that the technology of the physical devices was abstracted without losing any of the features of these devices. Finally, higher-level services or applications exploited the just described virtualization capability to build complete end user solutions or other services without having to deal with the underlying network complexities. In fact, any of the service groups presented in the architecture was extended to support new technologies or to provide new capabilities in the virtualization layer.

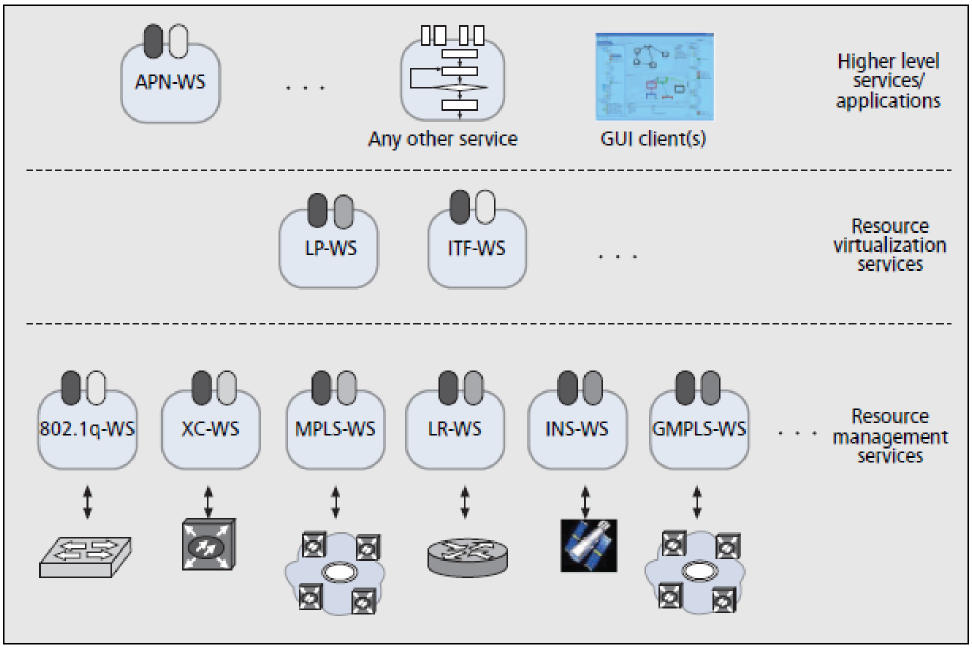


Figure 18. UCLPv2 service oriented architecture

6.3.1. Network element web services

Network element Web services (NE-WS) were used to control physical devices based on the device technology. Each WS behaved like a proxy that sat on top of one or more network element(s) and partitioned the element(s) into several management domains that could be handed off to different users. The following is a list of Web services that were implemented:

• Cross connect Web services (XC-WS): Managed devices that created *cross connections* such as SONET/SDH based equipment, wavelength switching...

• 802.1q-WS (or VLAN Web services): Configured devices that used virtual local area network (VLAN) technology to multiplex connections belonging to different users.

• Logical router Web services (LR-WS): Managed logical routers [71].

At that time, some routers could partition a single router into multiple logical devices that performed independent routing tasks (each logical instance had its own routing table and protocols).

• MPLS Web services (MPLS-WS): Enabled UCLPv2 to control multiprotocol label switching (MPLS)-based devices.

• INStrument Web services (INS-WS): Presented a simple interface to control instruments such as sensors, data sinks/sources, and storage devices.

• GMPLS Web services (GMPLS-WS): Enabled UCLPv2 to trigger switched generalized-MPLS connections using the optical user to network interface (O-UNI).

The service WSDL interface was divided in two port types, a set of abstracted operations and the abstracted messages involved in these operations. The configuration port type contained methods to add, delete, and modify the configuration parameters of the equipment to be managed; whereas the operational port type provided different operations to control the hardware depending on the technology of the underlying physical devices. Each NE-WS was designed to control multiple devices from different vendors at the same time. Thus, the support for new equipment was added by writing an XML file that encapsulated the behavior of the device (communications and transport protocols, commands, and responses). This XML file acted as a driver for new equipment, so the ultimate goal would be for equipment manufacturers to create their XML files to UCLP-enable their equipment. In UCLPv2 each user could decide to use one or more of these services depending on their requirements. UCLPv2 provided an advance GUI (Figure 19) with the following set of key features:

• Network element abstraction: The specifics of each physical device were hidden inside the network element Web service, and only a set of high-level capabilities (e.g., makeXC, createVLAN, createLSP, query, getNeighbors) were exposed to clients.

• Service reusability: The GUI was a client of the NE-WS, but the functionality offered by the NE-WS could be used at the same time by other clients, such as other services or other network management/control applications.

• Enhanced security: NE-WSs acted as broker between the users of the GUI and the physical equipment. This way the physical equipment was accessed only by one entity (the NE-WS) and could remain in a private IP network, whereas users could access the NE-WS from the public Internet. XML signature and proxy certificates to authenticate and authorize its clients was used, and secure sockets layer (SSL) in order to encrypt the communication.

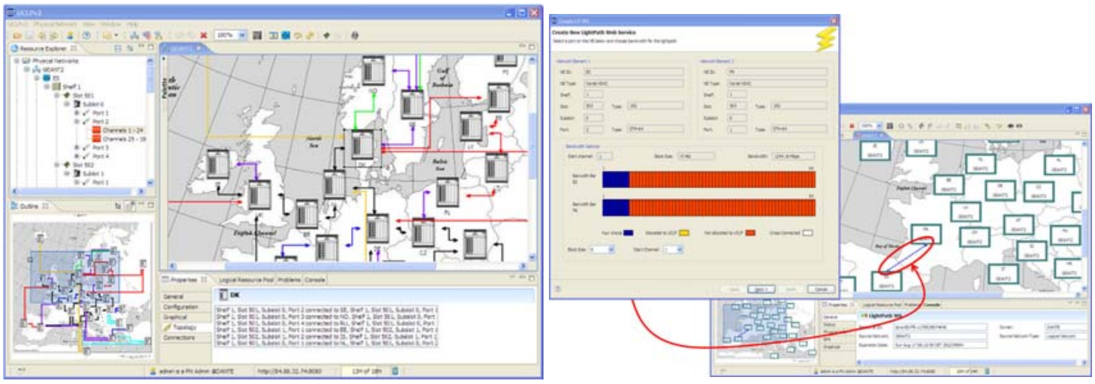


Figure 19. Examples (screenshots) of the GUI developed to partition and compose APN based on the services and functionalities described at the architecture. This GUI was also known as the Resource Management Center (RMC)

6.3.2. Lightpath and interface web services

The fundamental building blocks in the system were the lightpath Web services (LP-WS) and the interface Web services (ITF-WS). These services provided a layer of abstraction between higher level or user-defined services and the NE-WS, thus hiding the details about the underlying physical networks (Figure 20). A lightpath was a reserved, private, communication link between two network interfaces of two network elements. Examples of lightpaths could be a wavelength in a WDM system, a SONET circuit or an Ether-(LSP). The lightpath Web service was an abstraction of a lightpath represented as a Web service. This service provided the following operations: create, createSuper, delete, query, partition, bond, and lease. These operations are described below:

• Create: Created a new instance of a LP-WS.

• CreateSuper: Created a super LP-WS, which is a concatenation of several LP-WSs and ITF-WSs.

• Delete: Destroyed the LP-WS instance. In case it had a finite lifetime, the LP-WS instance would destroy itself upon lifetime expiration.

- Query: Provided data about the LP-WS properties (bandwidth, end points, expiration date, etc.) and status (in use, available, faulted).
- Partition: If possible, the LPWS was divided into N LP-WSs with smaller bandwidth.
- Bond: If possible, a group of LP-WSs with the same network end points was combined into a single LP-WS whose bandwidth would be the sum of the bandwidth of all the original LP-WSs.
- Lease: Changes the ownership of the LPWS.

An interface was a single network port on a network element. It could be both an add/drop port or a network port. Thus, an interface Web service was an abstraction of an interface represented as a Web service. The ITF-WS provided the same operations as the LP-WS except createSuper Lightpaths. Lightpath and interface Web services provided a virtualized view of the network segments; users did not see network elements anymore, but saw a set of network links (lightpaths) and end points (interfaces) with highlevel parameters, like bandwidth, delay, jitter, or start/end time. Therefore, they could be seen as a kind of "network lego blocks" that could be manipulated and integrated into applications or used to create higher-level services (such as bandwidth on demand or reservation services), with the advantage that these new services would be technology independent. The LP-WS would validate the connection and make the required calls to the NE-WSs that in turn would call the network elements.

	Without UCLP	**With UCLP**
Service negotiation	User tells his/her requirements to the NOC operator	User tells his/her requirements to the NOC operator
Service provisioning	NOC operator uses different tools (CISCO CTC, Nortel SiteManager) to create the connections	NOC operator uses a single tool (UCLP) to create the virtual resources and lease them to the user (no connections created yet). The user decides what connections he/she wants to create with the resources
Service reconfiguration	User has to tell the NOC operator his/her new requirements. NOC operator has to tear down the "old" connections and set up the new ones using different tools	User reconfigures the service, tears down the connection he/she no longer wants and sets up new ones (using his/her virtual resources) with just UCLP. If he/she needs more virtual resources, he/she asks the NOC operator for them. Moreover, the user may have created an innovative service that runs on top of his/her virtual resources; the NOC operator does not even notice it
Service termination	NOC operator tears down the connections using different tools	Connections are automatically torn down and virtual resources are automatically returned to the operator upon service expiration
Service monitoring	NOC operator uses different tools to monitor the resources utilization and to receive alarms/error notifications. If a user's connection has problems, the NOC operator has to notify the user	NOC operator uses UCLP to monitor the resources utilization and receive alarms/error notifications. If a user's connection has problems, both the user and NOC admin are automatically notified
Service isolation	Handled by the NOC admin, who is responsible for making sure that the resources assigned to a user's service are not impacted by another user's service (vulnerable to human errors)	Handled by the UCLP software. UCLP partitions the physical network into multiple "atomic" virtual resources (not vulnerable to human errors)

Table 3. Comparative analysis of lightpath/optical VPN service operation with and without UCLP

Another benefit of using LP-WS and ITF-WS was that resources (lightpaths and interfaces) could be assigned to different owners, enabling resource trading between different organizations. This way an organization could acquire resources from other organizations and integrate them into their own management domain in a recursive way; so that resources could be manipulated without the participation of the original owner. The higher level service was the APN. Table 3 shows a comparative analysis of lightpath/optical VPN service operation with or without UCLP.

L1VPN emerged as an Articulated Private Network service for multiple user networks over a common carrier transport network. This new L1VPN management service allowed network providers to manage physical network infrastructures, service providers to manage L1VPN services (by composing individual network resources into L1VPNs), and end users to invoke L1VPN management services to configure operational L1VPNs. This new architecture service was useful for Network providers to partition resources at the L1VPN level, and assign these resources, together with the corresponding WS based management services for the resources, to service providers. Service providers could use it to receive resource partitions from multiple network providers and partner service providers. Further resource partitioning or regrouping was conducted on the received resources, and leasing or trading resources with partner service providers was also supported.

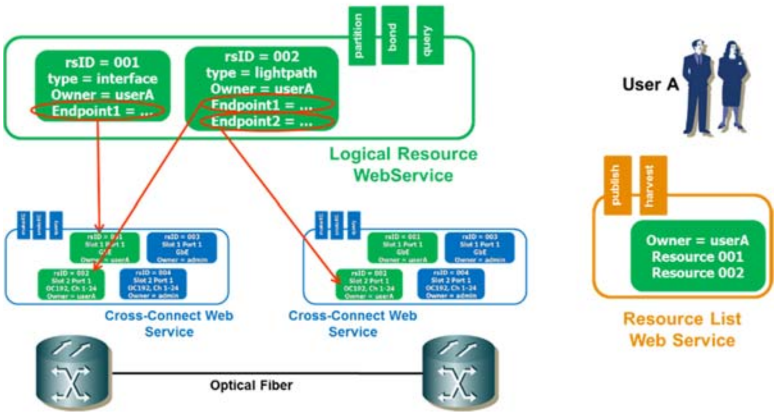


Figure 20. Example if the creation of a Logical Resource WebService defined in a Resource list and composed of abstracted resources from Cross-Connect WebServcies

6.4. Management of Optical Virtual Articulated Private Networks

Virtual Private Network (VPN) enabled the coexistence of multiple user networks over a common infrastructure. Thus, Layer 1 VPN (L1VPN) would offer virtually dedicated transmissions between groups of users over a common transport network. The L1VPN extended layer 2/3 packet-switching VPN concepts to circuits witching networks [72][73] e.g., Wavelength Division Multiplexing (WDM), Time Division Multiplexing (TDM) networks, etc.

With the L1VPN technology, the management of physical network infrastructures, L1VPN services, and application specific L1VPN reconfigurations could be separated, enabling a new network operation and business innovation model. It directly followed the trend of offering increased flexibility, extend and provide management functions closer to the end users, while maintaining proper resource access right controls [74][75]. Therefore the separation of the management of network infrastructures and L1VPN services created a new business opportunity for L1VPN service providers. In that sense, service providers were able to construct L1VPNs by composing resources from different sources and lease or trade resources with each other. Moreover, they could partition or bond resources, create or delete end-to-end connections, and create complex topologies of interconnected L1VPNs [76].

This separation allowed L1VPN end users to configure an L1VPN topology, and add or remove bandwidth. In that sense, it reduced time-consuming service orders to network, since it increased L1VPN end users' capability to manage their own leased resources. So that, an L1VPN end user could be a human operator, or an application program itself. The use of a Service-Oriented Architecture (SOA), together with Web services (WS), was suitable for coordinating the management tasks among those three players - network provider, service provider, and end user. SOA provided loose coupling among interacting programs and applications [77][78][79]. This L1VPN service was built as UCLPv2 service, emphasizing the end users' capability of configuring application-specific L1VPNs.

6.4.1. Roles of network providers, service providers and end users

The success of delivering L1VPN management to end users relied on seamless coordination of three management systems. The interaction of their management systems could be conceptually represented as shown in Figure 21, assuming the network provider used the Automatic Switched Transport Network (ASTN) architecture. The service plane consisted of service management systems that belonged to different service providers, which interacted with network providers through either a User to Network Interface (UNI) towards the control plane, or a Service Provider Interface (SPI) towards the management plane.

Actually, a service provider offered additional business values by negotiating and collecting network resources, although a service provider was not the real owner or operator of network resources. A service provider acted as a physical network broker, partitioning and bonding resources, leasing and trading with other service providers. The service provider maintained L1VPN servers (Figure 22), which interacted with multiple network providers' management servers. Within the scope of UCLPv2, L1VPN servers stored resource lists, and ran L1VPN-WS.

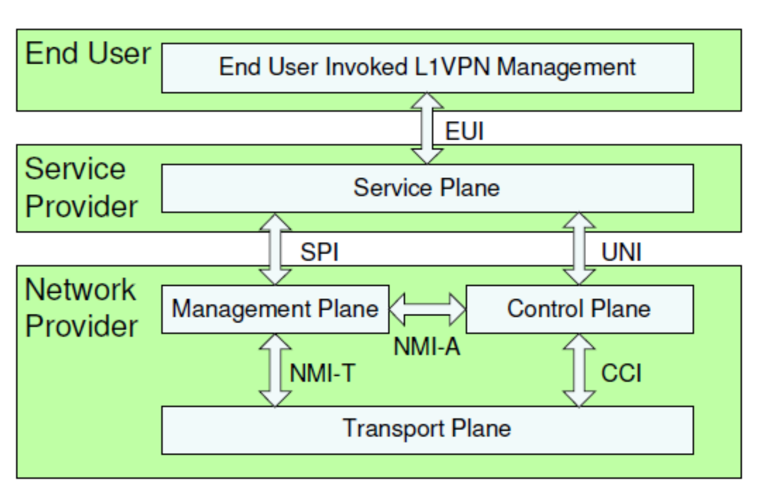


Figure 21. Interaction of the management systems of network provider, service provider and end user

The L1VPN management functions were implemented as WS. And so, a service provider imported resource lists to compose L1VPNs, while creating new L1VPN-WS to manage the composed L1VPN. The service provider could also create super LPs by concatenating a chain of LPs. When an L1VPN end user activated a L1VPN-WS, the manipulations of the resources were executed, and an operational network was created. The roles that network provider, service provider and L1VPN end user had is summarized

	Network Provider	Service Provider	End User
Create a physical network (NE-WS and link topology)	✓	✗	✗
View statistics of owned switches	✓	✗	✗
Create or delete LP-WS and I-WS	✓	✗	✗
Lease or advertise resource lists (LP-WS and I-WS)	✓	✓	✗
Import resource lists (LP-WS and I-WS)	✓	✓	✗
Create or dismantle super LPs	✓	✓	✗
Partition or bond LPs	✓	✓	✗
Create or delete end-to-end connections	✓	✓	✗
Create or delete L1VPNs	✓	✓	✗
Modify L1VPN topology	✓	✓	✓
Deploy or undeploy L1VPN-WS	✓	✓	✓
Activate or deactivate L1VPNs	✓	✓	✓
Query owned resources	✓	✓	✓
View statistics of owned LPs	✓	✓	✓
Receive alarms	✓	✓	✓

in

Table 4.

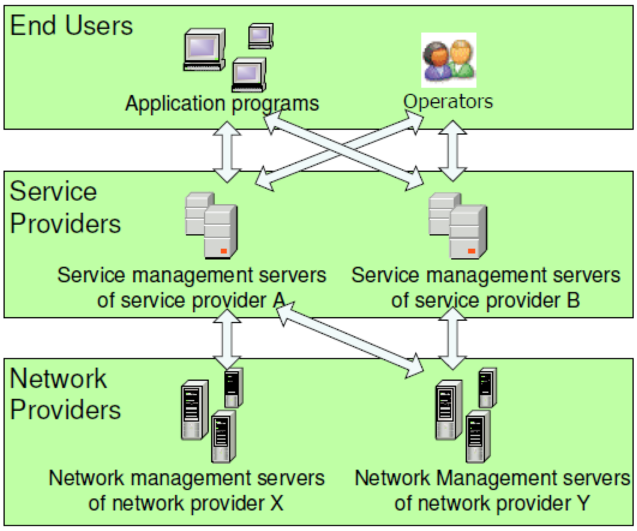


Figure 22. A service provider maintains L1VPN service management servers, which interact with the multiple network providers' management servers

	Network Provider	Service Provider	End User
Create a physical network (NE-WS and link topology)	✓	✗	✗
View statistics of owned switches	✓	✗	✗
Create or delete LP-WS and I-WS	✓	✗	✗
Lease or advertise resource lists (LP-WS and I-WS)	✓	✓	✗
Import resource lists (LP-WS and I-WS)	✓	✓	✗
Create or dismantle super LPs	✓	✓	✗
Partition or bond LPs	✓	✓	✗
Create or delete end-to-end connections	✓	✓	✗
Create or delete L1VPNs	✓	✓	✗
Modify L1VPN topology	✓	✓	✓
Deploy or undeploy L1VPN-WS	✓	✓	✓
Activate or deactivate L1VPNs	✓	✓	✓
Query owned resources	✓	✓	✓
View statistics of owned LPs	✓	✓	✓
Receive alarms	✓	✓	✓

Table 4. Roles of network provider, service provider and end user

6.4.2. APN life cycle

The definition of the APN lifecycle will help to the APN understanding. At the beginning, users were given a set of lightpath and interface Web services that they used to design and create their dedicated APN scenarios, each scenario being a representation of a network topology. After the APN was created (using the UCLPv2 GUI-RMC), the user could deploy it to a Web server and interact with the APN through a Web services interface. The methods offered by this interface were:

• Init(userID): Initializes the APN. Validation actions are performed to ensure the correctness of the device configurations.

• SetConfig(scenarioID, userID, usageTime): Performs all the device configurations specified in the scenario named scenarioID. When the usageTime is over, the scenario is automatically unset.

• UnsetConfig(scenarioID, userID): Clears all the device configurations specified in the scenario name scenarioID.

• QueryStatus(userID): Returns the status of the APN (i.e., provides information about the scenario being executed).

• Stop(userID): Destroys the process instance of the APN.

An example of the interaction of the APNWS with the other services is shown in Figure 23. This example involved different services plus the APN-WS. After the APN had been created, the user (if a human, through the GUI; or if an application, directly) could call the set-Config method of the APN-WS, specifying the scenario that had to be configured. Then, the APN-WS could issue a createSuper call for each connection that was defined in the scenario (connections could be multipoint to multipoint).

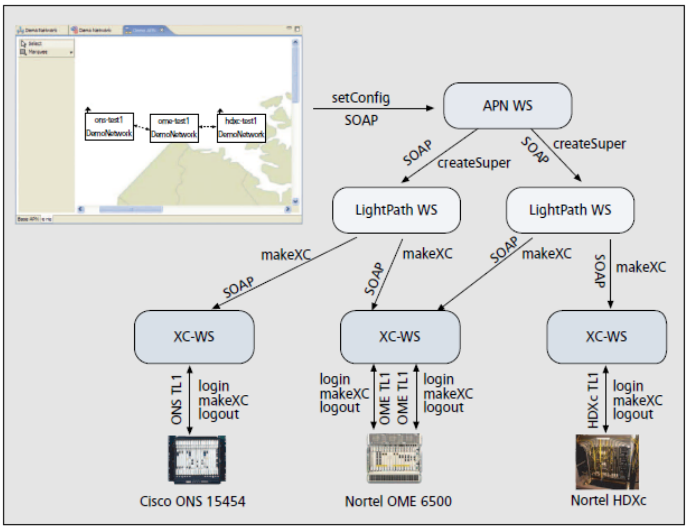


Figure 23. Example of the sequence of the calls required to set up one scenario of an APN. In this example the APN scenario has two end-to-end connections.

This approach of APN to provide VPN was very useful for supporting Grid applications. A Grid user would be able to change the topology of his switched VPNs, so that he could control the connectivity between different sites in his Grid, to trial different Grid application scenarios. Similarly, in a routed VPN, a Grid user would be able to control routing tables and policies, and achieve desired traffic engineering effects.

6.5. Infrastructure service for Optical Networks

Network operators were focused mainly on providing and selling network services on top of the network infrastructures they own and manage, while the end user had no control over how these services were provided. However, as already commented, specific requirements were coming from emerging applications that requested network and computational resources. These new requirements were difficult to accommodate with the existing telecommunication operational models because telecom companies had full control over the infrastructure. However, through virtualization, it was possible to allocate isolated instances from networking devices to different users or applications, and so, use new techniques to virtualize and manage networks in order to decouple the services from the underlying infrastructure. The benefit of these techniques was demonstrated in terms of network operation, emerging services, and industry impacts. In fact, the main rationale behind optical network virtualization was the capability to offer infrastructure services to end users by exposing to them the control of the infrastructure, allowing themselves to assemble networks. Moreover, these techniques validated how the "infrastructure as a service" (IaaS) concept brought a new role to infrastructure providers as it allowed the deployment of dynamic services in optical networks and the federation of their underlying infrastructures. Research on this area provided a solution for the next-generation Internet that would satisfy the needs of (a) users that required network and computing resources simultaneously; (b) users that required dedicated optical network services due to its capacity, deterministic behaviour, quality of service (QoS), and low latency; (c) users that required management and control capabilities over the physical infrastructure; (d) service providers that needed a dedicated infrastructure to offer services on top of it; and (e) service providers that wanted to federate resources obtained from various service or infrastructure providers.

With regard to optical networks technology, the mentioned services put in evidence the capabilities and evolution of photonics networks [80]. Current systems, typically designed for transmissions at 10 Gbits/s per wavelength, had to be upgraded to support long-haul transmissions up to 40 Gbits/s per wavelength, even though the improvement of transmission capabilities would not be enough to support the emergence of new user services.

New management paradigms were needed because service and infrastructure providers had to ensure operational simplicity in planning, engineering, deployment, and operation of services and networks in order to reduce associated operations expenses while improving network performance. Thus, addressing all of these challenges with current infrastructures and technology practices would have led to unmanageable networks for operators aiming to maximize fiber capacity and service re-configurability. The commented technical and management complexities, with the need for highly qualified employees, would have caused operational complexities and limit the operator's ability to support new services for their customers; therefore, QoS and bandwidth could no longer be used as a competitive differentiator. Thus, it was crucial for service providers to be able to deploy and support new services quickly and easily, with minimum manual intervention, hardware deployment, or complex engineering processes. Therefore, taking into account these considerations, a solution was the development of a middleware layer, together with the application of virtualization to optical networks.

Within the scope of next generation networks [81], it was crucial to have a middleware that could provide infrastructure as a service (IaaS) solutions and control, manage, maintain, federate, and configure virtualized slices of optical networks. A virtualized slice was a set of infrastructure resources segmented and virtualized to be federated together in a virtual network domain. Hence, this approach allowed service providers to create their underlying service infrastructures by acquiring resources from different providers on an as-needed basis.

The approach used came from UCLPv2 (User Controlled LightPathsv2) and directly affected the deployment of new services and the operation methodology that optical infrastructure providers had. As commented in previous chapter, UCLP allowed users to define their own packet—or switched-based network architecture including topology, routing, virtual routers, switches, and protocols based on the concept of many separate, concurrent, and independently managed infrastructures operating on top of one or more network substrates across different ownership domains.

This work provided enhancements to UCLPv2 concepts to more efficiently manage virtual instances and provide dynamic services. To do so the Infrastructure as a Service (IaaS) paradigm became the first approach dealing directly with optical networks, while the Argia system (a beta-commercial UCLPv2 implementation for optical networks) was enhanced to improve the existing features and reduce its operational complexity.

6.5.1. Infrastructure as a Service paradigm and Virtualisation

While optical transport network architectures provided capabilities to transport, multiplex, route, manage, and supervise the network, also ensuring its survivability, this was not the case for the networking domain. In wide-area networks, networks/infrastructure/service providers needed to negotiate between them to be able to provide end-to-end services to users, and it became difficult to have adequate prices and control over the whole infrastructure used to provide those services. The problem at that time was that acquiring fiber in the long haul was very expensive to obtain and light. So, the trend was on acquiring the so-called “dim fiber,” which were point-to-point wavelengths across a provider network. The biggest drawback to this approach was that these users wanted to do their own configuration schemas and change management features as they would do with their own dark fiber. Hence, a solution to this problem was to provide virtualized logical networks to these types of users, and here is where Argia came into the picture.

The challenge was to bring virtualisation down to the infrastructure. So far, it was understood as infrastructure segmentation. Segmentation referred to networking infrastructures that could be segmented into multiple parallel dedicated networks for different users, while virtualisation was based on the usage of different hypervisors to virtualise servers and routers. Therefore, the challenge was on identifying the equivalent requirements to bring virtualisation to the network infrastructure. Virtualization and Segmentation techniques allowed users to feel as if they had a dedicated infrastructure and full control over the resources their own. Segmentation allowed hardware resources to be isolated from each other by dedicating them to a process/user (i.e. Multiplexing, Switched Circuits) while virtualization helped to optimize the hardware by allowing multiple processes to use the same hardware resources (i.e. Virtual Machines, Virtual Routers).

Virtualization was currently a key research topic; and a lot of research on virtualization-related topics was being performed. However, virtualization had different meanings based on the underlying devices being virtualised. Network virtualization consisted of different types of virtualization techniques: (a) device virtualization was a technique that represented a physical device or substrate as a software entity, (b) another technique was sharing the resources via an hypervisor; this technique had been widely applied to PCs where tools like VMware, Xen, or others provided several virtual machines on the same hardware, (c) Data-path virtualization consisted in offering an isolated data path across the network. Although a variety of techniques existed to provide such virtualization based on the layer (Layer 2, Layer 3) approaches, L1VPNs [82] were classified as data-path virtualization and offered either dedicated time division multiplexing (TDM) via SONET/SDH technology or WDM channels over a shared infrastructure.

IaaS was the equivalent of software as a service (SaaS) for hardware devices. In IaaS, users did not buy hardware but instead paid a third party (infrastructure provider) for the use of it for a period of time. During the lifetime of the service, the user owned and controlled the infrastructure as if he were the real hardware owner; this business model was targeted toward long-term resource use compared with on-demand services as currently found in Grid Networks.

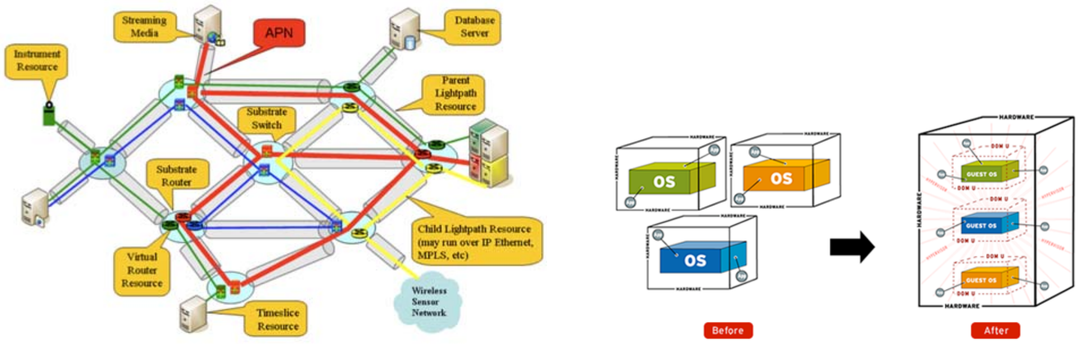


Figure 24. Left picture is an example of Network segmentation and right picture an example of virtualisation (Source: Techlahore Blog)

There were many advantages to the IaaS concept and relevant network virtualization. It mainly brought new business models and a potential optimization of the network infrastructure resources to users. Virtualization of a network device or resource was different from offering overlay networks like virtual large-area networks (VLANs) or virtual private networks (VPNs). When applying virtualization to a hardware device we were creating a logical or abstracted view of it (it could be called a software object), and therefore it could be partitioned into several similar devices offering the same behaviour, but logically separated to provide isolation between users. Although each one of the partitions would had its administrative domain, the virtualization creator would keep the ownership and responsibility for the whole physical device, but not its services management or configuration.

Actually, since this approach allocated several administrative domains under a unique hardware device, most of the complexity was in the internal logic needed to share this resource. It had to prevent users from using resources that they were not allowed to, and therefore kept proper isolation.

The main interests in network virtualization were derived from virtualization applied to computers. It allowed the optimization usage of the hardware devices, and therefore it avoided having an infrastructure with many equal devices performing less than 100% just because they had to be under different administrative domains. Hence, it was expected that some operators would no longer be only service providers, but also infrastructure providers as already envisaged while presenting the state of the art. As a primary conclusion, network virtualization and abstraction of the network physical devices as software objects allowed applications to easily deploy new services on top of virtualized infrastructures. Some of these new services could also be deployed on networks that did not perform network virtualization, although their functionalities would be more limited. Some of the main reasons for applying network virtualization were as follows (some of the reasons have already been pointed out when describing the motivation and state of the art).

• **Reason 1**: *Lower infrastructure operational costs*

Increasingly more and more organizations such as universities, schools, hospitals, or businesses were acquiring their own fiber networks. The problem was that acquiring fiber in the long haul was very expensive to light and obtain. So, the current approach was usually to acquire so-called “dim fiber,” which was point-to-point wavelengths across a provider network. The biggest drawback to this approach was that these users wanted to do configuration and change management as they would do with their own dark fiber. These dedicated networks needed to be manipulated in the same way an application could be manipulated. With an IaaS network, providers selected what resources the users were able to control, and the users themselves did most of the network operation and control, thus reducing the operational costs of the shared infrastructures.

• **Reason 2**: *To become infrastructure providers*

IaaS allowed a new business model to be applied to the telecom market: current carrier roles could be split into infrastructure providers and service providers with no infrastructure. Therefore, current carriers could also make profits by offering unused parts of their infrastructure to increase their revenues Table 5.

• **Reason 3**: *Enable new and Innovative business models*

Innovative Business models could be built on top of virtualized resources and IaaS. The condominium (resource sharing) concept could be applied to communications equipment. Infrastructure brokering sites (eBay-like) could help broker the different resources as a common exchange point for infrastructure providers. The infrastructure resources could be used in house to improve resource utilization and energy efficiency of the IT infrastructure within the same organization, or multiple organizations could create pools of resources that could be used by people that fulfil the right policies. The main advantage however, was that virtualization could decouple the infrastructure ownership from the service offering. This way new business models could appear where specialized infrastructure owners could build, deploy, and maintain (just the physical network) communication networks that would be operated by specialized service providers that control partitions of the provider’s networks for a period of time. This type of business model was the equivalent of the software as a service model for software applications, and it was commonly referred to as “infrastructure as a service” (IaaS).

• **Reason 4**: *Federate infrastructures*

Infrastructure integrators could integrate resources from different management domains into a single management domain, thus federating disparate resources into a single infrastructure. The infrastructure resultant of the federation process would now be managed by one domain (service provider), although the domain would not have the ownership of the physical device. Therefore, there was no need for inter-domain complex solutions, since the whole set of resources would be under the same domain. The example presented in Figure 25 shows a European project that gets resources from several National Research and Educational Networks (NREN) and GEANT2 (the European Research and Educational Backbone) and creates their dedicated network for the project's testbed by federating the different resources. We can see in Figure 25 that the physical substrate is broken into different pieces. This process of segmentation provided isolated networks over the shared infrastructure.

Effect	Current Carriers	Infrastructure Providers
Liability	Provider's liability	User's responsibility
Financial impact	ROI hard to achieve for infrastructure when selling services only	Users pay for both infrastructure and service
User satisfaction	Users locked in service contracts and have no control over the network	Users feel empowered to perform required changes on the network at will as if they owned all of it
Operation expenses	NOC must perform all the changes; little time left to plan ahead or monitor	Users do the simple changes; NOC does the network planning

Table 5. Comparison of current carriers of the time with infrastructure providers

• **Reason 5**: *Integrate hardware in applications with SOA*

By exposing internal processes as web services, these services could be exported or connected directly to external service-oriented architecture (SOA) applications. SOA helped to eliminate the network of layers and allowed network components and services to be accessed horizontally as well as vertically and externally; in contrast, traditional telecom network management systems and operation support systems were vertical, centralized, and transparent to applications.

• **Reason 6**: *Scale infrastructure on demand*

The infrastructure should follow the business or projects needs that organisations had. IaaS allowed an infrastructure to start small and scale efficiently in a matter of hours or less: just going to the infrastructure resource online market and get the extra resources you would need for the amount of time needed. And once the project was over, you could bring the resources back.

• **Reason 7**: *Environmental Impact*

In most IT environments energy was wasted because infrastructure was highly underutilized. IaaS and virtualization had a direct impact on the environment and CO2 emissions: partitioned networks were simpler to operate, and the equipment energy needs were less; IaaS maximized resource usage by sharing a common infrastructure; and organizations who bought equipment would be able to rent out existing infrastructure (and its control rights) and have the same level of control of their resources.

By the year 2009, most network operators still had a vertical business model where the infrastructure provider was also the service provider. However, in some places (mostly in Europe), the law started to change with the functional separate regulation. It tried to establish an open and competitive market. This regulation established the separation for telecom operators, which had to split themselves between infrastructure and service providers, so the same company would no longer offer both services as a unique company.

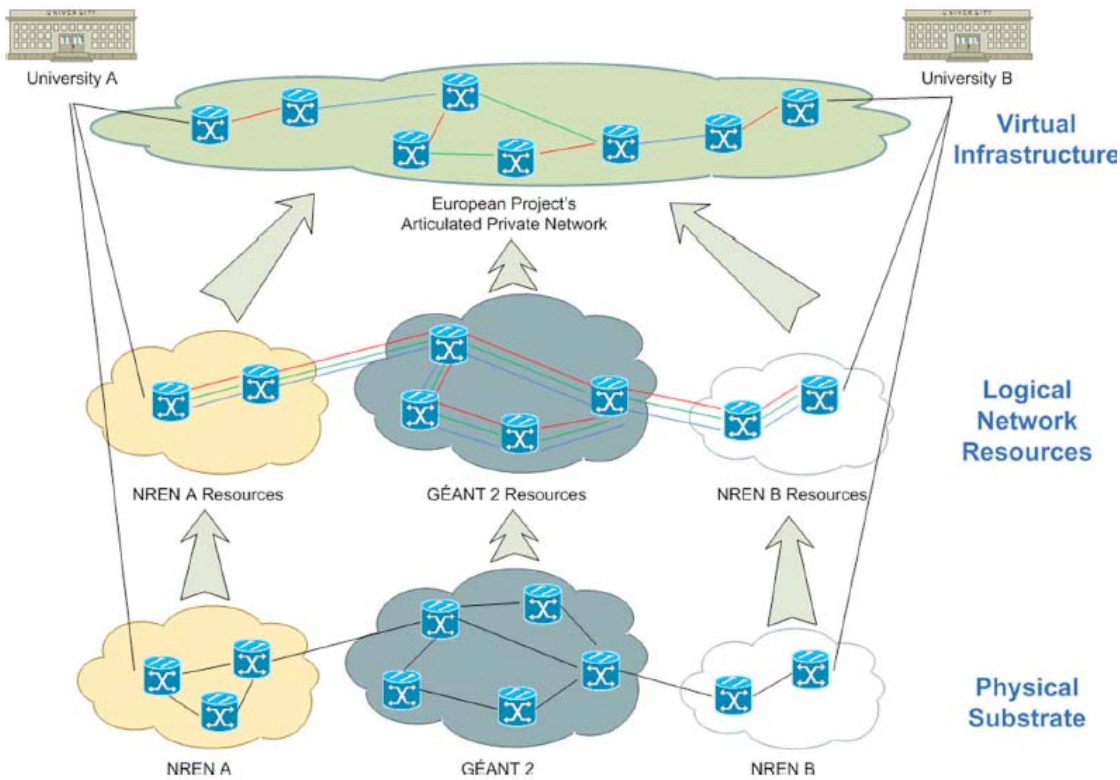


Figure 25. Example of network infrastructure federation

6.5.2. Virtualisation impact on network operators

Infrastructure providers were forced to rent his infrastructure to service providers that would want to offer services on top of it. Some infrastructure providers would no longer be allowed to sell telecom services and so, would be forced to concentrate their business on infrastructure services. Hence, the infrastructure would be considered a service. Virtualization helped infrastructure providers to rent out their resources, either with peer agreements or marketplaces for infrastructure resources. Actually, virtualization became critical for providing resource or virtual network isolation, configuring rights control, and service manageability among other functions, to different service providers while sharing the same infrastructure. Consequently its impact was reflected as follows.

6.5.2.1. Service Providers

Most network operators were making their profit by selling communication services on top of their network infrastructure. This was because services were higher up at the value chain and all the other layers considered a commodity. However, it started to be possible, for network operators, to differentiate among themselves only by offering special application services. Therefore, as network operators had operated in vertically integrated environments, they were not well positioned to compete with more agile companies that had typically developed in highly competitive software businesses, especially big agile companies like Google or Amazon.

On the other hand, they were in an excellent position to provide the high-speed pipes and connectivity facilities to enable these applications.

6.5.2.2. Infrastructure Providers

Demand for capacity fluctuated wildly, the equipment became outdated quickly, and expertise in maintaining high-availability systems was hard to maintain. These were reasons that justified the ability to rent out the infrastructure on an as-needed basis. Users (third parties) did most of the network operation and control, thus reducing the operational costs of that shared infrastructure. Meanwhile, traditionally providers stopped offering their infrastructure to turn to higher services in the value chain. So, with that fragmented value chain network operators had to go back at providing their infrastructure to allow quick allocation and user empowerment of infrastructure resources.

6.5.3. IaaS Framework

The IaaS Framework [83] was a generalized approach to the outcome of years of research under the aforementioned UCLP research programs funded by CANARIE. As a reminder, UCLP goal was to provide end-to-end paths across domains; while UCLPv2's [84] goals were to create reusable and configurable network blocks. The UCLPv2 concepts evolved into many different Physical to Virtual (P2V) products and research projects that were built on the IaaS Framework.

The IaaS Framework architecture was based on a set of resources, libraries and tools licensed under the Apache Software License version 2 that enabled developers to quickly create new infrastructure as a service solution based on the Framework programming model. The functionalities provided by these tools allowed a developer to choose which web service stack would be used to expose the physical infrastructure as a service (supported SOAP engines include Axis2, CXF, and Spring- WS), and to provide a series of modules to plug-in capabilities like security, reservation management and data persistence to the infrastructure service. The Framework also provided libraries to speed up the development of drivers to communicate with the physical devices, like protocol parsers (TL1, NetConf), transport handlers (TCP, SSL, SSH), and a driver architecture called the IaaS Engine. Figure 26 shows an overview of the components of the IaaS Framework and the framework based products and research projects. Device Controller Services were a group of services that acted both as physical device controllers, i.e. they communicated with the physical device using the required protocol, and as virtualization/partition factories; they provided the ``virtualize'' operation that caused a new resource representing a part of the physical device to be created (for instance a group of TDM channels of a port in a SONET switch or an Ethernet interface of a router). The resources created as a result of the ``virtualize'' operation belonged to the Network Resource Services group.

These services represented a part of a physical device: Ethernet Port Resources represented Ethernet interfaces, TDM Timeslot Resources represented groups of channels in TDM interfaces, WDM Resources represented wavelengths in WDM interfaces, and so on. These resources were the ones being exchanged by organizations to provide the infrastructure as a service. For instance, if organization A owned a SONET switch with 8 TDM interfaces and 16 Ethernet ports, the administrator of organization A could partition 4 TDM interfaces and 8 Ethernet ports and give them to organization B, so that organization B could control these ports as if it was the real owner. Furthermore, organization A could partition the other 4 TDM interfaces and 8 Ethernet ports and give them to organization C, so that both organizations B and C would control parts of the same optical switch as if they were the real owners. Network Application Services were the services that provide an end-user over the partitioned infrastructure (or infrastructure as a service). An example of a network application service was Argia's optical connection service, which allowed a user to perform point-to-point and point-to- multipoint connections over TDM, WDM, or Fibre Resources.

Resource Lists Services provided the means of exchanging resources between organizations. When organization A wanted to give permission to organization B to access some of organization A's resources, organization A created a resource list populated with device virtualization resources that represented all the physical infrastructure that organization A could access, and sent the resource list to organization B. When organization B received the resource list, it could assign the network resources it had received to one or more of the network application services organization B had deployed (for instance, organization B could say: these 4 Ethernet ports can be used by the IP Network service and these 8 TDM ports can be used by the Optical Connection Service). This resource exchange process was recursive, as shown in Figure 26, meaning that organization B could create another resource list and give access to some of the resources that organization B temporarily owned to another organization. Organizations and their users were managed by the User Workspace Service.

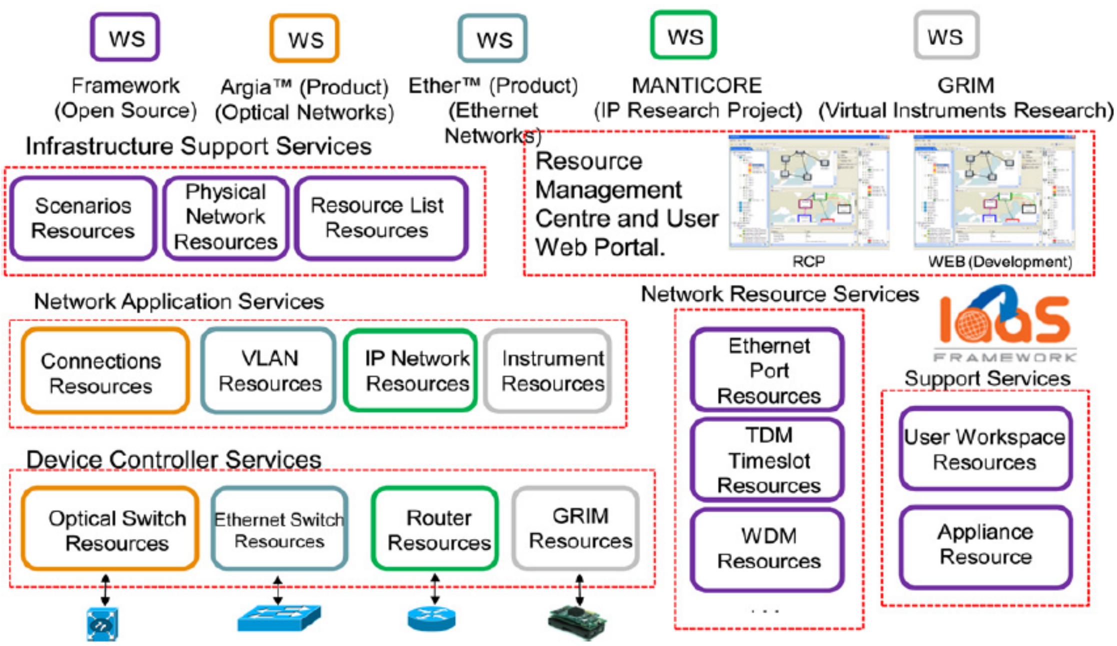


Figure 26. Architecture of the IaaS Framework and the framework based products and research projects

Each user in an organization had a user account and its associated credentials that would be used to interact with any of the services. Users could have three different user roles:

- Physical Infrastructure Administrator. Owner of a physical infrastructure. This type of user could partition the physical resources he owned and assign permissions for other users to control them (by creating and exporting resource lists).
- Virtual Infrastructure Administrator. This type of user got the resources from one or more Physical Infrastructure Administrators or other Virtual Infrastructure Administrators. He could also assign the resources he could control to different network application services.
- End-User. This type of user was the typical ``dumb'' end user that just wanted to use a service (like an end-to-end connection service or an IP Network service).

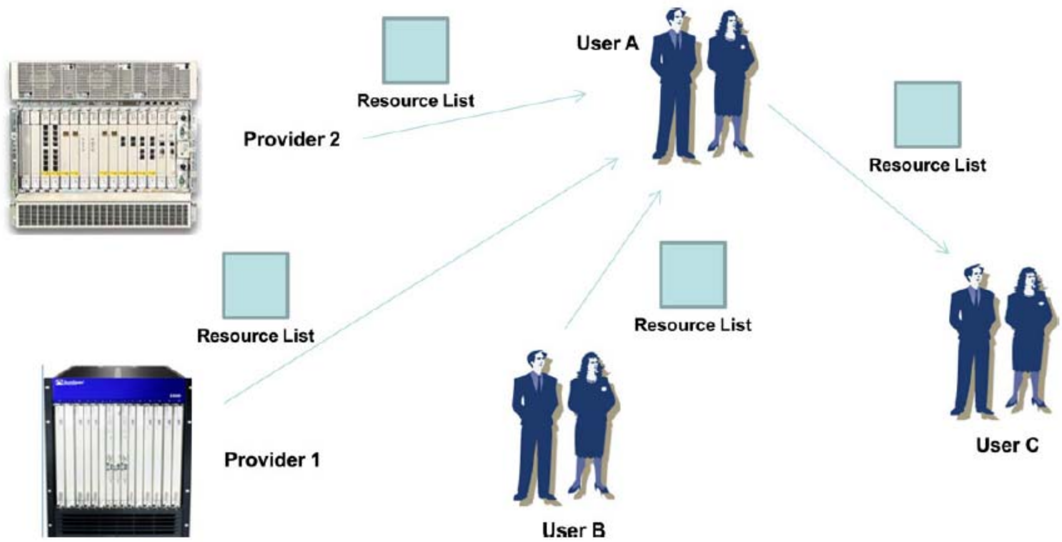


Figure 27. The resource trading model enabled by the IaaS Framework

6.5.4. The Argia system

Being the main goal behind UCLPv2 to provide a virtualization solution for optical networks by creating reusable and configurable network blocks. UCLPv2, which was built on top of the IaaS Framework, evolved into different physical-to virtual products. Argia software architecture, depicted in Figure 28, was based on the IaaS Framework software. Argia was the beta-commercial implementation for optical networks (TDM, WDM, and fiber technology), Ether would be the system for Ethernet and MPLS (multiprotocol label switching) networks, while OpenNaaS would appear later on as a solution for creating logical IP networks.

The IaaS Framework was a composition of a set of resources, libraries, and tools that were open source based which facilitated the development of IaaS solutions. The IaaS Framework, as evolution of UCLPV2, evolved into a new architecture that offered a better technology than its origin, based on Globus Toolkit 4 [85]. Even though it provided many functionalities and WSRF, WS-N, and WS-Security implementations, its programming model was too tightly coupled with WSRF, so it was impossible to separate the business logic from the web service code. Which was an important issue considering the importance of building new innovative business models that would make use of virtualisation, since it prevented the developers from exposing hardware resources using different remote technologies (like plain HTTP, REST-Style [86] Web service, or others) and business logic is polluted with WSRF variables and statements.

Therefore, the new IaaS framework was built on Spring Framework (a Java Framework to create enterprise applications) and OSGi [87].

Argia (Figure 28) was the IaaS-Framework-based system to create IaaS solutions for optical networks. The main goal of Argia was to enable infrastructure providers to partition their physical networks–infrastructure and to give the control of the partitioned infrastructure to third parties (infrastructure integrators or APN administrators) during a period of time. These third parties would use the partitioned infrastructure in house or would deploy some intelligent software on top of the resources (like a resource reservation service) to provide services for their end users, or they would even further partition the infrastructure and rent it out to other users.

Argia became the facto the evolution of the UCLPv2 software; it pursued an ongoing effort toward creating a commercial product that could be deployed in production optical networks. Table 6 shows the network elements supported by the Argia system. Table 7 illustrates the networks and testbeds where Argia was deployed and validated. Argia's particular software modules were the Optical Switch Web Service (WS, a device controller service), the Connection WS, and the APN Scenarios WS (End User Services).

The Optical Switch WS interface provided a series of high level operations that encapsulated the physical device functionality. For instance, using the "invoke" operation and passing the appropriated operation identifier and parameters, an Optical Switch WS client could create a new cross connection (one to one unprotected or protected), undo a cross connection, or refresh the physical device state information by polling it. It was interesting to note that the parameters the operations had to work with could not be for any particular physical device, but they ideally worked for all the optical switch devices, because one of the main features that the Optical Switch WS had to deliver was multiprotocol and multivendor support.

Vendor	Model	Technology
Cisco	ONS 15454	SONET and SDH
Nortel	OME 6500	SONET and SDH
Nortel	HDXc	SONET and SDH
Nortel	OPTera Metro 5200	DWDM OADM
Calient	FiberConnect PXC	Photonic Cross Connect (PXC)
W-onesys	Proteus	DWDM ROADM
Cisco	Catalyst 3750 (basic support)	Ethernet (only port switching)
Allied Telesis	8000s 48 cls (basic support)	Ethernet (only port switching)

Table 6. Network elements suported by Argia

Multivendor support was accomplished mainly through the use of the IaaS Engine, a Java-based framework to create drivers for physical devices using a model driven approach. The Engine's interface provided a Java-based model of the physical device's state that satisfied two needs:

- Engine to Optical Switch WS communication: the engine filled the model attributes with the information of the physical device, allowing the Optical Switch WS to get the latest physical device information.
- Optical Switch WS to engine communication: the Optical Switch WS filled some model attributes to request the Engine to perform some actions over the physical equipment; such as making a cross connection.

The engine had to ensure that the resources of the internal model state would keep synchronized with the hardware device it was controlling. The Connection WS was a service that managed one or more connection resources (connections could be one to one, one to many, or loopback). Each connection resource had pointers to the set of network resources that were connected together. To create a connection, first the Connection WS classified all the resources belonging to the same connection per optical switch; next it extracted the relevant parameters from the network resources (like the slots/ports/channels, the bandwidth, a cross-connection description); then it issued all the required messages to the Optical Switch services, and finally it updated the state of the network resources. The procedure to undo a connection was symmetric.

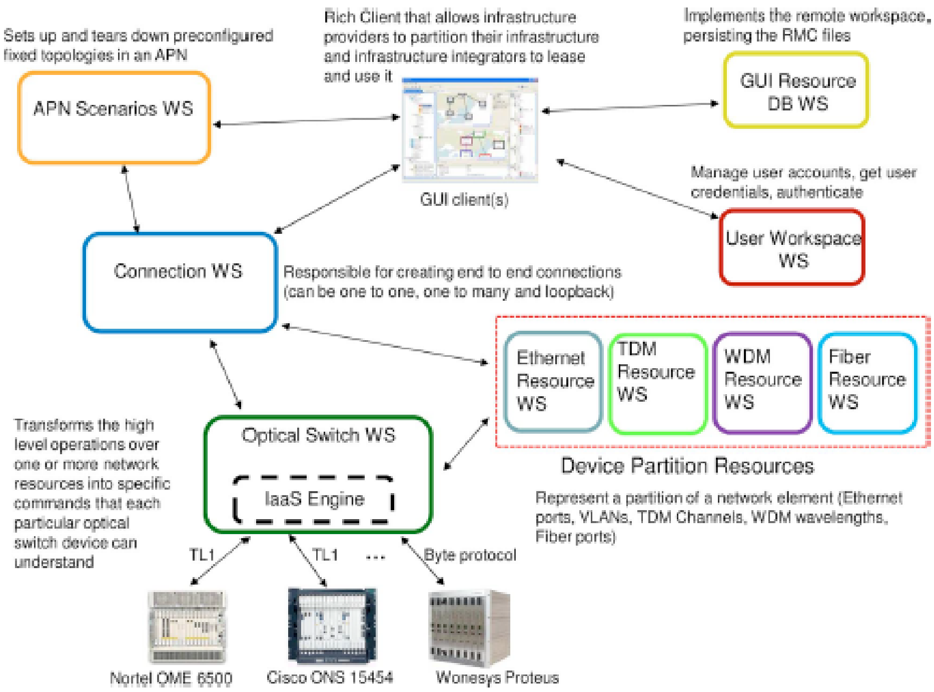


Figure 28. Argia service oriented architecture

Finally, the APN Scenarios WS was the evolution of the Custom APN Workflow [54]. This service could setup and tear down preconfigured topologies consisting in a set of connections in an APN. To achieve its goal, when the setup operation was called on an APN Scenarios Resource, the APN Scenarios WS called the Connection WS to create all the connections required by the scenario. Tearing down a scenario was a similar process: the Scenarios WS called the destroy operation on each of the connection resources that had been created in the setup operation.

Network or Testbed	Being Used for
CANARIE network	Beta testing for use in production network
	HPDMnet research project
STARlight (GLIF GOLE)	HPDMnet research project
PacificWAVE (GLIF GOLE)	HPDMnet research project
KRlight (GLIF GOLE)	PHOSPHORUS research project
	HPDMnet research project
CRC Network	PHOSPHORUS research project
I2cat Network	PHOSPHORUS research project
University of Essex testbed	PHOSPHORUS research project
Poznan Supercomputing Center	PHOSPHORUS research project
DREAMS Project testbed	DREAMS research project

Table 7. Argia deployment in research projects and NREN

6.5.5. New Argia services

This section introduces two of the main services for optical networks that were designed and developed for Argia, based on the IaaS Framework, and used on their respective testbeds and pilots. These services were an advance reservation functionality applied to optical network resources and an optical multicast service. Actually, many services could be built on top of virtualized optical networks under the IaaS Framework Figure 29 while being validated and deployed in different research projects Table 7. The services introduced provided a better integration of the network with the application and helped on the network resource usage optimisation.

This developments were tested and demonstrated during the SuperComputer 2008.

6.5.5.1. Advance Reservation Service

The advance reservation service (ARS) worked on top of Argia and was used to allow users to request resources for a future period of time. Although it is a scheduled on a future time basis, it was also very interesting for service operators in order to optimize the resource usage on their network planning.

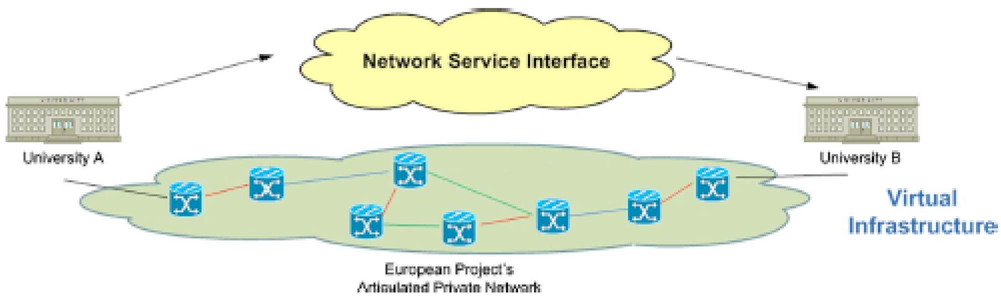


Figure 29. Network service interface over virtual infrastructures

The ARS was an optical network service that, besides guaranteeing the availability of optical network resources to service providers, allowed service providers to manage network resources more efficiently and thus offer a better QoS to users. Hence, it could be used as a planning tool for future availability of network resources. This service was extended, deployed and tested within the NRPS-Harmony system of the PHOSPHORUS EC project (this extension is presented within section 4.1.2). Harmony was a network brokering system implemented at the service layer. It provided advance reservation features for network resources in a heterogeneous environment and across different administrative domains. Moreover, Harmony had a service interface that communicates with the grid middleware and therefore allowed grid applications or end users to request end-to-end paths with specific delay and limited bandwidth.

In order to integrate the UCLP/Argia-ARS as a Harmony service, a web service interface was designed. This interface contained some functionalities that allowed the user to check the availability of a reservation, get the reservations done previously, create a new one, and cancel a reservation done before as seen in the schematic use case of

Figure 30. In fact, an advance reservation was composed by a set of services, and the services were composed by one or more connections. While connections were of a topological nature, services added time constraints to a group of connections. When these functionalities were called, the service processed the requests, computed the results, and then returned the response to the requester.

Generally, two types of resource reservations could be distinguished: immediate reservations, which were made in a just-in-time manner, and advance reservations, which allowed reserving resources for a future period of time. As an example, this was the case in web caching or distributed multimedia applications where large amounts of content such as video files had to be transmitted up to a certain, predefined deadline [88]; or in grid computing, where typical computations on the distributed parallel systems resulted in large amounts of data that also had to be transmitted in time between different specified machines. As commented above, this advance reservation service brought to users the capability to request the future availability of network resources needed to build the optical paths between two specified nodes. In order to allocate the resources that would create the optimal path through the network, the routing strategies of the ARS were based on Dijkstra's shortest path (DSP) algorithm.

From the reservation point of view, this service allowed users to create, delete, and query advance reservations. From the topology perspective, it stored, retrieved, modified, and deleted the resource-related information according to the topology of the network reservations.

Figure 30. ARS's Unified Modeling Language (UML) use case diagram

In ARS, a connection was defined by a unique identifier of the service, a source endpoint and one or more destination endpoints, some bandwidth-delay constraints (minimum and maximum bandwidth, which may be equal, the actual bandwidth, and the maximum latency), the amount of data to be transferred in megabytes (required only for malleable reservations, where both the starting time and the duration were variable), the status of the connection, the EPR (endpoint reference) necessary to create the connection physically when the service contacted to the UCLP system, and finally the directionality of the connection, which could be:

- Unidirectional tree: There were unidirectional connections from the "source endpoint" to each of the "target endpoints". Each of these connections fulfilled the given bandwidth/delay constraints.
- Bidirectional tree: The "source endpoint" had a bidirectional connection fulfilling the bandwidth/delay constraints in each direction to every "target endpoint".
- Full mesh: Every endpoint was connected to every other endpoint with the given bandwidth/delay constraints.

While "connections" were of topological nature, "services" added time constraints. A service was defined by a service type and consisted of one or more connections. All connections grouped in a service were characterized by exactly the same time constraints. Therefore, the information model was as follows:

The parameters of a service were:

- The service identifier. It was unique within a set of services aggregated to a single reservation

- The type of reservation, which could be fixed (the starting time and the duration were both fixed), deferrable (a deferrable reservation also has a fixed duration, but a variable starting time) or malleable (reservation that finally has a variable starting time, a variable duration and, as a result, variable bandwidth). These three reservation types are visualized in Figure 10.2.

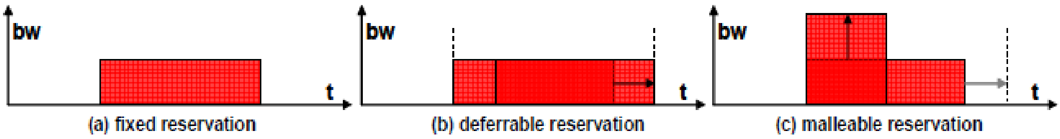


Figure 10.2 Types of reservations

- Automatic activation: a boolean that indicated whether the service would be set up automatically at the start time or required the client of the advance reservation system to explicitly signal the activation of the service
- A date that indicated the expected starting time for pending services or the expected ending time for active services
- Status: the status of the service
- The time constraints for fixed reservations: start time, for deferrable reservations: earliest start time, duration, and deadline and duration; and for malleable reservations, earliest start time and deadline (time by which data completely has to be transferred).

Reservation: A reservation was defined by the following parameters:

- Reservation ID: unique identifier of the reservation.
- isPre: indicates if the reservation is a pre-reservation (true) or a permanent one (false).
- Notification consumer URL: URL of a Notification Consumer that is to be notified when any of the services' status changes.
- Owner: the owner of the reservation.
- Services: the set of services.

Job: Several independent reservations could be grouped to a single "job". The following parameters define a job:

- Job ID: unique identifier of a job.
- Pre-reservations: a set of pre-reservations grouped into the job.

When the connections were to be activated automatically, a scheduler was needed. So, when a reservation was created, the scheduler was programmed at the requested time. When the requested start time arrived, the end-to-end path needed for the connection was physically created by means of Argia. In the same way, when the end time arrives, the Argia system was called to delete the end-to-end path. To implement this scheduler the Quartz [89] library from Open Symphony was used.

6.5.5.2. Dynamic Optical Multicast Service

A growing trend for data networking consisted in the need to transport data-intensive high performance and high-quality digital media traffic over long distances [114], which required transport profiles included point to point, point to multipoint, and multipoint to multipoint. However, these type of traffic could be well supported with traditional data networking architectures and available techniques. At best, available architectures provided performance for low-quality and less demanding solutions that did not scale to high-definition digital media. This behaviour made optical networks suitable for them, although it was very important to optimize the use of the resources since their availability was somehow scarce. The main requirements for high-definition digital media precluded the use of common L3 techniques. These requirements were those that defined the metrics to define high quality, performance, data volume, scale, and network optimization. The requirement characteristics of these parameters could not met with packet routing techniques or through specialized methods for optimizing routed performance. These streams did not have the attributes that edge platforms had, those based on L3 multicast techniques with QoS standards methods. Consequently, alternative techniques were investigated. These requirements had several challenges, from the dynamic allocation of core network resources to advance reservation features (which could be achieved by means of ARS), integration of L1 and L2, path duplication, edge device addressing, path identifications, path monitoring, and new management techniques.

To address these challenge, a research consortium created an international testbed, the High Performance Digital Media Network (HPDMnet) where the research outcomes could be validated. This consortium addressed the creation of a global dynamic optical multicast service (DOMS). The DOMS was built as a service extension for Argia. It provided the user the capability of receiving multiple high bandwidth streams from the network and/or sending the same stream or group of streams to one or more users in the network using L1 multicast (or optical multicast) techniques. The HPDMnet initiative demonstrated the functionality of streaming multiple high-performance, high-quality digital media streams among multiple sites around the world from three continents simultaneously. The optical multicast configurations were established dynamically by using the Argia dynamic optical multicast service.

The DOMS service was initially targeting the TDM technology (SONET and SDH). In particular the Nortel Optical Multiservice Edge 6500 (Nortel OME6500) and Nortel Optical Cross Connect HDXx (Nortel HDXc) [1] platforms. Both platforms provided the drop-and-continue functionality that allowed data streams in an input port to be replicated up to N−1 times and forwarded to N outputs ports (the N−1 copies plus the original stream). The OME platform allowed the creation of up to three copies (so the split ratio couild be up to 1:4), and the HDXc platform allowed the creation of only one copy (split ratio of 1:2). Data replication were performed in real time, so the end-to-end delay and jitter experienced by the data stream are not affected.

Initially, Argia already supported the command set of the Nortel OME6500 and the Nortel HDXc platforms, but the core Argia services and its resource management center could handle only one-to-one connections. With these enhancements Argia could also support multicast functionalities.

In order to help the user to create connections with their resources, the resource manager center of Argia was enhanced. It initially supported only point-to-point connections; therefore two new algorithms were implemented to handle point-to-multipoint connections. The first algorithm was the equivalent to Dijkstra's shortest path algorithm and was used to select the minimum number of resources that would create a valid multicast tree given a source interface, a list of destination interfaces, and the desired bandwidth. The shortest path algorithm created a graph with all the nodes and links and calculated all the possible routes between the source node, 's', and each one of the target nodes, '*ti*'. So for each pair {*s,ti*} a list with all the possible routes between '*s*' and '*ti*' was generated. The complete list of routes between all source and destination pairs was used to create a solution tree by using a backtracking algorithm. The solution tree had the source as a root and as many levels as targets existed in the multicast connection. The root element had '*n'* children leaves, '*n*' being the number of possible routes between '*s*' and '*t1*'. Each one of the '*n1'* leaves of the first level of the tree stored a list of the resources used in that particular route between '*s*' and '*t1*'. Each leaf of the first level of the solution tree had '*n2*' children leaves, 'n2' being the number of routes between '*s*' and '*t*2'. Each one of the '*n2xn1*' leaves of the second level stored the

[1] It is interesting to Point out how Internet is a fast moving sector. Nortel, of the ot biggest, one of the biggest optical manufacturer of that time, did not manage to cope with the growing demands of new services and ended up disappearing a few years later

resources resulting from the union of two sets: the resources used in route *n1* between *s* and *t1* and the resources used in route *n2* between *s* and *t2*. Further levels of the solution tree (if they would exist) would be computed the same way.

The optimal solution was the leaf of the last level that has fewer resources stored. An example of the solution tree of a multicast connection with one source and two destinations is given in Figure 31. As the number of leaves in each level grows exponentially, it is very inefficient to compute the complete solution tree. This was why a backtracking algorithm that used the branch-and-cut approach was used to generate the solution tree. The algorithm would keep developing a branch only if the number of resources stored in the current leaf of the branch would be less than the number of resources of a temporary solution (the temporary solution is the minimum number of resources of all of the already computed multicast trees). If this condition was not fulfilled, the algorithm would cut the branch and would continue developing a new one until there were no more branches to develop.

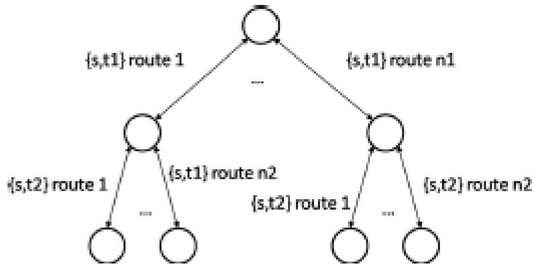


Figure 31. Example of the solution tree computed by the algorithm that finds the multicast tree with minimum number of resources. In the example there is one source, s, and two destinations, t1 and t2.

The second algorithm was very similar to the first one, the only difference was that the solution had to compute all possible multicast trees. Therefore, the same solution tree as in the first algorithm was created, but this time the backtracking algorithm did not cut any branch. Because the number of possible multicast trees could be overwhelming, the user interface displayed only a limited number of solutions (configurable by the user). This way the user could use the multicast service in a more intuitive and simple fashion, because some complexities associated with the provisioning of the multicast service, such as route selection and the modification of existing multicast connections without tearing them down, would remain hidden from the user interface.

Figure 32 shows the validation done in a real testbed (the HPDMNET infrastructure). This set up was used in particular for a demo during the 7th LambdaGrid workshop organised in Prague. Each coloured line represented a different stream, and each dashed square a different site. All the multicast connections were defined prior to the demo using the Argia Resource Management Centre, and were stored as an APN Scenario. During the demo, the APN Scenario was instantiated and Argia configured all the required cross-connections on the network elements to create the multicast trees. After a certain period of time, the APN Scenario was torn down and the Argia software undid all the cross-connections. This sequence of events was performed several times during the demonstration.

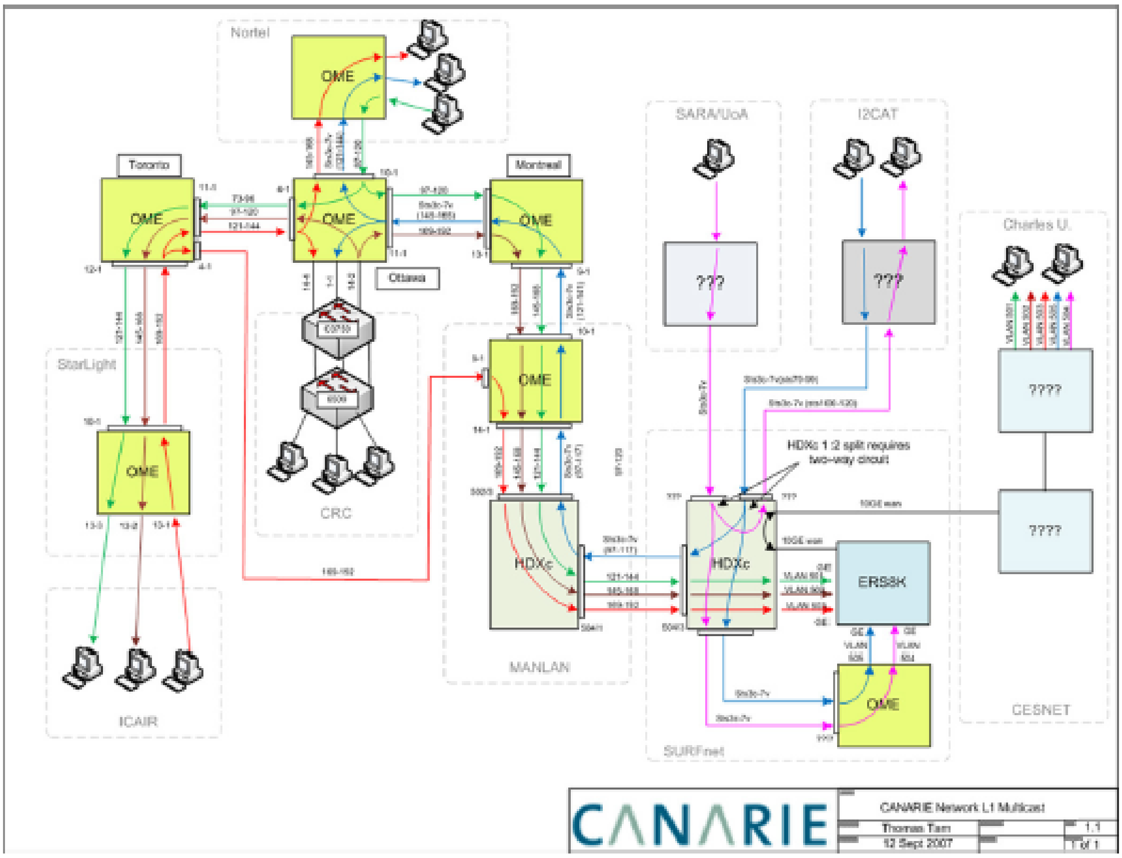


Figure 32. Layer 1 scenario of the dynamic optical multicast carried out at Prague

Applications like live high-definition multi-conferencing or super high-definition TV broadcasting, where virtualization could provide the needed optical network resources to provide the service for a fixed period of time, benefited from a multicast service such as DOMS. DOMS and ARS were both deployed along the HPDMnet initiative. This initiative was a major advance beyond such legacy services, which were based on architecture and technologies that had been in use for many years.

6.5.6. Converged optical network and infrastructure to address GHG emissions

In general, ICT technologies produced more Green House Gas (GHG) emissions than the airline industry, at around 2%–3% [90] of global emissions primarily through the consumption of electricity at coal fired power stations, which are the pre-dominant power sources in most countries around the world. Actually, it was expected that if no steps were taken to limit these emissions, the ICT contribution would double in the following 4–6 years [90]. The rate of growth of ICT was expected to further accelerate because ICT was also seen as a powerful tool to reduce GHG emissions in other sectors of society. Some studies indicated that overall emissions could be reduced by as much as 15%–20% through the application of ICT technologies [91].

Networks, where Internet traffic was doubling every 2 years [92], and data centers were major consumers [93] of this power. It was expected that next generation networks alone would consume 5% of a nation's energy as broadband networks move to speeds in excess of 100 Mbps [94].

With the advent of cap and trade in the USA and Europe the cost of electrical power from coal fired power plants would jump dramatically. Therefore equipment vendors and network operators started to look to a variety of solutions to reduce the GHG footprint of their equipment and networks. Energy efficiency was considered as part of the solution, but there were some concerns that given the rate of growth in demand for ICT products and services, an increase in efficiency would not be sufficient to counter balance the growth in the on-going deployment of new equipment and services. As well, phenomena such as the Khazzoom-Brookes postulate [95], also known as the Jevons paradox, would mitigate against any efficiency gains as it had been demonstrated that paradoxically increased efficiency results in increased consumption. So depending solely on increased equipment efficiency would not result in any significant reduction in GHG emissions from computer and network equipment.

In addition to the rate of growth of ICT and its concomitant electrical consumption, there was an increasing concern amongst climatologists that globally the tipping point in terms of climate change was already passed [96], [97]. Already many jurisdictions such as British Columbia and New Zealand [98] had already mandated carbon neutrality on public sector institutions and the UK required carbon accounting on all responses to Requests for Proposals from the UK government [99]. These mandated carbon neutrality and accounting requirements were expected to have significant impacts on network equipment manufacturers and operators when they bid for business from government and public sector institutions.

Using more renewable energy had long been a recognized goal by the ICT industry, but increasing the amount of renewable power as part of an energy mix was complicated by the fact that most renewable energy sources were in remote locations. Moreover, the current electrical grid of that time was not designed to be able to deliver this amount of power to consumers because of a lack of distribution capacity [100]. Thus, it would take years, if not decades to upgrade the electrical line transmission infrastructure to deliver the renewable power that was needed by the ICT and other industry sectors.

The ICT industry was competing with many of these same sectors for this same power which would inevitably drive up its cost. As well, electrical transmission line losses could be as high as 15% [101] which would further increase the carbon costs of using renewable electricity in cities.

The consequence of dramatically increased power costs due to cap and trade, lack of access to renewable power because of poor transmission line infrastructure, transmission line losses of delivering electricity, mandated carbon neutrality and accounting by government and the need to plan for climate disasters suggested that new business models and network architectures were required, especially now that ICT was being seen as a critical infrastructure. Because of the increased concern of climate catastrophes and high cost of coal fired power it seemed logical to build an ICT infrastructure that was independent of today's fossil fuel systems and that could be deployed using entirely renewable energy sources. But given that most renewable energy sources were at relatively remote locations with poor transmission line infrastructure and the significant loss of power to transmission line losses on any infrastructure that did exist, suggested that network nodes could be located at these remote locations.

The ICT industry, as opposed to any other industry sector, was ideally suited to have some of its infrastructure located at these renewable energy sites as long as these facilities were connected to high speed optical networks. Many large data centers deployed by Google and others had already adopted this strategy [93]. An ICT infrastructure deployed at renewable energy sites had a number of advantages in that there was now a completely independent redundant power infrastructure for the ICT industry made up of the traditional one provided by electrical utilities in cities and another based on independent renewable power sources located largely outside of cities but interconnected through optical networks. More importantly, by locating network nodes at renewable power sites network providers would have a greater certainty of price, as they would not be competing with other industry sectors for that same power.

Although parts of Canada and Norway had abundant and reliable hydro-electric power, experts believed that most future sources of renewable power in Canada and elsewhere around the world would come from solar or wind powered facilities[102].

Nuclear power was expected not to have a significant role because of its high opportunity cost due to the long time for planning and construction and the ongoing issue of how to dispose of nuclear waste. As wind and solar power was far less reliable than hydro-electric or traditional fossil fuel power plants, designing an ICT infrastructure for this type of renewable power would be much more challenging. In fact, there were a set of questions to which technology should provide answer: How do one design a network and computing Grid infrastructure where many nodes might not be on-line because of lack of sun or wind? How do different businesses independently would manage their own network and computational resources located in a remote data center thousands of kilometers away? How do users re-configure and manage multiple resources at different locations depending on the local availability and price of renewable power?

Grid architectures [103] from the very beginning had been designed to address many of these questions with the concept that computation resources were likely to be distributed globally, across independent managed domains and the availability of resources at any given site would vary depending on local demand, suitability of the resource, contractual relationship, etc. In terms of using renewable energy sources, all we were doing was adding another variable to the mix that would determine the availability of resources at a given site. This concept of using, managing and aggregating distributed resources for computation with grids could also be applied to optical networks. This was the original principle behind the development of User Controlled LightPaths (UCLP) wich could be more generally extrapolated into a broader concept of Infrastructure as a Service (IaaS) which applied many of the underlying principles of grids to an entire range of computer and network infrastructure including optical networks, grids, routers, instruments and so on.

IaaS was a business model that consists in offering infrastructure pieces through a service middleware instead of giving direct control over the hardware devices. Having all changes to hardware go through this software layer made possible to efficiently share the infrastructure and handle all previously identified challenges of identifying, managing and re-configuring resources as required.

A range of traffic grooming techniques using IaaS allowed for the optimum placement of routers, add-drop multiplexers and cyber-infrastructure computation nodes at renewable energy sites. It would also enable a range of different network topologies and traffic grooming strategies to meet demands of high end applications given the availability of nodes with sustained and plentiful renewable power.

Optical networks using Wavelength Division Multiplexing (WDM) provided high bandwidth and low processing loads and had significantly less power consumption than traditional electronic networks. Moreover modern optical networks had much great flexibility and could respond quickly to changes from users demands in terms of topology or outages and therefore were ideal for IaaS grooming techniques.

Routing and computing platforms at renewable energy sites might serve as common resources to a number of network operators and institutions. The platform might support a host of virtual or logical routers and optical cross connects, each independently connected to separate international or enterprise networks. Similarly, computation Grid platforms supporting a number of Virtual Machines, each running a different application for different users would be possible.

Infrastructure integrators could already integrate resources from different management domains into a single virtual management domain; thus federating despair resources into a single infrastructure. The resultant infrastructure as a result of the federation process was managed by one domain (service provider), although it did not have the ownership of the physical device. Therefore, traditional complex inter-domain issues were avoided, since the integrated resources would be under the same domain.

IasS virtualization capabilities allowed the optimization of the whole network resources, from L1 to L3 and allowed the composition of specific networks for a given application or service. So the usage of needed resources from each service provider brought optimization of resource usage and reduction of costs, since service providers only used resources needed for certain period of time.

Another feature that the IaaS Framework brought to this topic was that service providers could also be able to exchange resources among them when not used, and therefore not having active resources that are not used. By exposing internal processes as web services, these services could be exported or connected directly to external Service Oriented Architecture (SOA) applications. SOA helped eliminating the network of layers and allowed network components and services to be accessed horizontally as well as vertically and externally; in contrast, traditional telecom Network Management Systems (NMS) and Operations Support Systems (OSS) were vertical, centralized and transparent to applications.

6.5.7. Conclusions

As more devices were connected to the Internet, more organizations and individuals sought to share these hardware resources between users. This drove new market opportunities with the IaaS business model and had a definite impact on the way networks were operated and managed today. However, at this period in time, the separation that existed between the underlying hardware and software services, was still a barrier and additional research was needed in order to improve performance, reliability, and maintainability of these middleware systems to be used in production environments and to fully replace traditional management systems.

The work done provided for the first time an architecture solution capable to virtualise the optical network and offer it under an Infrastructure as a Service (IaaS) model. It also provided a first level of optical isolation in terms of management, although not for the physical non-linear characteristics of the fiber. These developments and the fact that they allowed the delegation of rights of use over the assigned wavelengths under the IaaS model, which meant that resources could be partitioned, abstracted and leased; opened up a new management model and approach where telecom operators could play different roles, service or infrastructure providers, or both. From now and on, third party, users or applications, would be able to take over the request provisioning of optical services while sharing the same physical substrate.

The work done on the IaaS Framework opened the door to new services and technology solutions. After years of activity behind the IaaS Framework, it evolved towards an open source framework capable to deliver similar services to what Openstack offers to the IT sector on IT services provisioning, but based on network services. Thus, a new framework architecture emerged which allowed the virtualisation of the network (independently of the technology) and so extending the concept of optical IaaS to the IP layer by providing IP Networks as a Service. It was called OpenNaaS and offered a new way to easily deploy IP networks across different IP domains. IP Network as a Service (IP Network Service) ended up being a key enabler for future flexible and stable e-Infrastructures. It extended the myriad of tool prototypes that provided point-to-point links to researchers. The tools and/or frameworks available at that time, while providing high bandwidth pipes to researchers, only addressed one side of the problem. Those researchers that wanted to create a virtual communities to address scientific problems were still connected to each institution's networks, and so, it was complex to directly connect them with high bandwidth pipes. It caused a number of issues such as security or routing integrity. Thus, IPNaaS solved it by creating a logically separated IP network (on top of the high bandwidth pipes and optical networks), or by using separate instances of virtualised routers, or a combination of both, and dedicating them to the virtual research community. So, in order to maximize the flexibility and convenience of IP Network Service, the users of the virtual community were able to modify the characteristics of their IP network by themselves (such as the addressing, dynamic routing protocols, routing policies, quality of service and so on).

Later on, by 2012, and a couple of years after OpenNaaS, a new open source framework emerged from the industrial community, which somehow covered the goals of OpenNaaS and many more (it is interesting to pint out the similar architecture they had). This framework is the so called OpenDayLight, and it is the framework currently supported by most of the telecom vendors (although with lots of cross-interests) for virtualizing the network under a SDN approach. Besides this, openstack also provides an API (called Neutron) to offer Network Service to the cloud, although it does not offer the flexibility that OpenNaaS had, since it offered network services and not the network as a service, which is rather different in terms of conceptualisation.

Moreover, it was clear that virtualisation was the trend to go and also the fact that energy consumption would reconsider and impact the research in ICT. The threat of Climate Change was very real and serious. Although the ICT sector's footprint was still small it was expected to grow significantly in the coming years if steps to reduce or eliminate its footprint were taken. Technologies like IaaS that enable a converged optical and grid network promise an architectural solution that not only reduced the ICT carbon footprint, but had the potential to enable a zero carbon infrastructure.

7. Service plane for resource virtualisation and provisioning

As described in the previous chapters, the demand for dynamic, user-controlled networks led in the past to the development of several Network Resource Provisioning Systems (NRPSs). While these systems served their purpose well in a single-domain scenario, many cases involved connections through multiple domains managed by different types of NRPS.

The context of multi-domain environments introduced several challenges that were not fully solvable by these NRPSs. Although there were different attempts to a solution, most of them did not consider obstacles such as heterogeneity of the environment, independency and privacy of different administrative transport domains, inter-domain topology abstraction, sophisticated types of reservations or a coupled integration with a Grid middleware. Based on this observation, the Harmony system explained in current chapter, was developed within the PHOSPHORUS project (funded by the European Commission by means of the 6th Framework Programme [104]). This research project lasted for two and a half years, starting October 2006. It aimed to interconnect existing single-domain NRPS solutions while focusing on issues in the context of a multi-domain environment

Two well-known examples among the whole set of related NRPS systems of the period and besides the aforementioned ARGIA, were: on the one hand, the Nortel proof-of-concept middleware called Dynamic Resource Allocation Controller (DRAC) –not available anymore- that allowed for an application-initiated configuration of transport network resources on an end-to-end basis. While on the other hand there was the Allocation and Reservation of Grid-enabled Optical Networks (ARGON) [105] system that enabled the integration of metro and wide area network resources into a Grid environment for both, the intra- and inter-domain provisioning of packet and circuit switched network resources. DRAGON used a peer inter-domain model supporting abstracted topology information sharing and inter-domain path computation, equivalent to the Path Computation Element Architecture.

Besides those three, there were others research activities aiming at achieving the challenges of multi-domain dynamic circuit provisioning. Originatd from the GÉANT2 project, the Auto-mated Bandwidth Allocation across Heterogeneous Networks (AutoBAHN) [106] architecture targeted at the needs of a multi-domain, multi-technology research community. Based on the so-called Inter-Domain Manager (IDM), the Auto-BAHN architecture defined an inter-domain network reservation mechanism based on a decentralised architecture for peer domain signalling.

As an achievement of the DANTE-Internet2-CANARIE-ESnet (DICE) collaboration [107], a web-based, inter-domain control plane was developed where Inter-Domain Controllers (IDCs) communicated in a decentralised way to provision end-to-end multi-domain network paths.

In Japan, the main goal of the G-Lambda project [108] was to establish a standardised web service interface between Grid resource management systems and the network resource management systems that also supported advance reservations.

In US the EnLIGHTened Computing [109] project focused on dynamic optical light paths between supercomputing sites that were created upon application needs. A domain manager allocated network resources by setting up circuits using GMPLS.

Although the previous citations represented a set of projects aiming at similar challenges, in real environments network resources would often be heterogeneous in type and independently controlled and administrated. Moreover, the integration of malleable advance reservations and co-allocation into such an environment was not addressed by any of these systems

Thus, and as previously discussed, by 2007 there was already a new generation of scientific applications emerging that coupled scientific instruments, distributed data and high-end computing resources, often interconnected via high-speed optical networks. Developed by collaborative, virtual communities, these Grid-based applications and networks were a hallmark of 21st century e-science. Many of these applications had requirements of one or more of these constraints: determinism (e.g. guaranteed QoS), shared data spaces, large transfer of data, and latency requirements that were often achievable only through dedicated optical bandwidth (lambdas). High capacity optical networking would soon provide the capacity to carry large amounts of data, but software tools and frameworks addressing end-to-end user and application-level access as well as provisioning-on-demand of such bandwidth needed to be developed in coordination with other resources, such as CPU and storage.

Highly-dynamic e-science driving applications required access to the following to maximize scientific discovery:

- Application-level middleware providing the execution environment of generic e-science applications including all necessary service abstractions exposing advanced Grid-like functionality combined with network services.
- A new generation of management and control planes with strong interaction with near-real-time resource (compute and network) scheduling, allocation and reservation.

Provisioning network resources on demand across multiple administrative and network technology domains was actually a key enabler of Grid services. For example, an end-to-end connection from Grid application source to destination would require setting up layer 3, layer 2, layer 1 and layer 0 network elements along the end-to-end path.

The ultimate goal within the research community was on defining an architecture and a middleware that enabled authorized end-to-end dynamic service provisioning across the European and worldwide heterogeneous network infrastructure, and with the ability to treat the underlying network as first class Grid resource. And so, enhance the level of integration between application middleware and the optical transport networks by advanced interworking between heterogeneous network domains and their applications environments. To do so, interfacing solutions that could facilitate vertical and horizontal communication between applications middleware, existing Network Resource Provisioning Systems and the Grid-GMPLS Control (G^2MPLS) Plane where envisaged. Integration of AAA mechanisms at various network and management layers would also be essential to ensure that stakeholder interests could be represented and enforced.

So far there was a tremendous amount of research and development in the Grid community in terms of Grid services infrastructure and Grid application development. However, there had been very little work done in the area of using network as a first-class Grid resource. Up to that period, there was no existing implementation nor architecture that could demonstrate the power of exploiting the optical network as a first-class Grid resource, with the challenges that would arise the fact of provisioning end-to-end light-paths across different management and control plane technologies spanning multiple administrative domains. To do so, research should be focused on providing a unified network/Grid infrastructure that could flexibly adapt to the demands of applications having strong, combined requirements on CPU, memory and storage resources as well as on the communication network. Therefore, applications could rely on a network infrastructure capable to adapt to the application, rather than having the application to adapt to the network.

Thus, focus was put on developing a new network Service and Control Plane where the network (lightpath) and Grid (computational, storage) resources could be provisioned in a single-step: network and Grid-specific resources controlled and set-up at the same time and with the same priority, with a set of seamlessly integrated procedures. From a user's perspective, this resulted in a real, node-to-node deployment of on-demand Grid services.

7.1. A SON architecture to use optical networks as a first-class grid resource

The architecture presented in Figure 33 and developed within the scope of the Phosphorus research project put the optical networks as a first-class grid resource. It addressed some of the key technical challenges that enabled on-demand end-to-end network services across multiple domains. This architecture [110] allowed the applications to be aware of their complete Grid resources environment -computational and networking- and its capabilities. It enabled and validated dynamic adaptive and optimised use of the heterogeneous network infrastructures interconnecting various high-end resources. With full support to on-demand service delivery across access-independent multi-domain/multi-vendor networks. To do so, the architecture enhanced solutions that facilitated vertical and horizontal communication among applications middleware and the network resources across different domains. These domains were managed by already existing Network Resource Provisioning Systems (NRPS), or domains that integrated a new Grid-GMPLS (G^2MPLS) Control Plane, both under a new AAA architecture that supported policy based on-demand network resource provisioning. This G^2MPLS extended ASON/GMPLS provided part of the functionalities related to the selection, co-allocation and maintenance of both Grid and network resources, by exposing upgraded interfaces at the UNI and E-NNI network reference points -i.e. G.OUNI and G.E-NNI-.

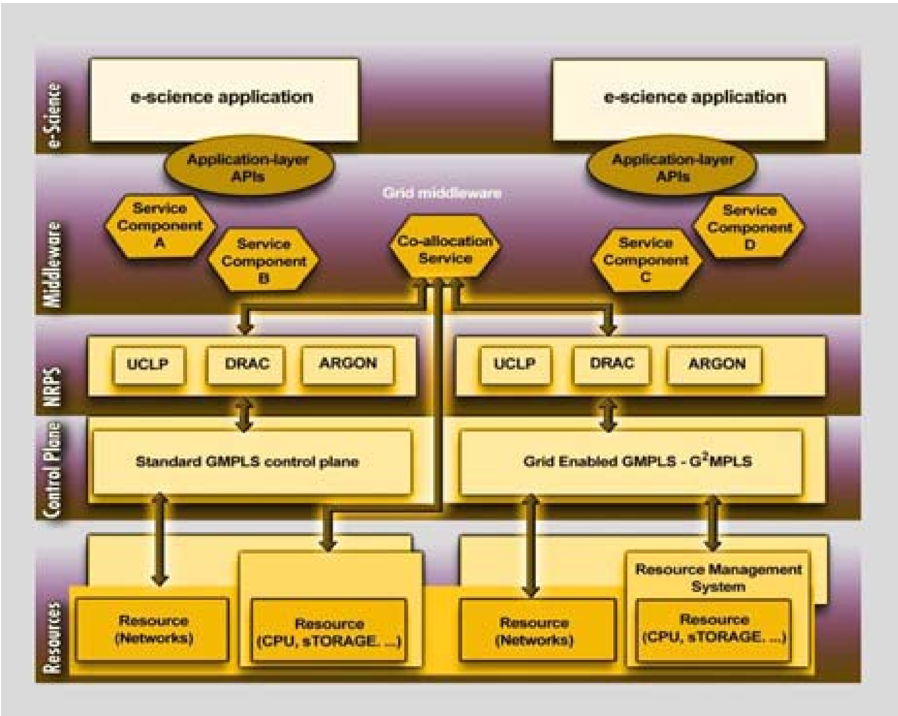


Figure 33. An architecture to provision optical networks as a first-class grid resource

Many of Grid applications had requirements such as determinism, shared data spaces and large data transfers, often achievable only through dedicated optical bandwidth. High capacity optical networking could satisfy bandwidth and latency requirements, and this architecture for end-to-end on-demand provisioning of network services, together and in coordination with other resources (CPU and storage), allowed the provisioning spanning multiple administrative and network technology domains.

This architecture concept made applications aware of their complete Grid and network resources, due to the fact the underlying network was treated as first class Grid resource. This concept of integrating applications, middleware and transport networks under a new AAA architecture was based on four planes: Service Plane, NRPS Plane, Control plane and the Data plane. These four planes facilitated the integration between the application middleware layer and the transport network layer. They had the following main characteristics:

Service plane: It contains middleware APIs to expose network and Grid resources and to create network connectivity services with advance reservations and policy mechanisms, in a global hybrid network infrastructure. Thus:

- Middleware extensions and APIs to expose network and Grid resources and make advance reservations.
- Policy mechanisms (AAA) for networks participating in a global hybrid network infrastructure, allowing both network resource owners and applications to have apart in the decision to allocate specific network resources.

Network Resource Provisioning plane: It contained the NRPSs and standard GMPLS control plane. Thus:

- An adaptation layer is required for interacting with the NSP and allowing multi-domain interoperability
- Implementation of interfaces between different NRPS to allow multi-domain interoperability with Phosphorus' resource reservation system.

Control plane:

- Enhancements of the GMPLS CP (G^2MPLS) to provide optical network resources as a first-class Grid resource.
- Interworking of GMPLS domains and G2MPLS with NRPS-based domains (UCLP, DRAC and ARGON).

Data Plane: All the transport network devices, Grid resources and links between all of them compose the data plane. The Harmony system was created to work over data plane infrastructures based on the circuit-oriented paradigm. For this reason, the network technologies considered were: Sonet/SDH, WDM or optical Ethernet for optical transmission systems; and Ethernet and MPLS for electrical systems.

This architecture allowed a serial of Grid applications to request by means of middleware a set of resources with its needed bandwidth and delay, depending on the application requirements (results were validated with the following past Grid applications: WISDOM – Wide In Silico Docking On Malasia-, KoDaVis – Collaborative Data Visualisation- and DDSS – Distributed Data Storage System-) [112].

The work and results presented below are the challenges addressed in the Thesis. This contribution was focused on the network provisioning part of the architecture with the main focus in the NSP and the NRPS layer. The NRPS systems to Integrate and manage within this architecture, were systems that had already been validated in other scenarios and provided by third parties. In fact, few changes were needed on those specific NRPS systems to make them operation within this framework architecture. These NRPS used were the following ones:

Network Resource Provisioning Systems (NRPS)	
ARGON ARGON (MPLS/GMPLS)	The Allocation and Reservation in Grid-enabled Optic Networks system was developed to manage resources of advanced network equipment as it is present in the German VIOLA test-bed. The advance reservation service of ARGON is able to operate on the GMPLS as well as on the MPLS level. It guarantees a certain QoS for applications for the requested time interval. This feature enables a Meta-Scheduling Service to seamlessly integrate the network resources into a Grid environment.
DRAC DRAC (Nortel)	The Dynamic Resource Allocation Controller system was developed by NORTEL and it is a commercial-grade network abstraction and mediation middleware platform, acting as an agent for network clients (users, applications, compute resource managers) to negotiate and reserve appropriate network resources on their behalf. DRAC uses client's QoS requirements and pre-defined policies to negotiate end-to-end connectivity across heterogeneous in support of just-in-time or scheduled computing workflows.
UCLP ARGIA ARGIA/ UCLP	The User Controlled LightPaths system was developed by CRC, Inocybe, i2CAT and UofO under the CANARIE support. It provides a network virtualization framework upon which communities of users can build their own services or applications. Articulated Private Networks (APN) are presented as the first services. The APN can be considered as a next generation Virtual Private Network where a user can create a complex, multi-domain topology by binding together network resources, time slices, switching nodes and virtual/real routing services

Table 8. NRPS description

7.2. Harmony: An Interdomain Broker

The Harmony[2] system came as a consolidated architecture for the provisioning of optical virtual infrastructures, and so it integrated functionalities of the NSP and the NRPS. It defined an architecture for a service layer between the Grid middleware and applications and the Network Resource Provisioning Systems (NRPS), and so, enabled users and applications to make dynamic, adaptive and optimized use of heterogeneous network infrastructures connecting various high-end resources with the ability to create point-to-point connections using network resources from several domains in a transparent way. It was a path provisioning architecture/system where both Users and Grid applications could book in advance paths and network resources over heterogeneous domains. It was defined and developed within the framework of the Phosphorus research project which demonstrated solutions that facilitated vertical and horizontal communication among application middleware, NRPSs, and an extended GMPLS control plane: the Grid-GMPLS (G^2MPLS). The project addressed some of the key technical challenges to enable on-demand, end-to-end network services across multiple domains.

The Harmony system was designed to meet a set of requirements. First, the system had to be multi-domain and capable of creating end-to-end optical paths in a seamless environment for the scientific personnel at the end points. The domains should be considered as a set of independent administrative transport domains controlled by different NRPSs within a heterogeneous environment. Also, they should or shouldn't accept the same policies when provisioning paths. Secondly, there were privacy and confidentiality reasons that forced not to share the internal topology to the other providers or to the public, also to avoid business disadvantages. Finally, considering again heterogeneity of the different involved NRPSs, there was the requirement of making signalling interoperable in or-der to provide on-demand, multi-domain circuit provisioning, given that each one of the NRPSs offered a different communication interface. Thus, the key features were:

- A **multi-domain path** computing and **provisioning system** where users and Grid applications could **book in advance** end-to-end paths and network resources with **AAI[3].** The Network Service Plane (NSP) performed on-demand path computation involving resources located at several independent, heterogeneous domains, with the capability of instantiating resources.

[2] It stands for harmonisation and orchestration

[3] Although AAI is not in the scope of this Thesis, the Harmony system implemented Authentication and Authorisation infrastructure based in the Generalised AAA Toolkit http://staff.science.uva.nl/~demch/projects/aaauthreach

- **Inter-domain network resource brokering** over Network Resource Provisioning Systems (NRPS).
- **Malleable multi-domain network resource allocation** Users or Grid Middleware could book two type of advance reservations: fixed or malleable. This last case provided a lot of flexibility to find a slot to serve reservations and maximizes the usage of the network resources.
- **Topology knowledge:** The knowledge of the global topology was restricted, due to reasons of confidentiality, to a set of basic information based on three main elements: the endpoint (user or border), the inter-domain link, and the domain-associated data.
- **Common language:** the Harmony Service Interface (HSI).

The proposed Harmony architecture (Figure 34), allowed the creation of complex resource reservations with in advance booking features, involving several NRPSs and/or a GMPLS control plane. A common Network Service Plane (NSP) for signalling was defined and hence, interoperability between NRPSs (DRAC, UCLP/Argia, and ARGON), the GMPLS control plane and the Grid applications/middleware was seamlessly achieved. The validity of the definition, design and implementation of the Harmony system was demonstrated in several international events in the period from 2007 to 2010, proving the feasibility to provide services across multi-domain and multi-vendor transport network test beds for research. The Harmony field test bed involved up to ten independent domains.

The Harmony system, apart from controlling multi-domain scenarios, introduced the network as a manageable resource in the Grid by means of the Harmony Service Interface (HSI). The Harmony system implemented resource co-allocation and scheduling capabilities (reservation service), able to reduce the probability of resource blocking, and providing inter-domain topology awareness services (topology service) by restricting the intra-domain topology information.

Due to the successful tests and public demonstrations performed, some of the work of the HSI was taken into consideration within the NSI (Network Service Interface) Working Group of the OGF (Open Grid Forum) [113].

The Harmony architecture was built over a SOA architecture and was compliant with the Web Service Resource Framework (WSRF) version 1.2. The depicted architecture of the system shows its main layers and components. Harmony was a multi-domain system that had the capability to create end-to-end optical and layer 2 paths in a seamless environment for scientific users at the end points. The integration between application middleware and optical transport networks was based on three main layers (planes), the NSP (Network Service Plane), the HAL (Harmony Adaptation Layer) and the Harmony NRPS Layer which integrated the three NRPS presented in the previous section.

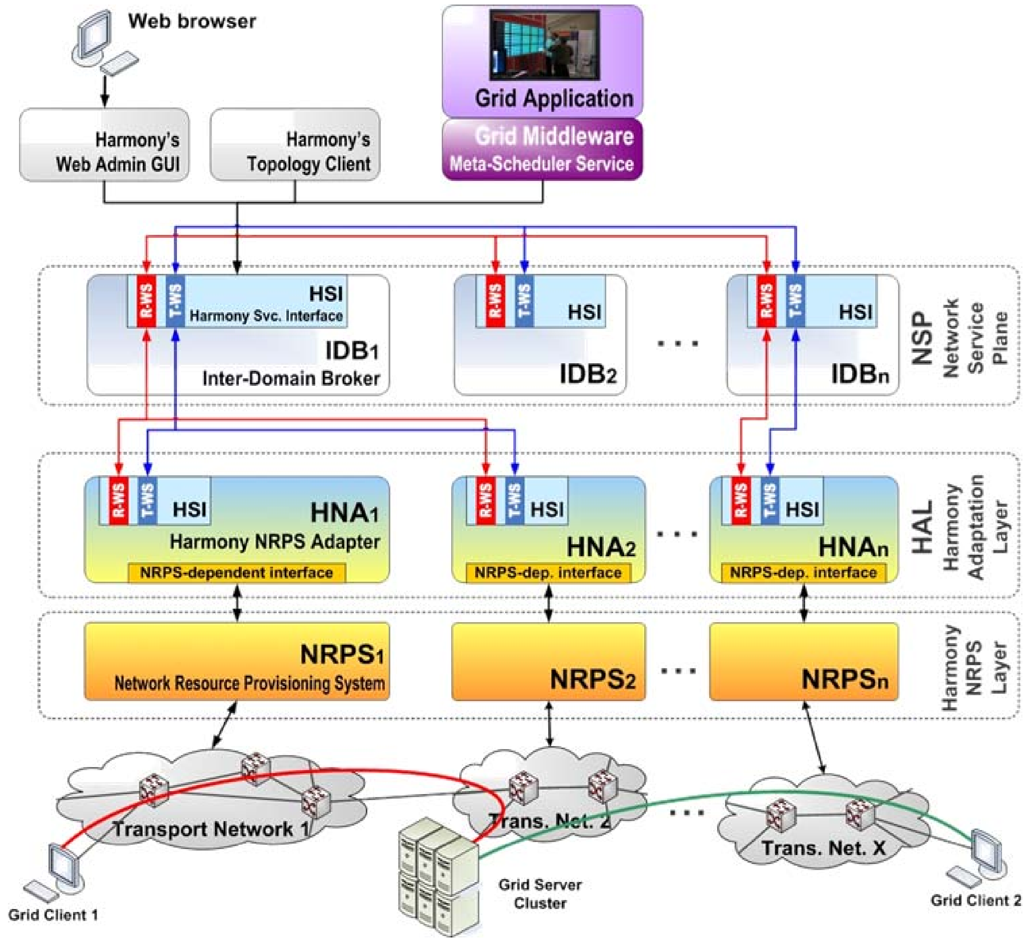


Figure 34. Harmony architecture

Moreover, and although it is not included on the architecture figure (Figure 34), but explained along the chapter, Harmony also provided a Thin NRPS entity which was able to communicate with the Optical User to Network Interface (OUNI) of a Generalized-MPLS Control Plane (GMPLS-CP) – not considered an NRPS itself– and performed signalling operations and path management functions.

Figure 35 shows a simple Harmony set up with one IDB at the NSP and one NRPS which was connected to the NSP by means of one HNA (only one HNA allowed per NRPS). This simple workflow facilitates the understanding of the actions performed by the different building blocks. Names in the arrows are detailed hereby:

- a1. Resource reservation requests Client-to-IDB (administrator or normal user or middleware).

- a2. Topology requests Client-to-IDB (administrator only).
- b1. Resource reservation requests to NRPS (normal operation).
- b2. Topology requests IDB-to-IDB within the NSP (topology exchange).
- b3. Resource reservation requests IDB-to-IDB within the NSP (topology exchange).
- c1. Topology requests HNA-to-IDB within the NSP (topology exchange).
- c2. Resource reservation requests HNA-to-IDB (request forwarding).
- d. NRPS-dependent interface.
- e. Network device dependent interface

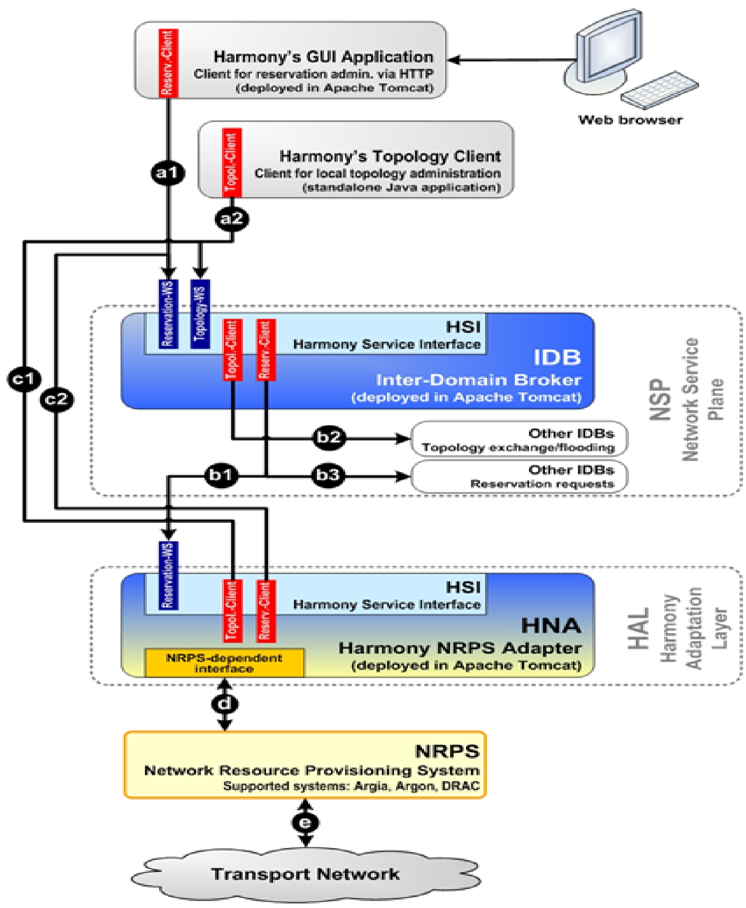


Figure 35. Harmony system set-up

Figure 36 exposes the Harmony architecture but putting emphasis on the different types of possible architectures within the Network service Plane.

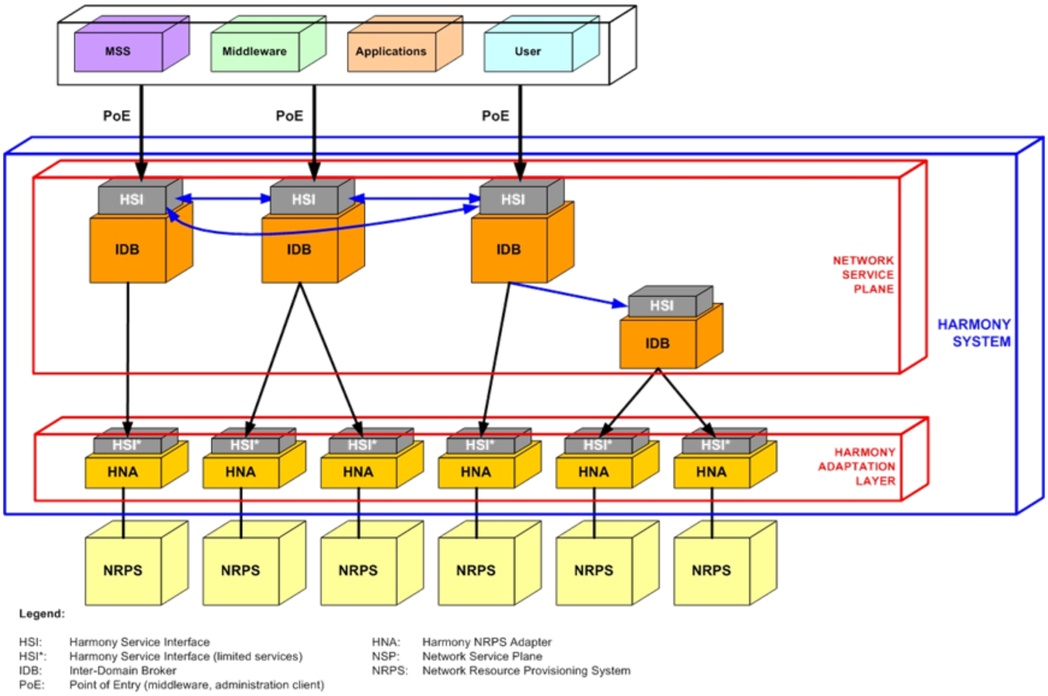


Figure 36. Harmony architecture interactions

To be more precis, Figure 37 shows the different type of architectures supported by Harmony. It started with a centralized approach and evolved towards a distributed approach. A more specific detail and analysis of its performance analysis is presented in section 7.2.3

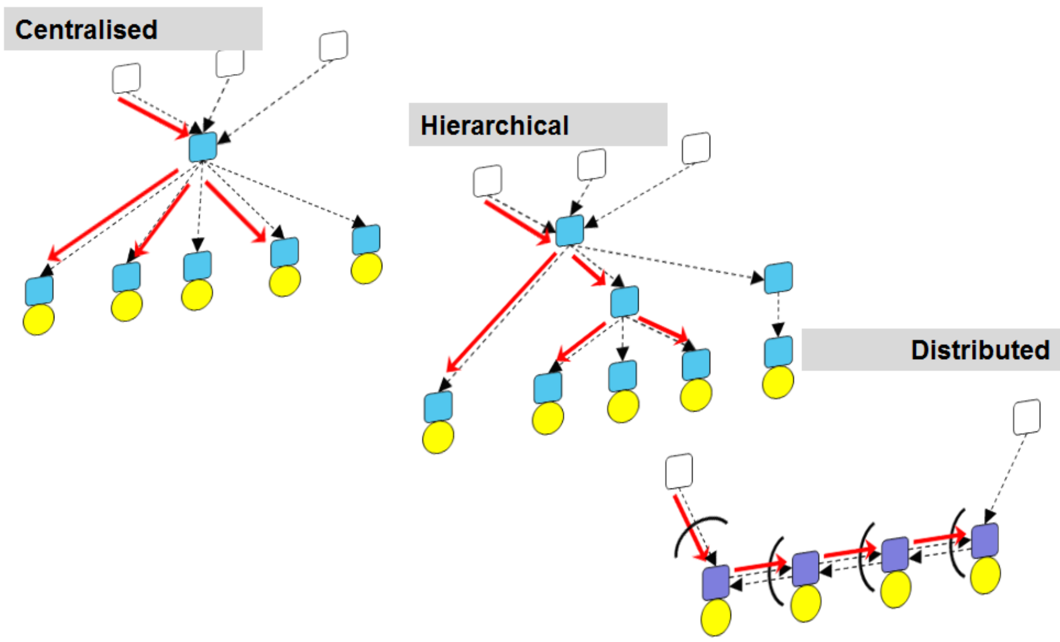


Figure 37. Types of possible architecture deployments for the Harmony system and its NSP

7.2.1. The NSP and the HAL

7.2.1.1. The Network Service Plane

The Network Service Plane (NSP) was the highest plane in the architecture of Harmony. Administrative users, Grid middleware or applications accessed Harmony by invoking the services offered by this plane. The NSP was populated with Inter-Domain Broker (IDB) software entities. Each IDB was responsible of managing and brokering the network resources offered by its underlying administrative domains. Thus, one IDB could control one or more Network Resource Provisioning Systems, also called simply domains in Harmony. As domains were represented only by their endpoints and the intra-domain topology was hidden, the path computer of the NSP could not calculate intra-domain paths. Rather, it generated a list of endpoints defining an inter-domain path and received, from the NRPS, the following results when an intra-domain path was requested: 0-path available, 1-source occupied, 2-detination occupied, 3-source and destination occupied, 4-no path between source and destination. The namespace of endpoints was chosen in such a way that the corresponding domain could be identified from the name of the endpoint. Interdomain connections (connections between domains) were administrated within the NSP.

The NSP was technology agnostic, i.e. paths were stitched at the domain boundaries and all endpoints were considered to operate at a common technology layer (Ethernet). Multilayer and multi-region issues were topics for further work. Regarding the path computation, it only took into account the interdomain links defined in the NSP by an automatic process. So, the calculated end-to-end path was split up into several intradomain paths. For each intradomain path, the reservation web service of the corresponding domain was called. Then, the NRPS calculated possible paths within its domain and returned the availability of the requested resources of the path. In case at least one of the domains returned that there was no path available, the occupied or unavailable resources were pruned from the graph of the path computer and a new end-to-end path was computed. To minimize the number of calls to the reservation web service of each domain, it was assumed that for each path, each domain was traversed at most one time. This assumption was fair as the interdomain links were scarce. The first prototype implemented assumed that the full bandwidth of an interdomain link was not shared among different paths.

Finally, when performing the co-allocation of the desired resources for a connection, it was straightforward to see the need for “advance reservations”, that was to say, reservations which had the start time in the future.

Actually, an efficient resource co-allocation required the possibility to query the availability of resources (as stated before) without making a (hard) reservation. In order to achieve this functionality, an availability request for the resources involved in the reservation was previously performed. A negative response to an availability request also included an alternative start time for the queried resources. This allowed the inter-domain scheduling process to be more effective by querying the involved management entities with a new start time which indeed increased the probability of success in the co-allocation.

Regarding the authorization and accounting mechanisms, the NSP itself did not provide complex authorization or accounting mechanisms. Rather it forwarded attributes that were contained in the incoming message from the middleware to the involved NRPS adapters and vice versa. The authorization process was in the scope of each NRPS and the middleware authorization tickets were transparently communicated through the NSP. It was assumed that each domain had its own policy and attribute database. The NRPS adapter mapped the global attributes to local ones or local user accounts. In case the global and local attributes were identical this mapping was reduced to the identity function.

Within this architecture, a Meta-Scheduling System (MSS) was responsible for the co-allocation of Grid and network resources managed by different administrative entities. A MSS did not distinguish between Grid and network resources. It treated the network in the same way as a grid resource, whose availability could be queried and reserved in the same way as processing power on a computing cluster. In a single-domain environment, a Network Resource Provisioning System (NRPS) offered these services to the MSS. A NRPS was a system that had full knowledge about the underlying network's topology and the utilization of resources at different points in time. It became a global broker responsible for creating the end-to-end service connections through different NRPS systems. The NSP was responsible for dealing with the NRPSs in order to provide end-to-end paths, manage AAA issues, and keep track of the resource utilization and to coordinate the different actions done. Moreover, the NSP provided interoperability towards the G^2MPLS architecture developed by other research activities in parallel with this Thesis.

The key points and benefits provided by the NSP were:

- *Its ability to create point to point, or point to multipoint connections* using resources from several domains in a transparent way. The proposed solution speeded up the creation of complex connections with advance reservation features involving several systems by making them interoperable.

- *Its simplification of AAA management:* once the user was authenticated, he could use any of the services offered by the Network Service Plane. Moreover, his credentials were automatically translated to the local credentials of the systems involved in the service.
- *The introduction of the Advance Reservations concept*. Users or Grid applications were able to program fixed, deferrable or malleable resource reservations with one or more connections.

The interoperability performed by IDBs, in conjunction with the new interfaces developed, was understood as the capability to create advance reservations. An advance reservation was already defined (as commented on chapter 3.3.5.1) as a reservation of network resources for a future time. This allowed users (Grid applications) to programme/request connections in the future that would be set up automatically as scheduled. These reservations were between two end points and could be located within the same or in a different domain at the network level, and were also defined by a specific bandwidth and a minimum delay time.

The consideration of a centralised approach in the NSP positively contributed to enabling multi-domain reservations as far as: 1) an abstracted image of the different domains and the inter-domain links/connection points among them was stored in only one entity in the NSP; and 2) the NSP kept track of the occupation and resource reservations of the different domains and their inter-connections. Additionally, this central entity allowed requests from the Grid middleware, completing in a simple way the management architecture from Grid applications to high performance (optical) transport networks.

However, the distributed approach for the NSP brought to a premier position the reliability of the whole NSP, since the failure of a single IDB did not turn down the full NSP, but only the part of the network (NRPSs) this IDB was in charge of. Moreover, the distributed NSP model reduced the overall signalling load of the NSP as far as each IDB entity could sort out where the asked resources were located, and thus, directly forward the request to the correct IDB. By contrast, a centralised NSP did not allow direct request forwarding, since every IDB had only knowledge about its underlying NRPSs (the NRPS were not aware of the request addressing in the NSP) and its unique parent IDB. This way, a child IDB had only the possibility to forward the request to its parent and let this higher level of hierarchy decide what to do with the request: serve it or forward it to upper layer again, recursively.

7.2.1.2. NSP services

7.2.1.2.1. Advance reservations

Harmony's NSP provided different types of advance reservations, so that the service was provided on top of the inter-domain topology abstracted by the service plane. A basic form of an advance reservation was defined as follows: The request received at $t_{arrival}$, is admitted and starts at t_{start}. Furthermore, the usage phase (duration) is limited by t_{end}.

Harmony supported three types of Advance Reservations for the NSP:

- *Fixed Advance Reservations:* in this type of reservation the user had to specify the bandwidth along with the reservation start and end times. A fixed advance reservation request for a single connection was depicted on the left side in Figure 38 and was defined as tuple (t_{start}, t_{end}, s, d, C) where $t_{start} < t_{end}$. The reservation started at t_{start} and ended at t_{end}. The endpoints of the connections were specified by s and d. C represented additional resource constraints, which were the required capacity and delay.
- *Deferrable Advance Reservations:* in this type of reservation the user had to specify the bandwidth, the duration of the connection and the earliest and the latest point in time when the connection was useful. This type of reservation helped to find gaps to serve reservations at a fixed bit rate. A deferrable advance reservation request had a degree of freedom in the time domain. In particular, time-related parameters defined a range of possible values to establish the reservation. The lifecycle of a deferrable reservation was given on the right side in Figure 38 and was defined as tuple ($t_{release}$, $t_{deadline}$, Δt, s, d, C) where $t_{release} + \Delta t < t_{deadline}$. The reservation could start at $t_{release}$ and had to end before $t_{deadline}$. The usage phase was specified by the duration $\Delta t > 0$. Compared to a fixed advance reservation, the parameters t_{start} and t_{end} ($\Delta t = t_{end} - t_{start}$) could be determined by the NRPS.
- *Malleable Advance Reservation:* in this type of reservation the user had to specify the maximum and minimum bandwidth allowed, the amount of information to be transmitted and the earliest and the latest point in time when the connection would be useful. This reservation provided a lot of flexibility to find a slot to serve reservations at a constant bit rate between the minimum and the maximum allowed throughput. A specification of the exact transmission rate could be omitted when a fixed amount of data had to be transmitted. By joining the time and resource constraints, the reservation system could find the most efficient solution for the requested transmission. A malleable reservation request was defined as tuple ($t_{release}$, $t_{deadline}$, s, d, S, C) where $t_{release} < t_{deadline}$. The endpoints of the connections

were specified by *s* and *d*. *S* determined the data size (trans-mission rate and time product) and *C* represented additional constraints. Typical constraints were lower/upper boundaries for the transmission rate.

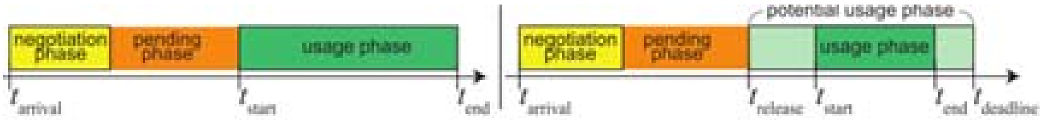


Figure 38. Life cycles of a fixed (left) and a deferrable (right) advance reservations.

7.2.1.2.2. Multidomain Path Computer

The Network Service Plane in Harmony provided multi-domain path computing over the abstracted topology based in border points (or endpoints) of the network domains and the inter-domain links. The path computation element in Harmony implemented Dijkstra's algorithm in order to find the shortest end-to-end path between the involved resources, based on this abstracted view. The chosen approach was suboptimal, since Harmony only sought the shortest path in the abstracted topology view of the multi-domain scenario, regardless of the internal topology or limitations in each domain. In case an NRPS reported any resource was not available, this resource was blocked in the service plane and the path computing tried to find another path.

The interfaces of the Path Computer are shown in

Table 9. When a new instance of the Path Computer was created, it read all border Endpoints of all domains and all inter-domain links. After that, one or more services, with begin and end time, were added. Path computation requests were grouped on a per service basis. When calculating a path, blocking of resources (i.e. resources that are in use by another service that is at least partly overlapping in time) was taken into account. All paths belonging to a specific service were calculated at the same time but could be selectively read from the Path Computer one by one. Each path was returned as a list of tuples of Endpoints. Each tuple consisted of two Endpoints of the same domain. The calculation of the intra-domain parts of a path was left to the Path Computer of the domain's NRPS. If an NRPS returned that one or both Endpoints were not available, or no connection between the endpoints was possible for this request, the unavailable resources could be pruned for this path computation and a new set of paths was computed.

The Path Computer had the following interfaces:

Interface	Parameters	Exceptions	Result type	Description
addService	long startTime long endTime int serviceId	InvalidServiceIdException	void	Add a service to the path computer's state. The start and end times can be given in arbitrary time units since they are only used to calculate which services overlap in time and which do not.
addConnection	Endpoint source Endpoint destination int serviceId int connectionId	EndpointNotFoundFaultException DatabaseException InvalidServiceIdException InvalidConnectionIdException	void	For a specific service add a connection to the path computer's state from Endpoint source to Endpoint Destination.
computePaths	int serviceId	PathNotFoundFaultException InvalidServiceIdException	void	Compute all paths for all connections added for a specific service.
getPath	int serviceId int connectionId	InvalidConnectionIdException	List<Tuple< Endpoint, Endpoint>>	Get shortest path for a certain connection of a specific service.
pruneEdge	int serviceId int connectionId Endpoint src Endpoint dst	EndpointNotFoundFaultException InvalidServiceIdException InvalidConnectionIdException	void	Prune an intra-domain edge from the internal topology graph of this path computer instance.
pruneEndpoint	int serviceId int connectionId Endpoint endpoint	EndpointNotFoundFaultException InvalidServiceIdException InvalidConnectionIdException	void	Prune an Endpoint from the internal topology graph of this path computer instance.

Table 9. Path computation interface

7.2.1.3. The Harmony Adaptation Layer (HAL)

The NRPS Adapter (

Figure 35), called the HNA (Harmony NRPS Adapter), is the adaptation layer between the NSP and the NRPS itself, and it is located at the HAL (Harmony Adaptation Layer). It contained a common communication part for all the NRPSs that consisted of the required Web Services to receive the calls to the operations at the NRPS level. Each adapter translated the incoming calls to the invocations implemented by the NRPS, and redirected them to the underlying system. The NRPS Adapter was located at the top of each NRPS. It implemented all the required operations to invoke the reservation functions of the NRPSs. The Adapter had a common part for all the NRPSs and another specific one for each kind of NRPS (ARGON, DRAC and UCLP). The common part mainly consisted of the communication interface for the NSP. Therefore, each kind of NRPS implemented its own adapter to translate the common requests of the reservation WS to the corresponding function of the NRPS.

The way to communicate between the layers, either in one way or the other, was through Web Services. The NRPS Adapter implemented an interface based in Web Services that was accessible by the NSP. The NSP, in turn, implemented the Web Service to allow the notifications. The NSP knew detailed information about all the controlled NRPSs, including the location of the Web Services in order to be able to invoke their operations.

The Web Service at the adapter consisted mainly of the operations implemented by the Reservation-WS, which offered the following functions [119]: Availability request, Reservation request, Cancel reservation, Status request, Bind request, Activation request, Complete Job, Cancel Job and Retrieve features.

Moreover, the adapter implemented a registration service in order to register the Domain and its endpoints automatically in NRPS start time. It was performed through the "AddDomain" operation of the Topology Web Service of the Network Service Plane instance that connected the domains. It registered its domain's border Endpoints ("push model") by calling the "addEndpoints" method of the same WS of the NSP. This service provides the NSP with all the required information about the Domain, the NRPS and the physical resources. The Adapter could get the Endpoints registered in the NRPS periodically and send them to the NSP in order to have an updated version of the underlying topology.

The resource provisioning capacities of an existing domain controller were (much) less or dissimilar compared to the NSP. In those cases developing the Adapter was much more a matter of implementing the reservation interface operations in contrast to the translation of existing operational features. It contained the same operations and communication functionalities, but it was implemented from the NSP point of view. Furthermore, some extended features were supported only by a subset of the systems. The supported features of a system could be queried using the "getFeatures" operation of the Reservation Web Service. The NRPS adapters interacted with the NSP (IDB) in order to enable interoperability among IDB.

7.2.1.3.1. Thin NRPS

The thin NRPS was not depicted in the architecture of the Harmony system, but it allowed Harmony to interoperate with standard GMPLS-CP. A specific interface was built. This was the so called Thin NRPS and GMPLS-WS within the HNA (also called GMPLS driver) (Figure 39). This interface was realized by two modules:

- The GMPLS driver acted as an interface between an NRPS and the GMPLS control plane. It offered a general web service, which was used to create, delete and monitor paths for different GMPLS implementations.
- The Thin NRPS was a network resources provisioning system for domains with a GMPLS control plane. It provided a reservation web service, which was used by the NSP to reserve, create and delete network connections via the GMPLS driver.

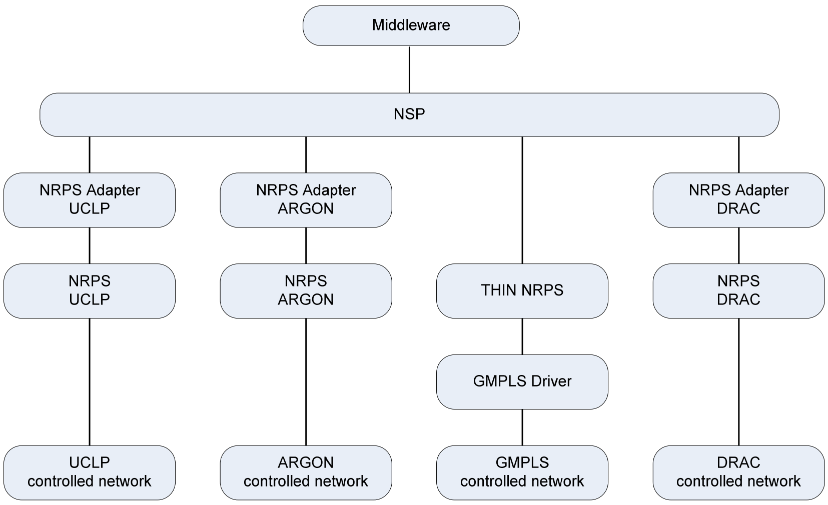


Figure 39. Thin NRPS and GMPLS driver within the PHOSPHORUS context

The Thin NRPS was a network resources provisioning system for domains with a GMPLS control plane. Since this interface had to deal with a standard GMPLS CP without advance reservation functionalities; this functionality was already provided by the Thin NRPS. However, it assumed the constraint that resources in a GMPLS domain were always available, unless another reservation was overlapping.

The Thin NRPS (Figure 40):

- Provided a reservation web service, which was used by the NSP to reserve, create and delete network connections. It used the same web service interface as the NRPS adapters.
- Provided a notification receiver interface, which enabled the GMPLS driver to send update information about endpoints and paths.
- Acted as a client to the topology manager web service of the NSP, e.g. it will add a domain and provide information about endpoints.
- Acted as a client to the GMPLS driver web service for creating and deleting network connections and for obtaining information about endpoints and existing connections.

Internally, the Thin NRPS provided the following functions:

- To handle reservation requests from the NSP; check, if the reservation could be granted and keep track of the reservations in a database.
    - As the Thin NRPS currently got endpoint information from the underlying GMPLS driver, it had no knowledge about internal links and their usage. This meant, that in case of advance reservation no checks for availability of bandwidth on the internal links could be performed. Therefore the Thin NRPS would only check, if endpoints were available and if there were conflicts concerning the usage of endpoints within overlapping reservations.
- To schedule creation and termination of network connections using the GMPLS driver.
- To register its domain at the NSP domain manager, retrieve user and border endpoints from the GMPLS driver and forward border endpoints to the NSP domain manager.
- To Handle endpoint updates from the GMPLS driver and inform the NSP domain manager.

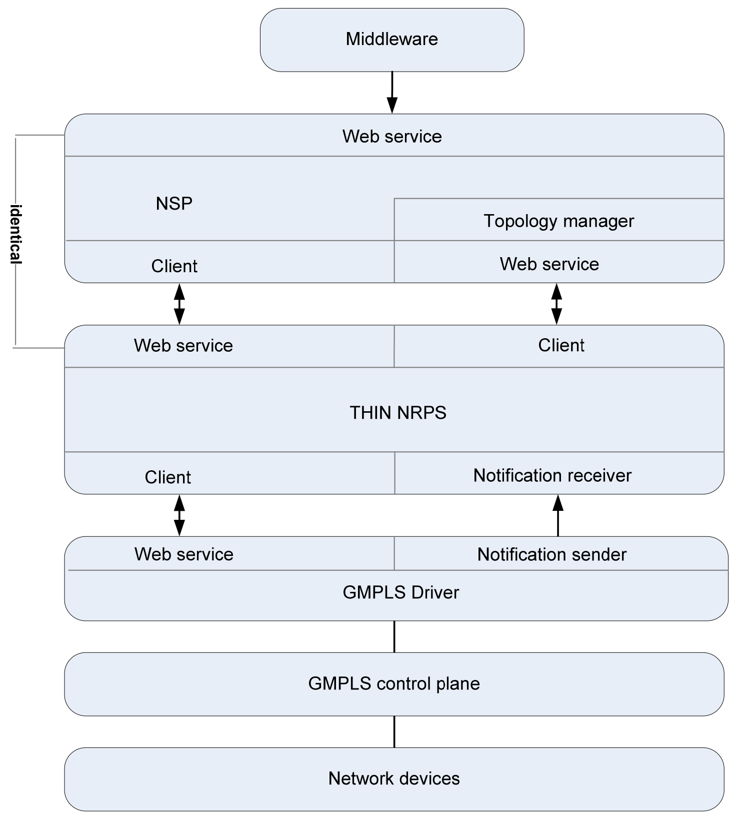


Figure 40. Thin NRPS and other PHOSPHORUS modules

In contrast to a normal NRPS the Thin NRPS had only a limited functionality. Only fixed reservations were supported, because no topology information about links and their usage was available. Therefore the following functions were currently not supported:

- Deferrable reservations, Malleable reservations, Unidirectional or point-to-multipoint connections, MaxDelay, DataAmount.

If one of these functions was specified, an exception was thrown. The Thin NRPS also provides a notification receiver interface, which enabled the GMPLS driver to send update information about endpoints and paths. Moreover, it acted as a client to the topology manager web service of the NSP (e.g. it will add a domain and provide information about endpoints) and as a client to the GMPLS driver web service for creating and deleting network connections, and for obtaining information about endpoints and existing connections. The GMPLS driver provided the following services and notifications described in

Table 10.

Path creation service	Creates a point-to-point path between two endpoints specified by TNA addresses.
Path termination service	Tears down a point-to-point path which has been set up by an NRPS.
Path monitoring service	Provides status information about the specified path.
Path discovery service	Retrieves any established point-to-point connection in the controlled network. It is needed to refresh the NRPS view on the network in case of losses or reboots.
Endpoint discovery Service	Retrieves information about any endpoint in the controlled network. It is needed to deliver the available endpoints to the NRPS during system initialization and to refresh the NRPS view on the available endpoints in case of memory losses through system failures or reboots.
Registration service	Registers the NRPS to receive messages from the web service in case of path status changes or endpoint changes.
Path delete notification	Informs all registered and authenticated systems, if a path is no longer available.
Endpoint update notification	Informs all registered and authenticated systems when endpoints have been removed or added to the controlled network domain.

Table 10. Basic definition of the GMPLS driver services and notifications

The GMPLS services were offered via a web service to the clients. A client could be an NRPS or a user for test purposes, and could access the services via a Java application. All Information about paths, endpoints and devices were kept in a MySQL database, which was only accessed by the core component of the GMPLS driver. It could be administered through any database front-end like phpMyAdmin. GMPLS operations were initiated and controlled by vendor specific modules.

7.2.2. System interfaces

One of the key developments for the NSP were the system interfaces, which allowed and facilitated the system interoperability between the different layers.

7.2.2.1. The HSI interface

The Harmony service interface (HSI) was the element of the system that enabled interoperability between the capabilities, functionalities and services being offered by the resource broker and the outer world. It also enabled the communication between entities within the service plane. Moreover, the functionalities provided by this interface were also being considered for standardization purposes within the NSI (Network Service Interface) working group of the OGF. The HSI was responsible for offering to the Grid resource: co-allocation, scheduling and network topology related services. Thus, Grid applications could access directly the NSP and its services through the service interface. This service interface was common for both the network service layer and the adaptation layer and it was web-service based[4].

The design of the interface took into account the modularity and the services nature in order to build an easily-maintained module. Thus, the HSI (

Figure 41) was composed of three main modules: Reservation Web Service (R-WS), Topology Web Service (T-WS), and Notification Web Service (N-WS). Moreover, there was an extra module within the HSI, called Common Types, which contained all the common data types used by the other modules.

The HSI was built over an SOA model and was WSRF v1.2 compliant. Therefore, the messages within each module were defined in the Web Service Description Language (WSDL), while the data types were defined in the XML Schema language or definition (XSD).

Figure 41 depicts the modules designed within the service interface. A more detailed description of the three modules is as follows:

Reservation-WS: Reservation Web Service was the element of the service interface that enabled the Grid Middleware to create, cancel and query advance reservations,

[4] Web services are software implementations of abstract service interface descriptions and they are defined within WSDL (Web Service Definition Language) description files. A WSDL describes the web service's functionality by means of a specific set of operations and the data type these operations can handle. The data types are separately defined through an XML Schema language within XML Schema files and referenced by the WSDL files. Once the web services are deployed they can be accessed through a web service client which references the same WSDL and XML Schema files that the web service is based upon. Thus the communication channel between client and service is established. Software code for web services can for the greater part be automatically created using the WSDL and XML Schema files by specific software tools

both malleable and fixed, in the network. The functionalities offered to the middleware, by means of the methods that could be invoked, are presented in

Table 11. The reservation interface was essentially the same as the interface that allowed interoperability for the NRPS, i.e. the interface that integrated an individual NRPS into the NSP. This was the case because both the NSP and the NRPSs allowed resource provisioning, however, the functionality provided by the reservation interface of the NSP was of a higher level coordinative nature. Typically, within the reservation interface the individual domain resource reservations were handled and presented to the higher level entities as one set of reservations.

Functions	Description
availability request	Checked the availability of the resources required for connections with endpoints in the same domain or in different domains. The resources were not yet reserved, instead a response indicating whether the reservation would be feasible or not was sent.
reservation request	Reserved network resources for one or more services. The reservation could contain fixed constraints or malleable constraints. The requesting entity received a response containing the result of the request and an identifier of the reservation.
reservation status request	Queried one or more reservations' status. The possible status of one reservation previously made in the network were: pending, tear up in progress, active, tear down in progress, completed, cancelled by user, and cancelled by system.
cancel reservation request	Deleted a reservation. Once the reservation was cancelled, the NSP released all the resources involved in the connection and tag them as available again.

Table 11. Reservation-WS methods offered to the Grid Middleware

Topology-WS: Harmony required acknowledgement of the resources that were under the control of the system itself. Therefore, the system had to store, retrieve, modify and delete the resource-related information according to the topology of the controlled network domains. Topology-WS enabled the system to carry out all these tasks. Thus, it enabled the system to know which resources were under the control of the system and how they were organized. The topology interface supported the adding of domains, domain endpoints and links between domain endpoints to the NSP thus enabling the NSP to compile a multidomain spanning topology.

Table 12 showed the methods available in this WS.

Functions	Description
add/edit/delete/get domain	These methods allowed the system to add a domain controlled by an NRPS, delete it, update the domain or get all the domains controlled by the system, that was, get all the domains present in the Inter-domain Broker database.
add/edit/delete/get endpoint	All of these methods were used in order to manage the endpoints within Harmony. Thus, all the endpoints, border or

	user, belonging to the different domains controlled by the system could be managed.
add/edit/delete/get link	These set of methods managed the links that inter-connect the domains controlled by the system.
addOrEditDomain	While the previously introduced operations were intended for manually editing the information stored in an Inter-domain Broker, this operation was used for the automated topology exchange. It was used by sub-domains to register with their parent domains by means of the flooding algorithm, which distributes topology information between peer Inter-domain Brokers when the NSP is operating in distributed mode.

Table 12. Topology-WS available methods

Notification-WS: Notification Web Service was the component of the HSI responsible of the event notification management. N-WS eliminated the need for the system to be polled periodically. When a connection was aborted, the higher layer entity that created the corresponding reservation would be notified.

During the development of this architecture there were several challenges beyond the end-to-end and the multi-domain path provisioning. Thus, and in order to make Harmony interoperable with other research activities from third parties, it was necessary to build a gateway. It was composed of the Harmony service interface on the one hand and the interface of the other system on the other hand. The principle behind the integration was a translation principle; this meant that the gateway maped the requests in one system-language to the other system-language, making possible the communication between the two different systems.

The NSP was a plane populated with one or more entities called Inter-Domain Broker (IDB). Other independent software entities existed on stage, such as the HNA (Harmony NRPS Adapter). HNAs were set on top of each NRPS system to interoperate with the NSP, they depicted a thin layer, the Harmony Adaptation Layer or HAL.

At the upper layer, but outside of the Harmony architecture, there were three types of clients. These clients were the following:

- *Harmony's HUI Application*: the client for reservation administration via HTTP.
- *Harmony's Topology Client*: the client for local topology administration.
- *Grid Middleware*: the client that would set-up the provisioning requests.

The network topological information stored in the database was managed through the Topology-WS interface of the NSP. This interface could be accessed by any client. For a human user, the Topology Client, a graphical user interface (GUI) was developed.

The Topology Client allowed the user to create, query, modify and delete any topological information of the network (add, modify, query and delete domains, endpoints and interdomain links). The GUI was implemented in Java Swing and contained a proxy to communicate with the Topology-WS of the NSP.

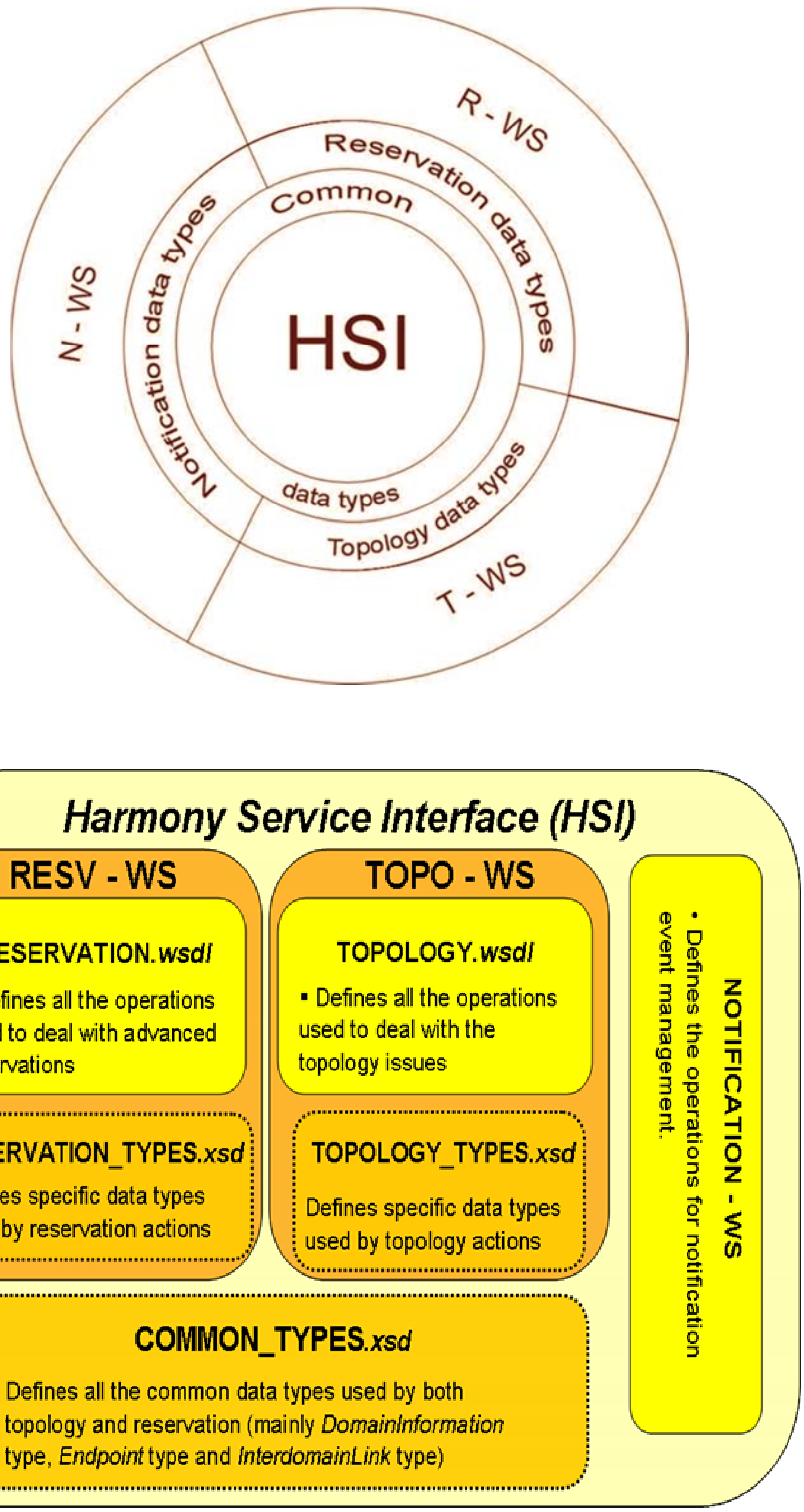


Figure 41. HSI internal architecture

7.2.2.1.1. Topology modification

The natural way to add the domains and endpoints was not through the GUI (although the user could do it this way as well), but by means of an automatic registration process located at the NRPS Adapter. When the adapter was initialized, a servlet was executed that contacted the NSP in order to send to it the information about the domain (identifier, description, WS endpoint references...). Once the adapter was registered, the NSP had the information required to send requests to the NRPS. After the registration, the adapter started a process that updated the NSP with the endpoints of the domain controlled by the adapter. This way, the NSP was updated periodically with the information of the local topology of each one of the NRPSs.

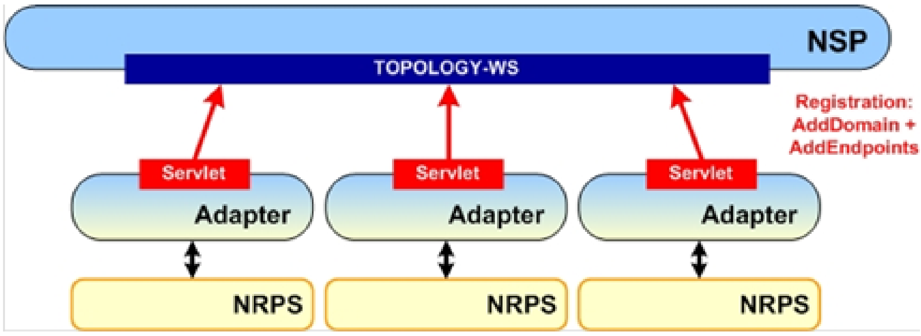


Figure 42. Figure 4.2: Domains registration

Physical interdomain links were not automatic, and had to be put manually through the Topology Client. Basically because the NRPSs did not have internal interconnectivity information. Thus, the user had to select the endpoints of the link manually and insert some related information to add the link to the topological database.

7.2.2.1.2. Network resource availability query

An availability query was triggered by the reception of an IsAvailable request by the Reservation-WS. Upon reception of an IsAvailable request, the corresponding routine in the ReservationSetupHandler class was called. Although a reservation would actually not be made, the task to be solved was very similar to an actual reservation, the only exception was that the queried resources would not be reserved. Also, alternative start time offsets could be returned, while an actual reservation would only either succeed or fail.

The requested services were used as input for the getAvailableServiceList routine that was also used for CreateReservation requests. This routine queried the

PathComputer for paths for all of the requested connections and split the single multidomain request to multiple single-domain requests, one for each of the involved domains. These requests were handed to the NRPSManager that took care of sending these requests to the NRPSs and collecting the corresponding replies.

If the requested resources were not available in one or more of the involved domains, they were pruned from the PathComputer instance (section 7.2.1.3). In this case, the domains replied with alternative start time offsets, the latest of which was recorded for later use if no suitable path could be found. Then, the PathComputer was queried again for an alternative path.

This process was repeated until either a suitable path was found, or until after pruning many resources no path was available for one or more connections. In the first case, the requestor was informed that the resources were available. In the latter case, the requestor was informed that they were not available, and the earliest alternative start time offset of those recorded as described above was reported as alternative start time offset.

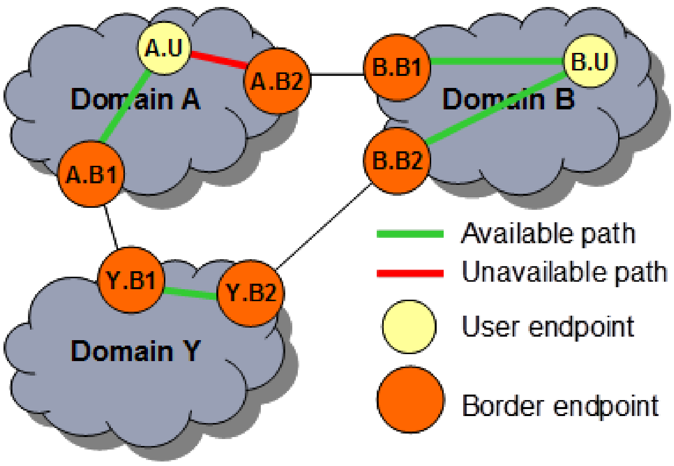


Figure 43. Example scenario for reservation setup

To illustrate this, consider the scenario sketched in Figure 43. The NSP is processing an IsAvailable message for a connection between A.U and B.U. The path computation first returns the two intradomain connections A.U-A.B2 and B.B1-B.U. According IsAvailable messages are then forwarded to domains A and B. A replies that A.U-A.B2 is not available, while B replies that B.B1-B.U is available. Therefore, the intradomain connection A.U-A.B2 is pruned from the PathComputer instance handling this request. If the path computer would not find an alternative path between A.U and B.U, then the NSP would reply with a negative IsAvailableResponse; an alternative start time offset returned by A would be included in the NSP's reply.

In this example however, there is an alternative path, so the path computation would yield three intradomain connections A.U-A.B1, Y.B1-Y.B2, and B.B2-B.U in the next iteration. All three corresponding IsAvailable requests to the domains A, B, and Y are answered positively. Thus, also the NSP's reply to the IsAvailable request for the interdomain path A.U-B.U is positive.

7.2.2.1.3. Network resource reservation

The reservation of network resources was internally handled similar to the availability query described in the previous section. Before sending CreateReservation messages to the NRPSs, the availability of the requested resources was checked. This was to prevent a series of CreateReservation and CancelReservation messages that would be necessary if one or more domains in a multidomain path were not able to fulfill the reservation.

Alternative start times reported by the NRPS adapters were however discarded. The availability of intradomain paths along different routes returned by the PathComputer was checked merely with the constraints specified in the CreateReservation message. In case a path consisting of domains that all gave a positive reply to the availability query was found, the final reservations for all intradomain paths were established.

Considering the exampled sketched in Figure 43 again, upon a CreateReservation request for the connection A.U-B.U, the NSP would first check for an available path just as described in the previous section. Only then the three CreateReservation requests for the intradomain paths A.U-A.B1, Y.B1-Y.B2, and B.B2-B.U would be sent to the corresponding domains.

7.2.2.1.4. Reservation status query

A GetStatus was mapped to a set of single-domain GetStatus queries in a straightforward way. From the reservation ID that was part of this message, the domains and the reservation IDs used for this reservation inside each of the domains were retrieved from the database, and a set of corresponding GetStatus messages were constructed and passed to the NRPSManager. Practical considerations during first testing of the code led to a slight modification of the GetStatusResponse messages. In addition to an overall status code for each connection that was generated from the set of status codes for this connection received from the participating NRPSs, the GetStatusResponseType optionally contained DomainStatus elements, each of which contained a domain name and an element of type ConnectionStatusType, i.e. the connection status received from the specified domain. This was mainly interesting for debugging purposes in cases where the status values were not consistent. E.g., in case a connection should be established, all domains should return the status code active. If one domain returned a different status code, it was immediately visible in which domain the error has occurred.

7.2.2.1.5. Reservation cancellation / connection teardown

An already established reservation was cancelled by a CancelReservation message. For the NSP, it was not of importance whether the reservation contained services that were already active or whether all services were still waiting to be started.

To cancel a reservation, the NSP looked up the intradomain reservations that were made for the input reservation and sent a CancelReservation message with the corresponding ID to each of the domains.

7.2.3. Harmony Performance and Scalability Analysis

Harmony evolved from a centralized, to a hierarchical and distributed architecture. Therefore, and in order to validate Harmony capabilities, a set of analyses were performed that demonstrated the performance improvement of moving from a centralized to a distributed architecture. They measured its success reservations percentage. The performance study showed the successful responses to incremental and poison requests for the Harmony system.

7.2.3.1. Architecture, Assumptions and Limitations

As commented, Harmony was a multi-domain path provisioning system where users and Grid (grid-jobs) [114] applications could book in advance end-to-end paths. Harmony allowed heterogeneous domain interoperability by performing an inter-domain resource brokering over well-known network resource provisioning systems (NRPS). The tests described below considered a network consisting of interconnected and independent domains. A domain in Harmony was considered a high performance optical network controlled by an NRPS or a GMPLS control plane. Thus, the building blocks of the Harmony system performed a different roles in order to allow the creation of on demand or in advance paths for several users along domains. As a consideration, a NRPS did not reveal to Harmony's Network Service Plane (NSP) either its internal topology or its user endpoints. Each Inter-Domain Broker (IDB) entity within the Harmony's NSP offered its optical network reservation services to a given (scientific) user community – from now on called population– which was assumed to be infinite for the stress tests.

7.2.3.2. Request Arrival and Distribution

On one hand, in a real situation, users belonging to a given population had data plane connections to transport networks controlled by their associated IDB and, therefore, generated requests uniquely to the NSP entity. Consequently, when an IDB received requests from its population which did not involve any resource under its control, it performed forwarding of the request towards the correct IDB in the NSP. This effect was modelled by the forwarding probability (p_{fw}) which characterized a population. As a result, a request had a probability of 1-p_{fw} of being served by the IDB attached to the population. Otherwise, the IDB would forward it to the IDB controlling the source domain of the reservation, which in Harmony was the one in charge of computing the whole inter-domain path. Values considered for p_{fw} were 0.01, 0.1 and 0.5. Higher values were not realistic, since it would mean that populations would be requesting "far" resources more than 50%, indicating that a bad point of entry to Harmony would have been configured.

On the other hand, time elapsed between requests, the request inter-arrival time (IAT), was modelled following two different patterns: deterministic and Poissonian. The deterministic IAT allowed us to evaluate the response of the Harmony's Service Plane under stress conditions. The stimulus was created by generating a number of requests per second, from one to twenty, considering an incremental granularity of one request per second, where the highest threshold corresponded to busy periods in real Grid environments such as Kallisto site in HellasGrid (Patras, Greece) or BEGrid in Belnet (Ghent, Belgium).

On its turn, Poisson IAT distribution allowed us to introduce more realistic behaviour in the request arrival, as it was the basis for complex distributions like Pareto-exponential or Markov-Modulated Poisson Processes, proven too closely model IAT in the sites mentioned above [115].

7.2.3.3. Harmony Service Plane Scalability and Load analysis

To evaluate the scalability of the current service plane implementation, both centralized and distributed, the request response time was a function of the load measured

[116]. In Figure 44 the results of the create reservation call over the distributed architecture are shown. The measurements were executed on a system based on two Intel Core 2 CPUs with 2.66GHz each, and the population requests were generated from two PCs. To obtain statically sound results, 30 repetitions were made of each load step and the vertical bar over each symbol indicates the maximum deviation registered in the

samples per step. It shows that the system was stable up to 50req/sec. To obtain statically reliable results, 30 repetitions were made and the error bar indicates one standard error of the mean.

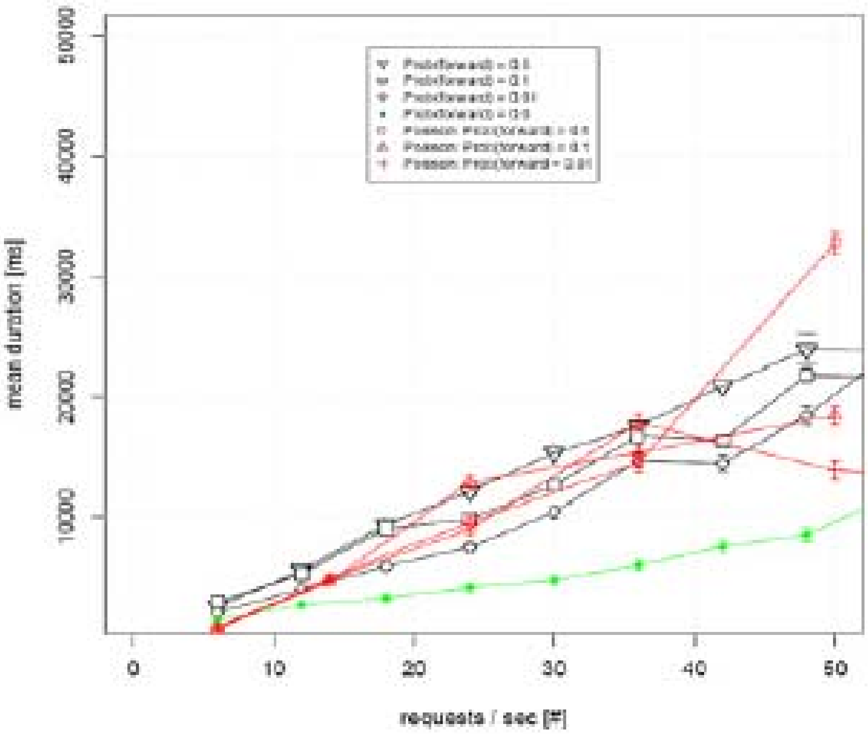


Figure 44. Request-response time while connecting to a distributed service plane composed of six IDB's

Based on a uniform distribution or a Poisson distribution the number of requests per second was increased, and a timeout of 50s between each IDB involved in the NSP and its corresponding emulated NRPS adapter was specified. Results (Figure 44) confirm that, as the probability of forwarding the request in one IDB increased, the time response of the whole service plane was affected by decreasing its performance.

Above, a forwarding probability(p_{fw}) was introduced and discussed to describe the involved IDBs in a reservation process. The decision of which systems came into consideration for a given request was mainly based on the chosen service plane topology. Apart from political choices (e.g. one IDB per administrative domain) its design could be influenced by robustness or scalability aspects. We defined the scalability of an NSP as the amount of required resources as a function of a given work load. Basically, the number of handled requests per second or the required signalling bandwidth could be taken into account as considered metrics. Then, the decision of which systems came into consideration for a given request was mainly based on the chosen service plane topology. Apart from political choices (e.g. one IDB per administrative domain) its design was influenced by robustness or scalability aspects. In Figure 45 the mean duration per request related to the load is depicted for different numbers if IDB layers [117][118]. It can be seen that the mean duration per request growth in a linear fashion with the number of incoming requests. Each layer amplified this effect and occurring timeouts prevented the operation in higher load scenarios. As a deduction a distributed service plane could help to improve the responsiveness behaviour by keeping down the number of involved IDB layers and by preventing an overload of single entities.

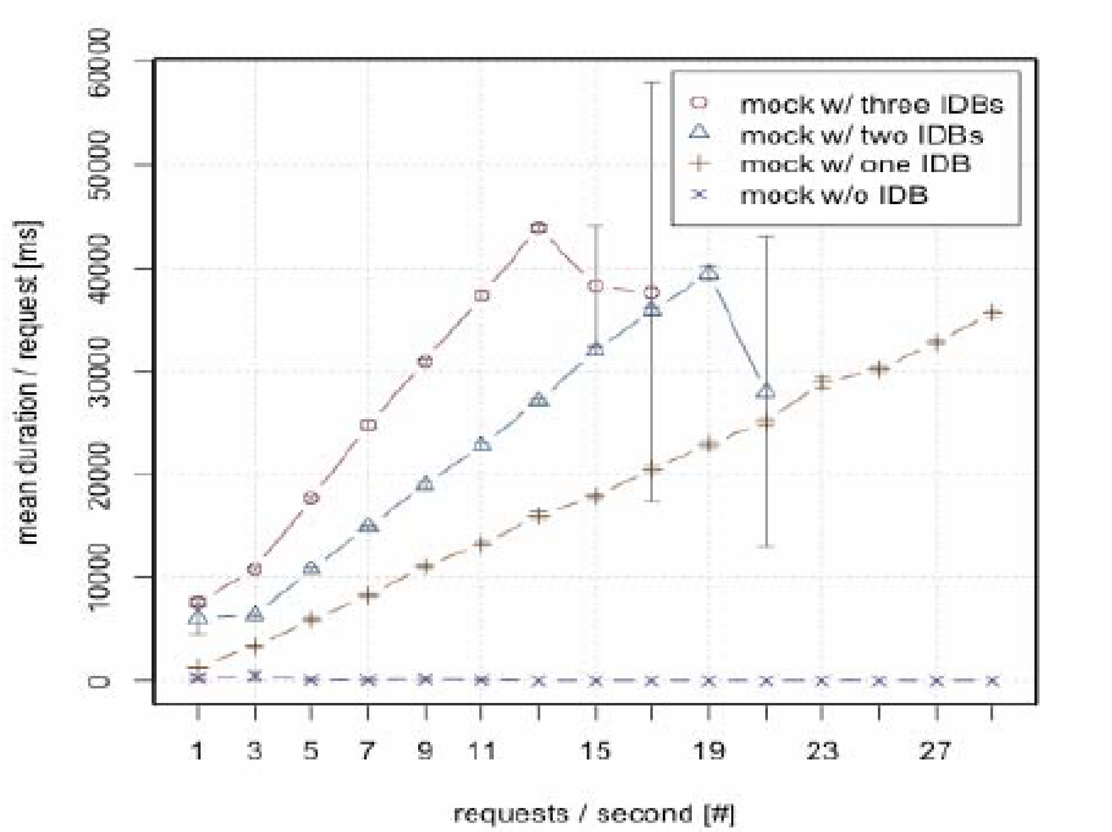


Figure 45. Request-response times while connecting to a mock Adapter through different layers of IDBs

In this context, a further issue influenced the design of the service plane topology. Different request and response types needed specific bandwidths. Based on measurements done in the Phosphorus test-bed, Table 13 shows the average size of the four most used requests. Given these values, it was stated that a complete reservation workflow required about 13 kb per reservation between two involved entities. The topology updates were sent every five minutes and therefore just consumed about 18 bytes/sec., again only between two entities. That implied that at a load of 30 req./sec. connections with at least 390 kb/sec. were needed between the IDBs and Adapters.

Description	request [bytes]	response [bytes]	∑ [bytes]
Availability	2768	2072	4840
Creation	2792	1966	4758
Cancellation	1572	1786	3358
topology update	3539	1760	5299

Table 13. Average sizes of different request and response messages

However, in comparison to the centralized approach, the results showed that the distributed NSP could handle successfully more number of simultaneous requests, since Figure 46 showed that the system was stable up to 28 requests per second in a centralized approach, while in a distributed approach was able to handle up to 50 requests per second, as shown in Figure 44, and assuming six distributed IDBs [116].

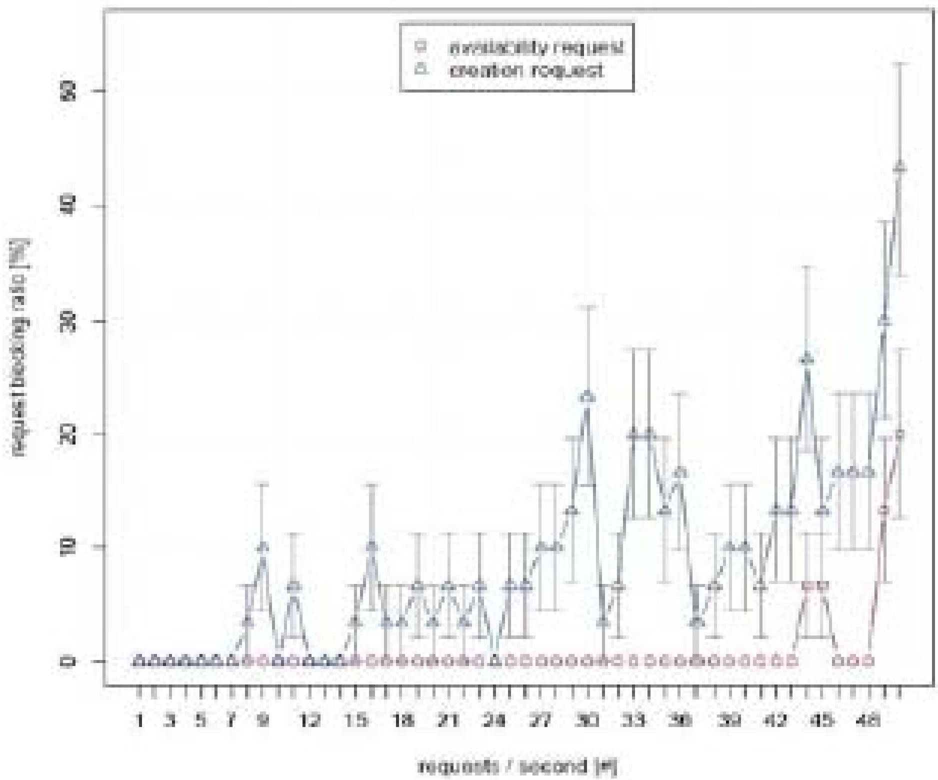


Figure 46. Request blocking ration while connecting to an emulated NRPS in centralized service plane architecture

Furthermore, the figure shows notable response time variations on higher load for the centralized approach. As shown in Figure 46, the cause of this observation was an increased amount of failed requests that in turn was caused by timeouts between the communicating entities.

Another analysis was performed with regards to the impact of a hierarchy architecture. As shown in

Figure 47.a the mean response time of each NRPS adapter (t_h) varied from 200ms to 410 ms and had a noticeably large number of outliers. The dummy adapter delay pointed out 10ms communication and processing overhead for each request while time response of the adapters did not exceed of 500 ms (average). As seen in

Figure 47.b, each additional hierarchical level increased the total response time (t_h^{total}) by approximately 500ms, so that the response time of the system increased linearly as the hierarchy level increases.

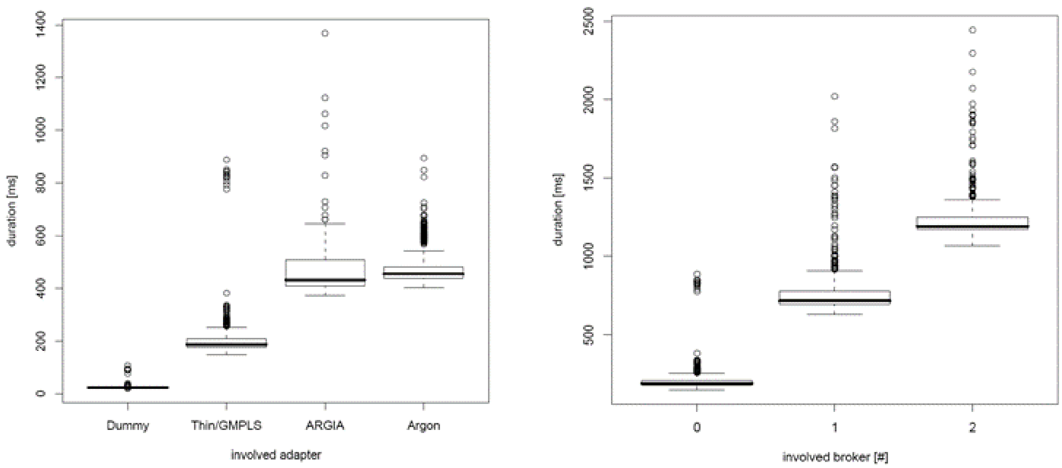


Figure 47. Processing of a reservation request within a single IDB (500 repetition). (a) Depicts the response time for each Adapter and (8) shows the delay that is added by every hierarchy level.

Thus, under the given assumptions a Poisson distribution improved the request-response time when Harmony performed the forwarding in a distributed architecture. We tested it in a realistic scenario under the Phosphorus emulated testbed, and the result were that even with a very limited knowledge of topology information and without an homogeneous control plane for the different domains, the service plane implemented by Harmony scaled according to the model deployed, centralised or distributed, and presented different behaviours under stress conditions with better performance under a distributed one. Based on these results, Harmony would allow to different National Research and Education Networks (NRENs), normal potential users of Harmony, to handle up to 50 requests per second with a maximum request response time of 22 seconds for the e2e provisioning, allowing the operators to not exchanging or exposing confidential information of their internal topology.

7.2.3.4. Phosphorus testbed

The following figure (

Figure 48) picture presents de Phosphorus EC project test-bed where Harmony was deployed, tested and assessed.

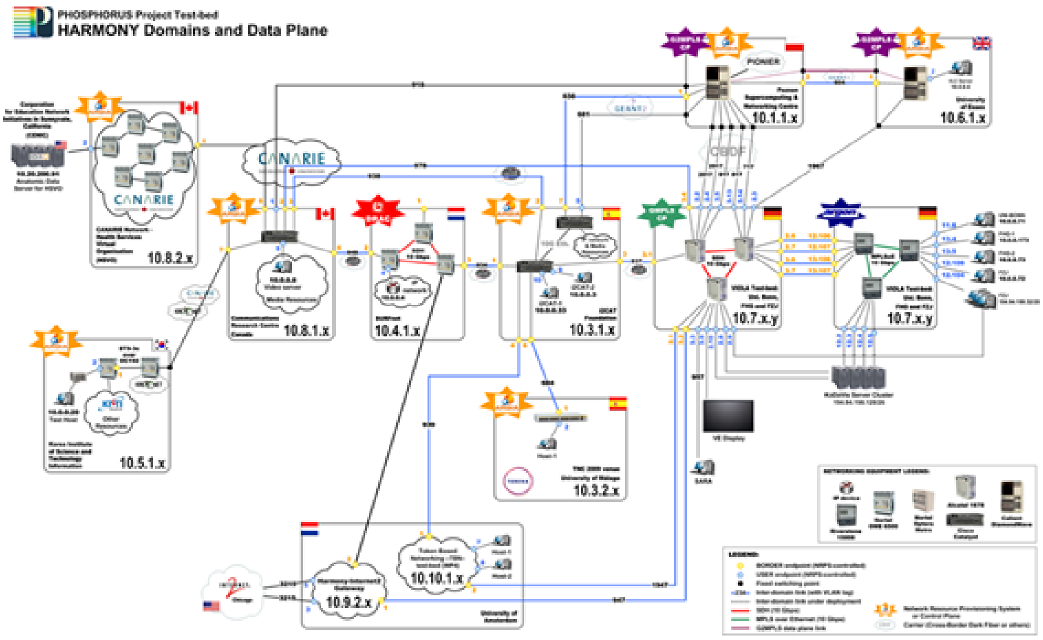


Figure 48. Phosphorus and Harmony testbed

7.2.3.4.1. Analysis conclusion

Harmony was a system able to accept multiple simultaneous provisioning requests directly from the Grid middleware under a centralized or a distributed architecture. Considering the results presented, a distributed approach was the most suitable for Harmony, since it improved the responsiveness behaviour by keeping down the number of involved IDB layers and by preventing an overload of single entities, making the system more efficient and scalable.

We also had to take into account that when the probability of forwarding the request in one IDB increased, the time response of the whole service plane was affected. As distributed systems increased the load traffic in the service plane due to signalling, it was interesting to point out that under high stress conditions (30 req./sec.) it had a maximum load of 390 kb/sec between the IDB and the NRPSs, which was acceptable.

7.2.4. G^{2}MPLS

One of the main goals and technical domains of the Phosphorus project concerned the architectural definition, software design and prototypal implementation of the Grid-enabled GMPLS (G^2MPLS) Network Control Plane, as an enhancement of the ASON/GMPLS Control Plane architecture that implemented the concept of Grid Network Services (GNS). In the PHOSPHORUS framework, GNS was a service that allowed the provisioning of network and Grid resources in a single-step, through a set of seamlessly integrated procedures.

G^2MPLS resulted in a more powerful Control Plane solution than the standard ASON/GMPLS, because it complied with the needs for enhanced network and Grid services required by network "power" users/applications (i.e. the Grids). Nevertheless, G^2MPLS was not conceived to be an application-specific architecture, and it supported any kind of endpoint applications by providing network "legacy" ASON/GMPLS transport services and procedures. This compliance fostered the possible integration of Grids in operational/commercial networks, by overcoming the limitation of Grids operating on dedicated, stand-alone network infrastructures.

The main rationale behind this new G^2MPLS architecture was made of different points: firstly, it was a dual approach with regards to grid brokers working with and configuring network resources. Secondly, grid nodes were modelled as network nodes with node-level Grid resources to be advertised and configured, and this was a native task for GMPLS. Thirdly, it could also inherit "useful" GMPLS native features,

Basically, G^2MPLS extended the ASON/GMPLS architectures in order to provide part of the functionalities related to the selection, co-allocation and maintenance of both Grid and network resources, by exposing upgraded interfaces at the UNI and E-NNI network reference points (i.e. G.OUNI and G.E-NNI).

The G^2MPLS Network Control Plane brought an innovation to the field of co-allocation of Grid and network resources, because of its faster dynamics for service setup, adoption of well-established procedures for traffic engineering, resiliency and crank back and uniform interfaces for the Grid user to trigger Grid & network transactions.

7.3. Conclusions

The development of a new service plane capable to provision on demand connectivity services from the application API, and in a multi-domain and multi-technology scenario based on a virtual network infrastructure, composed of resources from different infrastructure providers, brought a new type of services planes that facilitated the deployment of applications consuming large amounts of data under deterministic conditions. Moreover, and considering the user as Grid applications, it provided an enhancement and allowed networks behave as a Grid-class resource, and so offer different types of advances provisioning schemas at the network level.

The Harmony system became a game changer for the technology foreseen up to that date. It fitted the NREN (Network Research Educational Networks) ecosystem with the strict requirements and demands imposed by scientific users with large dataset grid applications. In fact, it was the first on-demand provisioning system that at lower levels (optical infrastructure) allowed the creation of one virtual domain composed from resources of different providers. Indeed it had some limitations in terms of availability, since the internal domain information from each provider was not populated to the rest, but the border nodes. However, the fact of creating a multi-domain virtual infrastructure composed by the border nodes of these individual domains (NRENs), with a set of systems capable to provide inter-domain connectivity services, allowed the Grid applications, or end users, to request advanced end-to-end services across different domains. Harmony allowed NRENs (the type of traffic services offered by NRENs, mostly deterministic, fit perfectly with Harmony, which was not thought for busty packets) to handle up to 50 requests per second with a maximum request time of 22 seconds for e2e provisioning, allowing the operators not to exchanging or exposing confidential information of their internal topologies. And this, compared to the times needed for NRENs, or GEANT , on establishing and end to end path, was a major outcome for the community, since not only was offering the capability to set-up dynamic e2e light paths, it also offered new advanced reservation mechanisms for the optical networks and full integration and interoperability with a GMPLS control plane.

8. A Generalized Architecture for Dynamic Infrastructure Services

8.1. Introduction

Over the last years, and around 2012, Internet had already became a central tool for society. Its large adoption and strength originated from its architectural, technological and operational foundation (a layered architecture and an agreed-upon a set of protocols for the sharing and transmission of data over practically any medium). The Internet's infrastructure was essentially an interconnection of several heterogeneous networks called Autonomous Systems, which were interconnected with network equipment called gateways or routers. At the same time, routers were interconnected together through links, which in the core-network segment were mostly based on optical transmission technology, but also in the access segments a gradual migration to optical technologies occurred. The current Internet became an ubiquitous commodity to provide communication services to the ultimate consumers: enterprises or home/residential users. The Internet's architecture assumed that routers were stateless and the entire network was neutral. There was no control over the content and the network resources consumed by each user. So far, it was assumed that users were well-behaving and have homogeneous requirements and consumption.

After having dramatically enhanced our interpersonal and business communications as well as general information exchange—thanks to emails, the web, VoIP, triple play service, etc.—the Internet started to provide a rich environment for social networking and collaboration and for emerging Cloud-based applications such were Amazon's EC2, Azure, Google apps and others. The Cloud technologies were emerging as a new provisioning model [120]. Cloud aimed for on demand access to IT hardware or software resources over the Internet. Clouds started to revolution the IT world [121], but treated the Internet as always available, without constraints and absolutely reliable, which was to be achieved. Analysts predicted that in 2020, more than 80 % of the IT would be outsourced within the Cloud [122]. With the increase in bandwidth-hungry applications, it was just a matter of time before the Internet's architecture would reach its limits.

The new Internet's architecture should propose solutions for QoS provisioning, management and control, enabling a highly flexible usage of the Internet resources to meet bursty demands. Thus, Internet's architecture needed to be redesigned, otherwise mission-critical or business applications in the Cloud would suffer, but even conventional Internet's users would be affected by the uncontrolled traffic or business activity over it.

At this period in time, best practices in optical network and IT infrastructure management were characterized by global services and delivery over generic infrastructures, driven by the ubiquitous presence of the Internet. As the scale of information processing was increasing, from Petabyes of Internet data to the projected Exabytes in networked storage at the end of this decade, Future Networks were required to provide new answers to support the Future of Internet and its new emerging applications[123][124]. This was becoming even more important considering the current development and technical enhancement of photonic networks, dynamic control planes, multi core processing, cloud computing, data repositories, and energy efficiency, which were driving profound transformations of optical networks and users capabilities. These technological advances were driving the emergence of ever more demanding applications such as UHD IPTV, 3D games, virtual worlds, and photorealistic telepresence in media. These were high-performance and high-capacity network based applications with strict IT (e.g. computing and data repositories) resource requirements, which the current Best Effort Internet intrinsically could not deliver. However, and besides these potential constrains and limitations, it was impossible to throw away what had made the enormous success of the Internet: the robustness brought by the datagram building block and the end-to-end principle which were of critical importance for all applications. In this context, the performance, control, security and manageability issues, considered as non-priority features in the 70s [125] had to be addressed [126].

The need of an improvement of the current Internet's architecture started to arise within the research community. The combination of Cloud-based resource provisioning and the virtualization paradigm with dynamic network provisioning as a way towards such a sustainable future Internet became the key research topics of the period. It meant an architecture for the future Internet that would provide the basis for the convergence of networks—optical networks in particular—with the Clouds while respecting the basic operational principles of today's Internet. It was important to note that for several years, to serve the new generation of applications in the commercial and scientific sectors, telecom operators already considered methods for dynamic provisioning of high-capacity network-connectivity services tightly bundled with IT resources. To realize this kind of networked-IT infrastructure service, envisioned to facilitate Future Internet, a next generation network architecture was needed. This new generation network architecture had to seamlessly integrate optical network technologies and IT resources, and provide customized infrastructure provisioning services to facilitate the seamless integration of optical network segments and technologies. Such an infrastructure service should be supported by a revolutionized service provisioning framework [127].

The requirements for resource availability, QoS guarantee and energy efficiency mixed with the need for an ubiquitous, fair and highly available access to these capacities became the driving force for a new architecture.

Nevertheless, during those years several research initiatives to support the Future Internet emerged around the world e.g. FIRE[128] in the EU and FIND[129] in the US. These initiatives were mainly focused on providing experimental infrastructures where network and IT resources could be exposed to the research community to perform disruptive network research. Important example were the FEDERICA[130] EU FP7 project, or the GENI[131] project in the US with the iGENI[132] initiative. iGENI infrastructure would integrate with GENI resources, and operate them in order them to be used by GENI researchers conducting experiments involving multiple aggregates (at multiple sites). In addition, the iGENI consortium planned to integrate its global infrastructure with the Open Resource Control Architecture (ORCA) [133] control framework, developed by GENI at RENCI (Renaissance Computing Institute) and Duke University, to enable GENI researchers to dynamically control international network services, associated transport resources, and GENI aggregates.

All these approaches were considerate together with a set of related initiatives that were taken into account for the work of this thesis and so the design of the new architecture.

- NGN Open Service Environment (OSE): The NGN reference model, according to ITU-T Y.2011 Recommendation [134], suggested the separation of the transport network and application services and defined them as NGN service stratum and NGN transport stratum consisting of User plane, Control plane and Management plane. The NGN Y.2012 architecture defined also the Application Network Interface (ANI) that provided an abstraction of the network capabilities. It was used as a channel for applications to access network services and resources.
- Composable Services Lifecycle Management: The Service Oriented Architecture-based technologies provided a good basis for creating composable services that, in case of advancing to dynamically re-configurable services relied on the well-defined Services Lifecycle Management (SLM) model. The architecture framework designed considered dynamic provisioning as a major issue, thus, dynamically provisioned and re-configured services would require re-thinking of existing models and propose new security mechanisms at each stage of the typical provisioning process.
- 4WARD Project[135]: EU-FP7 4WARD's goal was to make the development of networks and networked applications faster and easier, leading to both more advanced and more affordable communication services. According to 4WARD outcomes, network virtualisation was not only an enabler for the coexistence of

multiple architectures, but also provided a path for the migration towards more evolutionary approaches to the Future Internet. In 4WARD's vision, virtualisation could help to keep the Internet evolvable and innovation-friendly, particularly since it could mitigate the need to create broad consensus regarding the deployment of new technologies among the multitude of stakeholders that composed Internet. In 4WARD, the goal was to develop a systematic and general approach to network virtualisation. The problem space was divided into three main areas: virtualisation of network resources, provisioning of virtual networks and virtualisation management. The 4WARD roles model were used as state of the art for the new architecture design presented in this thesis.

- RESERVOIR Project: The FP7 RESERVOIR[136] project's goals were defined as to enable massive scale deployment and management of complex IT services across different administrative domains, IT platforms and geographies. The project considered virtualisation technologies to transparently provision distributed resources and services on-demand with the specified QoS based on Service Level Agreement (SLA). RESERVOIR was limited to server virtualisation, regardless of the transport networks, since it was focused on building the foundations of clouds services. The expertise generated on RESERVOIR for the management of IT resources was very useful and so considered along the design of virtual infrastructures.

Besides the different activities developed worldwide, there was not yet an architecture capable of seamless and coordinated provisioning of optical network and IT resources, and end-to-end service delivery with flexible, adaptive and dynamic association as well as integration of heterogeneous network infrastructures and IT resources.

8.2. Main research goals

There are various challenges that were driving Internet to its limit, which in turn had to be addressed by a new architecture [137]. Thus, and in order to priories them, the following six challenges needed to be addressed in a global approach:

1. **Enable ubiquitous access to huge bandwidth**: As of the period, the users/applications that required bandwidth beyond 1 Gbps were rather common, with a growing tendency towards applications requiring a 10 Gbps or even 100 Gbps connectivity. Examples included networked data storage, high-definition (HD) and ultra-HD multimedia-content distribution, large remote instrumentation applications, to

name a few. But so far, these applications could not use the Internet because of the fair-sharing principle and the basic routing approach. As TCP, referred to as the one-size-fits-all protocol, had reached its limits in controlling—alone—the bandwidth, other mechanisms had be introduced to enable a flexible access to the huge available bandwidth.

2. **Coordinate IT and network service provisioning**: In order to dynamically provision external IT resources and gain full benefit of these thanks to Cloud technologies, it was important to have control over the quality of the network connections used, which was a challenge in today's best-effort Internet. Indeed, IT resources were processing data that should be transferred from the user's premises or from the data repository to the computing resources. When the Cloud become largely adopted and the data deluge fall in it, the communication model offered by the current Internet of the period would break the hope for fully-transparent remote access and outsourcing. The interconnection of IT resources over networks would require well-managed, dynamically invoked, consistent services. IT and network should be provisioned in a coordinated way in the future Internet.

3. **Deal with the unpredictability and burstiness of traffic**: The increasing popularity of video applications over the Internet caused the traffic to be unpredictable in the networks. The traffic's bursty nature required mechanisms to support the dynamic behavior of the services and applications. Moreover, another important issue was that the popularity of content and applications on the Internet would be more and more sporadic: the network effect amplifies reactions. Therefore, the future Internet needed to provide mechanisms that facilitate elasticity of resources provisioning with the aim to face sporadic, seasonal or unpredictable demands.

4. **Make the network energy-aware**: as reported in the literature [138], ICT was at that time responsible for about 4 % of the worldwide energy consumption, and this percentage was expected to rapidly grow over the next few years following the growth of the Internet. Therefore, as a significant contributor to the overall energy consumption of the planet, the Internet needed to be energy-conscious. In the context of the proposed approach, this should involve energy awareness both in the provisioning of network and IT resources in an integrated globally optimized manner.

5. **Enable secured and reliable services**: The network's service outages and hostile hacks were receiving significant attention due to society's high dependency on

information systems. The current Internet's service paradigm allowed service providers to authenticate resources in provider domains but did not allow them to authenticate end-users requiring the resources. As a consequence, the provisioning of network resources and the secure access of end users to resources was a challenge. This issue was even more significant in the emerging systems with the provisioning of integrated resources provided by both network and IT providers to network operators.

6. **Develop a sustainable and strategic business model**: the business models deployed by telecom operators were only focused on selling services on top of their infrastructures. In addition, operators could not offer dynamic and smooth integration of diversified resources and services (both IT and network) at the provisioning phase. Network-infrastructure resources were not understood as a service within the value chain of IT service providers. Thus, a novel business model was necessary, which could fully integrate the network substrate with the IT resources into a single infrastructure. In addition, such business model would let operators offer their infrastructures as a service to third-party entities.

8.3. Architecture and infrastructure technical challenges

This section describes the specific technical challenges that the new architecture proposed along the chapter had to overcome. The achievement of these technical challenges provided solutions to the main/global research challenges discussed on section 8.2.

The architecture had to be based on an end-to-end principle, based on extending standards in order to create a new planning, provisioning and (ultimately) business framework for network infrastructure provider and network operators. To do so, the definition of a novel photonic network architecture, capable of provisioning 'Optical Network + Any-IT' resources to network operators for end-to-end service delivery was a must. Thus, a revolutionary vision under an evolutionary approach, following a network centric and bottom up strategy, was needed. This vision was based on partitioning the photonic network infrastructure to create specific logical infrastructures, composed by optical network and IT resources. This composition would overcome the limitations of networks and domain segmentation. Each logical infrastructure would be controlled by an enhanced Network Control Plane capable of provisioning Optical Network Services bundled with IT resources in an on demand basis. Furthermore, the logical composition of photonic networks would enable the GMPLS/PCE control plane to dynamically scale infrastructure resources based on the needs of the network operator.

All these challenges came after many brainstorming sessions that put the basis of the design of this new architecture. An architecture that would change the way operators would deal with the optical infrastructures and the services they could provide. And so, define new business models. Figure 49 shows the first draft discussion of the architecture presented in the coming sections. This discussion happened in Brussels (together with Nicola Ciulli, CTO from Nextworks) around 2007. This discussion brought up the new ideas I had on mind and set-up the basis for further discussions on creating a generalized architecture. The idea consisted on extending virtualisation from the resource up to the control plane. So, an enhancement for virtualisation frameworks for optical an IT infrastructure was needed. This architecture would allow the deployment of different virtual infrastructures, each one controlled by a different control plane. While these control planes would be serving different cloud applications, and creating customized virtual infrastructures according to the needs and SLA's of these cloud applications. In fact, all these virtual infrastructures would share the same physical substrate (physical infrastructures), which could be composed of logical (virtual) resources from different infrastructure providers. The main activity behind this Thesis was the design of this architecture and the definition of its services and functionalities.

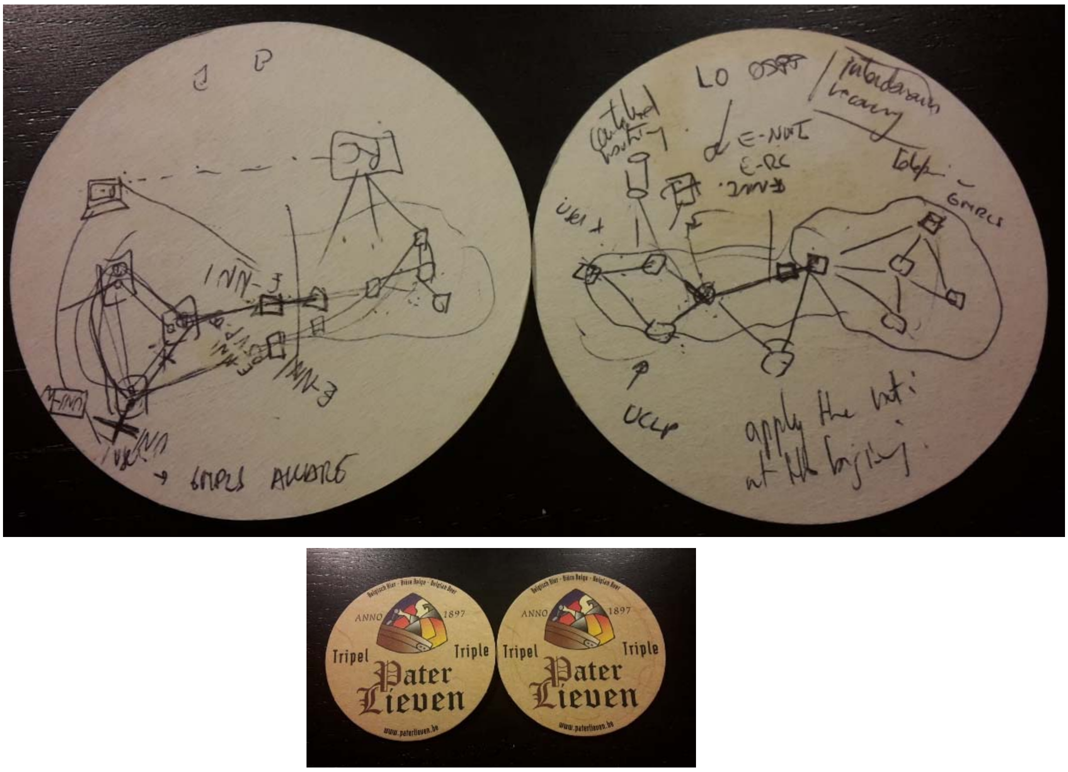


Figure 49. First drawn draft of a new architecture done in 2007 in a Brussels Pub.

With this new architecture concept, network infrastructure provider would be able to compose and offer part of their infrastructures (possibly combined with IT infrastructures) as logical infrastructure to network operators, which would act as virtual network operators. So, network operators would operate the hired logical infrastructures with advanced Control Plane (ASON/GMPLS and PCE) technologies and would offer coupled, optimized and dynamic Network + IT provisioning services (i.e. interconnections between end-user and IT resources) to allow service provider deploy applications over virtual networks SLA-application-aware. This novel architecture would facilitate the emergence of new business models in the future telecom market, where current telecom operators could split and extended their business by offering services as physical infrastructure provider and/or (overlay/customized) virtual network operators. Actually, telecom operators could also make profits from their infrastructure by carefully/better planning resource allocation to limit unused cycles or bandwidth, and improving the energy efficiency by optimizing their resource usage. The flexibility of a logical infrastructure composition mechanism would allow network operators to lease and operate multiple instances of application-specific networks towards service providers, each with its own specific IT and network requirements [139].

This architecture would overcome some of the limitations existing in current infrastructure services:

- New emerging applications with highly demanding network IT resource requirements, but still unable to exploit the potentials of the current optical network technologies.
    - The optical Network layer was unaware of applications' dynamic requirements.
    - The Net & IT resources were being controlled separately without any integration.
- Optical infrastructure providers and network operators were unable to provide enhanced and customized network services.
    - Telecom operators and service providers only offered simple connection services over generic optical network infrastructures.
- Opportunities for new business models.
    - Multiple virtual networks integrated with IT resource offered as a service
    - Use of latest advances in network virtualization and physical resource partitioning.

- Customized virtual networks supporting dynamic Optical Network + IT provisioning services.

8.3.1. Infrastructure Challenges

There were many challenges as driving forces to redefine the architecture of the Network of the Future and the services it would provide. The ones listed below were addressed by the architecture proposed on this chapter as the fundamental future challenges where current networks would reach its limits. It would:

• Define and develop a novel dynamic wavelength service provisioning mechanism that would enable network operators to efficiently manage their capacity and infrastructure.

• Define and implement a new mechanism for network operators to request and setup scheduled high bandwidth deterministic optical network connectivity in an on demand manner.

• Investigate the definition of - and evaluate - new inter-domain trust models and access control mechanisms that ensure guaranteed main service delivery/provisioning based on customized SLA, while providing consistent dynamically invoked security and access control services to network operators.

• Define and develop a novel end-to-end service provisioning mechanism that would automatically and efficiently bundle suitable IT resources with the required optical network connectivity services in a single step.

• Define and develop a novel mechanism where infrastructure providers could partition their infrastructure resources (optical networks and/or IT) to compose logical infrastructures and offer them to network operators as a service. The logical composition mechanism should support dynamic and on-demand changes of combined optical network and IT resources.

• Architect and develop a novel Network Control Plane, as an enhancement to well-established architectures (ASON/GMPLS and PCE) that would surpass the mentioned limitations, in a comprehensive and consistent functional scenario.

• Develop an energy aware optical network and IT service provisioning approach, in which the energy optimization objective would sought during the dynamic planning and allocation of resources (e.g. connection set-up) by the Network Control Plane.

The concept behind this architecture would result in a new role for telecom operators that owned their infrastructure by enabling them to offer their optical network integrated with IT infrastructures (either owned by them or by third party providers), as a service to network/service operators. This would enable business models where complex services (e.g. Cloud computing) with complex attributes (e.g. optimized energy consumption and optimized capacity consumption) and strict bandwidth requirements (e.g. real time and resilience) could be offered economically and efficiently to users and applications.

Thus, it would significantly improve the efficiency of the current operation of optical networks by extending standard network architectures. This new architecture would consider the above-mentioned challenges, and would address the full integration of the optical network substrate with the attached network edges and IT resources into a single infrastructure. And so, enable applications and users to deploy different styles of networks operations (in terms of, for example, granularity and dynamicity of transport bandwidth) on this infrastructure.

8.3.2. Technical approach/challenge

This infrastructure challenge would be achieved through two new technical layering approaches capable to map SLA from the application down to the network/IT infrastructure. The first technical approach would be based on defining and designing a new Logical Infrastructure Composition Layer (LICL) as a new powerful tool for infrastructure providers. It would provide a semantic framework for composition of the logical infrastructures and exchange of resources information between infrastructure operators, according to the Infrastructure as a Service paradigm. The second one would be based on the definition and design of an enhanced Network Control Plane (NCP) architecture and its protocol extensions. It offered a major provisioning tool for network operators, with backward-compatible superset of the ASON/GMPLS and PCE architectures, thus enhanced capabilities, still preserving the legacy and standard ones from ASON/GMPLS and PCE.

This NCP would implement an advance provisioning functionality (i.e. dynamic and/or scheduled) of end-to-end optical network (i.e. connections), for advanced Bandwidth on Demand (BoD) services over a partitioned physical infrastructure and at the user-network interface (thus extending the ASON/OIF UNI reference point). This would take into consideration the experience acquired on the development of IT-aware GMPLS control plane for grid resources from the Phosphorus [140] EU FP6 IP.

Thus, the combination of the LICL and the enhanced NCP would introduce a re-planning capability (triggered by the NCP) in the network Traffic Engineering procedures, in order to change the underlying controlled infrastructure when needed. Since the infrastructure would be a logical one, and offered dynamically by the LICL, the network operator would be able to lease/drop pieces of infrastructure, as needed by its technical operations or business requirements. GMPLS+/PCE+ would be the tool to implement these dynamics. Moreover, energy-awareness, for an energy-efficient routing and provisioning of transport network connection services would be an outcome of these technical challenges.

8.3.3. Business model impact

The aforementioned challenges and services would introduce a new architecture with a set of key business differentiating factors that would re-qualify the interworking of legacy planes by means of a logical infrastructure representation layer for network and IT resources. This would enable a new business model where Infrastructure providers' resources would be partitioned and offered on demand to network operators as a Service.

This research innovation aimed at establishing a breakthrough in the design and implementation of the Future Networks. It provided an evolved Network Control Plane that would enable European network operators to provide end-to-end services by seamlessly controlling optical network and IT resources, with dynamic wavelength allocation functionalities and capabilities to (re)plan infrastructure resources on demand. Thus, it would lead to an optimized cost-efficient use of infrastructures.

The Infrastructure as a Service (IaaS) approach would allow this model. The resources partitioning, and the dynamic and on-demand capability, would allow network operators to offer tailored made services to novel markets. And so, enable telecom operators to access new markets with new business models by moving their business towards high value application layers.

This new concept would allow the development of new actors in ICT environment (existing and emerging Network Operators) with large impact on CAPEX and OPEX optimization.

8.3.4. Rationale

The motivation behind all the research challenges presented had a clear rationale. The current evolution of services, applications and the emerging needs and trends of the telecom sector validated the need to research towards new architectures.

Cloud was clearly emerging as a key future technology that would require advances on the network side and how DC's could be interconnected for service and resource sharing. Since cloud would be the type of application driving the research at the infrastructure level, it is important to understand the main characteristics that cloud technology and its systems had on this period (Figure 50):

- A 'cloud' was an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality (of service).
- Cloud computing refered to the on-demand provision of computational resources (data, software) via a computer network, rather than from a local computer.
- Cloud infrastructure services, also known as Infrastructure as a Service (IaaS), delivered computer infrastructure – typically a platform virtualization environment – as a service.

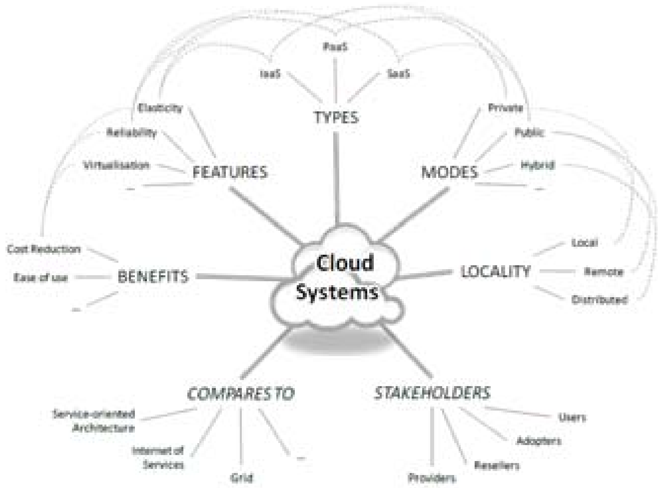


Figure 50. Non-Exhaustive view on the main aspects forming a cloud system (Ref. The Future of cloud computing – Opportunities for European Cloud Computing beyond 2010)

In the end, all these needs and trends targeted a convergence scenario between the Telco and the IT sector (Figure 51). This convergence should end up with services that would span their common boundaries and independently deal with network and IT resources. "Internet-scale" was misleading, as the sufficiency of the existing best-effort Internet as the infrastructure and engine that enabled these applications.

Thus, optimization of interactions between resource consumers, operators and providers (IT and Telecom) was a must. It meant that optimization could be achieved by means of virtualization, flexibility and dynamic behavior of systems, while circuit oriented networks and seamless provisioning of IT and Network resources seemed to be key. All these functionalities together represented one of the challenges behind cloud systems. Actually, there was an explosion of cloud technology. This explosion could be considered as an evolution from Cloud 1.0 former technologies, services and functionalities towards a Cloud2.0, where interconnection of resources would be a must (Figure 52).

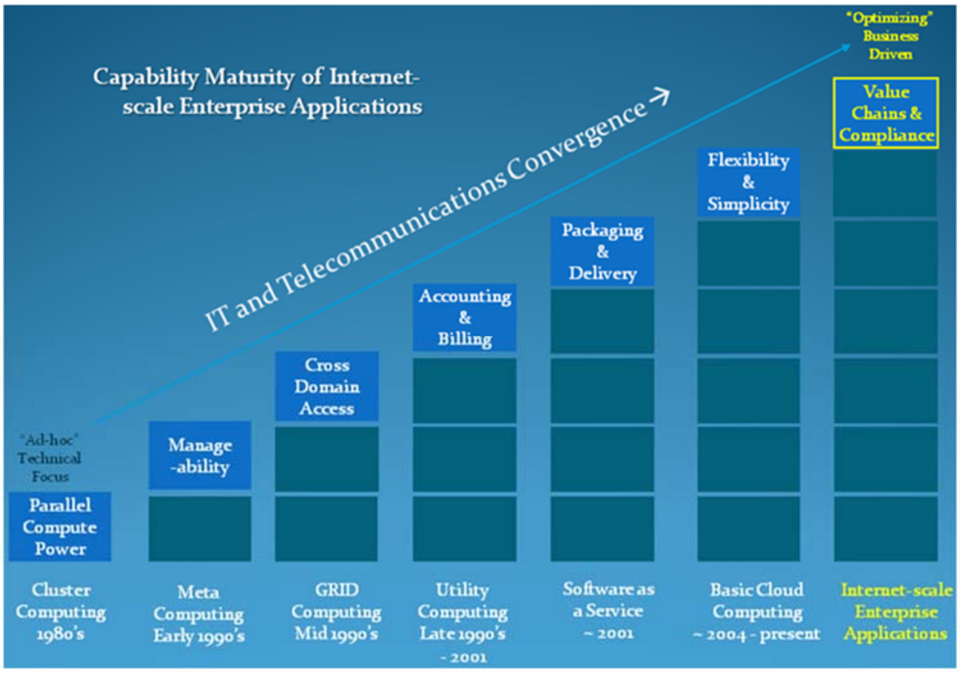


Figure 51. Capability Maturity of Internet scale Enterprise Applications

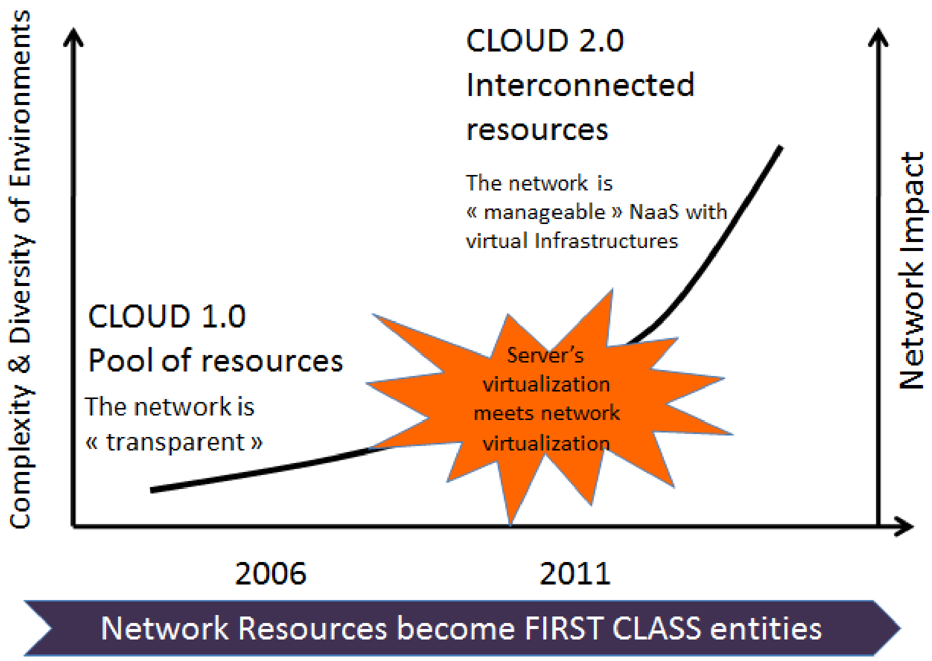
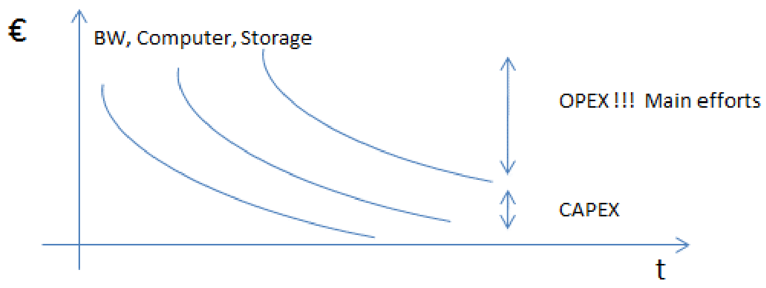


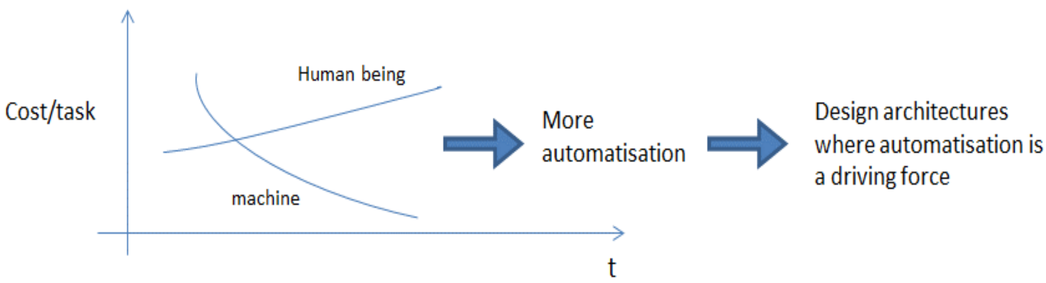
Figure 52. From Cloud 1.0 to Cloud 2.0

This cloud evolution towards a Cloud 2.0 scenario was encompassed by three main driving forces, being: economics, mobility and enabler (ref: talks with Juniper's CTO, Pradeep Sindhu, on 2010):

1. **Economics:**
    a. OPEX dominates ¾:1. Thus, efforts needed to be put on how to reduce the OPEX since CAPEX would keep decreasing with time.



    b. Cloud Computing: Centralize what you can, distribute what you must. There was no clear rule, it was a compromise, but automatisation would become a driving force for technology evolution in any type of platforms (network/IT).



2. **Mobility:**
    a. Desire or people/companies (…) to consume at anytime, anywhere, any device….
    b. Persistence of the data was very important, and Security was very important too.
        i. The data was into the cloud, so this features were a must.
3. **Enabler:**
    a. Networking technology was ‘finally’ powerful enough for the cloud computing model.
    b. The network, and its integration with IT, was a key enabler.

At the same time, these three forces together allowed the estimation of three main design principles for the cloud. These design principles were:

1. **The infrastructure should be built using small set of powerful programmable building blocks**.
    a. The way to get complexity was combining these building blocks (
    b. Figure 53).

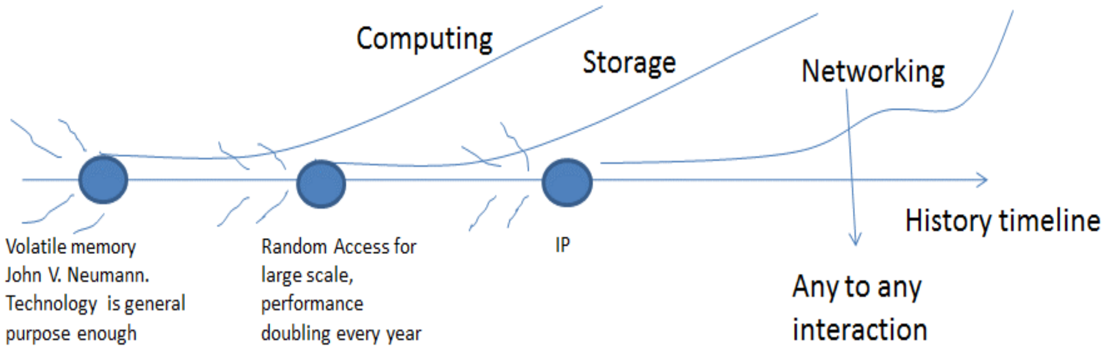
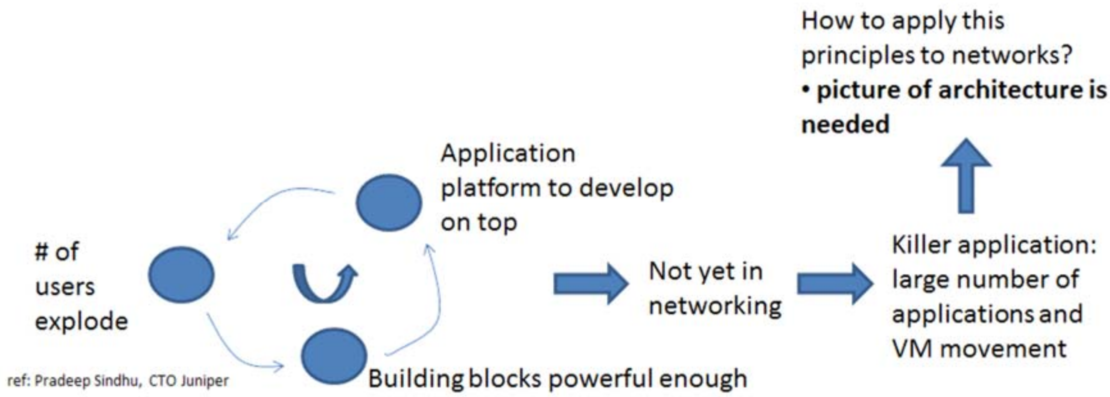


Figure 53. Cloud and network evolution

2. **Automate anything that can be automated:**
    a. Simplify
    b. Consolidate (from many to few)

    } Automate

3. **Open Source ecosystems**
    a. Platform effect (at that time these openness was not yet solved, however we can see it in nowadays in Openstack or Opendaylight platforms)



These deign principles aimed at understanding the needs to solve complexity and automation while developing open source technologies, in the contrary of what the manufacturing and telecom sector had been doing for the last period. Thus, the global conclusion towards the main challenges for networks in cloud were:

- Isolation and flexibility of circuit oriented networks
- Access to huge bandwidth
- Coordination of IT and network
- Sustainable business model

Moreover, the investment devoted in DC was growing, while new services were announced, as VirtualDC, and started to become key and pay the attention of future customers (broadcasters/CDNs and big data dissemination research bodies - e.g. ESA GMES data dissemination and circulation-). By 20011 Virtual Data Centers (Figure 54) were currently rolled out by many operator with the goal to offer operator-driven private clouds, however, only IT services were provisioned and inter-cloud services were not yet available either.

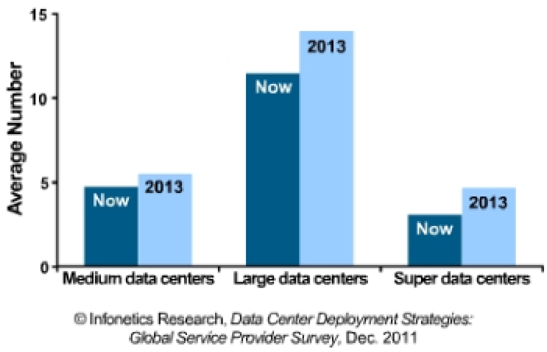


Figure 54. Operators increase of investments in data centers facilities across the board

Actually, this was the period were the network started to become a vital commodity for the cloud computing/distributed DC services (src. NYSE-Euronext keynote @ECOC_2011):

- An exemplar footprint for a distributed DC (Figure 55) is:
    - 6k+ network devices (routers, switches, etc); 200k+ 1G ports or below; **10k+ 10G ports;** 500k+ Routes in routing tables; 20k+ Telco circuits (170 lambdas deployed in USA & EU,); **18+ Pbps data center core bandwidth;** 10k+ servers; **60k+ cores (CPU); 10+ PB storage;** 22+ billion messages daily (vs. 1-3 billion searches daily at Google)
- Currently, all rely on *static over-provisioned services*
    - **No correlation** between IT resource dynamics in DC and the network

All these principle and movements of the industry directly impacted the evolution of DCs, which rapidly found the need to evolve towards:

- Very high throughput transport planes
  - All-optical TP for ultra-low latency and jitter (huge uncompressed data to move, min queuing/OEO)
  - Optical connections applied to both intra-DC and inter-DC (app/service/VM movement)
- Dynamic and flexible end-to-end Net+IT resource control
  - Resource and network elasticity (automated scaling up/down)
  - Tight coupling with virtualization technologies and any IT/network hardware
  - Seamless view and operation, location independent
- Mechanisms and tools for on-demand resource provisioning
  - allow pay-per-use/pay-as-you-grow models

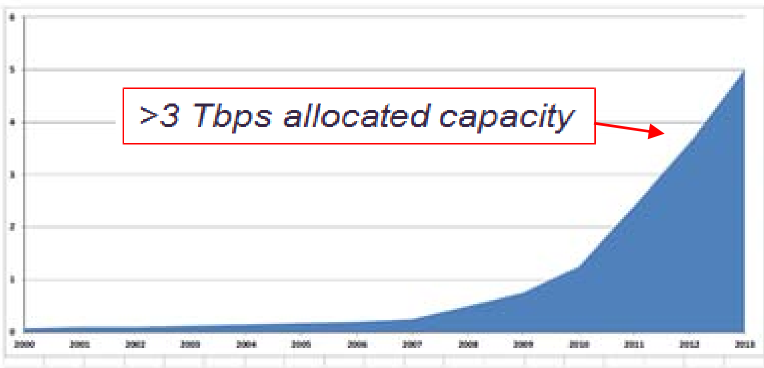


Figure 55. Typical capacity for a distributed DC

8.4. *The Generalized architecture*

The proposed architecture to cover the aforementioned challenges and new services is presented in Figure 56, and it was developed within the framework of the research project GEYSERS [139]. This architecture supported dynamic infrastructure services and unified network and IT resource provisioning. It re-qualified the interworking of legacy planes by means of a virtual infrastructure representation layer for network and IT resources and an advanced resource provisioning mechanism. It offered an innovative structure by adopting the concepts of Infrastructure as a Service (IaaS) and service oriented networking to enable infrastructure operators to offer new network and IT converged services.

On the one hand, the service-oriented paradigm and IaaS framework enabled flexibility of infrastructure provisioning in terms of configuration, accessibility and availability for the user. While, on the other hand, the layer-based structure of the architecture enabled separation of functional aspects of each of the entities involved in the converged service provisioning, from the service consumer to the physical ICT infrastructure. This was achieved through the novel Logical Infrastructure Composition Layer (LICL) already commented, which offered a framework for abstracting, partitioning and composing virtual infrastructures from a set of physical resources in an automated way. And an enhanced network control plane capable of seamlessly controlling the virtualized network and IT resources by means of provisioning dynamic end-to-end services on top of these virtual infrastructures.

Figure 56 shows the layering structure of the proposed architecture reference model. Each layer was responsible to implement different functionalities covering the full end-to-end service delivery from the service layer to the physical substrate. Central to the architecture and focus of the project were the enhanced Network Control Plane (NCP), and the novel Logical Infrastructure Composition Layer (LICL). The Service Middleware Layer (SML) represented existing solutions for service management and at the lowest level there was the Physical Infrastructure layer that comprised optical network and IT resources from different Physical Infrastructure Providers. Each of these layers is further described below.

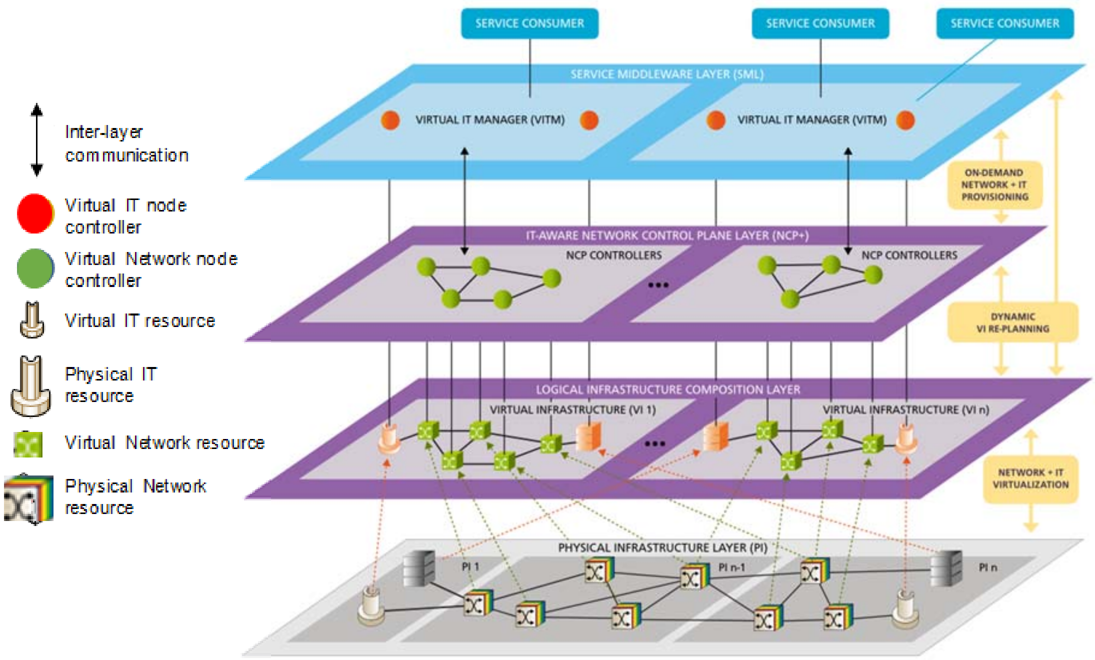


Figure 56. The generalized architecture for dynamic infrastructure services

This architecture delivered an overall service provisioning solution that addressed the following strategic requirements aligned with those already exposed in the previous chapter:

- **Scalability**: Applications designed for cloud computing needed to scale with workload demands so that performance and compliance with service levels remained on target. Moreover, the elasticity concept had to be ported to transport networks, in order to allow combined IT and network scalability.
- **Availability:** Through a synchronised combination of virtualisation, control and management techniques applied to network and IT domains.
- **Reliability:** Application data had to be processed properly and delivered with the minim losses.
- **Security:** Applications needed to provide access only to authorized, authenticated users, and those users needed to be able to trust that their data was secure.
- **Flexibility and agility:** the dynamic virtual infrastructure planning and re-planning processes introduced a new level of flexibility to the virtualisation services.
- **Serviceability:** Once a virtual infrastructure was deployed, it needed to be maintained. Proper synchronisation mechanisms were inherent to the infrastructure provisioning service, based on mixed synchronous-asynchronous models.
- **Efficiency:** In terms of energy or resource utilisation, several techniques for achieving efficient resource and service operation at different layers, ranging from physical infrastructure up to the control plane were implemented [141].

8.4.1. Architecture roles

The layered architecture enabled new service models associated with the way on which different actors (operators, service providers, etc.) would interact with the infrastructure resources. In particular, the term 'role' was defined as the behavior of an entity participating in a specific workflow in a particular context. The detailed definition of roles' functionalities and responsibilities ensured the correct operation of the architecture while at the same time fostered the appearance of different actors willing to adopt these new atomic roles, making the telecom industry much more flexible. This architecture identified three main different roles:

- Physical Infrastructure Provider (**PIP**), who owned a physical infrastructure and was willing to partition and virtualise it in order to make it available to others with the final aim of gaining revenue from it. As a result of these partition and virtualisation operations over the physical resources, the PIP generated Virtual

Resources, that were offered to the VIP as a service, relegating on this provisioning step the operational rights over the resource to the VIPs.

- Virtual Infrastructure Provider (**VIP**), provided VIs to the VIOs, and transferred to them the control rights over the provided VI, to be able to operate it. The VIP acted as an infrastructure broker, responsible for combining different IT and Net resources belonging to different Physical Infrastructure Providers. As a result, single VIs with coordinated control & management functionalities were offered.
- Virtual Infrastructure Operator (**VIO**): it was the entity in charge of operating efficiently the Virtual Infrastructure (VI) and aiming at providing the unified services to the SCs. A main asset of the VIO was the inherent knowledge about how to handle and size the VIs dynamically by means of configuring, reshaping and resizing them so that the service requirements were met in an efficient way.
- Application/service Provider (**A/SP**), the beneficiary of the service offered by the SML layer through the VIO (the so called unified IT + Network services). It only had usage rights.

Relationships between them were structured in a chain-like manner, where the different services provided by each role were consumed by the next role in the chain, until the final service consumer. This scenario led towards a design of the architecture that should be capable of being sliced in layers. Therefore, this design allowed to assign each layer to a given role. The description of the interaction between the roles involved gave the faculty to determine the architectural modules needed by each role and the required functionality that each module has to implement. It also gave an insight on the necessary interfaces between the layers in the architecture and the information that had to flow between them, which was used as a starting point for its design Figure 57.

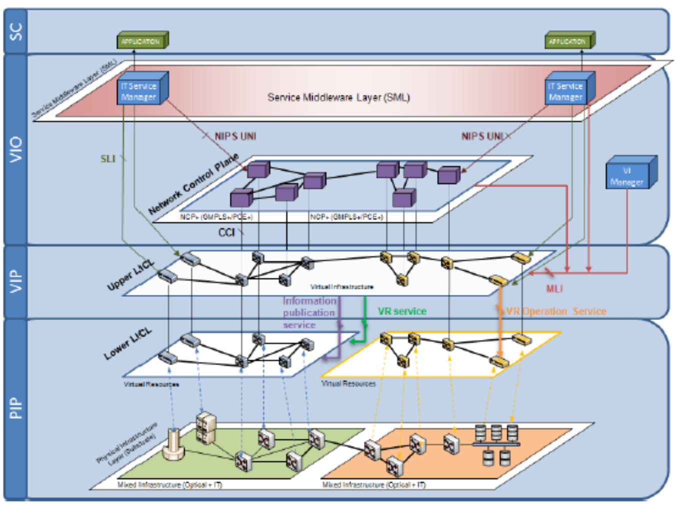


Figure 57. Roles within the architecture

Moreover, the roles could be further split if an entity worked over a single type of resource (network or IT). To denote such cases, suffixes –IT or –N were added to the terms PIP, VIP and VIO. Figure 58 shows the architecture roles services and the evolution offered by this new architecture, where the traditional carriers role were split among Physical Infrastructure Providers (PIP), Virtual Infrastructure Providers (VIP) and Virtual Infrastructure Operators (VIO). PIPs owned the physical devices and rented partitions of them to VIPs. These, in turn, composed VIs of the virtual resources rented at one or several PIPs, and leased these VIs to VIOs. VIOs can efficiently operate the rented virtual infrastructure through the enhanced Control Plane, capable of provisioning on-demand network services bundled with IT resources.

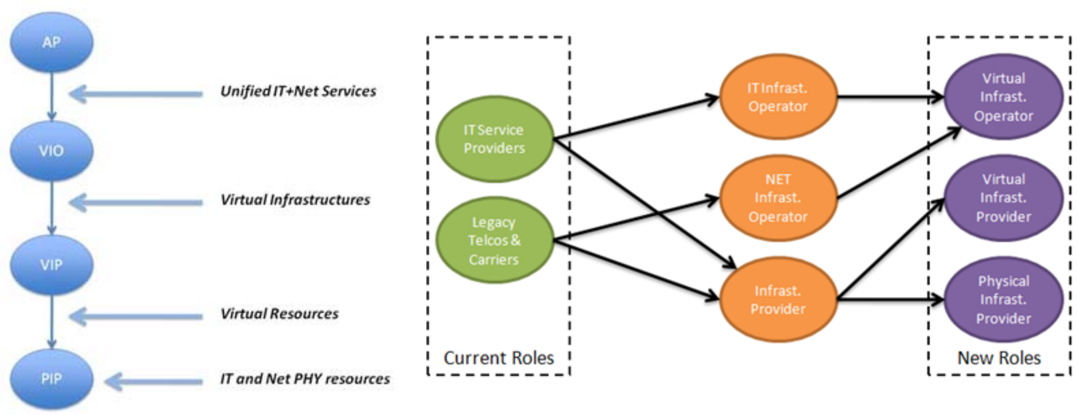


Figure 58. Architecture Roles. a) Type of service consumed by the role, b) evolution of current roles towards the new roles offered by the new architecture

Actually, decoupling the traditionally integrated roles towards infrastructure providers and infrastructure operators brought:

- increased flexibility
- improved manageability
- allowed the exploitation of new revenue streams
- allowed the emergence of new business models
- facilitated the entrance of new "players"
- reduced the barriers for entering the market
- targeted to enable reduction of capex (due to infrastructure multiplexing)

This new roles scenario created new market opportunities for all the different actors addressing: infrastructure providers, infrastructure operators and application providers cooperated in a business model where on-demand services were efficiently offered through the seamless provisioning of network and IT virtual resources. The following use case example describes this new model. A company hosts an Enterprise Information System externally on a Cloud rented from a Software-as-a-Service (SaaS) provider. It relies on the resources provided by one or more IT and network infrastructure providers. It also connects heterogeneous data resources in an isolated virtual infrastructure. Furthermore, it supports scaling (up and down) of services and load. It provides means to continuously monitor what the effect of scaling will be on response time, performance, quality of data, security, cost aspect, feasibility, etc. So, the presented architecture would result in a new role for telecom operators that own their infrastructure to offer their optical network integrated with IT infrastructures (either owned by them or by third party providers) as a service to network operators. This, on the other hand, will enable application developers, service providers and infrastructure providers to contribute in a business model where complex services (e.g., Cloud computing) with complex attributes (e.g., optimized energy consumption and optimized capacity consumption) and strict bandwidth requirements (e.g., real time and resilience) can be offered economically and efficiently to users and applications.

8.4.2. Architecture interfaces

The communication between the abovementioned four main layers (SML, NCP, LICL, and PI) was done through interfaces (Figure 59) between layers (mainly vertical interfaces), with different types of functionalities in each interface [143]. This interfaces were totally implemented in terms of functionalities, messages, key fields, workflows, and information exchange models (IEM), although not part of this Thesis [144][145]. Following, there is a brief review of these interfaces and their functionalities:

- **SML to LICL Interface (SLI)** described the messages and interactions necessary to operate over the IT virtual resources within a Virtual Infrastructure (VI). The SLI had the following functionalities: Information Retrieval, Virtual Resource Configuration, Virtual Resource Monitoring and Notification, Runtime Control, Virtual Resource Creation and Virtual Resource Destruction.
- **SML to NCP interface (NIPS UNI)** enabled a joint and on-demand provisioning of network and IT resources. The NIPS UNI was defined between a NIPS client located in the SML and a NIPS server located in the NCP+. The main functionalities offered by the NIPS UNI were: NIPS service discovery, IT resource advertisement, Service setup and tear-down, Service modification and monitoring.

- **LICL to PHY interface (LPI)** was responsible for the discovery of physical resources and their characteristics, the provisioning and configuration of resources, the monitoring and the decommissioning of resources. The main functionalities offered by the LPI were: Synchronization, Provisioning, Monitoring and Release.
- **Management to LICL Interface (MLI)** provided all the management functionalities over a Virtual Infrastructure in its lifecycle, that was, from its request to its decommissioning. The functionalities that had to be supported by the MLI were: VI Requesting, VI Instantiation, VI Decommissioning, VI Re-planning.
- **Connection Controller Interface (CCI)** was used by the NCP+ to communicate with the virtual network resources in the LICL. The main functionalities offered by the CCI were: Synchronization, Configuration and Monitoring.
- **Common Security Service Interface (CSSI)** allowed the application to consume AAI security services. The main functionalities offered by the CSSI were: Authentication, Delegation and Authorization.

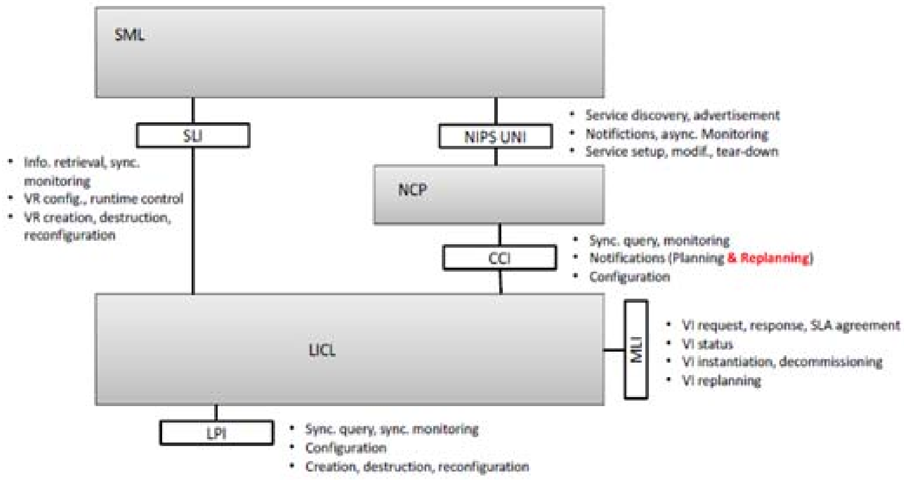


Figure 59. Interfaces between architecture components

8.4.3. *Physical infrastructure*

The physical infrastructure layer was composed of optical network and IT resources [146]. These were resources that belonged to one or more physical infrastructure providers and so could be virtualized by the LICL. The term infrastructure referred to all the physical network resources (optical devices/physical links) used to provide connectivity across different geographical locations, and the IT equipment providing storage space and/or computational power to the service consumer. From the network point of view the architecture was designed to be based on L1 optical network infrastructure. But with the aim to be generic enough to cover most of the technologies used in existing optical backbone infrastructures and those offered by infrastructure providers/operators of the time. Considered technologies were Fiber Switch Capable (FSC) and Lambda Switch Capable (LSC) devices. From an IT point of view, IT resources were considered as service end-points to be connected to the edge of the network. IT resources referred to physical IT infrastructures, such as computing and storage.

The physical infrastructure provided interfaces to the equipment that allowed its operation and management, including support for virtualization (when available), configuration and monitoring. Depending on the virtualization capabilities of current physical infrastructure, physical infrastructure providers implemented different mechanisms for the creation of a virtual infrastructure. In terms of optical network virtualization, it considered optical node and optical link virtualization. Moreover, the virtualization methods included partitioning and aggregation. Following there is a clarification of the meaning for partitioning and aggregation (Figure 60).

- Optical node partitioning: It entailed dividing an optical node into several independent virtual nodes with independent control interfaces by means of Software and Node OS guaranteeing isolation and stability.
- Optical node aggregation: It entailed presenting an optical domain or several interconnected optical nodes (and the associated optical links) as one unified virtual optical switching node with a single/unified control interface by means of Software and Control/Signalling Protocols. The controller of the aggregated virtual node should manage the connections between the internal physical nodes and show the virtual node as a single entity.
- Optical link partitioning: It entailed dividing an optical channel into smaller units. Optical fibres could be divided into wavelengths and wavelengths into sub-wavelength bandwidth portions that could be performed e.g. using advanced modulation techniques. The latter was a very challenging process especially when the data rate per wavelength was >100Gbps.
- Optical link aggregation: Several optical wavelengths could be aggregated into a super-wavelength with aggregated bandwidth ranging from wavelength-band, to fibre or even multi-fibre level.

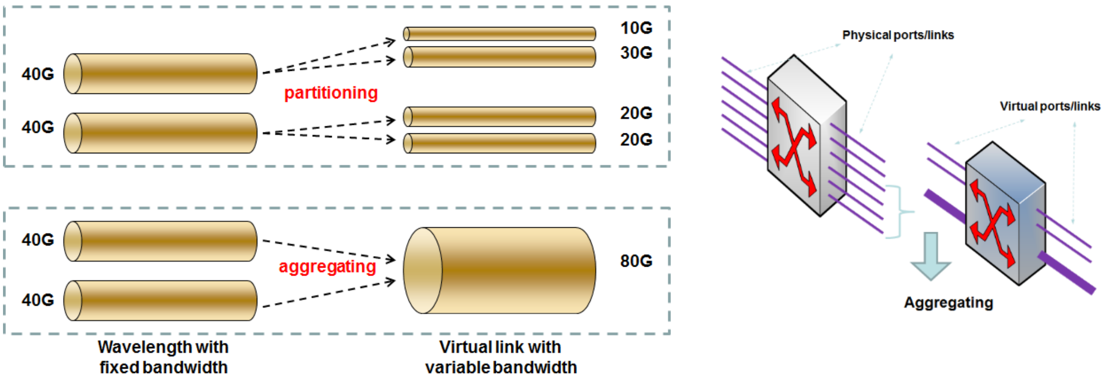


Figure 60. Partitioning and aggregation functionalities

After partitioning and aggregation, optical virtual nodes and links were included in a virtual resource pool used by the LICL to construct virtual infrastructures (Figure 61); thus, multiple virtual infrastructures can share the resources in the optical network. This meant that isolation between the partitioned virtual resources had to be guaranteed at both data (physical isolation) and control level.

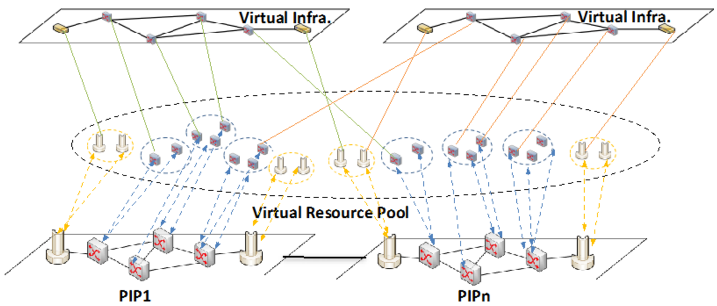


Figure 61. Virtual infrastructure resource pool

8.4.4. Logical infrastructure composition Layer

In terms of infrastructure management, the main objective of the LICL [147][148] was to provide a mechanism for virtualisation and associated techniques such as uniform resource description, resource abstraction and composition. It was located between the physical infrastructure and the upper layers, NCP and SML. The LICL (Figure 63) relied on a solid resource description framework, which allowed applying a common set of procedures and signals to both network and IT resources, so that, it was responsible for the creation and maintenance of virtual resources as well as virtual infrastructures. Infrastructure virtualisation was the creation of a virtual representation of a physical resource (e.g., optical network node or computing device), based on an abstract model that was often achieved by partitioning or aggregating. A virtual infrastructure was defined as a set of virtual resources interconnected together which shared a common administrative framework. Within a virtual infrastructure, virtual connectivity (virtual link) was defined as a connection between one port of a virtual network element to a port of another virtual network element. Moreover, the LICL allowed novel on-demand planning and re-planning actions to be invoked over the virtual infrastructure in a coordinated action between the VIO and the LICL, for better performance, optimal resource usage and efficient service provisioning.

The LICL utilized a semantic resource description [148] and information modelling mechanism for hiding the technological details of the underlying physical infrastructure layer from infrastructure operators. Consequently, the LICL acted as a middleware on top of the physical resources and offered a set of tools that enable IT and Optical Network resource abstraction and virtualization. The LICL managed the virtual resource pool (Figure 61) where virtual resources were represented seamlessly and in an abstract fashion using a standard set of attributes, which allowed the enhanced Control Plane to overcome device dependency and technology segmentation. The LICL also brought the innovation at the infrastructure level by partitioning the optical and IT resources belonging to one or multiple domains. Finally, LICL supported the dynamic and consistent monitoring of the physical layer and the association of the right security and access control policies. LICL mainly supported the following functionalities (Figure 62):

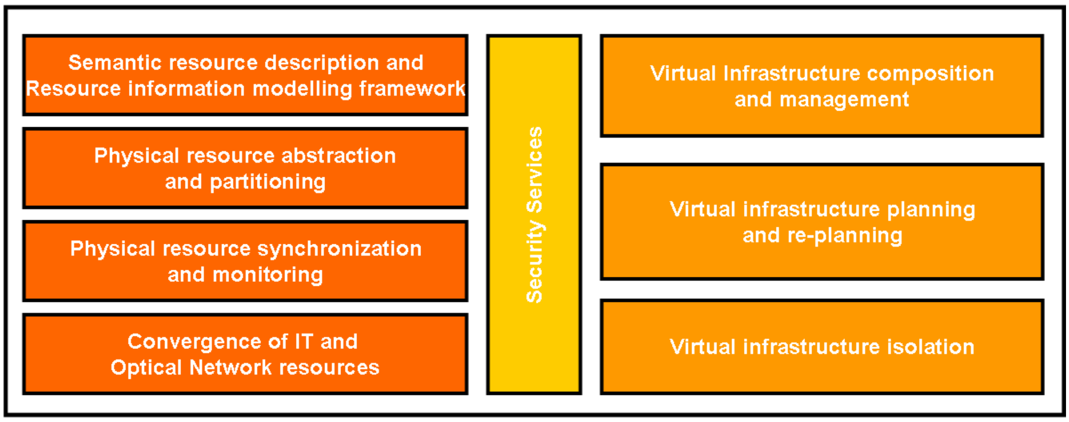


Figure 62. LICL functionalities

As depicted in Figure 62, and in order to facilitate the virtual infrastructure provisioning, the physical resources that were to be abstracted and partitioned had to be exposed in order to enable the subsequent processes. Thereafter, the Physical resource abstraction and partitioning functionality in LICL were responsible for abstracting the physical infrastructures and representing them as a set of attributes, characteristics and functionalities. Hence the unnecessary characteristics from the resource itself could be hidden.

The Semantic resource description and resource information modelling defined and implemented a resource information model that could be used for the definition of both optical and IT infrastructures. The Semantic resource description and resource information modelling was based on semantic techniques for resource description. Moreover, it included the energy efficiency properties to enable the novel energy aware NIPS service provisioning. After the physical resources had been abstracted and partitioned, the synchronization and monitoring functionality were required in order to ensure reliability and coherence at every stage of the resource lifecycle.

The Virtual infrastructure planning and re-planning was considered as a key enabler in the service provisioning architecture. On the one hand, it enabled the automated virtual infrastructure planning that was triggered by the requests from a VIO. On the other hand, the dynamic re-planning of virtual infrastructures happened when the VIO required more resources or releases leased resources during the operation. Above all the functionalities, a set of tools were also provided in LICL in order to compose and manage the virtual infrastructure, which was referred to as the Virtual resource composition and management in Figure 62. This component enabled VIPs to compose and manage the virtual infrastructures already provisioned for the different VIOs. Convergence of IT and optical network resources functionality allowed the LICL to build virtual infrastructures composed of both optical network and IT resources. This functionality facilitated later the NIPS service provisioning process.

Finally, the Virtual Infrastructure Isolation functionality guaranteed that there was no overlapping between different virtual infrastructures that could cause inconsistencies at the physical layer. Moreover, it allowed different VIOs to operate different virtual infrastructures keeping privacy between them. All the resources had to be handled in a secure manner, due to the multi-infrastructure provider nature of the LICL, where virtual infrastructures were composed by means of virtualisation of resources belonging to different domains. Security handling was considered a transversal functionality, which was dimensioned over the whole LICL layer, including access control to resources and virtual infrastructures, data protection and policy enforcement. The functionalities provided by the LICL solved needs from both PIPs and VIPs; which led to a division of the LICL in two sub-layers: the lower LICL and the upper LICL (Figure 63). Each one of these two components comprised the functionalities related to one of the architecture roles. The lower element was the one related with the PIP, and its main task was to provide an abstraction from the physical infrastructure to upper layers. While the main task of the upper sub-layer was to create virtual infrastructures from a collection of virtual resources, which corresponded with the VIP's main functionality.

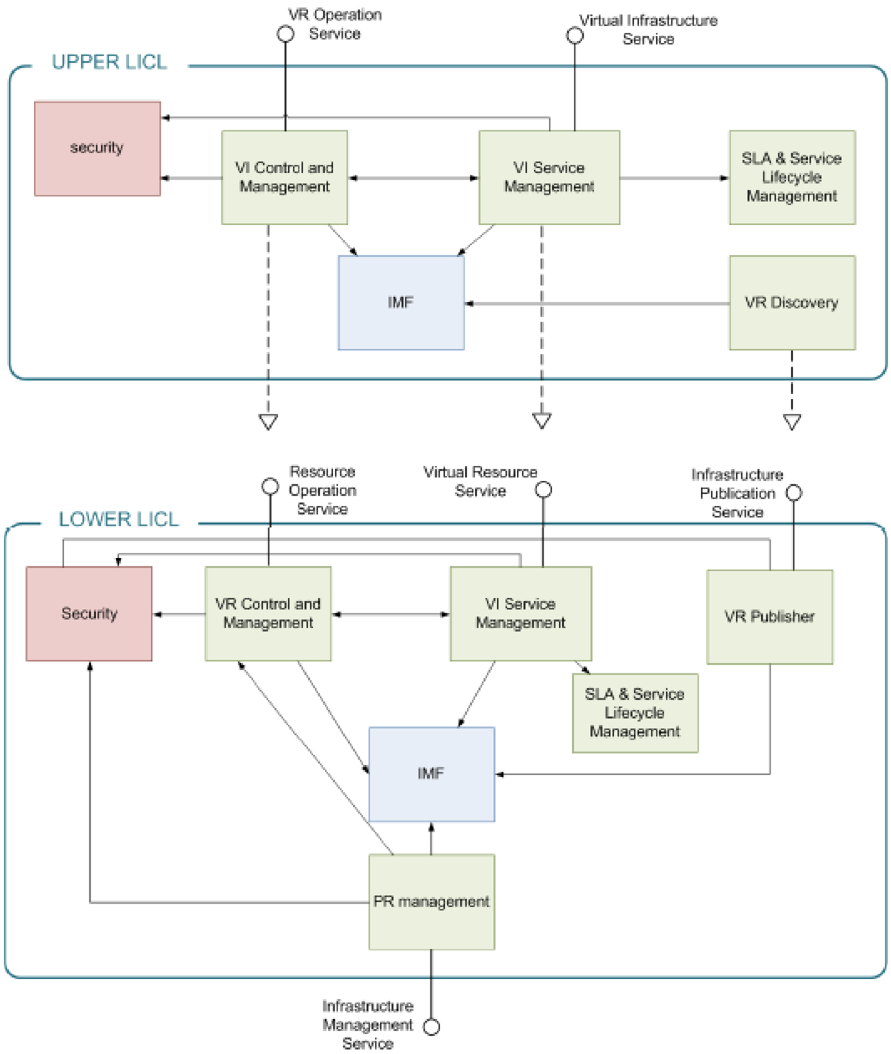


Figure 63. LICL architecture

One of the key elements from the LCL architecture was de Information Modelling (IMF). To enable the different LICL components to interact using a common vocabulary, an information Model was needed. As such, the information model was mostly an internal model for the LICL, although some LICL interfaces and other system components could also use it. A number of requirements for the LICL information model were also described [149]. These requirements were described in terms of the types of information, and what the information Model described: physical resources (IT and network aspects), virtual infrastructure and virtual infrastructure requests, energy related aspects, quality of service and security aspects. Based on this state of the art overview we select two existing information models (NDL and VXDL) [157][150] as the basis for the LICL information model.

The LICL required privileged access to the physical infrastructure resources in order to implement isolation in an efficient manner. It also worked as a middleware that forwarded requests and operations from the NCP to the physical infrastructure native controllers. This was achieved by using a Virtual Infrastructure Management System (VIMS). The VIM was composed by a set of tools and mechanisms for control and management of its resources. As already commented, the LCL enabled the optical network virtualization. Virtual resources were obtained by means of different types of paradigms:

- **Abstraction** (1 PHY to 1 Virtual)
- **Partitioning** (1:N) – Direct scope (IaaS)
- **Aggregation** (N:1) – Indirect scope (NaaS)

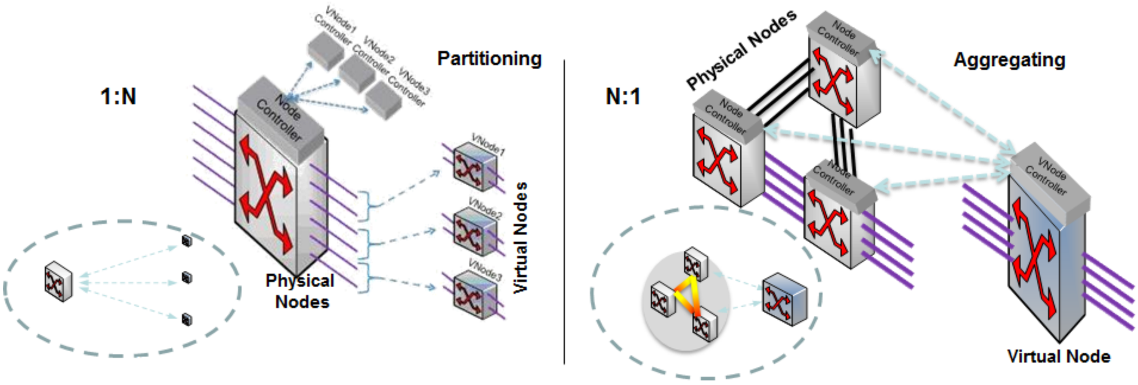


Figure 64. LCL and resource virtualisation

8.4.4.1. *Resource virtualisation*

In the context of network and computing infrastructure, virtualisation was the creation of a virtual version of a physical resource (e.g. network, router, switch, optical device or computing server), based on an abstract model of that, which was often achieved by partitioning (slicing) and/or aggregation.

Resource virtualisation was a critical enabler for the LICL that was closely related to the subsequent VI provisioning and operation. Resource virtualisation in LICL could be categorized into four paradigms: aggregation, partitioning, abstraction and transformation, as shown in Figure 65 it contemplated these four paradigms and how they could be supported for different type of resources, e.g. IT and optical Netwok resources.

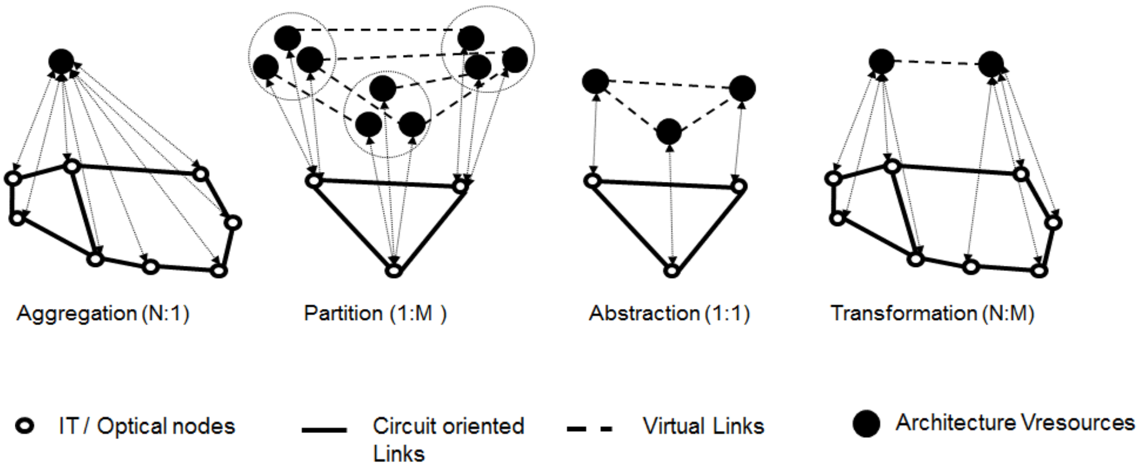


Figure 65. Virtualisation paradigms considered within GEYSERS.

8.4.4.1.1. IT Resource virtualisation

The IT resources considered were computing and storage nodes, running user applications, and being interconnected by a virtualized network infrastructure.

Users could request such IT resources that were in reality partitions or aggregations of real physical resources. An IT resource could be partitioned into N virtual IT resources using common virtualisation technologies, such were Xen [152], KVM [153], VMware [154], VServer [155], etc., where each partition was represented as a virtual machine (VM) with computing and storage resources. These technologies used different types of virtualisation to partition the resources. While OS-level virtualisation (e. g. VServer) offered interesting performance, it allowed only limited isolation and customization. On the contrary, performing emulation and hardware virtualisation (e. g. KVM, Xen), each VM had its own isolated execution environment where any OS could run.

As opposed to partitioning, aggregation consisted in exposing a set of physical IT resources as a single virtual IT resource to the user. Such aggregation was for example possible with vSMP [151] (Versatile SMP) which aggregated many physical servers and made them appear to the OS like one giant machine with many cores.

Regarding only storage nodes, it was possible to aggregate different disks into a common logical storage pool. This could also be done using SNIA technology, allowing not only sharing a device into several ones, but also aggregating several physical devices and making them appear as one single virtual device.

8.4.4.1.2. Network Resource virtualisation

Network virtualisation brought the concept of server virtualisation to the framework of communication networks. The first attempts for network virtualisation came from the

IP routing world (e.g.[156]), where routers were sliced into virtual routers and interconnected by virtual links such as VLANs. However, virtualising optical networks had major differences to such technologies. Optical networking traditionally relied on a strong involvement of manual planning, engineering and operation, due to the fundamental impact of physical impairments in network creation and service provisioning. Even with the widespread of GMPLS-based control planes, optical networks were still manually operated by most network operators. The first attempts to automate such networks were facing the challenge of how to automate the impairment-aware service provisioning and integrate it with the concept of dynamic control planes and zero-touch networking. Thus, optical network virtualisation had to address all these challenges inherent to optical networking.

Optical network virtualisation was the creation of virtual instances of optical network resources the behaviour of which was the same to that of their corresponding physical optical network resources. It relied on the abstraction of heterogeneous network resources, including nodes, links and segments comprising both nodes and links.

Optical network virtualisation techniques depended on the type of optical element to be virtualized and should enable the representation of the virtual optical resources inheriting the critical characteristics of the physical ones.

In this architecture, the basic optical elements to be considered were optical nodes and optical links. Each virtual instance of an optical node had its own ports and switching capability and the separation and isolation between the control of each virtual instance depended on the virtualisation capabilities of the device itself. Regarding optical link virtualisation, it consisted of abstracting optical data links as virtual instances by partitioning or aggregation.

The partitioning of optical data links was introduced by dividing the link capacity into smaller units, resulting in the granularities of subwavelength and wavelength while the aggregation results in a granularity of waveband (or fibre or group of fibres). Optical fibre partitioning was easily achieved in DWDM where the optical links (i.e. fibres) could be inherently split into individual wavelength channels. Highest bandwidth granularity allowed for more efficient bandwidth utilisation could be achieved by having access to even lower bandwidth units at the sub-wavelength level. The virtualisation capability of a link was related to the optical port characteristics of the associated optical node.

8.4.5. Network + IT Control Plane (NCP+)

The provisioning and control of the end-to-end reservation of IT and network resources was performed by the NCP, which also provided optimized path computation to minimize the energy consumption, realizing the energy efficient routing at the network level, and taking into account a variety of “green TE-parameters” as additional constraints, so achieving the optimal energy consumption at the IT level by selecting the most efficient IT end-points. The network and IT control plane (NCP+) operated over a virtual infrastructure. The virtual infrastructure was accessed and controlled through a set of interfaces provided by the LICL for operation and re-planning services. The NCP+ offered a set of functionalities towards the SML, in support of on-demand and coupled provisioning of the IT resources and the transport network connectivity associated to IT services.

The combined Network and IT Provisioning Service (NIPS) required the cooperation between SML and NCP+ during the entire lifecycle of an IT service. This interaction was performed through a service-to-network interface, called NIPS UNI [158]. Over the NIPS UNI, the NCP+ offered functionalities for setup, modification and tear-down of enhanced transport network services (optionally combined with advance reservations), monitoring and cross-layer recovery.

This architecture supported several models for the combined control of network and IT resources. The NCP+ assisted the SML in the selection of the IT resources providing network quotations for alternative pairs of IT end points (assisted unicast connections). Alternatively the NCP+ could select autonomously the best source and destination from a set of end points, explicitly declared by the SML and equivalent from an IT perspective (restricted anycast connections). In the most advanced scenario, the NCP+ was also able to localize several candidate IT resources based on the service description provided by the SML, and computed the most efficient end-to-end path including the selection of the IT end-points at the edges (full anycast connections). This was a key point for the optimization of the overall infrastructure utilization, also in terms of energy efficiency, since the IT and network resources configuration was globally coordinated at the NCP+ layer [159].

The NCP+ was based on the ASON/GMPLS [160] and PCE [161] architectures. It was enhanced with routing and signalling protocols extensions and constraints based route computation algorithms designed to support the NIPS and, on the other hand, to optimize the energy efficiency for the global service provisioning. Particularly the NCP layer implemented mechanisms for advertisement of the energy consumption of network and IT elements as well as computation algorithms which were able to combine both network and IT parameters with energy consumption information to select the most suitable resources and find an end-to-end path consuming the minimum total energy. Figure 66 shows a high-level representation of the NCP+: the routing algorithms at the PCE operated over a topological graph created combining network and IT parameters with "green" parameters, retrieved from the SML (IT side) and the LICL (network side).

Finally, another key element for NCP+ was the interaction with the LICL in order to trigger the procedures for the virtual infrastructure dynamic re-planning on the network side. In case of inefficiency of the underlying infrastructure, the NCP+ requested the upgrade or downgrade of the virtual resources in order to automatically optimize the size of the virtual infrastructure. The involved algorithms took into account current network traffic, forecasts for resource availability and utilization in the medium and long terms, as well as specific SLAs established between provider and operator for dynamic modifications of the rented virtual resources.

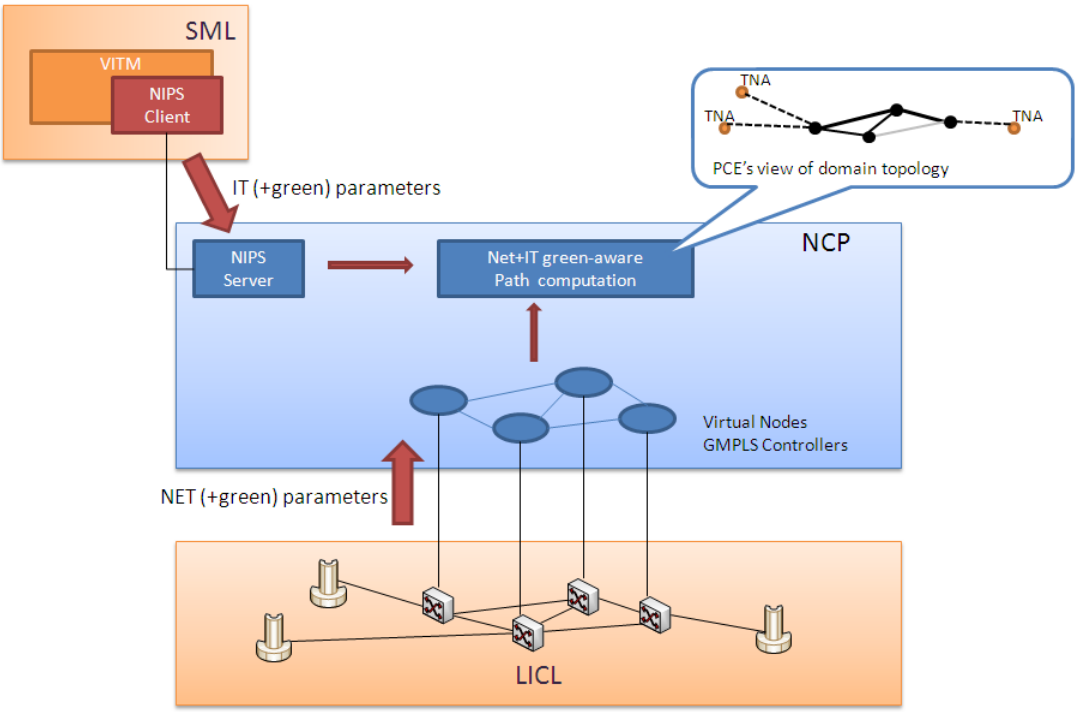


Figure 66. NCP Network and IT energy-aware provisioning high level architecture

8.4.6. Service Middleware Layer (SML)

The Service Middleware Layer (SML) was a convergence layer for coordinating the management of IT resources that belonged to an aggregate service (AS). An AS was a collection of heterogeneous services from different providers used to deliver a single capability to a specific customer or a market segment. This architecture primarily focused on the Infrastructure as a Service (IaaS) concept when referring to services that supported specific applications.

The SML was hence responsible for the following tasks from an application-level perspective, where applications were deployed by Service Consumers:

- Matching application requests to infrastructure resources as specified in Service Level Agreements (SLAs).
- Monitoring and maintaining a "landscape model" of the infrastructure and applications under the management of the SML instance. A landscape model was a specification of a collection of managed elements and their associations. These managed elements could include software, data-base instances, virtual machines and physical machines.
- Triggering the infrastructure provisioning process with application deployment or adaptation requests. These requests contained properties and constraints to be satisfied by resources.
- Accounting of resource usage on a customer or application basis.
- Authentication and authorization processes for access to virtual resources on a customer or application basis.

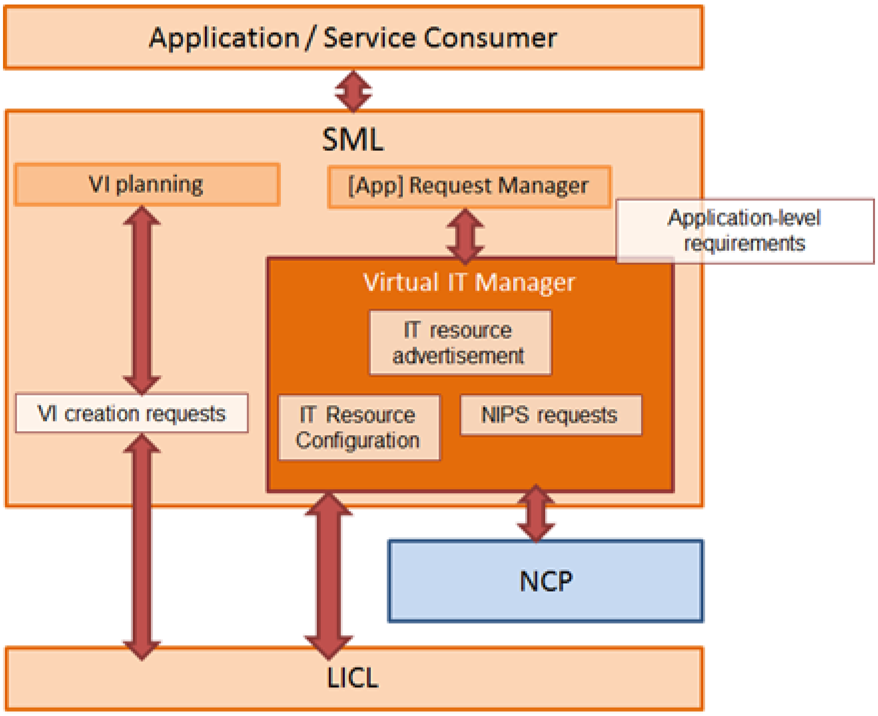


Figure 67. SML internal architecture

The SML exposed an interface to application providers and customers, such that the complexity of network and IT provisioning was transparent to them. Business objectives [162] for a specific application scenario were declared to the SML and translated into provisioning requests understood by a Virtual IT Manager (VITM). The VITM was in charge of the end-to-end IT service management and the virtual IT resource configuration. The SML also maintained a registry of assets (local physical or virtual infrastructure), IT infrastructure providers and network providers. The registration entry for each of them included the path to the respective control planes of these resources (VITM for IT and NCP for network). For example, the SML maintained entries of network providers, including information about the service access points of their respective NCP, such that networking resources can be reserved for the purpose of the application.

8.4.7. Service consumer or application

The service consumer or application was the user of a virtual infrastructure. A virtual infrastructure could be offered to a single application or user (super user) or shared among several users or applications.

8.5. Architecture services

The architecture presented provided two main services; namely, the virtual infrastructure service and the on-demand provisioning service. The former comprises the process of requesting and creating a VI, whilst the latter deals with the provisioning of the connectivity required between the different virtual IT elements involved in a specific application landscape. Both correspond to the planning [142] and operation phases of the service lifecycle. Also, the roles model with respect to role interactions, described the services that were consumed between different roles within the business model to fulfil the general service provisioning. Below (Table 14) there is an overview and a description of the main services considered for the presented architecture, which facilitates the achievement of the challenges presented at the beginning of the chapter [144]. Table 14. Summary of the basic services of the architecture

- **Virtual Infrastructure Service:** The virtual infrastructure service provided virtual infrastructures on-demand. This service took as inputs requirements for the VI composition and delivers a slice of the physical infrastructure as virtual IT resources interconnected by a virtual network topology.

- **Infrastructure Information Service:** The infrastructure information service involved the VIP and the PIP, by means of the upper-LICL and the lower-LICL respectively. This service enabled the information exchange between the PIP(s) and the VIP, and allowed the VIP to receive information from the different PIPs about the resource kinds they hold, the minimum granularity in which those could be partitioned, and the connections towards other PIPs (i.e. the inter-domain links). This service was periodically updated in order to maintain the information at the VIP level as accurate as possible.

- **Virtual Resource Service:** The virtual resource service involved, equally to the Infrastructure information Service, the VIP and the PIP. In this case, the VIP acted as consumer, while the PIP was the provider. This service enables the VIP to request for VR reservation, and instantiation. It was the service that allowed the VIP, during the planning phase, to request for VR reservation to the corresponding PIP.

- **Enhanced Network Connectivity Service:** The enhanced network connectivity service involved the VIO-IT, acting as consumer, and the VIO-N, acting as provider. It offered the on-demand provisioning of optical network connectivity between IT end-points attached to the virtual network infrastructure. The enhanced network connectivity service was characterized by the features defined in the architecture in support of the Unified IT + Net Service (e.g. provisioning of network quotation, automatic selection of the IT resource at the NCP+, support for advance reservations…).

- **Unified IT + Net Service:** The unified IT + Net service provided the on-demand reservation of IT resources and connectivity services in order to interconnect them in a seamless way. The global set of resources were selected taking into account the application requirements and could be modified (e.g. changing the bandwidth or adding new IT resources) during the service life-cycle according to the dynamics of the application.

Service	Consumer	Provider	Interface	Service type
Virtual Infrastructure Service	VIO	VIP	MLI	Planning
Infrastructure Information Service	VIO, VIP	VIP, PIP	Internal LICL interface	Information
Virtual Resource Service	VIP	PIP	Internal LICL interface	Planning
Enhanced Network Connectivity Service	VIO-IT	VIO-N	NIPS UNI	Operation
Unified IT + Net Service	Application provider	VIO (-IT)	Appl. layer <-> SML	Operation

Table 14. Summary of the basic services of the architecture

8.5.1. Service delivery and VI life cycle

The VI Provisioning service [146] consisted of several phases that include both automated and engineer/human assisted procedures Figure 68.

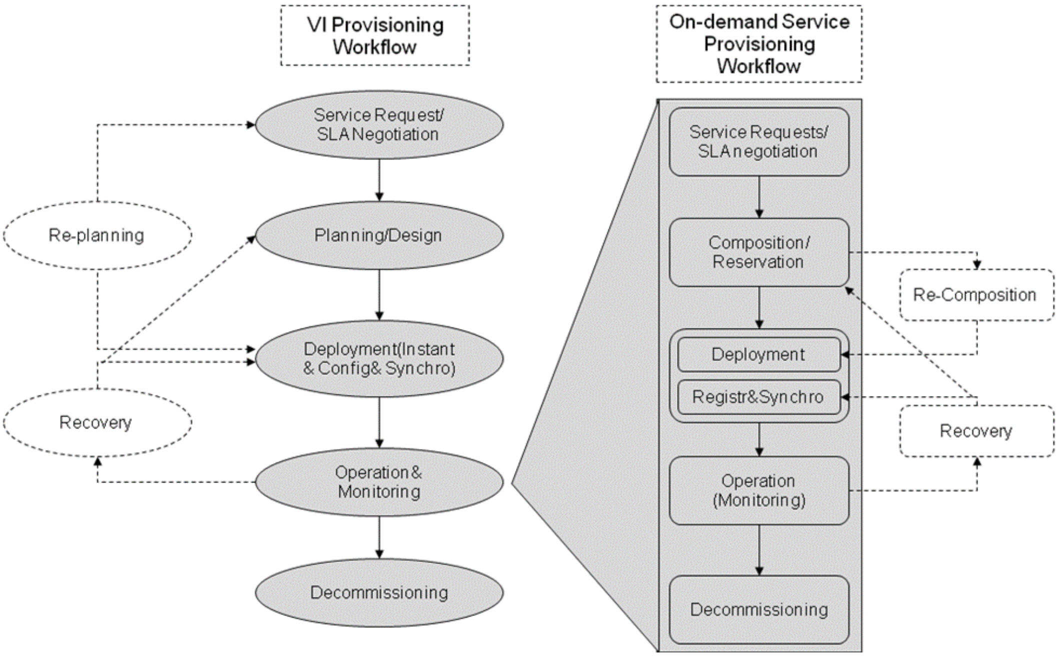


Figure 68. VI Provisioning workflow

In the Figure 68, circles represent the different stages composing the whole workflow and arrows represent transitions between the different phases. The VI service provisioning request started with a Service Request/SLA Negotiation. In this phase, the VIO defined the requirements for the desired VI and initiated the SLA negotiation with the VIP. The SLA defined in this phase provided a set of basic requirements, which included QoS requirements, security policies and robustness requirements among some others. The security policies defined in this phase were used in the planning, deployment and operation phases. Additionally, SLAs would contain trust anchors in a form of public key certificates. It was an initiation point of the VI lifecycle (Figure 69).

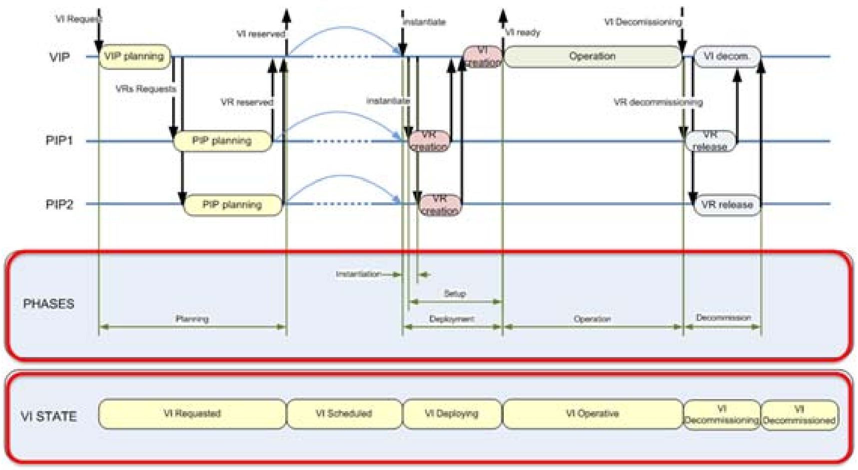


Figure 69. VI life cycle

Just after the service request phase, there was the Planning/Design phase. This service could be decoupled into the following sub-phases:

- Virtual Infrastructure design: The Virtual Infrastructure design was carried out by a VIP based on the requirements received from the service request and SLA negotiation phase. These requirements were in form of SLAs that were expected to be fulfilled for the service provisioning. These SLAs were decomposed into more technical constraints in the SML in a semi-automated fashion. The SML was a rule-based expert system which could also support a human interaction for planning. Fully automated planning schemes were also considered in the SML architecture.
- Virtual resources selection and composition: In this sub-phase, the VIP searched/negotiated the virtual resources offered by one or multiple PIPs. Dynamic Algorithms were used compose IT and network resources and to produce as a result, a blueprint of the virtual network was available, and ready to be included in a contract between the VIO and the VIP.
- Virtual resource reservation: In this sub-phase, each selected virtual resource was associated with a common reservation ID (to be hereafter referred to as Global Reservation ID (GRI)) that also bound the reservation session/instance with the SLA initiated at the provisioning process. The reserved resources needed to be configured and initiated in the deployment phase.

Once the Planning/Design phase was finished, the process entered into the deployment phase. During the VI deployment phase, the reserved infrastructure instances were instantiated, configured, registered, and initialised. This phase allowed the review and approval from the network/IT engineers. The deployment phase could also be decoupled into the following sub-phases:

- NCP and IT controller's instantiation and deployment: in this sub-phase, the VIO deployed its NCP by taking consideration of the VI specifications. The software modules were then deployed/installed to control the different virtual network nodes composing the VI. Similarly the IT controllers related to the IT virtual resources were deployed/installed.
- Configuration of the NCP and IT controllers: The network controllers and PCE modules deployed in the NCP were configured with the network topology

information and policies. Similarly, the IT controllers were configured with information about virtual IT resource availability and properties.

- NCP initialization: The NCP modules were initialized and started, and it is when the network auto-configuration process took place (e.g. neighbour/UNI discovery, initial flooding of TE parameters and routing protocol convergence). In this phase, the IT controlled also injected the capabilities of the IT site under its control into the NCP.
- Instant Network+IT service/infrastructure registration and initialisation: this sub-phase allowed a new service to be registered in the VIO and put into operation. It also allowed binding security and provisioning sessions with the service ID and (underlying/implementing) platform runtime environment. The importance of specifying this phase was defined by the need to address such scenarios as infrastructure re-planning and failure restoration.

As a result of the instantiation phase, the VIO had configured the virtual resources and had deployed its control plane over the virtual infrastructure. The virtual infrastructure was up and running. At this point is when the service entered into the Operation and Monitoring phase. This phase included all the processes for the provisioning of network + IT services (NIPS) to users. During the operation phase, the VIO ran its own virtual infrastructure provisioning service that was targeted to deliver the necessary infrastructure resources (both network and IT) to users, project or applications. It was intended that this provisioning process was automated and allowed using the same business model as traditional physical operators although behaving under different roles, depending on the model role. The on-demand service provisioning happened in this phase.

Along the whole provisioning service the Re-planning and Recovery phases could take place. These were additional phases triggered by special events during operation, or on the request process of any of the actors. Re-planning was a special VI stage in which the LICL implemented a change in the VI. This phase is further detailed in the next section. A recovery phase/process took place when the running virtual/provisioned service failed (e.g. because of hardware failure). Depending on the type of failure, restoration could require just restarting/redeploying the virtual service or involve new planning/design/reservation processes.

The last phase of the VI provisioning service was the Decommissioning, which was triggered whenever a virtual infrastructure was no longer in operation and had to be terminated. This usually happened when the leasing contract between VIP and VIO ended and the VI was no more suitable for other VIO customers of the VIP. The termination phase ensured that all the authorization right of the VIP for access to the PIP resources were inactivated as well as the authorization right of the VIO for access to the virtualized physical resources. Once a VI was decommissioned, the physical resources of the PIPs became available for planning and instantiation of new VIs.

In essence, the VI provisioning service consisted in creating virtual infrastructures upon request and in on-demand basis. From a business perspective the VI provisioning service involved the participation of several of the mentioned architecture roles. In this architecture the virtual infrastructure operator (VIO) was considered an entity generating the request. Nevertheless, anyone in need of a virtual infrastructure could issue a VI request (e.g. application provider). Figure 70 shows the most basic workflow diagram for the VI provisioning service. The service started with the VIO requesting the creation of a VI to a virtual infrastructure provider (VIP). The VIP processed the requests and interacted with the Physical Infrastructure Provider (PIP) to request the creation of virtual instances of the physical resources (by partitioning or aggregation). Once the required virtual resources (VR) had been created, the VIP used them to compose a virtual infrastructure and offer it to the VIO.

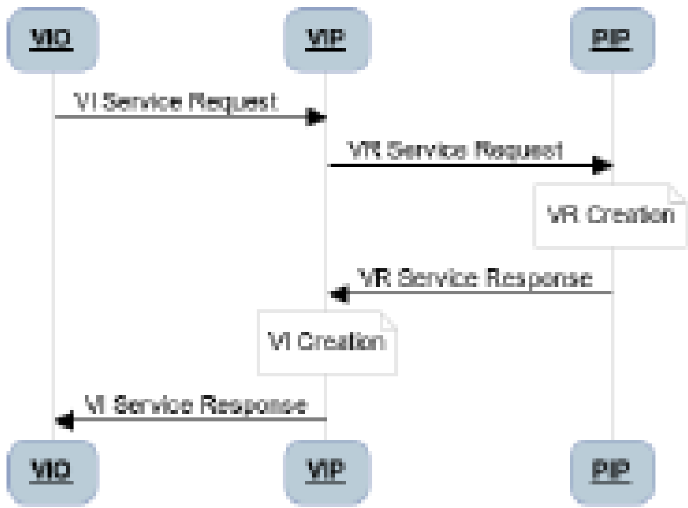


Figure 70. Service provisioning from a business perspective.

During this process, negotiation between the different roles is required when different actors were carried out. Figure 71 show a use case example of the provisioning of an Anycast Network in virtual infrastructure:

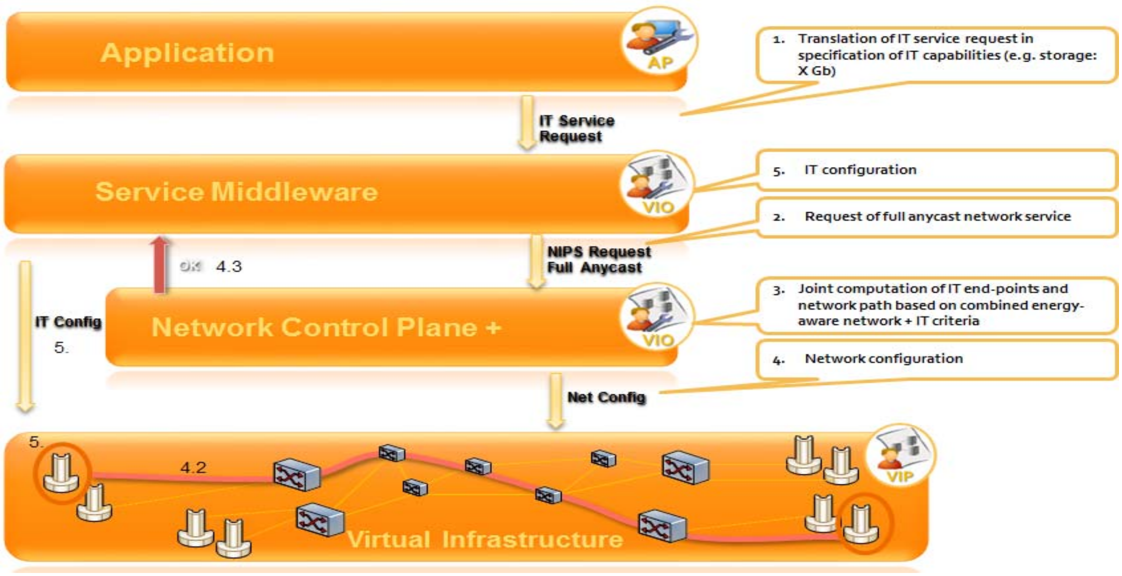


Figure 71. Example of an Anycast Networks in virtual infrastructures service

8.5.2. VI Re-planning

Re-planning was a VI provisioning phase that was triggered by special events or upon request by an actor (specific roles) during the operation phase of an already planned and provisioned VI [164]. Therefore, VI re-planning was a stage in which the LICL would be responsible for implementing the required changes in the virtual infrastructure. The requirement for VI re-planning could be triggered at any time once the virtual infrastructure had been instantiated.

Within this scope of the research done, the VI re-planning (Figure 68) procedure could involve a modification of a network infrastructure, IT infrastructure or both network and IT infrastructures. When the procedure described modifications applied to the network infrastructures, it was referred to as VI network re-planning. Analogously, when the procedure affected IT infrastructures only, it was referred to as VI IT re-planning.

From an operator perspective, infrastructure re-planning usually took place in long timescales and was usually human-driven, although supported by dimensioning tools, as in some cases it could involve the investment of CapEx for the acquisition of new physical equipment and its installation in the infrastructure. However, it was relevant to consider a VI re-planning that could take place in short timescales, so it would be useful for infrastructure operators to adapt their infrastructures dynamically with the aim to improve the efficiency of resource utilisation and increase service availability for end users.

It is important to take into account how the whole process of modification of the existing virtual infrastructure was performed. It had been identified as a key functionality for the VI re-planning to be performed in short timescales and be supported by a dynamic procedure for the implementation of changes in the virtual infrastructure. This procedure was named automatic VI re-planning. The automatic procedure allowed applying the required changes immediately, with a limited or preferably no human intervention with supporting tools. When the process of VI re-planning took place in long timescales, there was no need from an infrastructure operator point of view to have full automation of the whole VI re-planning procedure. This procedure was named manual VI re-planning.

The manual VI re-planning workflow is shown in Figure 72. The whole procedure involved the three roles: VIO, VIP and PIP. Once the VIO issued a request for VI re-planning, the VIP checked the consistency of the request. Once the request was positively validated, the VIP sent a confirmation (VI Re-planning response) back to the VIO. In the following step the VIP initiated the procedure of implementation of the request in the VI to identify a list of virtual resources (VRs) capable of satisfying the VIO request. Once the list of VRs was ready, the VIP issued the request to the PIP to instantiate them and attach to the existing VI.

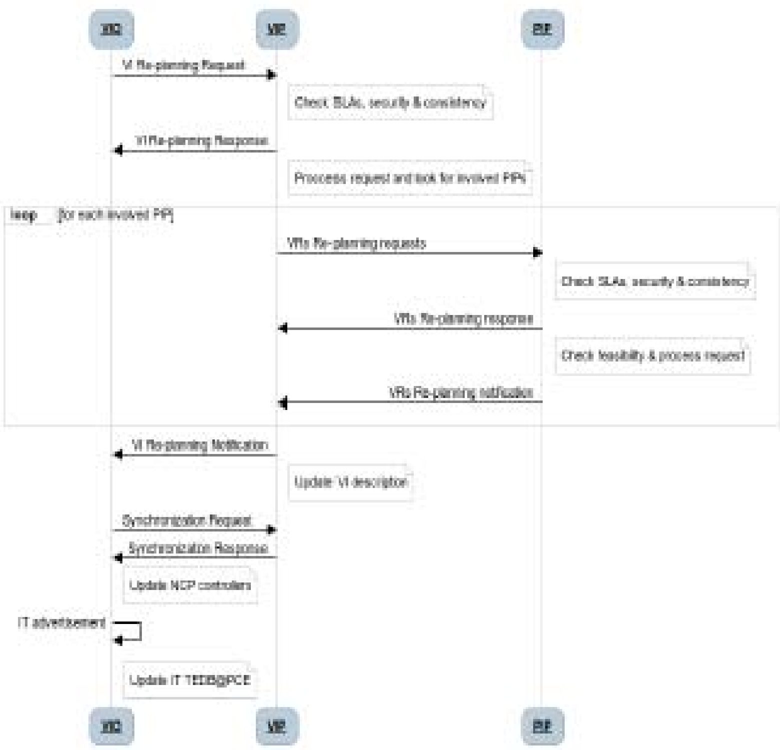


Figure 72. Manual VI re-planning workflow

Once the VI was changed, the notification was sent to the VIO and VIO instantiates (or reconfigures) relevant NCP and VITM controllers (located at the SML), to manage the network and IT resources, respectively. During the initialization of the new NCP controllers, the status of particular virtual resources was exchanged between VIO and VIP, and finally, based on received information, the NCP controllers were synchronized with relevant information.

The manual VI re-planning request was considered a management operation. Thus, the functional module within the VIO responsible for the manual re-planning was located at the SML.

8.5.2.1. VI re-planning requests

Following the initial VI planning request, the VI re-planning requests could vary with regards to the amount of specific information that was provided as part of the request itself. In this context, VI planning requests could be generated by the VIO and communicated to the PIP through the VIP. It was important to note that to perform VI planning and hence re-planning the VIP had to rely on the physical resources information provided by the PIP, being the entity that had full knowledge of the physical infrastructure. The types of re-planning requests were:

- Service-driven VI request: They allowed the VIO to request a VI with the maximum degree of flexibility. The information provided as part of the request involved the prediction of the volume and type of infrastructure expected to support the required services, as well as other service related information as determined by the associated SLAs, including availability etc.
- Constrained VI request: They generally followed the "service driven VI request" described above, but could also impose some additional constraints associated e.g. with the location of some or all involved IT or network resources in the space of an area/country/continent or possible energy consumption requirements.
- Specific VI request: They included specific information regarding the IT resource requirements and processing capability in addition to the usual service specific requirements including availability etc. Taking into consideration the above, the "specific VI request" therefore resulted in a request for a specific virtual topology that was already capacitated with regards to the IT and network resources required.

In the specific VI requests, the virtual topologies of the VI needed to be indicated, including the virtual nodes and the virtual links. The virtual nodes were partitioned from a single physical node in GEYSERS, and the geo-location could also be specified to get the location of the physical node to be mapped. Finding the optimum mapping between the VRs and the PRs was part of the VI planning that applied optimisation with specific objectives.

8.5.3. Re-planning and service procedure performance

The performance of the re-planning procedures [164] was evaluated in terms of the time required to perform the different steps of the workflow. Assuming that performance on a re-planning analysis would validate the function of the architecture, since it is one of the most complex cases. As shown in Figure 73a, the most time consuming function was associated with the VI modification at the LICL that took around 3.505s, also including the NCP+/LICL communication time. The initial computation of the VI upgrading actions took around 8.4ms, while the time needed at the NCP+ to receive and distribute the new infrastructure capabilities was less than 12s. In order to evaluate the impact of the initial re-planning computation on the entire procedure, the re-planning algorithm from a functional perspective on a real prototype was verified. It was tested over different emulated topologies with size ranging between 3 and 10 nodes, and nodal degrees varying between 2 and 3.3. Figure 73b shows that the computational cost of the re-planning algorithm scales linearly with the number of nodes.

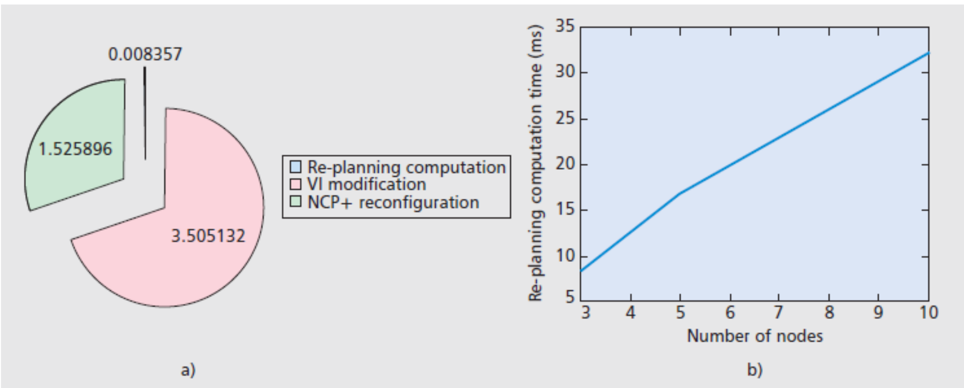


Figure 73. a) Duration of re-planning steps in seconds; b) re-planning computation time for different topology sizes

This architecture was properly implemented into the GEYSERS testbed (Figure 75). A testbed distributed across different countries and composed of different network links (some optical some based on VLAN tagging) with different IT compositions on the border nodes. The deployment of the presented architecture, with all the elements prototyped in software thanks to the work done of dozens of engineers and researchers within the GEYSERS project, demonstrated that this architecture could become a key driver in the telecom sector, since for the first time there was an architecture capable to provide dynamic on-demand provisioning of optical and IT resources in one single-step. Figure

74 exposes the time response impact that this architecture and its implementation Geysers modules) had with respect to current standards of the time being. Moving the provisioning time from the order to months to minutes.

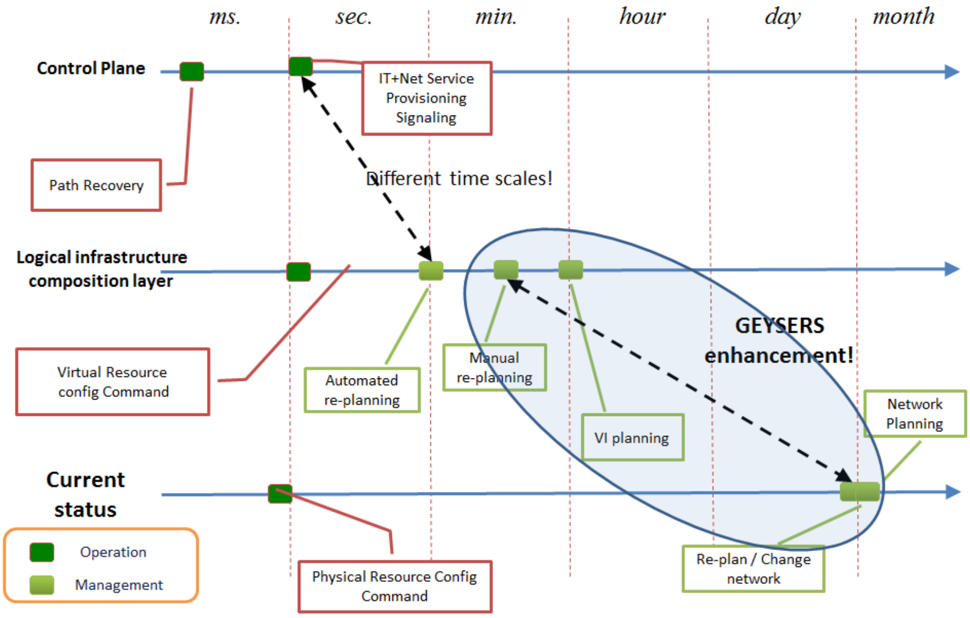


Figure 74. Time response impact of the GEYYERS architecture

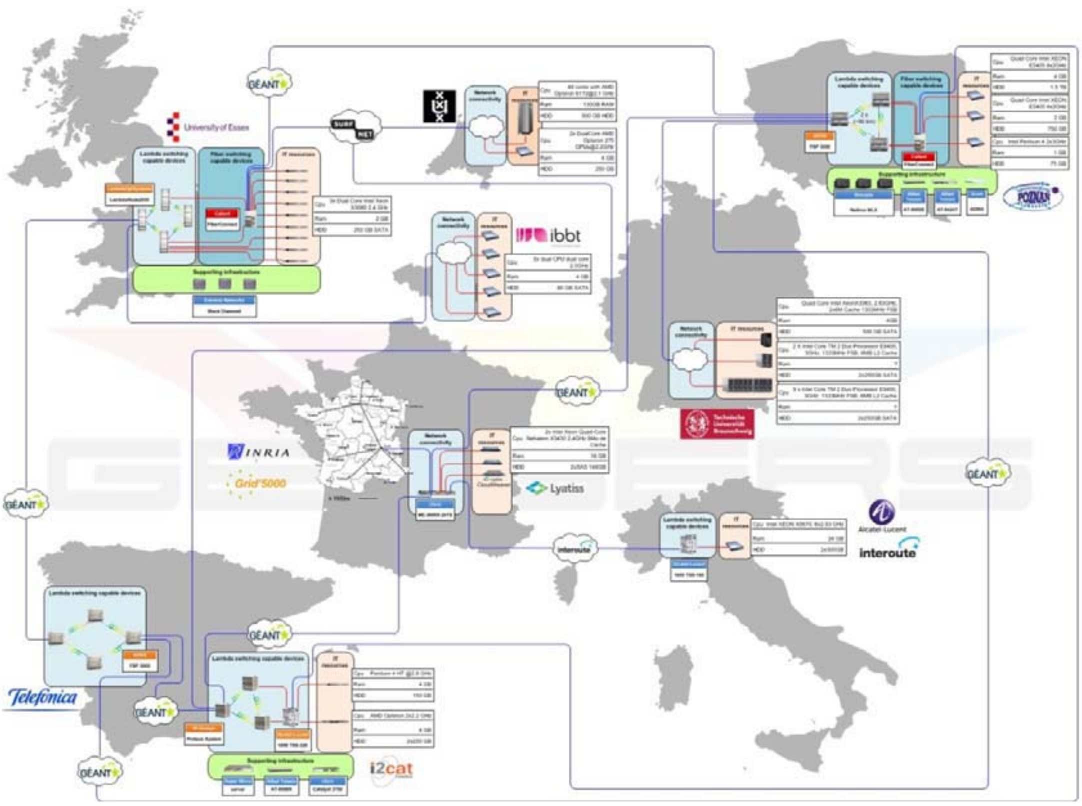


Figure 75. Large scale GEYSERS test-bed

8.6. Energy-efficient Network + IT provisioning

The new generation of carrier-class network equipment available on the market just provided CPUs and electronics that stabilized the required energy budget. They used a reduced and fixed amount of energy, e.g. regardless whether they are forwarding packets or not, or they are using internal memories for forwarding table lookups. Moreover, the continuous miniaturization in optical and electronic components allowed packing equipment in less cumbersome racks, thus reducing the structural costs (e.g. server rooms) and the resulting OPEX, both direct (e.g. for equipment power on) and indirect (e.g. for cooling systems).

An energy efficient networking approach was emerging in which the energy optimization objective was sought also during the dynamic operation of the (mostly IT) infrastructure. A few research projects around the world already started designing architectures and procedures for the minimization of energy consumptions. For example, the NSF GreenLight [165] projects used the automation of the selection of compute and RAM power based on the measure of relevant parameters (e.g. power consumption, temperatures, etc.).

Analogous initiatives in the pure network infrastructure were still missing. Some attempts to "energy efficiency" could be found in the network planning phase, during which the network topology was created in a way that could allow to minimize the number of interfaces/equipments, and to ensure connection redundancy at the same time. Other researches focused on the energy efficiency aspects in Internet routing or Green IT, e.g. by analysing mechanisms for shutdown/standby of the inactive routers' cards, or identifying and moving data process across data centres fed with renewal energies.

The architecture presented on this Thesis leveraged on its unique positioning of a network-centric end-to-end control of both IT and network resources to add a third orthogonal control dimension: the energy efficiency. Thus, the information on power consumption was modelled for the different resources (IT servers, network equipments, links), measured and published at the different layers (i.e. in the LICL and the NCP routing plane) for their consideration as a further constraint in service computation (as long as the available bandwidth on a link or the number of CPUs in a cluster server).

So, the end-to-end seamless Network + IT transport service became also "end-to-end energy efficient" by minimizing the energy consumption metrics along its end-to-end path. Energy efficiency was not only a matter of optimized network routing, but also (and above all) the selection of the more energy efficient IT endpoint.

In fact, most of the energy consumption occurred at the IT level rather than at the network level, due to the use in the servers of more general-purpose CPUs instead of customized processors. Moreover, the continuous migration towards the full optical networks contributed to the minimization of the energy budget in the network segment, because of its reduction of the per-bit power consumption when compared to IP based networks.

However, this new architecture significantly contributed to the energy efficiency dimension with the augmented information about "green" (i.e. energy related) metrics circulating throughout the routing plane of each domain (by means of the enhanced OSPF) and being used by the PCE+ to add energy-aware constraints to the path computation procedures. Nevertheless, the dynamic procedures for re-planning the network operator Infrastructure between the NCP and the LICL also contributed to the Green objective, by optimizing the release/acquisition of IT and network resources based also on energy consumption performance markers monitored along with end-users' SLA fulfilment.

Figure 76 provides the energy efficiency/saving approach considered when designing the architecture.

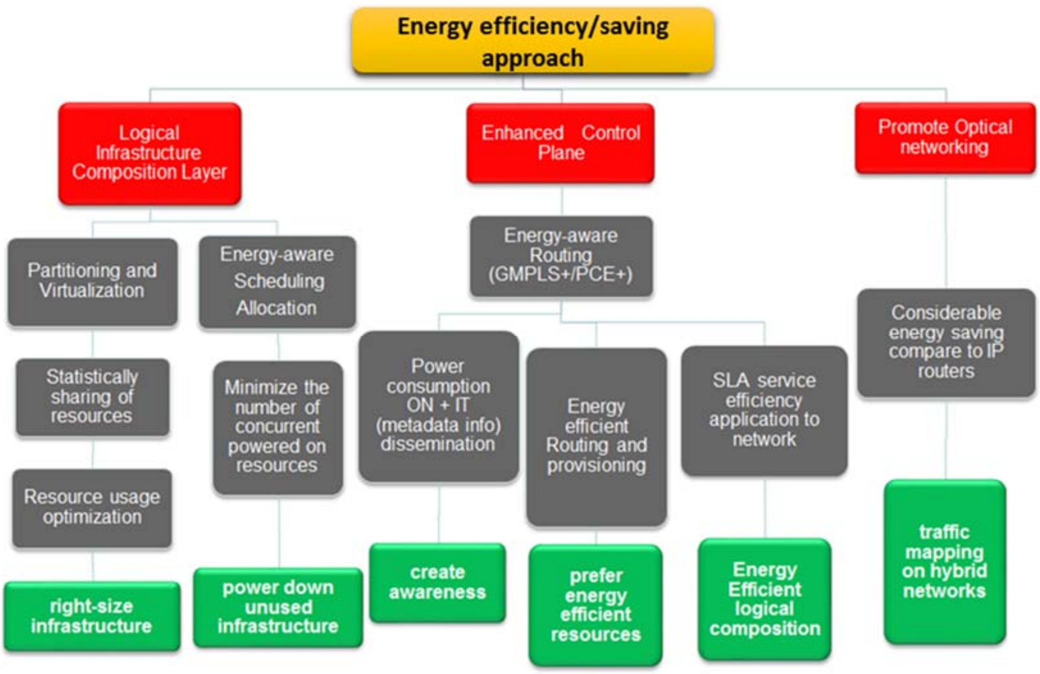


Figure 76. Energy efficiency strategy facilitate by the architecture design

8.7. Conclusions

This architecture was the first on its class. Cloud computing in essence has emerged thanks to the increased availability of network connectivity and bandwidth. However, despite the crucial role that networks play in making cloud services possible, network resource provisioning to date are not yet fully, operationally talking, an integral part of the cloud service provisioning process. Thus, this architecture provided the first architecture and implementation of a system that ensured the provisioning of network resources in order to meet the specific characteristics of cloud-bases applications. It was the first holistic architecture, handling both IT and network resources in a converged manner, while exploiting virtualisation of both of them in order to maximize their efficient utilisation in an infrastructure as a service model.

This architecture has brought into the scene many new services and functionalities, and a first approach to what later on become the SDN and NFV paradigm. The presented architecture already decoupled the data plane from the control plane by means of a novel Logical Infrastructure Composition Layer, while providing full convergence (net + IT) under advanced control planes (NCP+) with novel TE features (i.e. considerations of energy efficiency for the resource provisioning among many others, like different types of connectivity services – uniycast and anycast-). Moreover, its vision fully integrated the service provisioning with the cloud applications, while providing elasticity at the network level, which means upgrading or downgrading the network and IT requirements depending on the 'real time' application needs. This was a major contribution to the research and telecom community, since it was the first generalized architecture capable to map down to the infrastructure the cloud application SLA's.

Actually, all these functionalities, and the fact of decoupling the infrastructure from its control, allowed the creation of new business roles and so a new potential market for telecom operators and new players, while providing cloud applications infrastructure-aware and applications gateways to network. So, facilitating an operator-driven private cloud model. In this architecture, infrastructure services were of paramount importance not only for the IT resources but also for the network resources required to interconnect them. Infrastructure services allowed the possibility of leasing physical resources, releasing the burden of having to purchase physical infrastructure for application providers.

So far, provisioning services over hybrid infrastructures (managed networks and IT), composed of both IT resources (i.e. compute and storage) and high capacity, networks had been considered. Thus, there was a strong need of a unified management and provisioning procedures for infrastructure services across the whole set of resources involved in a cloud computing scenario. This means the usage of core and metro cognitive, flexible, elastic and adaptive technologies for optical networks, with dynamic control plane functionalities, and software defined networks (SDN) for the whole integration with the datacentre network and IT infrastructure services. SDN gives owners and operators of networks better control over their networks, allowing them to optimize network behaviour to best serve their and their users' needs. However, current disjoint evolution has ended up with totally decoupled solutions for each type of resource and infrastructures, those under the network operator domain and those under the datacenter administrator domain. Therefore, there is a key technical challenge towards this ICT convergence and hence, be able to optimize the (i) infrastructure sharing for lowering OpEx/CaPex costs, and (ii) the (dynamic) services and applications deployed on top of these hybrid infrastructures with energy efficiency considerations. In this context, convergence also considers the trend toward infrastructure resource virtualisation and federation, thus providing full flexibility at the infrastructure level.

Therefore, IT & network resource management and control convergence is required as a must for future-proof, and internet-scale enterprise applications. Distributed applications, consuming resources spread all over the world, require datacentres and network core/metro convergence in order to optimize the service workflow and overall performance for cloud computing applications. Dynamic provisioning of one type of infrastructure resources only considers part of the problem, and typically leads to a waste of resources due to over-provisioning, mostly in networks, and sharing limitations in all kinds of resource usage. It must be noted that as time goes, hardware is increasing its power (switching, computing, storage, etc.) and embedding degree, which means that a higher control in granularity is needed too, both at the network and IT level. In the end, the challenge is on providing a common and transparent infrastructure able to integrate different technologies and services, where virtualisation is not the end solution but an adequate technique for overcoming many limitations.

Some future research considerations are:

- Keep the IT/Telco converged Infrastructure provisioning service (IaaS) time at a minimum.
- Unified and converged resources description languages and frameworks.
- Multi-granular, cognitive, elastic and flexible adaptive optical networks (e.g. HW configuration).
- Isolation and flexibility of circuit oriented networks (virtualisation).
- Definition of the impact of these new technologies on legacy business models.
- Inter-administrative domain issues between networks and datacentres.
- Non-standard service provisioning (Alien wavelength services).
- Carrier grade cloud and datacentre integrated infrastructure services.

To sum up, the work done on this architecture demonstrated the feasibility of supporting dynamic infrastructure e2e services and unified network and IT resource provisioning based on the virtualisation of infrastructure resources. The dynamicity and flexibility of the proposed architecture allowed the provision of specific virtual infrastructures on demand while promoting the emergence of new business roles such as the virtual infrastructure provider. This architecture also enabled the runtime modification of the virtual infrastructure through an on demand re-planning service and so it opened up a new type of vision for the forthcoming research.

9. Conclusions

Since partial technical conclusions have already been presented in each cone of the chapters. This section will consider the personal conclusions of the author, and so a personal opinion of the work done, which have evolved under a strategic approach from a Network Resources Provisioning System, a Networks Service Plane to end up with the definition of a whole generalized architecture. Thus, starting from the basic and indivisible piece, to a holistic approach.

I personally think that the evolution of this Thesis has followed the evolution of technology and somehow impacted the path (and the mind) to new approaches, paradigms and architectures that are currently being supported by new open source frameworks. Open source frameworks that aim at full convergence between the IT and the Telecom sector. And so, new paradigms like SDN or NFV are coming into the scene, which aim at dealing with external controllers to manage generic purpose hardware with policy based mechanism through Network Function Virtualisation. I consider this as an evolution of the work presented and developed along the Thesis. This Thesis already presented solutions to decouple the data from the control, and allowed third parties to define the SLA policies to be applied at the infrastructure level. Thus, providing and infrastructure which was virtualised and so offered in an abstracted way, as a generic purpose HW, to be re-configured and controlled by means of software.

I personally feel proud of the contribution brought to the community, since I consider myself as an engineer and researcher that has put its grain of sand for the evolution of optical networks and the future of Internet.

10. Acknowledgements:

The research developed on this Thesis has been partially financed by the following research projects:

- CANARIE's directed research program on **UCLP** (User Controlled LigthPaths)
- **FP6 IST-034115 PHOSPHORUS** (lambda User Controlled Infrastructures for European Research)
- **FP7 ICT-248657 GEYSERS** (Generalised Architecture for Dynamic Infrastructure Services)

11. Other achievements:

- Since 2012: Co-chair of the CaON EC FP7 cluster (Converged and Optical Networks cluster)
- OFC 2014 and 2015 TPC member (N2 Dynamic software controlled and multilayer networks).
- Acknowledgement letter from Vice-Rectorate of Research and Technology Transfer of the UPC, for the effort in the signature of a MoU with the CRC (2004)
- Acknowledgement letter from the head of the PhD program from the TSC department of the UPC for the organization of a PhD course and 2 seminars (2004)
- Awarded with a 4 years PhD's fellowship (FPU) from the Ministry of Education & Science of Spain to develop a PhD in the Optical Communications Group of the UPC (2002)
- Project Coordinator of the Manthychore EC FP7 and IRATI EC FP7 research projects.
- Technical/Research Coordinator of the GEYSERS EC FP7 research project
- Co-leader of a JRA (Joint Research Activity) in the FEDERICA FP7 research project
- WorkPackage Leader of the Phosphorus EC FP6 Research project
- ONDM 2012: Workshop organisation (Network and IT convergence for Future Internet)
- Participation in several consultation meeting from the EC FP7
- Participation in more than 30 conferences

12. Publications

12.1. IEEE CommMag, Journals & Book Chapters

1. J. Ferrer, A. Tzanakaki, A. Antonescu, M. Anastasopoulos, J. A. García-Espín, E. Escalona, S. Peng, G. Landi, G. Bernini, B. Belter, D. Parniewicz, X. Hesselbach, **S. Figuerola**, D. Simeonidou, ”Virtual Infrastructures as a Service enabling Converged Optical Networks and Data Centres” Optical Switching and Networking 08/2014; Volume 14(Part 3): Pages 197–208. DOI:10.1016/j.osn.2014.05.017 · **1.07 Impact Factor**
2. A. Tzanakaki, M. Anastasopoulus, K. Georgakilas, G. Landi, G. Bernini, N. Ciulli, J. Ferrer, E. Escalona, J. A. Garcia, X. Hesselbach, **S. Figuerola**, S. Peng, R. Nejabati, D. Simeonidou, D. Parniewizc, B. Belter, J. Rodriguez; “Planning if dynamic virtual optical cloud infrastructures: The Geysers approach” IEEE ComMagazine, January 2014, Vol. 52, No. 1 DOI:10.1109/MCOM.2014.6710061 **4.01 Impact Factor**
3. **S. Figuerola**, Joan A. García, Jordi Ferrer, V. Reijs, E. Kenny, M. Lemay, M. Savoie, S. Campbell, M. Ruffini, E. Grasa, A. Willner “Optical Networks” Book Chapter, Cross-Layer Design in Optical Networks, 01/2012; Springer.
4. J. Mambretti, M. Lemay, S. Campbell, H. Guy, T. Tam, E. Bernier, B. Ho, M. Savoie, C de Laat, R. van der Pol, J. Chen, F. Yeh, **S. Figuerola**, P. Minoves, D. Simeonidou, E. Escalona, N. Amaya Gonzalez, A. Jukan, W. Bziuk, D. Kim, K. Jong Cho, H. Lee, T. Liu “High Performance Digital Media Network (HPDMnet): An advanced international research initiative and global experimental testbed” Journal of Future Generation Computer Systems (FGCS) 27 (2011) 893–905 **2.79 Impact Factor**
5. Szegedi, P., **Figuerola, S.,** Campanella, M., Maglaris, V., Cervello-Pastor, C. "With evolution for revolution: managing FEDERICA for future Internet research," CommMagazine, IEEE , vol.47, no.7, pp.34-39, July 2009 doi: 10.1109/MCOM.2009.5183470 **4.01 Impact Factor**
6. **Figuerola, S**.; Lemay, M.; "Infrastructure Services for Optical Networks [Invited]," Journal of Optical Communications and Networking, IEEE/OSA, vol.1, no.2, pp.A247-A257, July 2009 doi: 10.1364/JOCN.1.00A247 **2.06 Impact Factor**
7. **Figuerola, S**.; Lemay, M.; Reijs, V.; Savoie, M.; St. Arnaud, B.; "Converged Optical Network Infrastructures in Support of Future Internet and Grid Services Using IaaS to Reduce GHG Emissions," Journal Lightwave Technology, vol.27, no.12, pp.1941-1946, June15, 2009 doi: 10.1109/JLT.2009.2022485 **2.97 Impact Factor**
8. Grasa, E.; **Figuerola, S**.; Forns, A.; Junyent, G.; Mambretti, J.; "Extending the ARGIA software with a Dynamic Optical Multicast Service to support high performance digital media," Optical Network Design and Modeling, 2008. ONDM 2008. International Conference on, vol., no., pp.1-6, 12-14 March 2008 doi: 10.1109/ONDM.2008.4578397 & (special issue) Journal Optical Switching and Networking Volume 6, Issue 2, April 2009, Pages 120-128. doi: 10.1016/j.osn.2009.03.002 **1.07 Impact Factor**

9. Grasa, E.; Junyent, G.; **Figuerola, S.**; Lopez, A.; Savoie, M.; "UCLPv2: a network virtualization framework built on web services [web services in telecommunications, part II]," CommMagazine, IEEE , vol.46, no.3, pp.126-134, March 2008 doi: 10.1109/MCOM.2008.4463783 **4.01 Impact Factor**
10. Grasa, E.; **Figuerola, S.**; Recio, J.; López, A.; Palol, M.; Ribes, Ll.; Díaz, V.; Sangüesa, R.; Junyent G.; Savoie, M. "Video transcoding in a Grid network with User Controlled LightPaths" Future Generation Computer Systems Vol. 22, Issue 8, October 2006, Pages 920-928 doi:10.1016/j.future.2006.03.00 **2.79 Impact Factor**
11. H. Zhang, M. Savoie, S. Campbell, **S. Figuerola**, G. Bochmann, B. St. Arnaud "Workflow-based Virtual Private Networks for Grid Applications" Journal on Future Generation Computer Systems, summer 2006
12. J. Wu, M. Savoie, S. Campbell, H. Zhang, **S. Figuerola**. "Management of Optical Virtual Private Networks". APOC07, Proceedings of SPIE - The International Society for Optical Engineering 12/2007; DOI:10.1117/12.741831 Wuhan, China, November 2007 **0.20 Impact Factor**
13. Vicart.P., **Figuerola, S.,** Chen, X., Liadi, G., Escalona, E., Develder, C., Tzanakaki, A., Demchenko, Y., Garcia-Espin, J.A., Ferrer, J., Lopez, E., Soudan, S., Buysse, J., Jukan, A., Ciulli, N., Brogle, M., Laarhoven, L., Belter, B., Anhalt, F., Nejabati, R., Simeonidou, D., Ngo, C., Laat, Cees de., Biancani, M., Roth, M., Donadio, P., Jimenez, J., Antoniak, M., Gumaste, A., "Bringing Optical Networks to the Cloud: An Architecture for a Sustainable Future Internet". Book Chapter Springer-Verlag Berlin, Heidelberg ©2011 Pages 307-320 ISBN: 978-3-642-20897-3
14. J. Prat, P. Balaguer, J.M. Gené, O.M. Díaz, **S. Figuerola**. "Fiber to the Home Technologies" ISBN 1-4020-7136-1. Kluwer Academic Publisher (Holland), September 2002

12.2. Articles & Conference papers

15. José Ignacio Aznar, M. Jara · A. Roselló, D. Wilson, **S. Figuerola** "OpenNaaS Based Management Solution for Inter-data Centers Connectivity" Proceedings of the 2013 IEEE International Conference on Cloud Computing Technology and Science - Volume 02; 12/2013
16. **S. Figuerola,** C. Bock, J. Ferrer, E. Escalona, J. A. Garcia-Espin "SODALES: An Integrated Wired-Wireless Open Access Architecture" Asia Communications and Photonics Conference; 11/2013
17. J. Ferrer, E. Escalona, J. Batalle, J. A. Garcia-Espin · **S. Figuerola** "Management of Virtual Infrastructures through OpenNaaS" 2013 International Conference on Smart Communications in Network Technologies (SaCoNeT); 06/2013
18. C. Bock, **S. Figuerola**, M.C. Parker, S.D. Walker, T. Mendes, V. Marques, V. Jungnickelm, K. Habel, D. Levi "Convergent radio and fibre architectures for high-speed access" Transparent Optical Networks (ICTON), 2013 15th International Conference on; 01/2013

19. Yuri Demchenko, C. Ngo, Cees de Laat, Juan Rodriguez, L.M. Contreras, Joan Antoni Garcia-Espin, **Sergi Figuerola**, Giada Landi, Nicola Ciulli "Intercloud Architecture Framework for Heterogeneous Cloud Based Infrastructure Services Provisioning On-Demand" Advanced Information Networking and Applications Workshops (WAINA), 2013 27th International Conference on; 01/2013
20. A. Tzanakaki, M.P. Anastasopoulos, K. Georgakilas, J.A. Garcia-Espin, J. Ferrer Riera, **S. Figuerola**, M. Ghijsen, Y. Demchenko, C.T.A.M. de Laat, P. Vicat-Blanc, S. Soudan, F. Anhalt, S. Peng, E. Escalona, R. Nejabit, D. Simeonidou "Virtual Infrastructure Planning: The GEYSERS Approach" Future Network & Mobile Summit 2012; 07/2012
21. J.F. Riera, J.A. Garcia-Espin, **S. Figuerola,** J. Buysse, M. De Leenheer, C. Develder, Shuping Peng "Converged IT and optical network virtualisation: The last (clean-)step towards the Future Internet infrastructure management" Networks and Optical Communications (NOC), 2012 17th European Conference on; 01/2012
22. A. Pages, J. Perelló, S. Spadaro, J.A. Garcia-Espin, J.F. Riera, **S. Figuerola** "Optimal allocation of virtual optical networks for the future internet" Optical Network Design and Modeling (ONDM), 2012 16th International Conference on; 01/2012
23. J.A. Garcia-Espin, J.F. Riera, **S. Figuerola,** E. Lopez "A multi-tenancy model based on resource capabilities and Ownership for infrastructure management" Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on; 01/2012
24. **S. Figuerola,** D. Simeonidou, J.F. Palacios, A. Di Giglio, N. Ciulli, J.A. Garcia, R. Nejabati, X. Masip, R. Muñoz, G. Landi, M. Yannuzzi, R. Casellas "Research trends on ICT convergence from the CaON cluster" Transparent Optical Networks (ICTON), 2012 14th International Conference on; 01/2012
25. C. Bock, **S. Figuerola,** M.C. Parker, T. Quinlan, S.D. Walker "Convergent radio and fibre access architectures using low-energy systems" Transparent Optical Networks (ICTON), 2012 14th International Conference on; 01/2012
26. J. A. Garcia-Espin, J. Ferrer, **S. Figuerola**, M. Ghijsen, Y. Demchenko, J. Buysse, M. de Leenheer, C. Develder, F. Anhalt. "Logical Infrastructure Composition Layer, the GEYSERS Holistic Approach for Infrastructure Virtualisation" TNC2012 conference
27. **S. Figuerola**, J. A. García-Espín, J. Ferrer, E. López, E. Escalona, R. Nejabati, S. Peng, D. Simeonidou "Coordinating IT and Network Virtualisation to provide Infrastructure as a Service: a GEYSERS approach for the Future Internet". Invited MONAMI 2011, Aveiro Portugal. IEEE conference
28. E. Escalona, S. Peng, R. Nejabati, D. Simeonidou, J. A. Garcia-Espin, J. Ferrer, **S. Figuerola**, G. Landi, N. Ciulli, J. Jimenez, B. Belter, Y. Demchenko, C. de Laat, X. Chen, A. Yukan, S. Soudan, P. Vicat-Blanc, J. Buysse, M. de Leenheer, C. Develder, A. Tzanakaki, P. Robinson, M. Brogle,T. M. Bohnert "GEYSERS: A Novel Architecture for Virtualization and Co-Provisioning of Dynamic Optical Networks and IT Services" FNMS 2011. IEEE sponsored conference
29. Bohnert, T.M., **Figuerola, S**., Ciulli, N., Simeonidou, D., Vicat-Blanc, P. "Optical networking for Cloud Computing," Optical Fiber Communication Conference and Exposition

(OFC/NFOEC), 2011 and the National Fiber Optic Engineers Conference , vol., no., pp.1-2, 6-10 March 2011

30. **S. Figuerola**, D. Simeonidou, N. Ciulli, R. Nejabati, T.M. Bohnert, J.A. Garcia-Espin, P. Vicat-Blanc, B. Belter, M. Biancani, J. Jimenez “GEYSERS architecture”. Future Network & Mobile Summit 2010, IEEE conference
31. **Figuerola, S.,** Nejabati, R., Escalona, E., Simeonidou, D., Garcia-Espin, J.A., Landi, G., Ciulli, N., Brogle, M., Belter, B., Primet, P.V., Jimenez, J. "Generalized architecture for dynamic infrastructure services," Optical Communication (ECOC), 2010 36th European Conference, vol., no., pp.1-20, 19-23 Sept. 2010 doi: 10.1109/ECOC.2010.5621242
32. **Figuerola, S.**, García-Espín, J.A., Ferrer Riera, J., Ciulli, N. "An optical network and IT infrastructure virtualisation and provisioning framework," Optical Internet (COIN), 2010 9th International Conference on, vol., no., pp.1-3, 11-14 July 2010 doi: 10.1109/COIN.2010.5546328
33. C. Bock, **S. Figuerola**, V. Jiménez “The LambdaCat Open Testbed Facility: An Innovation Facility to Test Multi-layer Devices on a Real Environment” Testbeds and Research Infrastructures. Development of Networks and Communities - 6th International ICST Conference, TridentCom 2010, Berlin, Germany, May 18-20, 2010, Revised Selected Papers; 01/2010
34. Willner, A., Riera, J.F., Garcia-Espin, J., **Figuerola, S.,** De Leenheer, M. Develder, C. "An analytical model of the service provisioning time within the harmony network service plane," GLOBECOM Workshops (GC Wkshps), 2010 IEEE , vol., no., pp.514-518, 6-10 Dec. 2010 doi: 10.1109/GLOCOMW.2010.5700373
35. X. Hesselbach, J. A. García-Espin, M. González, J. Gonzalo, **S. Figuerola** “E3MS: A traffic engineering prototype for autoprovisioning services in IP/DiffServ/MPLS networks”; IX Jornadas de Ingeniería Telemática JITEL 2010, Valladolid
36. **Figuerola, S.**; Garcia-Espin, J.A.; Ferrer, J.; Willner, A.; "Performance analysis of harmony: An optical, multi-domain network resource broker," Transparent Optical Networks, 2009. ICTON '09. vol., no., pp.1-5, June 28 2009-July 2 2009 doi: 10.1109/ICTON.2009.5185032
37. **Figuerola, S**.; Garcia-Espin, J.A.; Ferrer, J.; Zhang, H.; Savoie, M.;”Enabling Network Resources in the Grid: GridNets 2009 Third International ICST Conference on Networks for Grid Applications. Network interfaces for Grid applications, functionalities and services (invited), Athens; 01/2009
38. A. Willner, C. Barz, J. A. Garcia Espin, J. Ferrer Riera, **S. Figuerola**, P. Martini “Harmony: Advance Reservations in Heterogeneous Multi-domain Environments” Proceedings of the 8th IFIP Networking conference, Springer's LNCS; 01/2009
39. **Figuerola, S.,** Garcia-Espin, J.A., Ferrer, J., Willner, A. “Scalability analysis and evaluation of the multi-domain, Optical Network Service Plane in Harmony” Optical Communication, 2009. ECOC 2009. 35th European Conference. IEEE conference ISBN: 978-1-4244-5096-1
40. A. Tzanakaki, M. Anastasopoulos, K. Georgakilas, J. Buysse, M. De Leenheer, C. Develder, S. Peng, R. Nejabati, E. Escalona, D. Simeonidou, N. Ciulli, G. Landi, M. Brogle, A. Manfredi, E.

Lopez, J. Ferrer, J. A. Garcia, **S. Figuerola**, P. Donadio, G. Parladori, J. Jimenez, “Energy efficiency considerations in integrated IT and optical network resilient infrastructures”, International Conference on Optical Transport Networks. ICTON 2011 Conference, vol., no., pp.944-951, 9-13 July 2007 doi: 10.1109/ICWS.2007.160

41. **Figuerola, S**.; Garcia, J.A.; Sanchez, A.; de Waal, C.; Willner, A.; "The network service plane: An approach for interdomain network reservations," Transparent Optical Networks, 2008. ICTON 2008. 10th Anniversary International Conference on , vol.1, no., pp.13-15, 22-26 June 2008 doi: 10.1109/ICTON.2008.4598358
42. E. Grasa, **S. Figuerola**, A. Forns, G. Junyent, J. Mambretti “Extending the UCLP Software with a Dynamic Optical Multicast Service to support High Performance Digital Media” Optical Network Design and Modeling, 2008. ONDM 2008. International Conference on; 04/2008
43. **S. Figuerola**, N. Ciulli, M. De Leenheer, Y. Demchenko, W. Ziegler, A. Binczewski , “PHOSPHORUS: Single-step on demand services across multi-domain networks for e-science”. APOC07, Wuhan, China, November 2007
44. E. Grasa, **S. Figuerola**, A. López, G, Junyent, M. Savoie, B. St. Arnaud, M. Lemay “Articulated private networks in UCLP” Internet Research 10/2007; 17(5):535-545. DOI:10.1108/10662240710830226 · **1.66 Impact Factor**
45. J. Trapero, P. Novell, **S. Figuerola**, F. Raspall, C. Macian “i2CAT experiences in advanced SIP-based communications” Campus-Wide Information Systems 09/2007; 24(4):280-291. DOI:10.1108/10650740710834662 TERENA Networking Conference, Lingby (Denmark) 2007, May
46. Hanxi Zhang; Savoie, M.; Campbell, S.; **Figuerola, S.;** von Bochmann, G.; Arnaud, B.S.; , "Service-Oriented Virtual Private Networks for Grid Applications," Web Services, 2007. ICWS 2007. IEEE International
47. Sanchez, **S. Figuerola**, G. Junyent, V. Reijs, E. Kenny, M. Ruffini. “A user provisioning tool for EoMPLS services based on UCLPv1” TERENA Networking Conference, Lingby (Denmark) 2007, May
48. Recio, J.; Grasa, E.; **Figuerola, S**.; Junyent, G.; , "Evolution of the user controlled lightpath provisioning system," Transparent Optical Networks, 2005, Proceedings of 2005 7th International Conference , vol.1, no., pp. 263- 266 Vol. 1, 3-7 July 2005 doi: 10.1109/ICTON.2005.1505800
49. **S. Figuerola**, E. Grasa, M. Hurtado, J. Recio, G. Junyent, M. Savoie “Transmission of HDTV over IP on an Optical Connection Managed using UCLP”, , XV Jornadas técnicas RedIris, October 2004
50. K. Meynell, L. Deri, **S. Figuerola**, L. Florio, A. Gall, L. Gommans, P. Grosso, G. Karmous-Edwards, S. Leinen, A. Liakopoulos, D. López, C. Meiroşu, M. Sova, S. Venas, K. Wierenga “Report on Technical Issues: Easnet” ISBN:978-77559-14-7 Terena Secretariat 2008
51. E. Escalona, J. Perelló, S. Spadaro, J. Comellas, G. Junyent, **S. Figuerola**. “Experimental Optical Grid Interface for ASON/GMPLS Service Oriented Networks”, e-Photon/ONe+ Summer School 2006 on Optical Grid and Optical Network Resilience, Zagreb, Sept. 2006

52. E. Grasa, X. Hesselbach, **S. Figuerola**, V. Reijs, D. Wilson, Jean-Marc Uzé, L. Fischer, T. de Miguel. “MANTICORE project: Providing users with a Logical IP Network Service” Terena Conference TNC2008, may 2008
53. E. Grasa, **S. Figuerola**, J. Recio, G. Junyent “Optical networks managed by the user”, IGC 2005
54. **S. Figuerola**, E. Grasa, G. Junyent “UCLP- User Controlled Lightpath Provisioning”, IGC-INET 2004, Barcelona Juny 2004
55. Javier González, Pablo Arozarena, Javier G. Algarra, Gal Shabtay, Carolina Pinart, Gabriel Junyent, Jaume Comellas, Eduard Escalona, **Sergi Figuerola**, “PROMISE: provisioning and monitoring of optical services”, Comunicaciones de Telefónica I+D 38 (2006): 205-218

13. Bibliography

[1] A. Singh, J. Ong, A. Agarwal, G. Anderson, A. Armistead, R. Bannon et all, Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network SIGCOMM '15 August 17-21, 2015, London, United Kingdom I: http://dx.doi.org/10.1145/2785956.2787508 ACM ISBN 978-1-4503-3542-3/15/08.

[2] http://opennaas.org/

[3] https://www.openstack.org/

[4] https://www.opendaylight.org/

[5] CaON Converged and Optical Cluster http://caon.i2cat.net/

[6] http://www.ietf.org/proceedings/52/slides/plenary-2/sld006.htm

[7] H. G. Perros, An introduction to ATM networks. Wiley, 2002. 4

[8] H. G. Perros, Connection-oriented networks: SONET/SDH, ATM, MPLS, and optical networks. Wiley, 2005. 4

[9] K. Thompson, G. Miller, and R. Wilder, Wide-area Internet traffic patterns and characteristics (extended version), IEEE Network magazine, vol. 11, no. 6, pp. 10_23, 1997.4

[10] E. Osborne and A. Simha, Traffic engineering with MPLS. Cisco, 2002.

[11] I. Rec, _G. 709: Network Node Interface for the Optical Transport Networks, International Telecommunication Union, ITU-T, 2003. 6, 11

[12] I. Tomkos et all, The Evolution of Optical Networking Vol. 100, No. 5, May 2012 | Proceedings of the IEEE

[13] E. Escalona, Solutions for Building Service Oriented Optical Networks, PhD Thesis, UPC 2010

[14] Architecture and functional requirements of control planes for automatic switched optical networks: experience of the IST project LION Antonio Manzalini*, Katsuhiro Shimano**, Carlo Cavazzoni*, Alessandro D'Alessandro*

[15] G.8080/Y.1304, Architecture for the automatically switched optical network (ASON)

[16] G.807/Y.1302, Requirements for automatic switched transport networks (ASTN)

[17] B. Rajagopalan, D. Pendarakis, D. Saha, Ramu S. Ramamoorthy, Krishna Bala (2000) "IP over optical networks: architectural aspects". IEEE Communications Magazine, pages 94-102

[18] F. Ricciato, S. Salsano, A. Belmonte and Listanti, "Integrated Dynamic IP and Wavelength Routing in IP over WDM Networks", Proceedings of IEEE INFOCOM 2001, April 2001.

[19] "A Management Architecture for IP-over-WDM integration", Cisco Systems White Paper, 2001.

[20] B. Rajagopalan, et al. RFC 3717 IP over Optical Networks: A Framework, March 2004

[21] J. Aracil, F. Callegati. Enabling Optical Internet with Advanced Network Technologies. (ed) Computer communications and Networks. Springer ISSN 1617-7975, 2009

[22] E. Rosen, A. Viswanathan, and R. Callon, IETF RFC 3031: Multiprotocol Label Switching Architecture,tech. rep., 2001. 15

[23] L. Berger, IETF RFC 3471: Generalized multi-protocol label switching (GMPLS) signaling functional description, tech. rep., 2003. 15

[24] E. Mannie, IETF RFC 3945: Generalized Multi-Protocol Label Switching (GMPLS) Architecture, tech. rep., 2004. 15, 69

[25] J. Lang, IETF RFC 4204: Link Management Protocol (LMP), tech. rep., 2005. 16, 21

[26] Y. Rekhter and K. Kompella, IETF RFC 4203: OSPF Extensions in Support of Generalized Multi-Protocol Label Switching (GMPLS), tech. rep., 2005. 15, 21, 51

[27] L. Berger, IETF RFC 3473: Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions, tech. rep., 2003. 15, 18, 32, 45

[28] W. Kellerer (1998) "Dienstarchitekturen in der Telekommunikation . Evolution, Methoden und Vergleich." Technical Report TUM-LKN-TR-9801.

[29] M. Blanchet, F. Parent, and B. St-Arnaud, "Optical BGP (OBGP): IDRA lightpath provisioning," IETF draft, ietf-draft-parent-obgp-01, Mar. 2001.

[30] Shinn-Jyh Jan and Wen-Fong Wang, "Study on Migrating BGP to OBGP"

[31] M. Yannuzzi, X. Masip-Bruin, and O. Bonaventure, "Open issues in interdomain routing: A survey," IEEE Network, Vol. 19, no. 6, pp. 49-56, 2005.

[32] "UNI 1.0 Signaling Specification OIF-UNI-01.0—User Network Interface (UNI) 1.0 Signaling Specification". Optical Internetworking Forum. [Online]. Available: www.oiforum.com

[33] Iannoce Eugenio, Telecomunications Networks, CRC Press ISBN 978-1-4398-4636-0

[34] R. SchollMeier, "A definition of Peer-to-peer networking for the classification of peer-to-peer architectrues and applications", Proceedings of the First International Conference on Peer-to-Peer Computing (P2P.01).

[35] B. Rajagopalan, J. Luciani, D.O. Awduche (2003), "IP over Optical Networks: A Framework," IETF Internet draft.

[36] J. Wu, J. M. Savoie, Bill St. Arnaud (2001) "Peer-to-peer optical networking –architecture, functional requirements and applications"

[37] E. Noam (2003) "The Internet: Still wide open and competitive?" TPRC

[38] A.Vukovic, H. Hua, J. Wu, J. M. Savoie (2002) "Architecture, routing and signalling for peer-to- Other peer inter-domain optical networking", NFOEL.

[39] "FIND, Future Internet Design". [Online]. Available: http://www.nets-find.net/

[40] J. B. Evans, D. E. Ackers: "Overview of GENI Overview of GENI and Future Internet in the US", 22 May 2007. [Online]. Available: http://www.geni.net/

[41] "Future Internet Research and Experimentation: Anoverview of the European FIRE Initiative and its projects", 1 September 2008. [Online]. Available: http://cordis.europa.eu/fp7/ict/fire/

[42] "Federated E-infrastructure Dedicated to European Researchers Innovating in Computing network Architectures". [Online]. Available: http://www.fp7-federica.eu/

[43] L. Peterson: "PlanetLab: A Blueprint for Introducing Disruptive Technology into the Internet", January 2004. [Online]. Available: http://www.planet-lab.org/

[44] T. Erl, Service Oriented Architecture: Concepts, Technology and Design, 792 pages, Hardcover, ISBN: 0131858580, September 2005.

[45] Maier, G. ; Girardi, R. ; Ferraris, Giuseppe "Towards the future european research network: the MUPBED project" Optical Network Design and Modeling, 2005. IEEE 10.1109/ONDM.2005.1426992

[46] Callejo-Rodrigitez, M.A.; Enriquez-Gabeiras, J. ; Burakowski, W. ; Beben, A. et all "EuQoS: End-To-End QoS over Heterogeneous Networks" Innovations in NGN: Future Network and Services, 2008. K-INGN 2008. IEEE 10.1109/KINGN.2008.4542264

[47] I. Foster, H. Kishimoto, A. Savva, D. Berry, A. Djaoui, A. Grimshaw, B. Horn, F. Maciel, F. Siebenlist, R. Subramaniam, et al., The Open Grid Services Architecture (OGSA), version 1.5,OGF GFD.80, 2006. 25, 42

[48] S. Tuecke et al., Open Grid Services Infrastructure (OGSI) Version 1.0, Global Grid Forum. GFD-RP.15. 25

[49] Bill St. Arnaud (2003), Customer controlled and managed optical networks

[50] B. St. Arnaud. "Frequently Asked Questions about Customer Owned Dark Fiber, Condominium Fiber, Community and Municipal Fiber Networks". [Online]. Available: http://whitepapers.techrepublic.com.com/abstract.aspx?docid=12105

[51] Grasa, E.; Junyent, G.; Figuerola, S.; Lopez, A.; Savoie, M.; "UCLPv2: a network virtualization framework built on web services [web services in telecommunications, part II]," Communications Magazine, IEEE, vol.46, no.3, pp.126-134, March 2008 doi: 10.1109/MCOM.2008.4463783

[52] D. Waldron. "Canadian School Board Investments in Private Fiber Optic Networks, a Cost Benefit Analysis". [Online]. Available: http://www.canarie.ca/conferences/advnet2000/presentations/waldron.pdf

[53] DeFanti, T., De Laat, C., Mambretti, J., Neggers, K., St. Arnaud, B. "TransLight: A Global Scale Lambda Grid for E-Science," Special Issue on "Blueprint for the Future of High Performance Networking," Communications of the ACM, Volume. 46, No. 11, November 2003, pp. 34-41.

[54] Grasa E. "Software Architectures for the Management and Control of User Owned and Sliceable Networks" PhD Thesis, UPC 2009

[55] Bill St. Arnaud, et all (2003), Web Services architectures for user control and management of optical internet networks.

[56] Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Larry G. Roberts, Stephen Wolff. "End-to-end principle of internet" http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet

[57] Barnes, C.R. Best, M.M.R. Bornhold, B.D. Juniper, S.K. Pirenne, B. Phibbs, P. , "The NEPTUNE Project - a cabled ocean observatory in the NE Pacific: Overview, challenges and scientific objectives for the installation and operation of Stage I in Canadian waters". Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies, 2007.

[58] H. Zhang, M. Savoie, S. Campbell, S. Figuerola, G. Bochmann, B. St. Arnaud,"Workflow-based Virtual Private Networks for Grid Applications" Journal on Future Generation Computer Systems, summer 2006

[59] M. J. Francisco et all (2003) Inter-domain routing in optical networks

[60] Bill St. Arnaud et all (2001) Optical BGP Networks

[61] J. Wu, Scott Campbell, J Savoie H. Ahang, G. Bocham, Bill St. Arnaud (NFOEC 2003) User-managed end-to-end lightpaths provisioning over CA*net 4

[62] Bill St. Arnaud (2004), Applications empowered networks, OnVector optical Networking Workshop

[63] G. Pavlou, On the Evolution of Management Approaches, Frameworks and Protocols: A Historical Perspective, Springer, Oct. 13, 2007.

[64] J. Schönwalder, A. Pras, and J-P. Martin-Flatin, "On the Future of Internet Management Technologies," IEEE Commun. Mag., vol. 41, no. 10, Oct. 2003, pp. 90–97.

[65] R. Boutaba and J. Xiao, "Network Management: State of the Art," IFIP World Comp. Cong., 2002.

[66] V.Reijs and E. Grasa, "The MANTICORE Project: Providing Users with a Logical IP Network Service," TERENA Net. Conf., May 2008.

[67] Szegedi, P., Figuerola, S., Campanella, M., Maglaris, V., Cervello-Pastor, C. "With evolution for revolution: managing FEDERICA for future Internet research," CommMagazine, IEEE vol.47, no.7, pp.34-39, July 2009 doi:10.1109/MCOM.2009.5183470

[68] https://www.geni.net/

[69] https://www.onelab.eu/

[70] N.M. M. K. Chowdhury, F.-E Zaheer, and R. Boutaba, "iMark: An Identity Management Framework for Network Virtualization Environment," 11th IFIP/IEEE Int'l. Symp. Integrated Net. Management, Long Island, NY, June 2–4, 2009.

[71] M. Kolon, "Intelligent Logical Router Service," Juniper Networks White Paper, 2004; http://www.juniper.net/solutions/literature/white_papers/200097.pdf

[72] T. Takeda, I. Inoue, R. Aubin and M. Carugi. "Layer 1 Virtual Private Networks: Service Concepts, Architecture Requirements, and Related Advances in Standardization". IEEE Communications Magazine. vol. 42, no. 6. June 2004. pp. 132-138.

[73] S. French and D. Pendarakis, "Optical Virtual Private Networks: Applications, Functionality and Implementation," Photonic Network Communications, vol. 7, no. 3, May 2004, pp. 227-238.

[74] G. Nunzi, M. Brunner and I. Kazuhiko. "Customer-Oriented GMPLS Service Management". 2004 Workshop on High Performance Switching and Routing (HPSR 2004). April 19-21, 2004. Phoenix, Arizona, USA. pp. 32-37.

[75] J. Wu, M. Savoie, S. Campbell, H. Zhang and B. St. Arnaud. "Layer 1 Virtual Private Network Management by Users". IEEE Communications Magazine. Vol. 44, No. 12. Dec. 2006. pp. 86-93.

[76] J. Wu, M. Savoie, S. Campbell, H. Zhang, S. Figuerola. "Management of Optical Virtual Private Networks". APOC07, Wuhan, China, November 2007

[77] J. Bih. "Deploy XML-Based Network Management Approach". IEEE Potentials. Vol. 24. No. 4. Oct.-Nov. 2005. pp. 26-31.

[78] F.L. Verdi, R. Duarte, F.C. de Lacerda, E. Cardozo, M. Magalhaes and E. Madeira. "Provisioning and Management of Interdomain Connections in Optical Networks: A Service Oriented Architecture-Based Approach". 10th IEEE/IFIP Network Operations and Management Symposium (NOMS 2006). April 3-7, 2006, Vancouver, Canada.

[79] D. Skvorc, S. Srbljic and M. Podravec. “Virtual Network for Development and Execution of Service-Oriented Applications”. 2006 International Conference on Networking and Services (ICNS 2006). July 19-21, 2006, Silicon Valley, California, USA.

[80] Infinera white paper, “Bandwidth virtualization enables a programmable optical network,” http://www. infinera.com/solutions/whitepapers.html.

[81] N. Morita and H. Imanaka, “Introduction to the functional architecture of NGN,” IEICE Trans. Commun., vol. E90-B, no. 5, pp. 1022–1031, 2007

[82] H. C. Liao and C. Q. Tao, “An anatomy to SaaS business mode based on Internet,” in Int. Conf. on Management of e-Commerce and e-Government, 2008. ICMECG’08, Jiangxi, China, Oct. 17–19, 2008, pp. 215–220.

[83] Grasa, E.; Figuerola, S.; Forns, A.; Junyent, G.; Mambretti, J.; "Extending the ARGIA software with a Dynamic Optical Multicast Service to support high performance digital media," Optical Network Design and Modeling, 2008. ONDM 2008, vol., no., pp.1-6, 12-14 March 2008 doi: 10.1109/ONDM.2008.4578397 & (special issue) Journal Optical Switching and Networking Volume 6, Issue 2, April 2009, Pages 120-128. doi: 10.1016/j.osn.2009.03.002

[84] E. Grasa, S. Figuerola, A. López, G. Junyent, M. Savoie, B. St. Arnaud, M. Lemay, Articulated Private Networks in UCLP, in: TERENA Networking Conference, May 2007 [Online]. Available: http://tnc2007.terena.org/programme/presentations/.

[85] I. Foster and C. Kesselman, “Globus: a toolkit-based grid architecture,” in The Grid: Blueprint for a Future Computing Infrastructure, San Fransisco, CA: Morgan-Kaufmann, 1999, pp. 259–278.

[86] S. Vinoski, “REST Eye for the SOA Guy,” IEEE Internet Comput.vol. 11, no. 1, pp. 82–84, Feb. 2007.

[87] “OSGi—the dynamic module system for Java,” http://www.osgi.org.

[88] E. Grasa, S. Figuerola, A. Forns, G. Junyent, and J. Mambretti,“Extending the UCLP Software with a dynamic optical multicast service to support high performance digital media,” in Int. Conf. on Optical Network Design and Modeling, 2008. ONDM 2008, Vilanova i la Geltru, Spain, March 12–14, 2008, pp. 1–6.

[89] Open Symphony, “Quartz—Quartz overview,” http://www.opensymphony.com/quartz/.

[90] “An Inefficient Truth,” Global Action Plan, London, U.K., Dec. 2007 [Online]. Available: www.globalactionplan.org.uk/upload/resource/ Full-report.pdf

[91] “Smart 2020: Enabling the low carbon economy in the information age,” The Climate Group, on behalf of the Global eSustainability Initiative (GeSI), London, U.K., 2008 www.theclimategroup.org/assets/resources/publications/Smart2020Report_lo_res.pdf

[92] Minnesota Internet Traffic Studies (MINTS). Univ. Minnesota, Minneapolis, MN, 2002–2007 [Online] http://www.dtc.umn.edu/mints/

[93] McKinsey Report on DC Efficiency. McKinsey&Company, Washington, DC, 1996–2009 http://www.mckinsey.com/clientservice/bto/pointofview/pdf/Revolutionizing_Data_Center_Efficiency.pdf

[94] R. S. Tucker, “A Green Internet,” ARC, CUBIN, Univ. Melbourne, Melbourne, Australia: http://www.ee.unimelb.edu.au/people/rst/talks/files/Tucker_Green_Plenary.pdf

[95] Overview of Khazzoom-Brookes: http://en.wikipedia.org/wiki/Khazzoom-Brookes_postulate

[96] J. Romm, “What are the near-term climate Pearl Harbors?” Climate Progress, A Project of the Center for American Progress Action Fund, Washington, DC, Nov. 24, 2008 [Online]. Available: http://climateprogress.org/2008/11/24/what-are-the-near-term-climate-pearl-harbors/

[97] P. U. Clark et al., “Abrupt Climate Change,” U.S. Climate Change Science Program and the Subcommittee on Global Change Research, U.S.Geological Survey, Reston, VA, Final Report, Synthesis and Assessment Product 3.4, Dec. 2008 [Online]. http://www.climatescience.gov/Library/sap/sap3-4/final-report/default.htm

[98] B. Penner, Bill 44–2007: Greenhouse Gas Reduction Targets Act, Victoria, BC, Canada, 2007 Legislative Session: 3rd Session, 38th Parliament, Third Reading, Nov. 26, 2007 [Online]. http://www.leg.bc.ca/38th3rd/3rd_read/gov44-3.htm

[99] “Climate Change: Valuing Emissions,” Dept. Environment Food and Rural Affairs (DEFRA), London, U.K., 2005–2008 [Online] http://www.defra.gov.uk/Environment/climatechange/research/carboncost/index.htm

[100] K. Batten and K. Manlove, “Identifying Hurdles to Renewable Electricity Transmission,” Center for American Progress,Washington, DC, Dec. 18, 2008 [Online] http://www.americanprogress.org/issues/2008/12/pdf/renewable_transmission.pdf

[101] Electric power transmission: http://en.wikipedia.org/wiki/Electric_power_transmission

[102] J. Romm, “Stanford study, Part 1: Wind, solar baseload easily beat nuclear and they all crush clean coal,” Climate Progress, A Project of the Center for American Progress Action Fund,Washington, DC, Dec. 14, 2008 [Online] http://climateprogress.org/2008/12/14/stanford-study-part-1-wind-solar-baseload-easily-beat-nuclear-and-theyall-best-clean-coal

[103] Grid Computing. [Online]. Available: http://en.wikipedia.org/wiki/Grid_computing

[104] Available online: http://ec.europa.eu/research/fp6/index_en.cfm

[105] Barz, C., Bornhauser, U., Martini, P., Pilz, M., de Waal, C., and Willner, A.: ARGON: Reservation in Grid-enabled Networks. In Proceedings of the 1st DFN-Forum on Communications Technologies, 2008.

[106] Campanella, M., Krzywania, R., Sevasti, A., and Thomas, S.-M.: Generic Domain-centric Bandwidth on Demand Service Manager: Deliverable GN2-08-129, GÉANT, August 2008. [Online]. Available: http://www.geant2.net/upload/pdf/GN2-08-129-DS3-3-4_Functional_Specification_and_Design_of_Generic_Domain-centric_BoD_Service_Manager.pdf.

[107] GÉANT2: DICE (dante-internet2-canarie-esnet) collaboration. [Online]. Available: http://www.geant2.net/server/show/conWebDoc.1308

[108] Takefusa, A., Hayahsi, M., Nagatsu, N., Nakada, H., Kudoh, T., Miyamoto, T., Otani, T., Tanaka, H., Suzuki, M., Sameshima, Y., et al.: G-Lambda: Coordination of a Grid scheduler and lambda path services over GMPLS. Future Generation Compuster Systems 22 (2006). pp 868-875.

[109] Battestilli, L., Hutanu, A., Karmous-Edwards, G., Katz, D., MacLaren, J., Mambretti, J., Moore, J., Park, S., Perros, H., Sundar, S., et al.: EnLIGHTened Computing: An Architecture for Co-allocating Network, Compute, and other Grid Resources for High-End Applications. Grid computing, high-perfromance and distributed Applications (GADA’07) (2007).

[110] S. Figuerola, N. Ciulli, M. De Leenheer, Y. Demchenko, W. Ziegler, A. Binczewski , “PHOSPHORUS: Single-step on demand services across multi-domain networks for e-science”. APOC07, Wuhan, China, November 2007

[111] Zervas, G.; Escalona, E.; Nejabati, R.; Simeonidou, D.; Carrozzo, G.; Ciulli, N.; Belter, B.; Binczewski, A.; Poznan, M.S.; Tzanakaki, A.; Markidis, G. “Phosphorus grid-enabled GMPLS control plane (GMPLS): architectures, services, and interfaces” Communications Magazine, IEEE Year: 2008, Volume: 46, Issue: 6 Pages: 128 - 137, DOI: 10.1109/MCOM.2008.4539476

[112] Hagemeier, B. KoDaVis - Collaborative Data Visualization Terena Networking Conference 2009, Workshop "On-demand network services for the Scientific Community,Malaga, Spain, (2009)

[113] “Network Service Interface WG (NSI-WG)”. ogf.org. (2010). Retrieved March 2011, from http://www.ogf.org/gf/group_info/view.php?group=nsi-wg

[114] J. Leigh, L. Renambot, A. Johnson, B. Jeong, R. Jagodic, N. Schwarz, D. Svistula, R. Singh, J. Aguilera, W. Xi, V. Vishwanath, B. Lopez, D. Sandin, T. Peterka, J. Girado, R. Kooima, J. Ge, L. Long, A. Verlo, T. A. DeFanti, M. Brown, D. Cox, R. Patterson, P. Dorn, P. Wefel, S. Levy, J. Talandis, J. Reitzer, T. Prudhomme, T. Coffin, B. Davis, P. Wielinga, B. Stolk, G. B. Koo, J. Kim, S. Han, J. W. Kim, B. Corrie, T. Zimmerman, P. Boulanger, and M. Garcia, “The Global Lambda Visualization Facility: an international ultra-high definition wide-area visualization collaborator,” FGCS, Future Gener. Comput. Syst., vol. 22, no. 8, pp. 964–971, Oct. 2006.

[115] A. Willner, C. Barz, S. Figuerola, J. A. Garcia-Espin, J. Ferrer Riera, and P. Martini: Harmony – advanced reservations in heterogeneous, multi-domain environments, in Proc. of the IFIP-TC6 Networking’09, Springer Ed. (2009).

[116] Figuerola, S.; Garcia-Espin, J.A.; Ferrer, J.; Willner, A. “Scalability analysis and evaluation of the multi-domain, Optical Network Service Plane in Harmony” Optical Communication, 2009. ECOC 2009. 35th European Conference. IEEE conference ISBN: 978-1-4244-5096-1

[117] Figuerola, S.; Garcia-Espin, J.A.; Ferrer, J.; Willner, A.; "Performance analysis of harmony: An optical, multi-domain network resource broker," Transparent Optical Networks, 2009. ICTON '09. vol., no., pp.1-5, June 28 2009-July 2 2009 doi:10.1109/ICTON.2009.5185032

[118] K. Christodoulopoulos, M. Varvarigos, C. Develder, M. De Leenheer, and B. Dhoedt: Job demand models for optical grid research, in Proc. of the ONDM’07, Springer Ed. (2007).

[119] Phosphorus EC Project D1.1 Requirements and specifications of interfaces and architecture for interoperability between NRPS, GMPLS, Middleware March 2007

[120] Buyya, R.: Market-Oriented Cloud Computing: Vision, Hype, and Reality of Delivering Computing as the 5th Utility. In: Proceedings of the 2009 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, Washington, DC, USA. CCGRID ’09, p. 1. IEEE Computer Society Press, Los Alamitos (2009), doi:10.1109/CCGRID.2009.97

[121] Rosenberg, J., Mateos, A.: The Cloud at Your Service, 1st edn. Manning Publications Co., Greenwich (2010)[121

[122] Nelson, M.R.: The Next Generation Internet, E-Business, and E-everything, http://www.aaas.org/spp/rd/ch20.pdf

[123] M. Handley,“Why the Internet only just works”, BT Technology Journal, Vol. 24, Nr 3, July 06

[124] B. Swanson, G. Gilder,"Estimating the Exaflood", http://www.discovery.org/a/4428, accessed 29th October 2008

[125] Clark, D.: The design philosophy of the DARPA internet protocols. SIGCOMM Comput. Commun. Rev. 18, 106–114 (1988), doi:10.1145/52325.52336

[126] Handley, M.: Why the Internet only just works. BT Technology Journal 24, 119–129 (2006), doi:10.1007/s10550-006-0084-z

[127] E. Desurvire, "Capacity demand and technology challenges for lightwave systems in the next two decades," JLT, Dec. 2006

[128] Anastasious Gavras et. al., "Future internet research and experimentation: the FIRE initiative", ACM SIGCOMM Computer Communication Review, Volume 37, Issue 3, Pages: 89 – 92, July 07

[129] http://www.nets-find.net

[130] www.fp7-federica.eu/

[131] GENI Planning Group,"GENI: Conceptual Design, Project Execution Plan, " GENI Design Document 06-07, January 2006. http://www.geni.net/GDD/GDD- 06-07.pdf

[132] http://groups.geni.net/geni/wiki/IGENI

[133] http://nicl.cod.cs.duke.edu/orca/

[134] https://www.itu.int/rec/T-REC-Y.2011-200410-I/es

[135] http://www.4ward-project.eu/

[136] B. Rochwerger, A. Galis, D. Breitgand, E. Levy, J.A. Cáceres, I.M. Llorente, Y. Wolfsthal, M. Wusthoff, S. Clayman, C. Chapman, W. Emmerich, E. Elmroth, R.S. Montero Design for Future Internet Service Infrastructures 10.3233/978-1-60750-007-0-227 pag.227 – 237 ISBN 978-1-60750-007-0

[137] Vicart.P., Figuerola, S., Chen, X., Liadi, G., Escalona, E., Develder, C., Tzanakaki, A., Demchenko, Y., Garcia-Espin, J.A., Ferrer, J., Lopez, E., Soudan, S., Buysse, J., Jukan, A., Ciulli, N., Brogle, M., Laarhoven, L., Belter, B., Anhalt, F., Nejabati, R., Simeonidou, D., Ngo, C., Laat, Cees de., Biancani, M., Roth, M., Donadio, P., Jimenez, J., Antoniak, M., Gumaste, A., "Bringing Optical Networks to the Cloud: An Architecture for a Sustainable Future Internet". Book Chapter Springer-Verlag Berlin, Heidelberg ©2011 Pages 307-320 ISBN: 978-3-642-20897-3

[138] Pickavet, M., Vereecken, W., Demeyer, S., Audenaert, P., Vermeulen, B., Develder, C., Colle, D., Dhoedt, B., Demeester, P.: Worldwide energy needs for ICT: The rise of power-aware networking. In: Advanced Networks and Telecommunication Systems, 2008. ANTS '08. 2nd International Symposium on. pp. 1–3 (2008)

[139] S. Figuerola, D. Simeonidou, N. Ciulli, R. Nejabati, T.M. Bohnert, J.A. Garcia-Espin, P. Vicat-Blanc, B. Belter, M. Biancani, J. Jimenez "GEYSERS architecture". Future Network & Mobile Summit 2010, IEEE conference

[140] http://www.ist-phosphorus.eu/

[141] Tzanakaki, A., Anastasopoulos, M., Georgakilas, K., Buysse, J., Leenheer, M. D., Develder, C. et al (2011). Energy Efficiency in integrated IT and Optical Network Infrastructures: The GEYSERS approach. In INFOCOM 2011 workshop on "Green Communications and Networking". Shanghai, China.

[142] A. Tzanakaki, M.P. Anastasopoulos, K. Georgakilas, J.A. Garcia-Espin, J. Ferrer Riera, S. Figuerola, M. Ghijsen, Y. Demchenko, C.T.A.M. de Laat, P. Vicat-Blanc, S. Soudan, F.

Anhalt, S. Peng, E. Escalona, R. Nejabit, D. Simeonidou "Virtual Infrastructure Planning: The GEYSERS Approach" Future Network & Mobile Summit 2012; 07/2012

[143] Geysers deliverable: D2.1 Initial Geysers architecture & interfaces specifications.

[144] Geysers deliverable: D2.2 Geysers overall architecture & interfaces specifications & service provisioning workflow

[145] Geysers deliverable: D2.6 Refined Geysers architecture, interface specification & service provisioning workflow

[146] E. Escalona, S. Peng, R. Nejabati, D. Simeonidou, J. A. Garcia-Espin, J. Ferrer, S. Figuerola, G. Landi, N. Ciulli, J. Jimenez, B. Belter, Y. Demchenko, C. de Laat, X. Chen, A. Yukan, S. Soudan, P. Vicat-Blanc, J. Buysse, M. de Leenheer, C. Develder, A. Tzanakaki, P. Robinson, M. Brogle,T. M. Bohnert "GEYSERS: A Novel Architecture for Virtualization and Co-Provisioning of Dynamic Optical Networks and IT Services" FNMS 2011. IEEE sponsored conference

[147] Geysers deliverable: D3.1 Functional description of the logical infrastructure composition layer

[148] Geysers deliverable: D3.4 Finvicat-blancal LCL prototype

[149] Geysers deliverable: D3.2 Preliminary LICL software release

[150] Primet, P.V.-B.; Soudan, S.; Verchere, D. "Virtualizing and Scheduling Optical Network Infrastructure for Emerging IT Services [Invited]" Optical Communications and Networking, IEEE/OSA Journal of Year: 2009, Volume: 1, Issue: 2 Pages: A121 - A132, DOI: 10.1364/JOCN.1.00A121

[151] vSMP: versatile SMP. Available Online: http://www.scalemp.com/smp

[152] Xen hypervisor. Available Online: http://www.xen.org/

[153] Kernel-based Virtual Machine. Available Online: http://www.linux-kvm.org

[154] VMware. Available Online: http://www.vmware.com/

[155] VServer. Available Online: http://linux-vserver.org

[156] EU FP7 Mantychore project, Available Online: http://www.mantychore.eu

[157] Van der Ham, J.; Grosso, P.; van der Pol, R.; Toonk, A.; de Laat, C., "Using the Network Description Language in Optical Networks" Integrated Network Management, 2007. IM '07. 10th IFIP/IEEE International Symposium on Year: 2007 Pages: 199 - 205, DOI: 10.1109/INM.2007.374784

[158] Geysers deliverable: D4.2 NIPS user-network interface and procedures

[159] Geysers deliverable: D4.1 GMPLS+/PCE+ Control Plane architecture

[160] ITU-T G.8080/Y.1304 Recommendations, "Architecture for the Automatic Switched Optical Network (ASON)", 2001

[161] A. Farrel, J.P. Vasseur, J. Ash, "A Path Computation Element (PCE) – Based Architecture", IETF RFC 4655, August 2006

[162] Geysers deliverable: D1.3 Technical specification of Service and system parametrization for Operation and Validation

[163] S. Figuerola, J. A. García-Espín, J. Ferrer, E. López, E. Escalona, R. Nejabati, S. Peng, D. Simeonidou "Coordinating IT and Network Virtualisation to provide Infrastructure as a

Service: a GEYSERS approach for the Future Internet”. Invited MONAMI 2011, Aveiro Portugal. IEEE conference

[164] A. Tzanakaki, M. Anastasopoulus, K. Georgakilas, G. Landi, G. Bernini, N. Ciulli, J. Ferrer, E. Escalona, J. A. Garcia, X. Hesselbach, S. Figuerola, S. Peng, R. Nejabati, D. Simeonidou, D. Parniewizc, B. Belter, J. Rodriguez; “Planning of dynamic virtual optical cloud infrastructures: The Geysers approach” IEEE ComMagazine, January 2014, Vol. 52, No. 1 DOI:10.1109/MCOM.2014.6710061

[165] GreenLight announcement http://www.calit2.net/newsroom/release.php?id=1342