034115

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Integrated Project

Strategic objective:
Research Networking Testbeds

Deliverable reference number D.2.6

Deployment Models and Solutions of the Grid-GMPLS Control Plane

Due date of deliverable: 2007-04-30
Actual submission date: 2007-05-31
Document code: PHOSPHORUS-WP2-D2.6

Start date of project:
October 1, 2006

Duration:
30 Months

Organisation name of lead contractor for this deliverable: **Athens Information Technology (AIT)**

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Abstract

This report details proposed deployment models of the Grid-GMPLS (a.k.a. G^2MPLS) Control Plane architecture. A high level description of the G^2MPLS reference points and procedures is provided in this document, as well as a summary of the current participating NRENs environments along with an analysis of the issues and solutions that arise from the deployment of the GMPLS/G^2MPLS architecture in these environments. This document addresses also issues like security considerations, AAA functionality and resource partitioning under the G^2MPLS Control Plane and presents network scenarios dealing with the coexistence of GMPLS and G^2MPLS architectures.

The main objective of the document is to facilitate the dissemination of GMPLS and G^2MPLS procedures and functionalities, in order to attract NRENs' interest on the proposed architectures. To improve the impact of this analysis and prepare the take up of PHOSPHORUS results, two questionnaires – one for NRENs and one for Super-Computing Centers – have been produced and circulated in order to evaluate their requirements and plans for GMPLS/G^2MPLS adoption in these research infrastructures.

A complementary objective of this deliverable is to drive the subsequent low-level designs and software developments of the identified Grid-GMPLS interfaces that will be detailed in deliverables D2.2 and D2.7.

List of Contributors

Anna Tzanakaki	AIT	Bartosz Belter	PSNC
Georgios Markidis	AIT	Damian Parniewicz	PSNC
Nicola Ciulli	NXW	Maciej Strozyk	PSNC
Gino Carrozzo	NXW	Radoslaw Krzywania	PSNC
Giodi Giorgi	NXW	Ferdinand Hommes	FHG-IAIS
Eduard Escalona	Fi2CAT/UPC	Dimitra Simeonidou	UESSEX
Artur Binczewski	PSNC	George Zervas	UESSEX
Maciej Stroiński	PSNC	Reza Nejabati	UESSEX
Jan Węglarz	PSNC	Wes Doonan	ADVA

Table of Contents

Table of Figures

0 Executive Summary

This document describes proposed deployment models of G^{2}MPLS Control Plane and provides a high level discussion of the identified network reference points and procedures. Also in the context of this document the current status of participating NRENs is investigated aiming to define the possibility of deploying GMPLS or G^{2}MPLS Network Control Plane.

In section 1 the objectives of the G^{2}MPLS deployment models are stated, as well as the scope of the document in the WP2 framework.

In section 2 the terminology relevant to the G^{2}MPLS architecture is presented by specifying the main source of information.

Section 3 provides a brief summary of NRENs environments participating in the PHOSPHORUS project and specific technologies and procedures for Bandwidth on Demand services and Grids are briefly introduced aiming to investigate possible deployment of GMPLS or G^{2}MPLS Network Control Plane.

In section 4 several issues concerning the GMPLS Control Plane deployment are presented. Topics such as resource partitioning and virtualization, security, ASON/GMPLS network reference points and procedures as well as possible deployment options for NRENs are discussed and analyzed.

In section 5 the G^{2}MPLS Control Plane deployment models are examined following the previous section and address issues related to the integration of Grid information into the proposed Control Plane.

Section 6 studies coexistence and interoperability issues of GMPLS and G^{2}MPLS Control Planes through network scenarios.

Appendix A presents case studies of PSNC and CESNET Test-Beds in the framework of PHOSPHORUS.

Appendix B provides the filled questionnaires from NRENs investigating their requirements and possible plans towards GMPLS/G^{2}MPLS adoption in their infrastructures.

Appendix C presents a similar questionnaire for super-computing centers.

1 Objectives and Scope

This document analyzes deployment models of Grid-GMPLS (G^2MPLS) Control plane, exposing benefits and enhancements compared to standard GMPLS Control plane. Deriving from these models, a set of network reference points and procedures is identified along with required signalling and accounting mechanisms. In addition AAA functionality is investigated under the G^2MPLS Control Plane while a framework providing resource partitioning and dealing with coexistence of GMPLS and G^2MPLS issues is described.

This document provides a brief summary of NRENs environments participating in the PHOSPHORUS project and specific technologies and the relevant procedures for Bandwidth on Demand services and Grids with the aim to identify possible deployment of GMPLS or G^2MPLS Network Control Plane. In addition, some investigation of the intentions of NRENs and Supercomputing Centers to adopt the framework of G^2MPLS Control Plane is attempted. Filled questionnaires from NRENs that demonstrate their interest for the deployment of the G^2MPLS framework into their local test-beds are provided and a template questionnaire for Supercomputing Centers has been also formed and reported, while case-studies regarding deployment issues into some NRENs test–beds are presented.

G^2MPLS aims to enhance the ASON/GMPLS Control Plane architecture by implementing the concept of Grid Network Services (GNS), providing single-step resource reservation, co-allocation and maintenance of both network and Grid resources. G^2MPLS aims to inherit the rights for full usage of the Transport Plane resources configured under its ownership from the GMPLS Network Control Plane and in some cases maintains some of these rights under the authority of the management plane.

This document will facilitate the dissemination of efficient G^2MPLS procedures and functionalities attracting NRENs interest on the proposed architecture. The provided questionnaires aim to identify current status and concerns on deployment approaches by NRENs participating in PHOSPHORUS project (i.e. PIONIER, CESNET, SURFnet), so that a variety of issues is addressed. The impact of the proposed solutions could be extended to other research networking test-beds (e.g. VIOLA and its follow-up) and GÉANT2 as well.

The support of Grids does not impose specific requirements on the G^2MPLS Signalling Control Network (SCN), and in general the same means available in GMPLS can be used to establish a Control Plane adjacency. This document aims to highlight the difference between GMPLS SCN and G^2MPLS SCN i.e. the existence of a Grid among the users attached to the Control Plane. This requirement, in the G^2MPLS case, implies that some middleware messages for Grid specific transactions should share the same SCN, used for the setup and maintenance of the Transport Plane connections among the performers.

This deliverable also aims to address some security related issues. Within this framework G^2MPLS Grid power-users represent a critical vulnerability point of the Control Plane and Transport Plane due to their capability to configure both Grid and network resources with seamless and automated procedures. Therefore as emphasized in this document, in addition to standard considerations for GMPLS SCN security in the Transport and Control Plane, the G^2MPLS users are supposed to access network and far-end Grid resources after a successful phase of authentication and authorization against an Authorization Entity.

Another issue that this document aims to address is the Introduction of the PHOSPHORUS deployment models and the required extensions to the standard interfaces and procedures. In the G^2MPLS architecture two Service and Control Planes layering models have been identified that have a different scope compared to the Overlay, Augmented and Peer models described in standard GMPLS: *"G^2MPLS Overlay"* and *"G^2MPLS Integrated"*, which refer to different layering solutions of Grid and network resources. The deployment of the enhanced G^2MPLS Control Plane sets analogous reference points with respect to the ASON/GMPLS ones. The resulting network interfaces (G.OUNI, G.I-NNI, G.E-NNI, SBI, NBI) are a Grid-aware evolution of the standard interfaces (UNI, I-NNI, E-NNI), with a set of procedures that maintain the backward compatibility with the original ASON references, but provide also the seamless and one-step control of both Grid and network resources. Procedures deployed by the GMPLS Control Plane require extensions to provide Grid information to the G^2MPLS framework and support its expectations. Also the Authentication and Authorization Infrastructure (AAI) must be extended to support Grid services as well as the SLA management system proposed for GMPLS in which several modifications are necessary.

In general, the G^2MPLS Control Plane has been defined as a superset of the original GMPLS functions and procedures, with the additions being devoted to support the specific additional functionalities required by Grids. One of the main objectives of this document is to identify and explore environments where the coexistence of the two architectures is necessary, or even required by the network operators. In this scope, some network scenarios are discussed and analyzed aiming in providing an evaluation of the achievable level of interworking in signalling and routing procedures.

2 Terminology

In this section some definitions particularly relevant in the scope of this deliverable are provided in case with reference to the originator document.

Keyword	Source	Definition
High Performance Computing	**[OGF-GFD.11]**	The technology that is used to provide solutions to problems that: • require significant computational power • either need to access, or process, very large amounts of data quickly • need to operate interactively across a geographically distributed network
Grid Scheduler	**[OGF-GFD.81]**	A service that manages a set of one or more *job* instances, which may be structured (e.g. a *workflow* or dependence graph) or unstructured (e.g. an array of non-interacting jobs). The Grid Scheduler encapsulates all aspects of job execution, including interacting with execution planning *services*, the *provisioning* system, *containers*, and monitoring services. It may also deal with *failures* and restarts, it may schedule *jobs* to *resources*, and it may collect *agreements*, *reservations* and job service data. It is analogous to the Job Manager definition in OGF-GFD.81
Execution Management	**[OGSA EMS]**	Services concerned with the problems of balancing the requirements of work to be done on the Grid with

Services		the constraints of the resource required to complete the work. Within OGSA-EMS the atomic unit of work is the task and can in its simplest cases refer to, for example, a legacy batch job, a database server, a servlet running in a Java application server container, etc.
Authentication Authorization Accounting	**WP4 D4.1**	A term used to refer to a framework for intelligently controlling access to computer resources, enforcing policies, auditing usage, and providing the information necessary to bill for services. These combined functions are considered important for effective network management and security. Authentication is the process of identifying a user or an access subject, based on identity credentials which examples are username and password, digital certificates, one-time-tokens, etc. Authorization refers to the confirmation that a user/subject who is requesting services is a valid user of the resources or services requested. Accounting is the process of keeping track of a user's activity while accessing the resources or services
Multi-Layer Network	**[IETF – gmpls-mln-eval]**	A network comprising transport nodes with different data plane switching layers (a collection of network resources capable of terminating and/or switching data traffic of a particular format e.g. LSC, TDM VC-11 and TDM VC-4-64c) controlled by a single GMPLS control plane instance
Multi-Region Network	**[IETF – gmpls-mln-eval]**	A network comprised of multiple switching types (e.g. PSC and TDM) controlled by a single GMPLS control plane instance. A GMPLS switching type (PSC, TDM, etc.) describes the ability of a node to forward data of a particular data plane technology, and uniquely identifies a control plane region.
Forwarding Adjacency - LSP	**[IETF – RFC4206]**	An LSP created by an LSR using GMPLS TE procedures and announce as a TE link into the same instance of the GMPLS Control Plane and therefore

		an FA is only applicable when an LSP is both created and used as a TE link by exactly the same instance of the GMPLS control plane. Also an FA is a TE link between two GMPLS nodes whose path transits zero or more (G)MPLS nodes in the same instance of the GMPLS control plane.
Explicit Route Object	**[IETF – RFC3209]**	It is used to encode the explicit path of a Label Switched Path through the network and is constructed and processed by LSRs. When ERO is present the Path message is forwarded towards its destination along a path specified by the ERO. Each node along the path records the ERO in its path state block and may also modify the ERO before forwarding the Path message. In this case the modified ERO should be stored in the path state block in addition to the received ERO.

2.1 Abbreviations

A full list of the abbreviations used in this document is provided in Section 9.

3 Current resource provisioning models in European Research networks

In this section a summary descriptions of the European Research networking framework is provided with the main focus on the NRENs participating in the PHOSPHORUS project. Network infrastructures, specific technologies and procedures for Bandwidth on Demand services and Grids are briefly introduced with the aim of defining the context for a possible deployment of GMPLS or G^{2}MPLS Network Control Plane.

3.1 PSNC

PSNC is an operator of the Polish National Optical Network – PIONIER. The PIONIER network topology, shown in Figure 3-1, includes over 6000 km of optical fibres (G.652 and G.655) and connects academic and metropolitan area networks and HPC centers in Poland. The transport DWDM system is provided by ADVA and supports up to 32 lambdas. The core of the PIONIER backbone network has capacity of 10Gbit/s and is based on Foundry Networks XMR switches and the main PIONIER router Juniper T320 placed in PSNC.

The PIONIER network based on Ethernet technology is supported by IP/MPLS mechanisms for service flexibility. Two main MPLS-based technologies, VLL (Virtual Leased Line) for point-to-point links and VPLS (Virtual Private LAN Service) for multi-point connection, are used in core of the network to satisfy users' requirements.

A possible PSNC network evolution is a migration from MPLS to GMPLS. Process of introducing GMPLS into already deployed PIONIER MPLS-based infrastructure, which is not currently GMPLS capable, is one of the key issues. Actually, only ADVA WDM system is equipped with proprietary GMPLS based Control Plane and the PIONIER upper layers equipment (i.e. XMR switches) does not provide any support for GMPLS. In the case, PSNC is very interested in development and implementation of GMPLS functionality for all PIONIER network layers in order to support user requests over UNI and to ensure efficient resource management in dynamic environment.

PSNC has also future plans for extension PIONIER network with GMPLS capable optical cross connects (OXC) in order to provide fully configurable on-demand end-to-end optical paths.

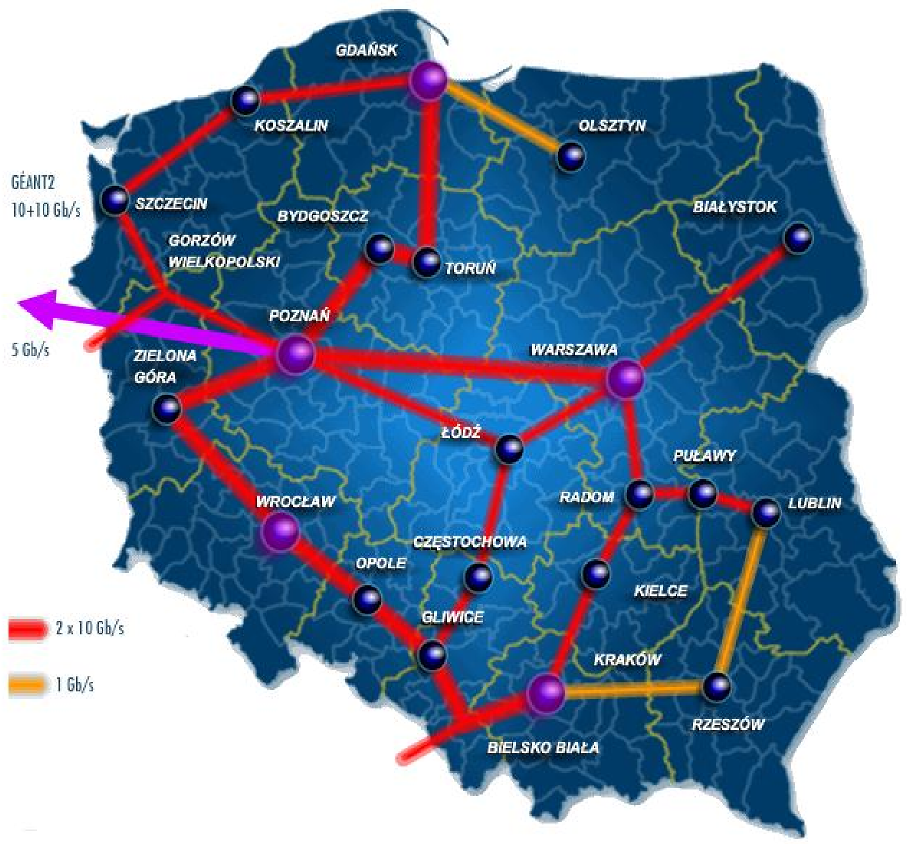

Figure 3-1: PIONIER topology

Since PSNC is network operator as well as HPC center and the PIONIER network connects other scientific HPC centers in Poland, the PHOSPHORUS G^{2}MPLS mechanism for combined reservation of network and grid resources is also in scope of interest. A scenario where a few HPC centers placed in one network domain, working together on one task and periodically exchanging huge amount of data could be an interesting G^{2}MPLS test-bed.

3.2 CESNET

The contemporary CESNET2 network (see Figure 3-2) is built around the DWDM core. Remaining backbone lines of the CESNET2 network are based on the mixture of Ten-Gigabit and Gigabit Ethernet and Packet Over SONET (POS). It ensures very high throughput (bandwidth of all core lines lies in the range of gigabits per second). The essential idea of CESNET2 design is to build the backbone network in small rings encompassing no more than four cities (points of presence). Such topology provides redundant network offering small hop counts and delays.

The CESNET network has following independent international connections:

- 10 Gbps to GÉANT, used for academic traffic
- 800 Mbps to Telia, used for commodity traffic
- 10 Gbps to SANET, academic network of Slovakia
- 10 Gbps to ACOnet, academic network of Austria
- 10 Gbps to PIONIER, Polish optical academic network
- 10 Gbps to NetherLight/GLIF for experimental traffic; it is not depicted in the map of CESNET2 above

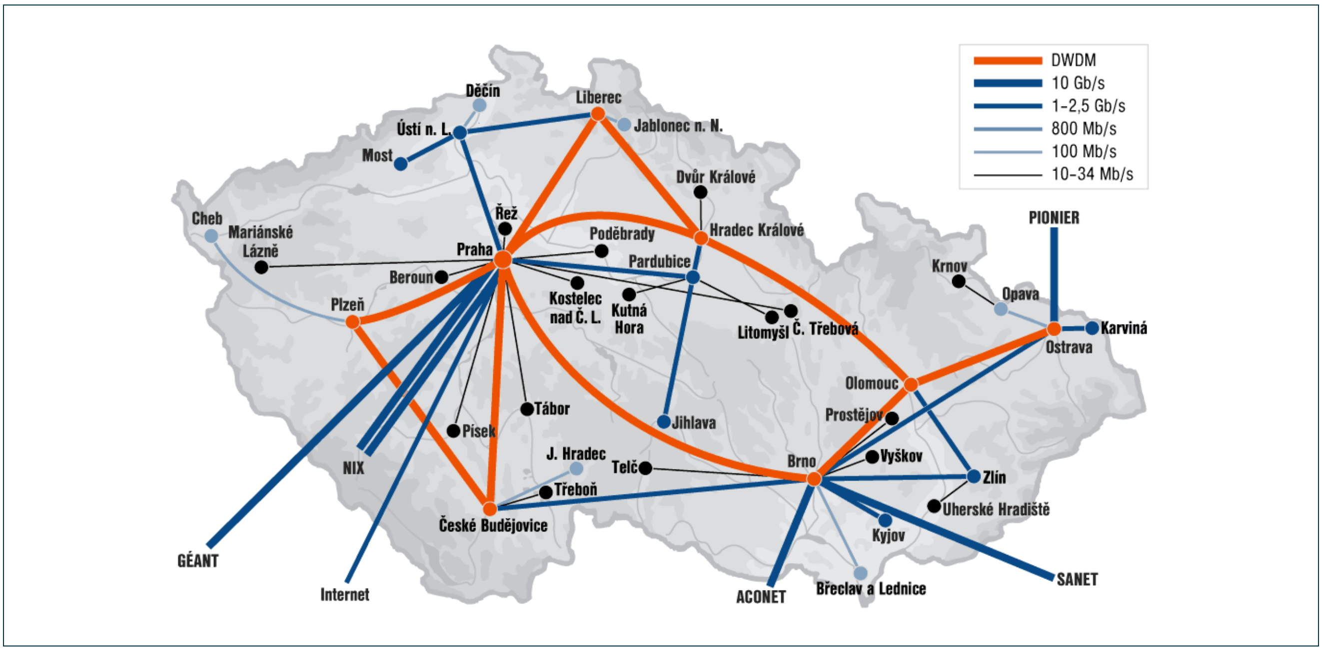

Figure 3-2: CESNET2 topology

CESNET infrastructure is currently not capable of using GMPLS protocols; however, network experts are willing to include this functionality into their network environment. The main objectives for this migration are:

- to implement automatic transport service restoration
- to give users access to dynamic transport service (through UNI)

CESNET is interested in introducing GMPLS protocols in very wide range of technologies, starting from IP layer, through Ethernet, SDH/SONET, to DWDM/ROADM and other optical layers. The implementation process should be focused on Cisco network devices ONS15454 MSPP and MSTP, which are already a part of CESNET infrastructure.

G^{2}MPLS enhancements of PHOSPHORUS project is also a point of interest, however a strong emphasis is placed on security issues. The idea of single-step for both network and Grid resources reservation is welcome and potential implementation points can be selected from linked HPC/Grid centers.

CESNET is involved in cooperation with EGEE and LHC-OPN projects, which require high-bandwidth circuits and utilize network and Grid resources above average level. EGEE HPC is integrated with CESNET network infrastructure and thus is indicated as potential implementation point for G^2MPLS features.

3.3 SURFNET

SURFnet is built on a DWDM core consisting of five national rings (see the Figure 3-3 below). All of them have a capacity of 72 wavelengths; at the moment about 65 wavelengths are lit with 10 Gbps rate each. Four of the rings connect to the two redundant core sites in Amsterdam, where traffic is terminated or interchanged between the rings. IP traffic is aggregated at L2 at fifteen major sites and from there transported to the two core sites where all routing takes place.

At one of the two core sites the network is linked to both GÉANT2 and NetherLight. Via these two international connections inside Europe and to the rest of the world via the GLIF can be provisioned. Currently NetherLight has 145 Gbps connectivity to different networks and locations, and carries links from Beijing, Moscow, Vancouver and several other locations (see Figure 3-4 below).

SURFnet provides two basic services to its connected organisations: firstly the IP service with interfaces at both 1 and 10 Gbps, and with a rich set of features ranging from IPv6 to multicast. The second service SURFnet has been providing - for more than a year now - its lightpaths (fully transparent point to point connections across the network). These are presented to the client with a dedicated GE interface, even when the provisioned bandwidth can be lower (standard rates currently offered are 1G, 600 Mbps and 150 Mbps).

By the end of 2007 this lightpath service will also be available on as a dynamic, user-managed service. (End) users will get the ability to create these lightpaths scheduled and on-demand with any bandwidth falling within their policy, both via a web GUI and a web services interface that allows for easy integration with applications. The technology that enables this service uses some Control Plane logic internally, but to make it into a tool that fits network operations and envisaged applications the development needed to go well beyond the currently available Control Plane technology.

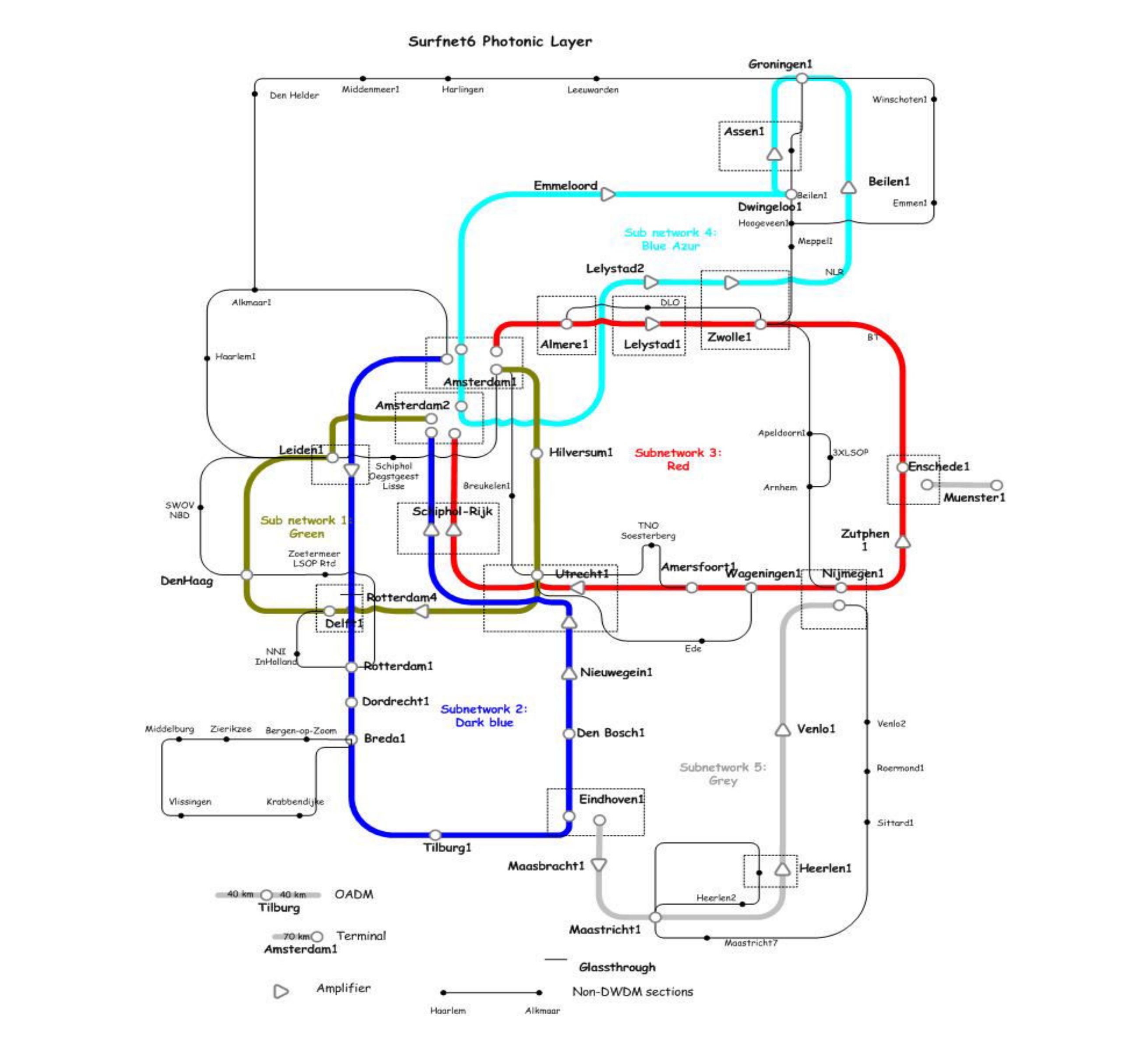

Figure 3-3: SURFnet topology

There are currently no plans within SURFnet to revert to standard Control Planes as the basis of this dynamic service, as this would reduce the richness of, and management over the service and its ways to interact with the network. As important is the fact that Control Planes do not have a user-focused approach for now and have far slower development cycles than what we expect to require for the service. At the same time there are no plans to implement Control Planes for the other SURFnet services. We will however continue to evaluate the progress in the Control Plane field.

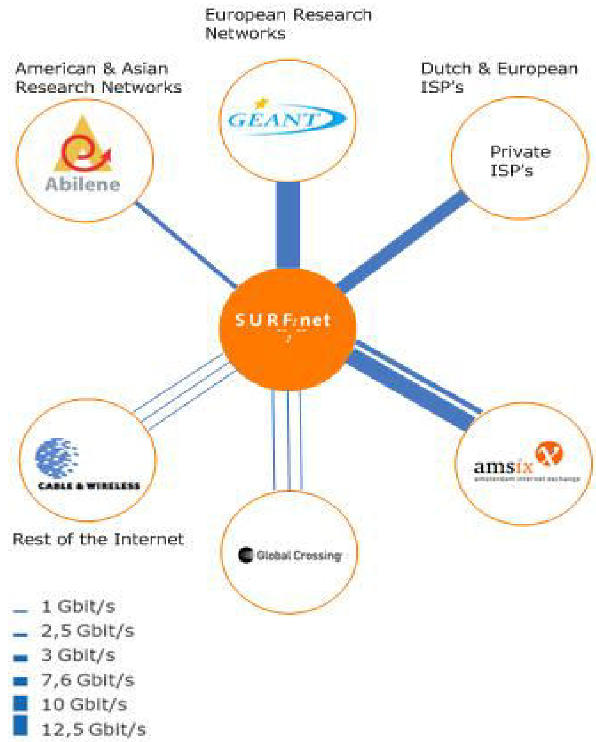

Figure 3-4: SURFnet external connections

3.4 GÉANT2

GÉANT2 [GEANT2] is a multi-gigabit research and education network, pan-European in scale and reach. It provides the European research and education community with a state-of-the-art data communications backbone network. The network provides the most advanced services and widest geographical reach of any network of its kind in the world, boasting leading-edge standards of reliability and innovation.

GÉANT2 network connects 34 countries through 30 national research end education networks (NRENs). The topology of the network is shown in Figure 3-5 below:

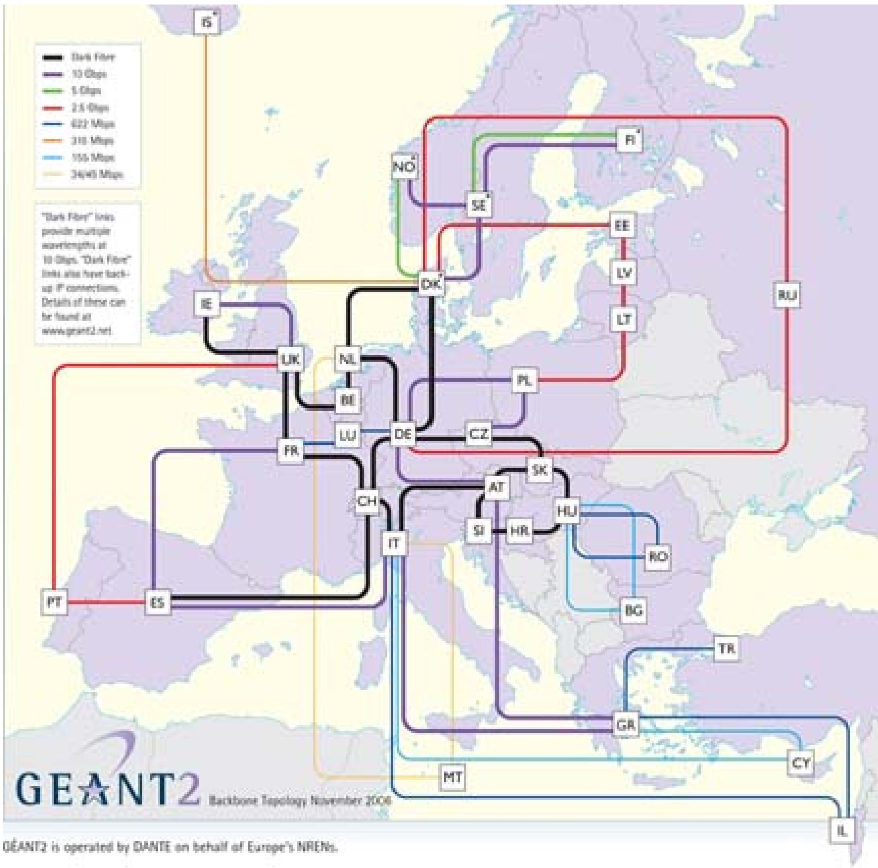

Figure 3-5: GÉANT2 Topology

3.4.1 Bandwidth on Demand services across the GÉANT2 network

The strong emphasis on the need for GÉANT2 to provide guaranteed, dedicated capacity to some of its users is a major driver for the development of the Bandwidth on Demand (BoD) research activity. The multi-domain nature of the networks on which the BoD service will be implemented makes it probable that more than one technology will be used to provide the service. Potential technologies and/or architectures that fall within the scope of this activity include:

- MPLS label-switched paths (LSPs), possibly enhanced with packet-based QoS standards
- Native (and emulated) Layer 2 channels (particularly Ethernet)
- Time-division multiplexing (TDM) channels (based on SONET or SDH transmission)
- Native Layer 1 wavelengths on fibre ("lambdas").

The ultimate goal is to create a dynamically-provisioned, end-to-end prototype BoD service. A prototype provisioning system has been developed and allows the reservation of a dedicated, manually provisioned connection within a single domain.

3.4.2 Multi-domain Premium IP provisioning on GÉANT2

The Premium IP model defined by SEQUIN project [QOS-SEQUIN] can be implemented via over-provisioning, DiffServ, ATM CBR PVC or any other technique providing reasonable assurances to QoS parameters. On GÉANT2, the DiffServ Expedited Forwarding Per-Hop Behaviour (EF PHB) was chosen in preference to over-provisioning because of the GÉANT2 topology and router functionalities.

GÉANT2 accepts Premium IP flows from NRENs, and polices these flows based on their source and destination address and the capacity requested by the NREN. This access-control is needed to protect the Premium IP packets against misuse, which could "degrade" Premium IP performance. This is based on the assumption that the NRENs are themselves able to isolate the authorised Premium IP flows from others and send only these to GÉANT2 tagged as Premium IP (the others having to be re-tagged as Best Effort). If the NRENs do not re-tag their unauthorised Premium IP packets, GÉANT2 can do it for a limited amount of Premium IP requests based on source and destination IP addresses provided by the NREN or projects.

Classifiers and queuing mechanisms have been configured on most of the GÉANT2 interfaces. An input filter is used to classify Premium IP packets into a dedicated Premium IP queue.

The GN2 Advance Multi-domain System (AMPS) enables authorized end-users to make a single reservation for Premium IP bandwidth (that is to say, a guaranteed un-congested path) that is effective across a chain of participating domains.

3.5 VIOLA

The German optical test bed VIOLA is an integrated test bed for applications and advanced network services, organised as a consortium with partners from industry, research laboratories, universities and the DFN association. Major goals are the test of different signalling mechanisms in a heterogeneous environment of routers and switches and the development of a user-driven dynamical bandwidth allocation.

Within the VIOLA test bed, but also in a joint activity with EU project MUPBED, experimental evaluations were performed regarding signalled, on-demand Ethernet services over SDH, spanning over multiple network domains. To enable this seamless inter-domain interworking OIF (Optical Internetworking Forum) interfaces, e.g. UNI 2.0 Ethernet (User Network Interface) and E-NNI (External Network-Network-Interfaces), were used. Figure 3-6 shows the test topology and the devices involved within the tests.

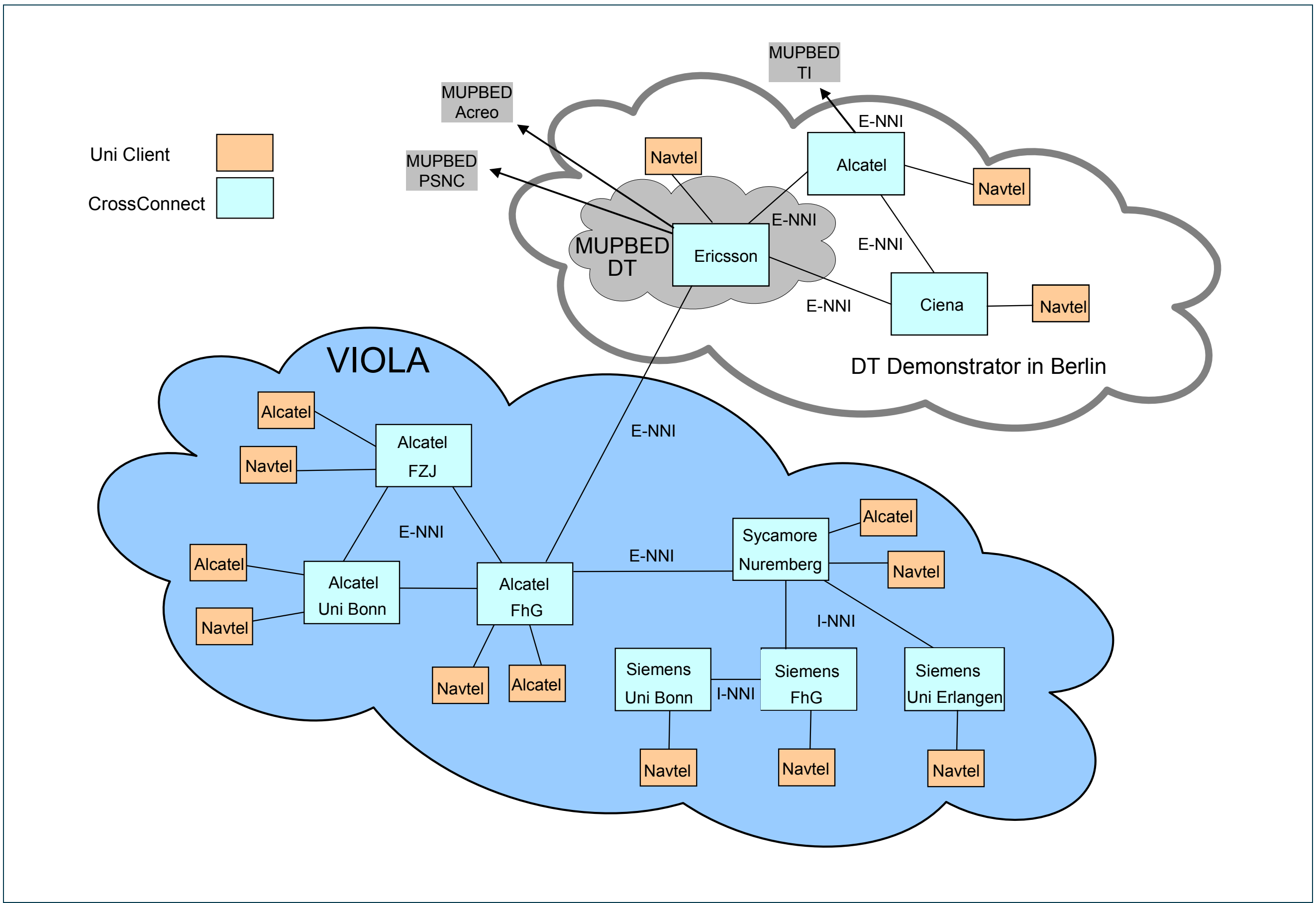

Figure 3-6: VIOLA ASON/GMPLS/OIF Tests

VIOLA has been started in summer 2004 and it will end at the end of April 2007. The VIOLA switching equipment will remain operational after the end of the VIOLA project and it will be used for PHOSPHORUS GMPLS/G^{2}MPLS deployment.

Meanwhile, a new project proposal called BODEGA has been specified. Its main focus is the establishment of a BoD service for GRID applications. It is based on all optical technologies and it will include service level agreements. It is planned, that within BODEGA the G^{2}MPLS Control Plane PHOSPHORUS could be deployed. Currently, no decision about the funding of BODEGA has been made.

The German NREN called X-WiN does not support any signalling functions. X-WiN is based on DWDM point to point connections. The DWDM lambdas connect routers, which are used for providing normal IP connectivity to the German research community. Besides this, the DWDM lambdas can also be used to provide point-to-point connections between two research institutes. These additional lambdas are static. They have to be ordered for longer time periods, e.g. 3 years.

4 Deployment modes for the GMPLS Control Plane

4.1 Resource partitioning and virtualization for GMPLS

GMPLS Network Control Plane is designed in order to operate autonomously on the Transport Plane resources under its ownership. In fact, GMPLS Control Plane components operate directly on the transport resources and in conflict with the configuration actions by other possibly co-existent managing entities on the same set of resources. In a pure GMPLS perspective, the Management Plane is conceived to interact with the Transport Plane under the mediation by the Control Plane.

In a network with pre-existing Management Plane infrastructure, it could be valuable to integrate GMPLS Control Plane on a disjoint partition of the available network resources, by moving some of them under the authority of the GMPLS Control Plane and maintaining some others under the management plane. There could be various reasons behind this solution, and they much depend on the Network Operator policy for the adoption of new technologies. For example, Management Plane full functionality could be maintained:

- in those legacy equipments for which no upgrade to GMPLS could be possible due to hardware limitations;
- in those sections of the network that are supposed to generate a high number of Control Plane transactions (e.g. in the access periphery) in a way that corresponding control traffic could be unsustainable by the existing SCN / DCN;
- for those network connections and resources with extremely demanding configuration and monitoring requirements that can just be inherited by the GMPLS Control Plane but not configured nor monitored (e.g. SDH/SONET ring protections).

In case of resource partitioning, it is not possible for the Control Plane to work on resources that are under the authority of the Management Plane as depicted in Figure 4-1. In fact, the atomic functions provided by the Transport Plane equipments towards the Management Plane are at the same hierarchical level of the GMPLS Control Plane ones and compete with them.

On the contrary, the Management Plane can modify the resources exported by the Control Plane and manage them as virtual resources, but according to the GMPLS information model. This model is the base for communications between Management and Control Planes and is a subset, though structured, of the classical management information model. Therefore, the configuration and monitoring functionalities exported by the Control Plane components (i.e. Network Call Controller, Connection Controller, etc.) are limited with respect to those available with traditional Telecommunications Management Network (TMN) infrastructures.

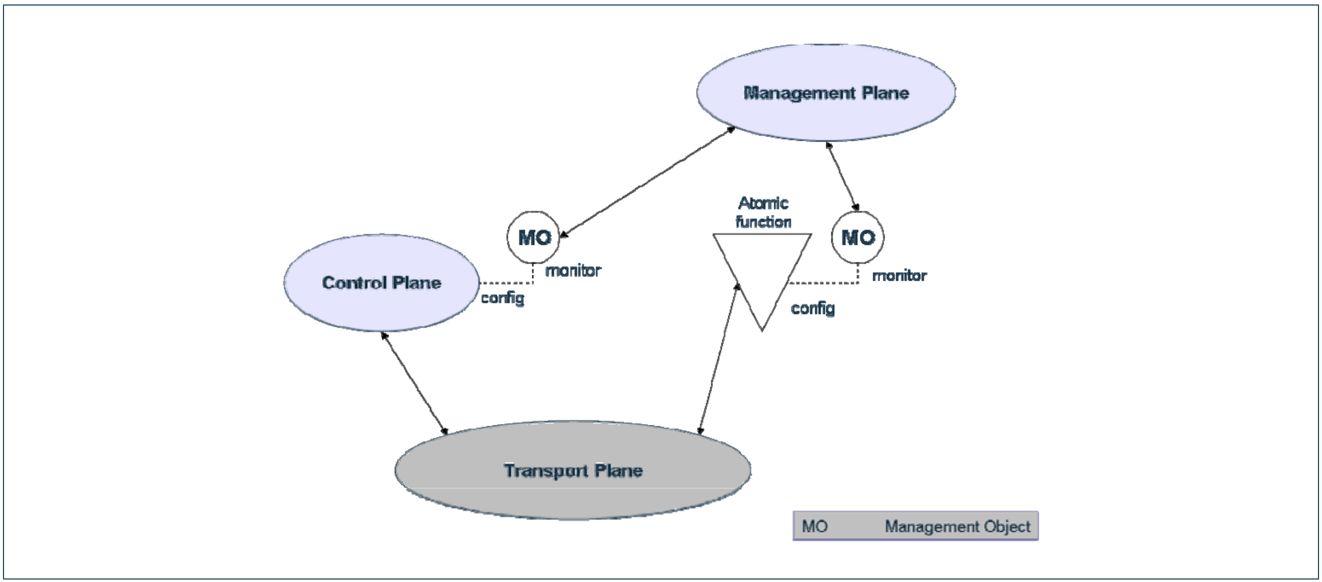

Figure 4-1: Representation of the interactions among Management, Control and Transport Planes.

The main actions that can be exported by the Control Plane to the Management Plane are:

- Configuration of protocols behaviours and Control Plane entities (e.g. identifiers for TE-links, interfaces, metrics, administrative colours, etc.)
- Request for calls/connections setup/tear-down in the Control Plane (Soft Permanent Connection in the ASON/GMPLS terminology)
- Some OAM functionalities on the installed connections.

These functionalities provide a virtualization of the Transport Network resources towards the Management Plane, which could treat them as unitary Transport Plane resources. For example, a Soft Permanent Connection (SPC) is the combination of two Permanent Connection segments at the source and destination user-to-network interfaces, plus a Switched Connection segment within the core network. The PC parts are solely owned by the Management Plane, while the SC part is directly owned by the Control Plane but exported to the Management Plane that can treat it as a direct forwarding adjacency (link) between the ingress and egress nodes (virtual resource).

4.2 Integration of ASON/GMPLS SCN in the existing DCN

As per ITU G.7712, ASON requires a communications network, the Signalling Control Network (SCN), to transport signalling messages between its components. The connectivity, i.e. signalling channels, among the SCN components is provided by the Data Communication Network (DCN), which is a separate network with respect to the Transport Network with its own communication stack and means [ASON-DCN].

Since GMPLS is designed natively for IP networks and explicitly requests for IP functionalities for the Control Plane infrastructure [GMPLS-ARCH], the SCN must support IP at Layer 3. GMPLS operation on top of legacy OSI DCN is a problem of interworking and deployment local to the specific implementations and solved during the DCN design and adaptation by the network operator. Some guidelines and specifications have been identified in [ASON-DCN].

Protocol communications between two GMPLS instances are possible through the SCN if an adjacency is established between them. The existence of such adjacency implies that an IP path must exist between the two communicating entities. According to [GMPLS-ARCH] this can be accomplished in different ways:

- IP tunnelling,
- GRE tunnelling,
- A TE-link with interface switching capability of PSC,
- A bidirectional LSP with interface switching capability of PSC.

Use of tunnelling (GRE or IP-IP) for GMPLS SCN adjacencies creation is one of the most common deployment options. In fact, with tunnelling:

- GMPLS control traffic can experience a virtual topology as close as possible to that of the transport network (e.g. adjacencies over the Transport Network are also adjacencies over the DCN, which is not necessarily true at the level the physical topologies). This facilitates planning and maintenance of the control plane communication;
- GMPLS control traffic could be secured with respect to other DCN traffic via use of encryption mechanisms at the far-ends of the tunnel;
- co-existence issues between GMPLS protocols and the different or similar protocols used in the DCN are limited (IP routers in the DCN could neither intercept signalling messages, nor participate in routing process in case they use the same IGP).

The deployment choice depends on the configuration pre-requisites of the adopted GMPLS stack and the existing DCN infrastructure. In any case, the SCN management layer in the GMPLS stack is responsible for abstracting the specific details of the underlying DCN, in terms of both Control Plane adjacency options and communication channels (e.g. in-fibre Data Communication Channel – DCC – in case of TDM networks, or dedicated out-of-band Optical Supervisory Channel – OSC – in case of G.709 networks, etc.). It is worth noting that in-fibre DCN communication channels bind the fate of the GMPLS Control Channels to the served TE-links, which could result in possible isolation of both Transport Plane and SCN in case of fault (at least temporary, i.e. until the DCN IGP reaches a new convergence status – 60 sec ca.). However, this SCN design option could be unavoidable in some deployment scenarios, depending on the pre-existing DCN infrastructure and the existence of alternative routes.

4.2.1 Security Issues

The automatic setup and maintenance of Transport Plane connections provided by GMPLS Control Plane raises a number of issues on the security of the network infrastructure, concerning attacks on the Control Plane and/or unauthorized usage of Transport Plane resources.

Transport Plane security is a major issue particularly in case of Switched Connections because, in this case, the full Control Plane functionality is used with an end-to-end scope, spanning from the source user to the destination user, and there is no mean for the network administrator to intercept or check the trust chain through its network resources. On the contrary, in case of Soft Permanent Connections, the inner Transport Plane is configured upon an internal trigger, i.e. the Network Management System or some equivalent planning tool, and a trust relationship can be natively assessed because the permanent connections configured at the end-sides are dedicated to users validated by the network administrator with the traditional means.

Control Plane security is a relevant issue that impacts GMPLS protocols and controllers, in particular. In fact, there are a number of security threats that may be experienced due to the exchange of messages and information. Some examples include interception, spoofing, modification, and replay of Control Plane messages, which are aimed at hampering the correct protocols procedures and/or snooping GMPLS control traffic. The seriousness of these risks depends on the level of trust between nodes that exchange GMPLS control messages, as well as on the implementation and physical characteristics of the control channels. For example, an in-fiber control channel over SONET/SDH DCCs is considered less vulnerable than a control channel realized over an out-of-fibre IP network relying on Ethernet segments. However, a number of mechanisms can be adopted against these threats:

- Signaling must be able to provide authentication, integrity, and protection against replay attacks. This could be done by using MD5 signatures, message identification mechanisms, etc.
- Authentication is required to ensure that the signaling messages are originating from the right place and have not been modified in transit.
- Privacy and confidentiality are not a basic requirement but it could be provided with IPSec mechanism

The consequences of poorly secured protection may increase the risk of triggering Control Plane actions (setup, tear-down, recovery, routing flooding) under false messages, with possibly catastrophic consequences. For example, recovery could be triggered for LSP identifiers that are not under failure, thus initiating a "false" recovery action that could involve nodes for which the failure condition was not learnt and lead to de-synchronisation between the control and the data plane, as well as increase the risk of misconnections. Analogously, tampering with a routing information exchange may also have an effect on traffic engineering, by filling in the TE-LSDB of the Control Plane routers until their overflow and thus leading to a block in pat computation process.

4.3 Choice of ASON vs. IETF network reference points

The set of interoperable interfaces between client (user) and network domains are based on an overlay model, specifying the signalling and routing information exchange between two domains. The exchange of information across these reference points is described by the multiple abstract interfaces between control components.

4.3.1 ITU-T ASON approach

In the ASON Control Plane three interfaces are identified:

- the UNI, i.e. the reference point between an administrative domain and an end user;
- the E-NNI, i.e. the reference point between domains
- the I-NNI, i.e. the reference point within a domain between routing areas and, where required, between sets of control components within routing areas.

These interfaces refer to the concept of Control Plane domain, i.e. the generalization of the administrative region concept (e.g. an Autonomous System in IETF) including administrative and/or managerial responsibilities, trust relationships, addressing schemes, infrastructure capabilities, survivability techniques, distributions of control functionality, commercial policies, etc. There is no link between the ASON domain and the AS domain, i.e. an ASON domain might include multiple AS, or – more likely – multiple ASON domains might be included in a single AS.

Definition of domain membership and partitioning is an operator policy and it could be based on grouping Transport Plane equipments (and related Control Plane instances) per vendor, geographical position, technology, political reasons, commercial practices, etc. As depicted in Figure 4-2, the different providers' domains interact through the E-NNI and the same interface could be used by a further partitioning of a single provider domain into sub-domains at a lower hierarchical level.

Definition of domains implies identification of service demarcation points.

ASON model is basically overlay-style because:

- the final end-to-end service (call) is obtained through the automatic stitching of the different service segments in the traversed domains,
- user administrative domains do not participate in routing information exchange
- the visibility of the inner structure of the administrative domain is controlled by the policy of the service provider.

- signalling protocols are specified for the boundary interfaces (UNI, E-NNI) and could be applicable to the I-NNI but without any assumption by the ASON recommendations.

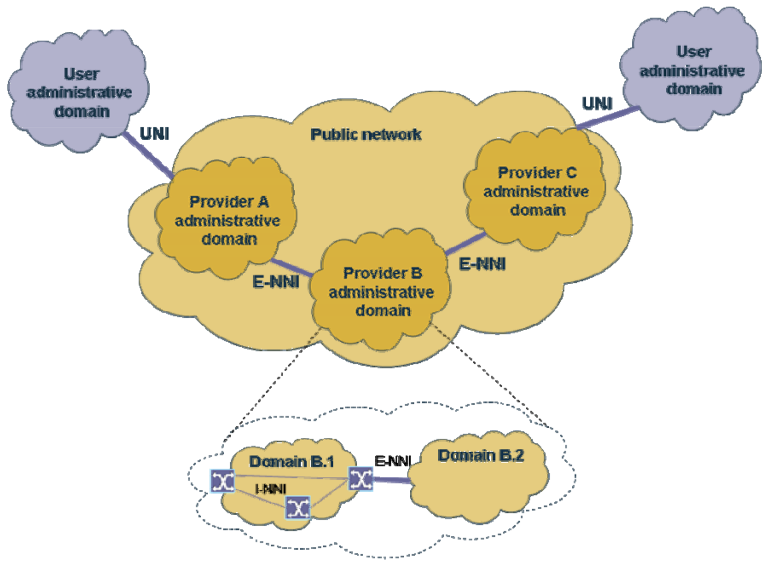

Figure 4-2: ASON identification of the network reference points between domains.

The implementation of the ASON reference points is specified in terms of the procedures and message extensions on top of GMPLS signalling and routing protocols [ASON-DCM, ASON-PNNI, ASON-RSVP, ASON-LDP, ASON-ROUT, ASON-OSPF]. The OIF Implementation Agreements [OIF-UNI2.0-COMM, OIF-UNI2.0-RSVP, OIF-ENNI-Sig-2.0, OIF-ENNI-Rtr-1.0] provide an ASON compliant implementation of the cited interfaces, though limited to Ethernet and TDM switching technologies and the inter-vendor, intra-carrier administrative partitioning. Inter-carrier issues are still under definition, because they involve critical issues like AAA and trust/security mechanisms for which it is hard to gain consensus of the networking community.

4.3.2 IETF GMPLS approach

GMPLS was developed as a unified Network Control Plane for different switching technologies at Layer 3 (PSC), Layer 2 (L2SC) and Layer 1 (TDM, LSC, FSC). Therefore, GMPLS can be applied to a number of contiguous switching layers from access network to core networks with the aim of a unified and automated service provisioning and maintenance (particularly for recovery).

IETF GMPLS does not specify separately network reference points as in ASON architecture. Edge nodes are connected to LSRs on the network side, and these LSRs are in turn connected between them, despite the possible administrative partitioning of the network infrastructure into areas and AS-s. Of course, the functionalities and protocols of an edge node are generally a subset of those of an LSR within the inner network (e.g. in terms of support for routing protocols).

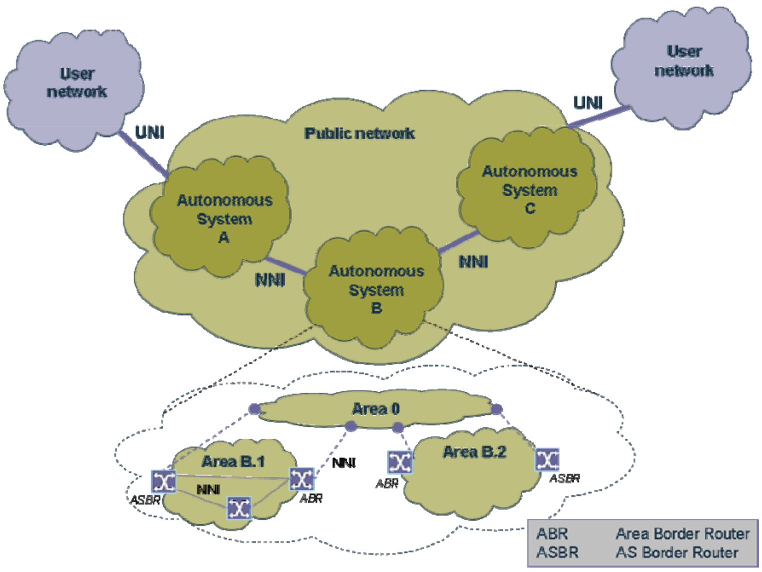

Figure 4-3: IETF GMPLS partitioning in Autonomous systems.

The GMPLS administrative partitioning relies on the IETF concept of Autonomous System managed by a single administrative authority (ref. Figure 4-3). According to the standardized GMPLS architecture [GMPLS-ARCH], each particular technology layer can be seen as a set of Autonomous Systems (AS) interconnected in an arbitrary way. AS can be sub-divided into different routing domains, each deploying an Interior Gateway Protocol for the dissemination of topological information. In case of OSPF, each routing domain can be further partitioned into areas. Intra-AS and inter-AS reachability information is disseminated among areas through the OSPF Backbone Area (Area 0) [OSPFv2].

In the IETF GMPLS multi-domain and multi-AS framework, a number of mechanisms are available for managing the connection setup, ranging from contiguous LSPs (the same LSP spanning all the domains) to stitched LSPs (more similar to the ASON call segmentation approach) and nested LSPs (obtained through the hierarchical LSP establishment procedures). All these mechanisms rely on proper extensions to the most popular GMPLS signalling protocol, i.e. G.RSVP-TE, thus implying the use of the same signalling protocol, often under a single end-to-end signalling session, across the different areas.

However, the main limitation of the AS-based partitioning approach is on routing side, because the inter-AS routing information exchange can be done via Exterior Gateway Protocols only (e.g. BGP-4) and GMPLS Traffic Engineering routing extensions are specified only for the intra-area scenario. This impacts the multi-domain constrained path computation process (necessarily loose and incomplete for the neighbouring domains to traverse, unless some bridging PCE means are used) and limits the applicability of recovery procedures and disjointness requirements.

In this framework, four different deployment models for GMPLS have been identified, namely the overlay, the peer (integrated), the augmented and the border peer model. The specific peculiarities of these models were not a core issue of the initial IETF CCAMP WG standardization activity, since the group was focused on developing consistent specifications for the I-NNI protocols in a single area and for a single switching capability. Consequently, their standardization is not complete and mostly a work in progress.

4.3.2.1 *GMPLS overlay model*

In the GMPLS overlay model [IETF-RFC4208], a GMPLS User to Network Interface is identified and strict separation of routing information between network layers is operated. As described in Figure 4-4, the topological view from an Area A is limited to the intra-area details, plus some reachability information about far-edge nodes, either statically configured (configuration-based reachability [GMPLS-ARCH]) or derived by some routing interactions (a kind on partial peering reachability [GMPLS-ARCH]). The core-nodes in Area B act as a closed system and the edge-nodes do not participate in the routing protocol instance that runs among the core nodes.

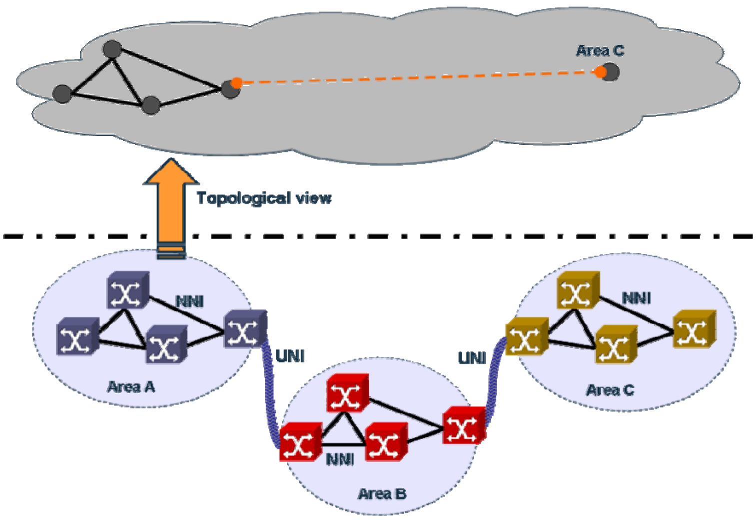

Figure 4-4: IETF GMPLS overlay model and related topological view.

IETF GMPLS UNI is based on mechanisms that are compatible with standard GMPLS signalling at the I-NNI [GMPLS-SIG, G.RSVP-TE], and just a single end-to-end RSVP session exists between the end-users including also the inner network signalling. The first and last hops constitute the UNI, and the RSVP session carries the user parameters end-to-end. It could be stated that this model avoids mapping (or piggybacking) of the user parameters to (in) the format expected by the I-NNI signalling protocol. Analogously, the GMPLS UNI model tends to limit the possible choices of the Service Provider, e.g. to use the same protocol or to let the signalling sessions go through its network and configure transport resources without a direct ownership from its ingress node to its egress node.

Concerning addressing, an edge-node and its attached core-node must share the same address space that is used by GMPLS to signal between the edge-nodes across the core network. When identifying an edge-node, two methods are used: a single IP address representing its Node-ID, or by one or more numbered TE links that connect the edge-node to the core-nodes.

An edge-node needs to know:

- Its own address
- An address of the adjacent core-node
- The address of any other edge-node to which it wishes to connect
- The addresses used in the overlay network island of which it is a part

A core-node needs to know (and track):

- The addresses on interfaces between that core-node and its attached edge-nodes
- The Node IDs of those edge-nodes
- The interface addresses and Node IDs of other edge-nodes to which an attached edge-node is permitted to connect

IETF UNI supports Explicit Route Object (ERO) processing and Explicit Label Control when setting up connections with Path messages, as further explained in [G.RSVP-TE]. It also supports Record Route Object (RRO) processing. ERO and RRO represent the major add-on with respect to OIF UNI Implementation Agreement, even if due to the limited topological knowledge about the core network they are useless for the purposes of source routing.

Moreover, IETF UNI defines two ways to delete connections:

- Alarm-Free Connection Deletion: an ADMIN_STATUS object must be sent in a Path or Resv message along the connection's path to inform all nodes en route to the intended deletion, prior to the actual deletion of the connection [RFC3473].
- Connection Deletion with PathErr: Using the Path_State_Removed flag to a PathErr message to indicate that the sender has removed all state associated with the LSP and does not need to see a PathTear [G.RSVP-TE].

Regarding security considerations, the trust model between the core and edge-nodes is different than the one described in [G.RSVP-TE], as the core is permitted to hide its topology from the edge-nodes, and the core is permitted to restrict the actions of edge-nodes by filtering out specific RSVP objects.

4.3.2.2 *GMPLS peer model*

In the IETF peer model, all the nodes run the same instance of GMPLS Control Plane, i.e. routing advertisements are distributed to the whole network and the nodes are expected to run the same signalling, even if they have different Switching Capabilities. This is the native GMPLS deployment model, whose scalability is critically threatened by the size of the controlled network and the coexistence of different switching technologies (Multi Region Network – MRN) or layering within a single switching capability (Multi Layer Network – MLN, as in case of High Order VC and Low Order VC in TDM networks).

Although the IETF peer model could be applied to cross administrative boundaries, it is mainly useful to handle the problem of controlling nodes with different switching granularities or capabilities within a single administrative domain (i.e. respectively, Multi-Layer or Multi-Region Networks). The application of the peer model to service provisioning across different administrative domains raises a number of issues in terms of service policy control (integration with policy-based authorization), service partitioning (i.e. integration with accounting, etc.), security (e.g. integration with authentication functions), and partitioning of topology and TE info between networks belonging to different operators.

In such a framework, which matches the requirements of many commercial and research networks, the basic construct is the Forwarding Adjacency LSP (FA-LSP). It is an LSP created either statically or dynamically by one instance of the Control Plane (at a given layer/region) and advertised as a TE link into the same instance of the Control Plane (for the use by upper layers or neighbouring regions). The topological view in a peer model is obtained by a mix of TE-link with different switching capabilities descriptors and dynamical virtual TE-links bound to FA-LSPs.

Signaling in the peer model may result in contiguous LSPs, stitched LSPs, or nested LSPs. From a signaling perspective, there are two alternatives to establish the lower layer FA LSP: static (pre-provisioned) and dynamic (triggered). Pre-provisioned FA-LSP will be initiated either by the operator or automatically using features like TE auto-mesh [GMPLS-MESH]. If such a lower layer LSP does not already exist, the LSP may be established dynamically upon reception of an upper layer signalling request.

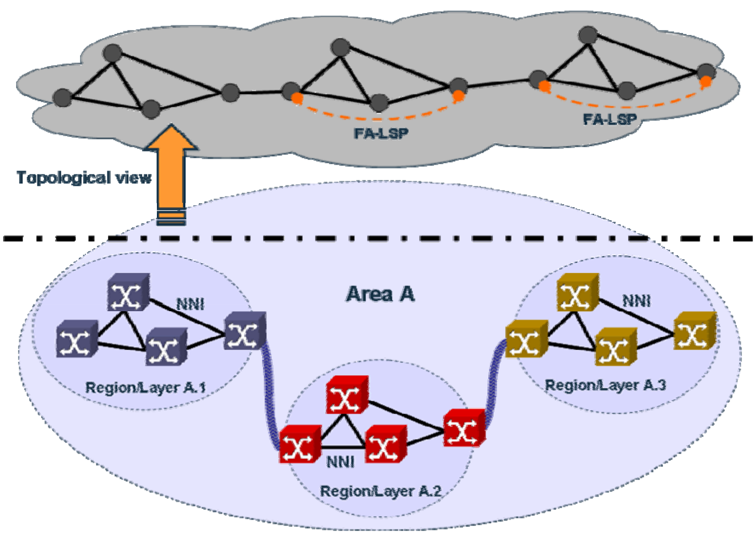

Figure 4-5: IETF GMPLS peer model and related topological view.

4.3.2.3 *GMPLS augmented model*

The augmented model allows limited routing exchange from the lower layer network to the higher layer network. This assumes that the border nodes on a region/layer provide some form of filtering, mapping or aggregation of routing information advertised from the lower layer network. The augmented model could be considered a reduced instantiation of the peer model, from which it inherits the main routing and signalling procedures. There is no formal specification of the augmented model in IETF GMPLS documents and, though it promises higher scalability performances due to the reduced/aggregated topology information (ref. Figure 4-6), it has no working implementation.

4.3.2.4 *GMPLS border peer model*

The border peer model is defined in [S-GMPLS] and is a modification of the augmented model. The border routers have visibility into both layers/regions they participate, but neither any routing information is exchanged between their routing protocol instances, nor the routing databases are flooded out of their native region/layer (ref. Figure 4-6). Signalling interworking is obtained as in case of the peer model. The border routers act as gatekeepers between the two routing instances and enable a segmented administrative boundary that helps ensuring management separation between the network layers/regions. This model is really part of a proprietary implementation by CISCO, i.e. the S-GMPLS, and is part of an ongoing standardization effort within CCAMP with the scope limited to MPLS-TE and GMPLS interworking.

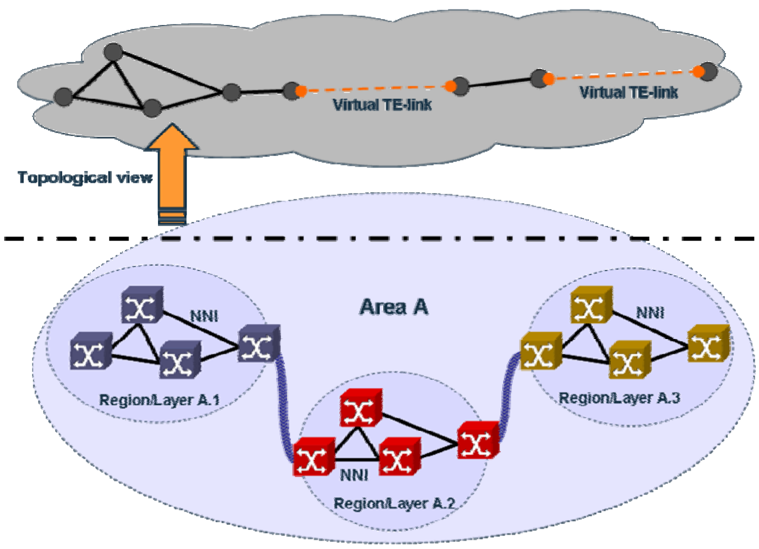

Figure 4-6: IETF GMPLS augmented model and related topological view.

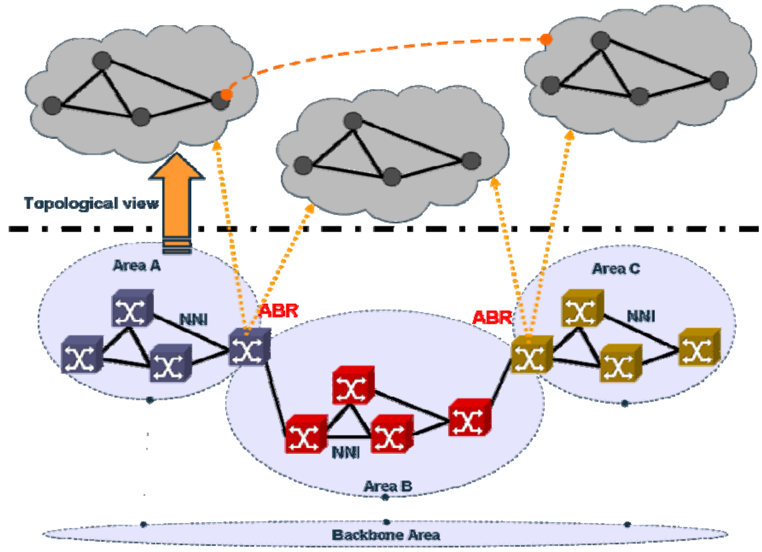

Figure 4-7: IETF GMPLS border peer model and related topological view.

4.3.3 Possible deployment options for NRENs

National Research and Education Networks (like many commercial infrastructures) that are willing to migrate under GMPLS control their infrastructures pose nowadays two possibly conflicting requirements:

- To preserve a strict administrative partitioning with respect to neighbouring domains, by controlling the service demarcation points (for maintenance and accountability)
- To support a structured infrastructure based on multiple and different switching technologies (e.g. L2SC + LSC, or L2SC + TDM + WDM, or PSC + LSC/FSC, etc.), in order to extend the automated Control Plane procedures from the inner core towards the access.

The first requirement is fully matched by the ASON approach. However, it would be prone to a proliferation of network reference points (nesting of UNIs) in case the generation of a new domain was bound to either technologies and physical location of the equipments and per-vendor groupings and commercial agreements with peering service providers.

The latter requirement is fully matched by the IETF GMPLS peer model in Multi-Region / Multi-Layer Networks. However, it would be inadequate for the deployment in inter-carrier scenarios (e.g. interfacing towards or across GÉANT2) in which Authentication and Authorization Infrastructures could exist and integration with them on a per-carrier base could be needed.

It appears that a mixed solution could achieve the best performances, i.e. by providing

- an ASON UNI (OIF UNI) for final NREN users,
- an ASON E-NNI (OIF E-NNI extended for inter-carrier) for the interworking with neighbouring carriers
- GMPLS MLN/MRN procedures in within at the I-NNIs.

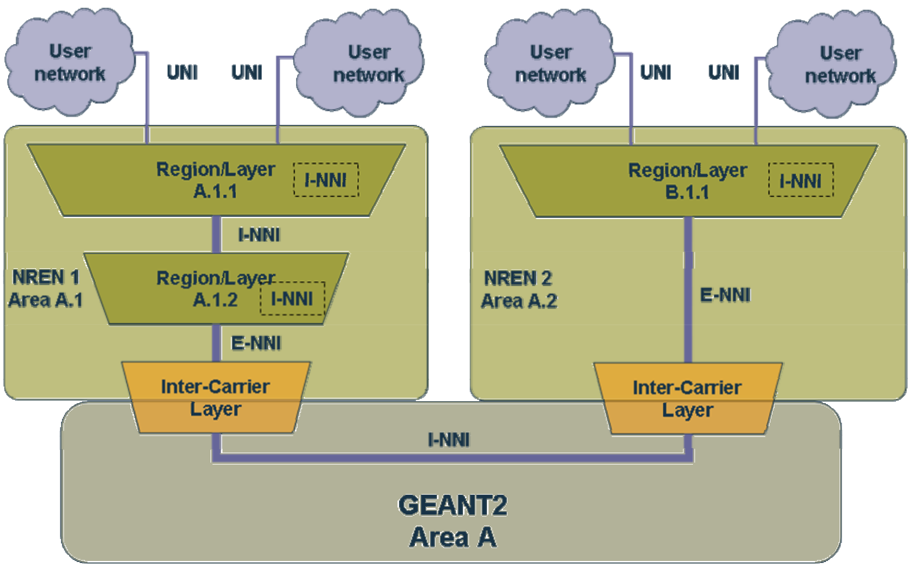

Figure 4-8: Possible ASON/GMPLS deployment in NRENs and through GÉANT2.

4.4 Choice of deployed ASON/GMPLS procedures

This chapter is aimed to provide a brief overview of main procedures provided by the ASON/GMPLS Network Control Plane.

4.4.1 Recovery

The concept of recovery refers to the required actions taken to maintain or retrieve the correct functioning of the network upon the detection of a failure. These actions consist in notification, fault isolation, repair of the failure and reconfiguration using survivability mechanisms (protection and restoration). In protection, the recovery paths are pre-planned and fully signalled before a failure occurs. For a restoration scheme, the recovery paths can be either pre-planned or dynamically allocated, but additional signalling is needed to establish the restoration path upon failure.

SCN involvement in failure detection

Mainly, two possibilities exist when deploying the SCN, depending on whether the Control Plane connectivity is implemented in-fiber or out-of-fiber. While in-fiber control channel failures could be detected by means of Loss of Light (LoL), in out-of-fiber configurations, more attractive since signalling communication continues alive upon a TP failure, other mechanisms must be provided. GMPLS Link Management Protocol (LMP), intended to maintain node resources, provides control channel connectivity maintenance functionalities by means of a fast keep-alive Hello protocol, whose parameters are negotiated in the control channel establishment phase. Additionally, LMP protocol can provide link property correlation, link verification and fault isolation functionalities, so being its implementation of critical importance.

If an out-of-fiber approach is chosen, different and specific recovery mechanisms have to be applied for the transport plane and the Control Plane.

Choice of Transport Plane recovery

Particularly, different mechanisms for protection and restoration are suitable depending on the situation:

Procedures	Pros & Cons	Schemes	Where & Why to use it?
Protection	Uses more resources than restoration Faster Protection switch time more constant than the restoration one Any kind of protection schemes guarantees recovery in the 50ms bound + availability of forced protection commands	1+1	• Rather expensive: requires resource duplication • Used for specific services that need a very high availability
		1:1	• Slower than 1+1 • Can be used to carry low-priority extra traffic in normal operation, if traffic pre-emption is allowed • Packet networks can pre-establish a protection path for later use with pre-planned but not pre-reserved capacity If no packets are sent onto a protection path, then no bandwidth is consumed
		1:n	• When multiple working entities have failed simultaneously, only one of them can be restored by the common protection entity, unless a different pre-emptive priority has been assigned to each working entity • Can be used to carry perceptible traffic in normal operation
		M:N	• Improves system availability with small cost increases • It has rarely been implemented or standardized
Restoration	More resource efficient than protection schemes, as no resources are committed until after the fault occurs and the location of the fault is known	Pre-Planned	• Not so much optimal in terms of routing (path may not be the shortest, signaling is used along the restoration path before the failure, etc), but faster and simpler when restoring

	But slower, since more tasks must be done following the detection of a fault Restoration time more variable than the protection switch time, hence it is essential to use restoration priority to ensure that service objectives are met cost-effectively	Shared-Mesh	• A case of Pre-planned LSP restoration, where the pre-planning of protection LSPs can also include resources already planned for other protection LSPs
		On-the-fly	• More optimal (the path is computed at the moment, no restoration bandwidth reservation, no signalling before the failure, etc), but the more time you take when computing the path, the more traffic you loose

Table 4-1: Recovery schemes

Restoration schemes are superior in terms of capacity efficiency compared to protection schemes, but the implementation of a mesh restoration scheme is quite complex and requires sophisticated algorithms. Restoration is usually also slower than protection. With restoration, capacity in excess of the working capacity needed to support the normal working traffic is provided in the network. This spare capacity, which is shared among the various working connections, will be used to recover from failures.

Recovery schemes in general can also be classified depending on the segment to be recovered.

- Span-level recovery
- Segment-level recovery
- Path-level recovery

Segment restoration requires typically more capacity than path restoration, because of the often suboptimal routes found with link restoration and because path restoration has a larger view on the network. Path restoration will typically distribute the backup routes of the affected connections over a larger part of the network than link restoration, allowing more opportunities to optimize the spare capacity needed in the network. Thus, with link restoration the route computation process is easier and the number of required switching actions is limited compared to path restoration (often resulting in a lower recovery time), but the capacity efficiency is lower.

In a deployment-oriented perspective, on-the-fly end-to-end restorations are more optimal in terms of network resource consumptions, though prone to experiment more traffic loss until a new path is established. For this purpose, pre-planned end-to-end restorations result in faster failure mitigations, though sub-optimal in terms of resource usage in the network. However, restoration schemes are always slower than any kind of protection scheme, for which the 50ms bound can be guaranteed nearly in any kind of meshed network. Discussion on possible escalation strategies (from protection to restoration) are implementation specific but they could be cited at least.

Choice of Control Plane recovery

The reliability of the Control Plane emerges as a key issue in next-generation optical networks, becoming the CP of crucial importance for the proper operation of the whole network. It is noteworthy that not only signalling and routing, but also resource management information and even some transport plane (TP) failure recovery mechanisms are transmitted through the CP.

SCN recovery typically relies on the neighbour liveliness detection by the routing protocol instance running in the DCN. In case of OSPF IGP, this is done by exchanging Hello messages between adjacent nodes. The mechanism is the same applied in IP networks and implies a loss of CP communication for periods in the range of 30 seconds. This would be unacceptable in a GMPLS environment, in which the occurrence of a failure in the SCN could derive from a Transport Plane failure and, thus, in that timeframe the Control Plane could be in the maximum of its activity for connection recovery. For this purpose, GMPLS Control Plane takes the benefit of the LMP functionality for Control Channel maintenance. In fact, LMP senses the status of its configured CCs and upon the occurrence of a CC down event, it associate a new CC (if any) to the affected TE-Links, thus negotiating what other CC(s) to use when sending CP traffic to a TN adjacent node. Therefore, the GMPLS CP mechanisms for SCN failure mitigation results more effective than those (IGP-based) of the standard IP networks, though relying on a preliminary planning of the SCN and CCs that can guarantee an adequate survivability of the SCN communications.

4.4.2 Crankback

Crankback is a signalling procedure aimed to improve the likelihood of LSP setup success. When crankback is applied, some setup failure information is carried back from the point of failure up to the repair point (e.g. ingress LSR), in order to facilitate new setup attempts. These attempts are now carried out with a finer knowledge of the blocked resources. It was originally introduced in ATM signalling [PNNI], and is also known as TE-LSP re-route in the MPLS community. When a signalling protocol attempting to set up a path on a calculated route encounters a problem, the path set-up attempt is rejected, and the reason and location of the error are then reported to the source node in order a subsequent second path set-up attempt to take place along a new route that is calculated taking into consideration the reported error.

The use of crankback in ASON/GMPLS architecture can improve significantly the setup process, in cases where one or more nodes along the path do not support label swapping (e.g. non Time Slot Interchange in TDM MS-SPRING, no wavelength converters in LSC, etc.) and, thus, setup requests are likely to be blocked more often than in a conventional MPLS environment. Furthermore, in this kind of networks, segment protection is much harder to achieve since a particular label may be already used in other links and therefore end-to-end protection is more adequate as it offers the choice of use of another wavelength. This requirement makes crankback rerouting particularly useful in an ASON/GMPLS network, in particular in dynamic LSP re-routing cases (no protecting LSP pre-establishment). Also in cases where both the working and protecting paths fail, crankback is necessary to re-establish the LSP on an end-to-end basis avoiding the known failures. Another case where crankback seems required is where many links or node failures occur at the same time. This condition might cause a significant divergence between the real network (topology and TE) status and the vision of it as perceived by any node in the network (and advertised by the routing protocol). Consequently, the source path calculated by a recovering node might, therefore, be based on outdated, inaccurate information, and might end with involving some failure points.

In general, the convergence of topology information (from which the ingress LSR could infer the re-routing conditions) using routing protocols is typically slower than the expected LSP setup times. Thanks to the crankback scheme, the information about failures may be applied to compute, more efficiently, an alternative path for the failed LSP. This information needs to indicate a number of “coordinates” that describe the blocking point, e.g. the upstream / downstream, incoming / outgoing TE link or link component, the node address, and other resources. PathErr or Notify messages are issued to this purpose, with an extended ERROR_SPEC object, from the point of failure to the ingress node or to a selected node (upstream) and also ResvErr messages are returned to the egress node (downstream) indicating “Admission Control failure” when resource reservation fails as the Resv is processed. Ideally, the repair node will be given also the reason for the failure since knowing the cause of the failure will lead to the proper corrective action i.e. “No route to Destination” will trigger a new path computation excluding the reporting LSR, “Temporary Control Plane Congestion” a simple retry after a suitable pause.

A number of expected crankback signalling properties are described in [IETF-RFC 4139]:

- Error information persistence: the entity that computes the alternate (re-routing) path should store the identifiers of the blocking resources, as indicated in the error message, until the connection is successfully established or until the node abandons rerouting attempts. Since crankback may happen more than once while establishing a specific connection, the history of all experienced blockages for this connection should be maintained (at least until the routing protocol updates the state of this information) to perform an accurate path computation that will avoid all blockages.
- Rerouting attempts limitation: to prevent an endless repetition of connection setup attempts (using crankback information), the number of retries should be strictly limited. The maximum number of crankback rerouting attempts allowed may be limited per connection or per node.

Controlling Crankback Procedure

Local policies may be applied to the repair point that receives crankback information to govern how it attempts repair of the LSP. For example it may prioritize repair attempts between multiple LSPs that have failed, and it may consider LSPs that have been locally repaired to be less urgent candidates for end-to-end repair.

When a request is made, the ingress LSR should specify whether it wants crankback information to be collected in the event of a failure and whether it requests re-routing attempts by any or specific intermediate nodes. Three categories of re-routing attempts are recognized according to the network topology:

In a flat network without partitioning of the routing topology, when the ingress LSR receives the error message it computes an alternate path around the blocked link or node to satisfy the QoS guarantees using link state information about the network.

On the other hand in a network partitioned into areas such as with OSPF, the area border LSP may intercept and terminate the error response, and perform alternate routing within the downstream area (Figure 4-9).

In a third scenario, any node within an area may act as a repair point.

For this purpose a re-routing flag field is added to the protocol setup request message with the following mutually exclusive values: No re-routing, end-to-end re-routing, boundary re-routing and segment-based re-routing.

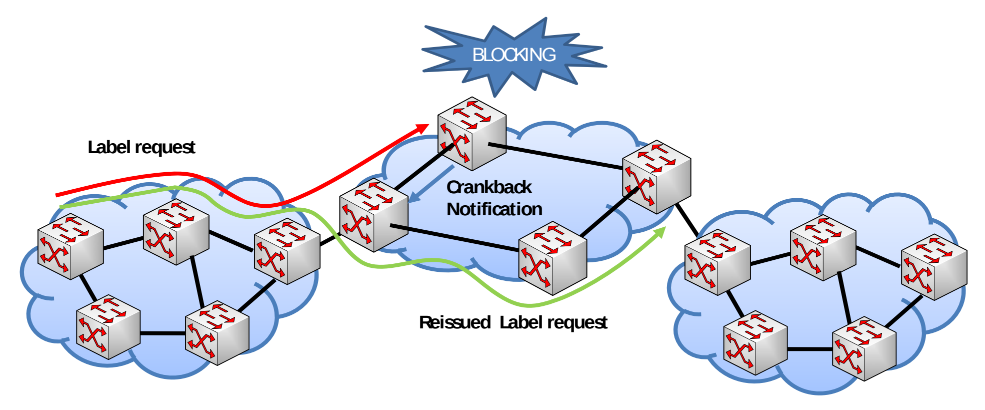

Figure 4-9: Crankback routing.

4.4.3 Claiming of existing circuits under the ownership of the Management Plane to the ownership of the GMPLS Control Plane

An interesting deployment scenario for GMPLS is the migration under the Control Plane responsibility of an operational Transport Plane previously fully configured by Management Plane. In this case the Network Operator could decide to let the GMPLS Control Plane operate just on a disjoint partition of the Transport Plane resources with respect to those just configured (i.e. GMPLS operates just on the spare free resources); or, more valuable, it could decide to move under the GMPLS control the permanent connections previously established by the Management Plane and transform them in SPCs. The rationale behind such a migration of resources is mainly the possibility to add restoration capabilities to those circuits. Nevertheless, the procedure could be used also for a gradual adoption of full-fledged Control Plane features and a progressive, traffic conservative, shut off of the conflicting Management Plane functionalities. A number of requirements have been defined in [GMPLS-PCSC] for the seamless and not-disruptive migration/adoption of resources configured by the Management Plane to the GMPLS Control Plane. Just to cite the relevant ones:

- Transport Plane consistency, i.e. the network resource used must be the same and GMPLS cannot apply route deviations, crankback, or just select different labels[1].
- non-disruption of user traffic, i.e. the procedures must involve just Control Plane entities and the configured cross-connections

[1] The selection of the same labels during connection migration is applicable only to those switching technologies in which a 1:1 relationship can be established between the Transport Plane resource representation by the Management Plane and that by the Control Plane. For example, in case timeslots in a SDH/SONET frame (TDM), wavelengths in a DWDM link (LSC), fibers in a photonic switch (FSC).

- synchronization among GMPLS nodes involved in the conversion process,
- roll-back of ownership in case of transfer failure.

All these requirements need to translate in proper extensions for the GMPLS signalling protocols, which are generally proprietary solutions of a GMPLS stack vendor since there is no standardization within IETF yet. In fact, the standardization effort towards the CCAMP WG draft document [GMPLS-PCSC] is currently limited to the requirements gathering phase and there is no wide consensus in that group on the generalized applicability of the proposed approach also in legacy networks non-NMS based (e.g. IP-based). However, the need for such a feature is one of the strongest requirements of those network operators that are willing to move their TDM and WDM infrastructures under GMPLS control. And this is one of the main applicability scenarios of current GMPLS operational implementations all around the world.

4.5 Integration with existing AAA and accounting mechanisms

AAA systems are related to security and resource abuse issues, which are one of the most important for production systems. Due to high complexity of solutions and approaches to tackle with Authorization and Authentication, it is recommended and efficient that PHOSPHORUS will integrate already available AAA systems. A common set of AAA system examples include:

- Shibboleth – http://shibboleth.internet2.edu
- PAPI – http://papi.rediris.es
- RADIUS based systems (below list includes open source software only):
  - FreeRADIUS – http://www.freeradius.org
  - GNU Radius – http://www.gnu.org/software/radius/radius.html
  - JRadius – http://jradius.sourceforge.net
  - OpenRADIUS – http://www.openradius.net
  - Cistron RADIUS – http://www.radius.cistron.nl
  - BSDRadius – http://www.bsdradius.org

In PHOSPHORUS project it is not decided yet which AAA system should be used. The system requirements are being specified for purposes of distributed environment and strongly depend on final architecture. Also more than one of the above systems can be used, with respect to deployment needs (e.g. local administrators, policy, etc.).

Most NRENs and network operators have already an AAA system implemented. In order to deploy a PHOSPHORUS GMPLS implementation in such environments, there is a need to adjust to those AAA systems and integrate authorization, authentication and accounting processes. It is expected that limited support and trust from administrators will be faced during integration, as the running AAA systems are very critical for network maintenance. Also the existing systems may not include all users (humans, organizations, or projects) that will be required for GMPLS usage. This issues needs to be solved individually with network and security personnel of particular NREN or organization.

Although AAA performance in single domain environment is a simple problem to solve, the solutions are not easily scalable for multi-domain infrastructure. The main problem is related to large number of users and permissions data which has to be shared between domains. There are already a number of solutions, which are based on federations of domain. This approach allows “single sign-on”, where user is authenticated once only, and domains can easily exchange AA information between each other. An example of such solution is work done in GÉANT2 project for AAI and EduGAIN infrastructure.

Since user is authenticated and his request is authorized by local AAA system, the information should be trusted by all GMPLS enabled nodes within single domain and in multiple domains depending on peering and policy. The easiest and most efficient way to implement this is to use [IETF-RFC 2750]. This standard extends RSVP message with policy handling information, which can be also used for transferring security, authentication and authorization data, or simply a token. The modification of the objects is prevented by integrity check, based on MD5 checksum calculation (according to [IETF-RFC2747]).

4.5.1 SLA operation and maintenance

SLA (Service Layer Agreement) as a part of AAA system is generally an administrative contract and does not contain technical service specification. The technical details like QoS requirements or traffic characteristic are described through SLS (Service Level Specification) therefore it is an instrument which can be use to support the procedure of making an advance reservations in GMPLS network.

The first step to improve SLAs is to specify a catalogue of accessible services and the resources access policies. The scope of available network resources and QoS parameters should be well defined and presented for users with a useful interface to support on-demand requests.

To provide an automatic process of negotiations and configurations network resources using SLA mechanism some transactional mediation system is required. Many parallel SLA requests that refer to the same resources in the same time can occur and the system should assure reliable realization of the transactions and it should be able to schedule such requests.

When the user specifies required resources, the SLS request record is sent to the SLA management system and the negotiation process is started. If it is a local, intra-domain request and all resources are available and there are no conflicts, the SLA is accepted. The SLS is translated into system-level reservation configuration format and it is stored in proper databases for advance reservations and devices can be configured. In case of inter-domain request, the SLA management system splits the SLS in two parts, one for the local domain and the other for the remaining domain. The latter part is forwarded over inter-domain interface to the next peer domain. It is assumed that each domain has its own SLA management system and the communication protocol between the systems is defined.

The SLA maintenance system in each single domain should have the capability to communicate with, or even should be a part of the advance reservations management system and should have access to the reservations calendar and SLA inventory.

In order to preserve SLA guaranteed user requirements monitoring system and capability of verification network parameters are also required.

4.5.2 Accounting mechanisms

On-demand resource provisioning in GMPLS introduces significant changes in the service management. In the traditional process structure, the customer has to negotiate the terms and conditions of the offer with the sales department of the particular network provider. The network provider must check whether the connection requests can be handled by the standard mechanism and infrastructure. In some case, for example, there is a need to order missing equipment causing additional effort and delay. The connection projecting results then define the price calculation, as well as the delay necessary to set up the service. After the contract has been accepted the service delivery process starts. The sales department handles the contract administration, and then the order is split into work packages according to the network domains involved. After providing the connection, and end-to-end test is conducted and customer care, accounting and alarm management are activated. Finally, a delivery report is issued by the sales department to the customer. The service is active and ready to be used by the customer. The accounting data are generated periodically by Administration and sent to the customer. In the traditional process structure, which is widely deployed, there is no relation between Control Plane and accounting mechanism. Some times there is no Control Plane in the operator's network.

The impact of GMPLS on the operational processes is related with the automation of some processes which significantly reduce costs and delays in service provisioning. However, for dynamic services, it is much more difficult to correctly assign costs to customer accounts. Calculating a new price for a new service is more expensive than just applying a traditional pricing scheme. Since the service delivery will now be automated, correct agreements and relations have to be negotiated by the sales department, and implemented well before in the form of Service Level Agreements (SLAs). After SLA framework has been set up, the service delivery process can be simplified due to GMPLS UNI interfaces. External signaling is forwarded to the call control that converts it into RSVP signaling for each domain. Manual intervention is necessary to set up the connection completely only if no positive responses were received. After database update, customer care is informed, and accounting and alarm management are activated. In the GMPLS on-demand provisioning the cessation request is also received via the UNI. The cease process then triggers the sales department to assess the cessation request and trigger the accounting and confirmation of cessation to the client.

The Control Plane is connected with the accounting system via Call Control module. The accounting system is informed when a connection is set up, when connection parameters are changing and when it is released which is related with receiving proper RSVP RESV messages by the Call Control module. Call Control informs the accounting system about connection initiator (to find proper customer account) and connection parameters (connection technology, bandwidth parameters, QoS parameters, affected domains, etc).

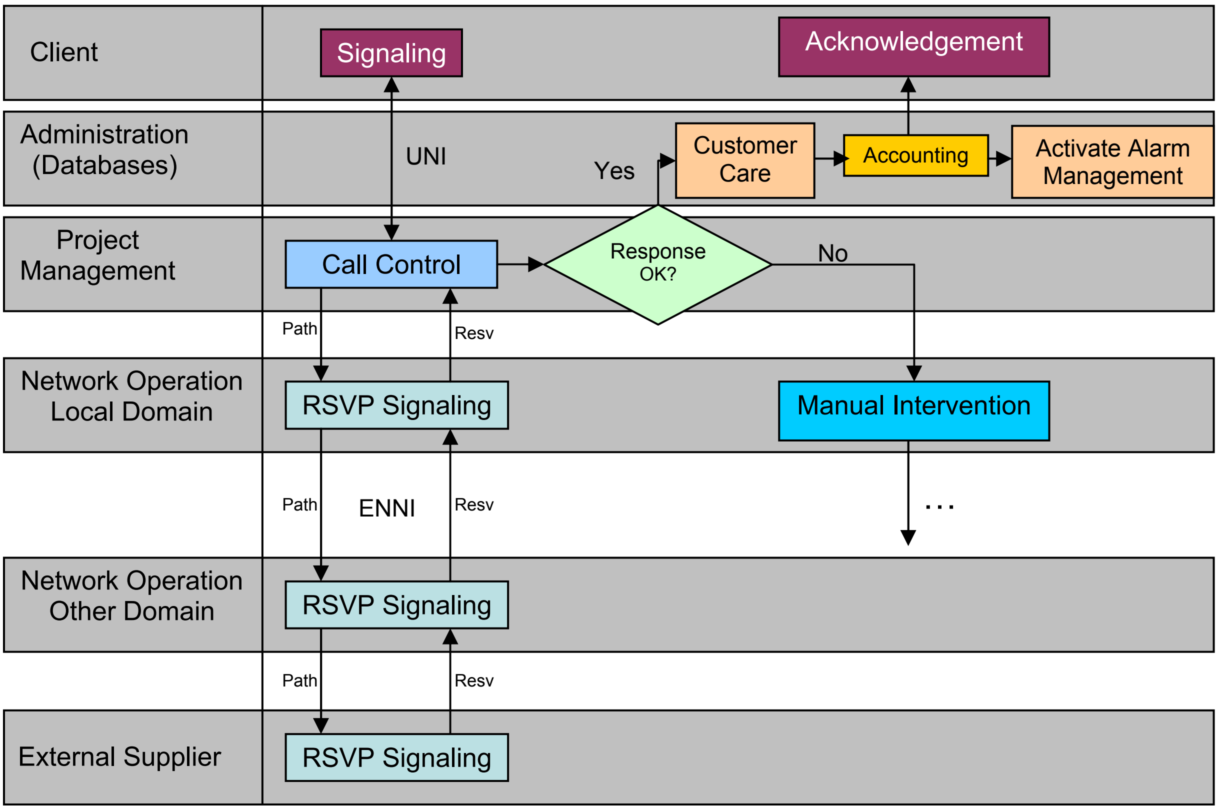

Figure 4-10: Service provisioning with GMPLS.

4.6 Integration of ASON/GMPLS management in the existing NMS

Prior to the advent of ASON/GMPLS, the circuit switched transport networks exhibited highly static behaviour and were managed and controlled through centralized management systems. Most of their limitations were due to the fact that they were operated manually or via complex and slow Network Management Systems. The major issues associated with centrally operated optical networks based on traditional legacy management systems and interfaces can be identified as follows:

- manual error-prone provisioning
- long provisioning times
- inefficient resource utilization
- complex network management
- difficult interoperability between multi-operator networks
- lack of protection in mesh-type optical networks.

ASON/GMPLS opens the path to a much more dynamic transport network. In particular, switched connections (SCs) introduce a radically different approach for connection provisioning, where the Management Plane (MP) may not be directly involved during the connection set-up phase. Following a signalled client request, the Control Plane (CP) alone takes care of all the path computation and the signalling which sets up the connection.

ASON/GMPLS allows certain functionalities to be implemented in a distributed fashion (each network element can make autonomous decisions, independently and without the intervention of the MP), and thus alleviates the MP of resource processing tasks. However, this approach does not imply that certain functions could not be implemented in a centralized manner but at the expense of the speed and scalability of the specific functionalities. For example, when centralized routing is implemented, restoration might take longer to get resolved due to the communication delay of the required messages between the centralized routing process and the involved network elements. In an attempt to overcome these problems a number of functions should be ideally decentralized and allocated to the Control Plane (e.g. call and connection admission control, path computation, etc.)

By introducing CP-based solutions MP-CP interworking becomes an important issue both for managing the CP as well as for the management and control of Transport Plane (TP) resources via the CP. Some requirements to ensure management and Control Plane interaction are:

- For intra-domain topology discovery, the MP must be capable of receiving notifications from the CP for the discovery of the addition or removal of a node or a link (using CP based neighbour discovery functions).
- For inter-domain topology discovery, the MP must support the ability to receive notifications from the CP for the discovery of the addition, removal, or properties changes for an inter-domain link that was created (using CP-based neighbour discovery functions).
- The MP must support the capability to query the CP for topological information that was learned in the discovery process.
- The MP must support the ability to receive notifications from the CP regarding the results of a successful service discovery procedure including service attributes for the UNI-C and UNI-N ports.
- The MP must be able to query the current status of a call which is related to the status of each related connection and the CP should be able to receive and respond to such call status queries from the MP.
- The CP should be able to receive call release requests from the MP.
- Upon the completion of a call release, the CP shall inform the MP of the outcome of the release operation, i.e. success or failure indication.
- The MP must support the ability to invoke a call modification request that includes all parameters necessary for the CP to complete the call modification according to the MP intent.
- The management system must receive sufficient information from the CP to allow it to do centralized path selection by computing an explicit route (loose or strict).
- The MP must support the ability to invoke the set up of soft-permanent connections. The set up request shall include all parameters necessary for the CP to complete its part of connection set up according to the MP intent. The CP shall be able to receive and respond to SPC set up requests from the MP.

- The MP must support the ability to invoke the set up of a connection using constraint-based routing. The CP in this case performs the final route calculation according to the conditions specified in the MP and after the connection set up the MP can query the actual route of the connection.
- Any failure within the CP should raise an alarm to the MP and should be reported in an event database. Recovery from the failure should withdraw the alarm and should be reported in the event database.
- The MP should be notified for a broken link through the CP which is able to detect the failure from the originating node of that LSP

4.6.1 Monitoring of ASON/GMPLS calls and connections

In ASON/GMPLS, functionalities are distributed in different planes:

- The Management Plane supports the capabilities of fault management, configuration management (including resource allocation/de-allocation), performance management, security management and accounting
- The Transport Plane supports capabilities of payload transport, performance monitoring, fault detection and protection switch.
- The Control Plane supports the capabilities of dynamic path computation, dynamic distributed call and connection setup/release and dynamic protection/restoration.

The capability of supervising and managing calls that are set up and released via distributed ASON/GMPLS mechanisms is provided by the Call Controller (CallC), the Connection Controller (CC) and the Link Resource Manager (LRM) components through communication with the MP. To support basic call capability and the logical call and connection separation in the scope of ASON, a call identifier is introduced to various message sets.

The separation between the call and connection control is one of the most service-related architectural functions of the CP. Call control embeds the association between a customer and the network through the end-to-end service session negotiation, while connection control corresponds to the administrative and Traffic Engineering validation of the connection request at individual NEs involved in the connection along the path. Call control allows bearing reciprocal relationship between service requests and network connection capabilities, configures service adaptation at the network ingress and egress points, and contributes, thereby, to service-fulfilment, assurance, and accounting. Therefore, connections depend now on the existence of the associated call session, and may be created, updated, and deleted within that call session using the same signaling protocol for both. Figure 4-11 shows a visualization of the separation between the call and connection separation in network, between a user and the network, and between two domains.

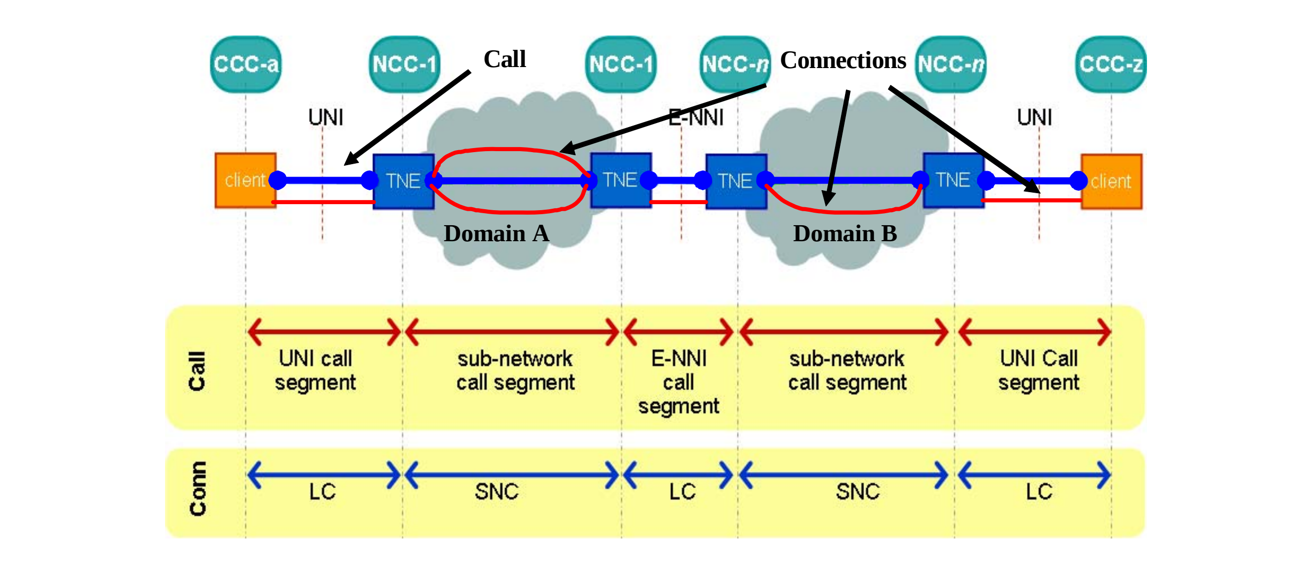

Figure 4-11: Call and connection separation in ASON CP.

A call is defined as “An association between endpoints that supports an instance of a service” [ASON-ARCH]. Multiple connections may be associated to a call. In Figure 4-11 this is illustrated in the connection control domain, where two sub-network connections have been established in Domain A to support the call. The call concept provides an abstract relationship between two users, where this relationship describes (or verifies) to what extent the users are willing to offer (or to accept) service to each other. Therefore, a call does not provide the actual connectivity for transmitting user traffic, but only builds a relationship by which subsequent connections may be made.

A property of a call is to contain zero, one or multiple connections. Within the same call, each connection may exist independently of other connections, i.e. each connection is set up and released with separate signaling messages. The call concept allows for a better flexibility in how end-points set up connections and how networks offer services to users. In essence, the separation of call and connections offers the following advantages:

- Each connection can travel on different diverse paths (also known as virtual concatenation), potentially making better usage of the available network resources.
- An upgrade strategy for CP operations, where a call control component (service provisioning) may be separated from the actual nodes hosting the connections (where the connection control component may reside).
- Identification of the call initiator prior to connection, which may result in decreasing contention during resource reservation.
- General treatment of multiple connections that may be associated for several purposes, e.g. a pair of working and recovery connections belonging to the same call.

A general description of the signaling flow for the setup procedure of SPCs and SCs is given according to [ASON-RSVP]:

- A Path message is sent from the source to the destination to request a connection.
- Upon reception of the Path message by the destination node, an RSVP session is set up between the source and destination.
- The destination node responds to the Path message via one of two messages sent in the upstream direction:
  - Resv (for normal setup response); or
  - PathErr (for error in the setup procedure); in this case the connection is not set up. If the Path state is not removed, then an explicit PathTear is needed to remove any extraneous states.
- Upon reception of the Resv message by the source node, an optional ResvConf message may be sent. This is dependent on the RESV_CONFIRM object within the Resv message.

Figure 4-12 [ASON-RSVP-SIG] shows an example signal flow for SPC request. For a SPC connection, it is assumed that the user-to-network link connection is provisioned i.e. it is setup by means beyond the distributed CP, and information is provided to the CP regarding the identity of the link connection. Setting up the switched portion of the SPC connection remains the same as that for setting up a switched connection.

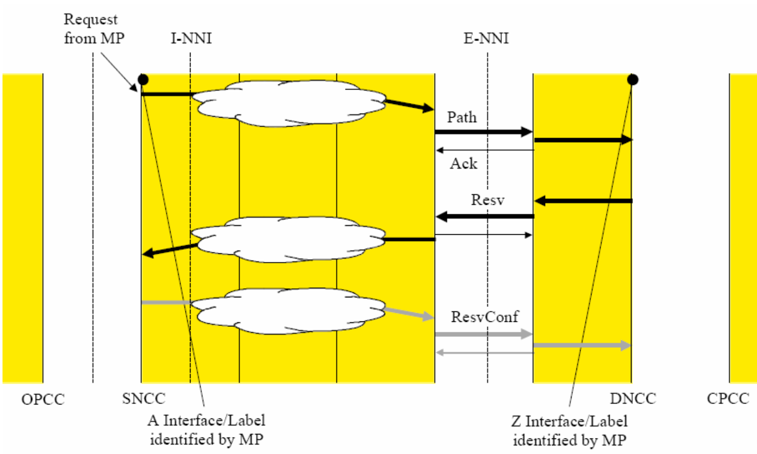

Figure 4-12: Soft Permanent Connection Setup – ref. recommendation ITU-T G.7713.2.

Figure 4-13 [ASON-RSVP-SIG] illustrates the setup of the SC. To set up a SC call, a user initiates the request across the UNI interface. The request is propagated across the network to the destination user. Upon verification/acceptance of the request, a positive indication is sent to the source user. Optionally, the source user may also transmit a final response. This third phase message is introduced to support explicit destination notification of completed connection setup.

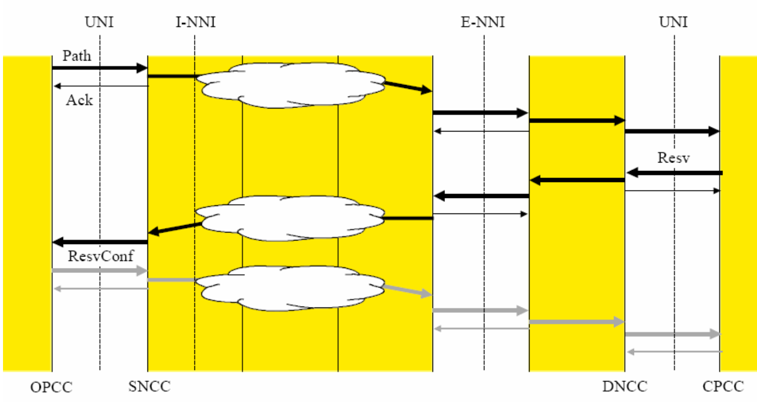

Figure 4-13: Switched Connection Setup – ref. recommendation ITU-T G.7713.2.

For the case of failures that impact the control or the data plane the CP needs to support appropriate behaviours for recovery. Some general guidelines for failure handling include:

- A Control Plane node should provide capability for persistent storage of call and connection state information in order to allow each CP node to recover the states of calls and connections after recovery. If connection/call states cannot be recovered, the Control Plane node may communicate with an external component (which may include neighbour Control Plane nodes or a persistent storage provided by the management plane) to attempt state information recovery.
- Control plane failures are notified to the management plane. The management plane may direct the Control Plane to take certain actions due to the failure. These actions may include entering into a self-refresh state, cleaning up of partial connections, release of certain connections, or other protocol-specific actions for state maintenance and recovery.
- A Control Plane node notifies the MP of the inability to recover (subset of) relevant information (e.g., inability to synchronize state of connections). Then the MP may respond with the following actions
  - Release the impacted connections.
  - Retain the impacted connections. In this case, a connection may remain non-synchronized from the Control Plane perspective; however, the connection may remain valid.

5 Deployment modes for the G^{2}MPLS Control Plane

As G^{2}MPLS is a superset of GMPLS, part of the consideration about the deployment modes for GMPLS also apply here. This section will only include considerations that are specific to G^{2}MPLS, by highlighting the main extension and improvements to the Control Plane architecture that can provide the seamless and one step reservation of both Grid and network resources.

5.1 Resource partitioning and virtualization for G^{2}MPLS

G^{2}MPLS is an enhancement of the ASON/GMPLS Control Plane architecture that implements the concept of Grid Network Services (GNS). As such, it inherits from the GMPLS Network Control Plane the rights for full usage of the Transport Plane resources assigned to its ownership.

Management Plane entities possibly competing for TN resources with G^{2}MPLS could be pre-existing NMS-es or, more specific in the PHOSPHORUS and Grids scope, some Network Resource Provisioning Systems (e.g. UCLP, D-RAC, ARGON). The same limitations on the activity of Management Plane on Control Plane resources described for GMPLS apply to G^{2}MPLS. Therefore, the Management Plane can modify the resources exported by the Control Plane and manage them as virtual resources, but according to the G^{2}MPLS information model.

It could be valuable to integrate G^{2}MPLS Control Plane on a disjoint partition of the available network resources, by moving some of them under the authority of the G^{2}MPLS Control Plane and maintaining some others under the management plane (including NRPS-es), as shown in Figure 5-1.

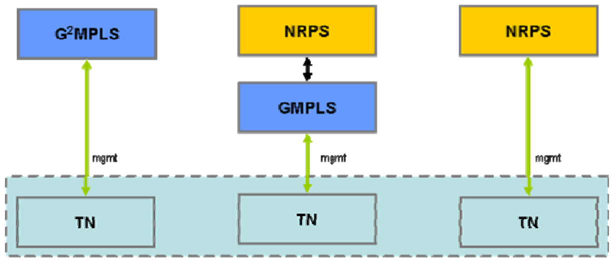

Figure 5-1: Resource partitioning among G^{2}MPLS, NRPS with GMPLS and NRPS.

Moreover, due to the availability of advance reservation services in G^{2}MPLS, the resource partitioning problem applies within G^{2}MPLS too, because of the temporal dimension of Transport Plane resources.

The simpler partitioning scheme that can be implemented in this case is to select resources to be used for bookings from a disjoint set with respect to those that could be used for immediate reservations. This approach simplifies the routing and signalling operations of the respective connection modes by guaranteeing

- a uniform semantic for resource availability specification in time
- the creation of complete and coherent set of resources.

The latter aspect, i.e. completeness of the set of resources, is particularly needed in case of non-label swapping for specific connections (i.e. non timeslot interchange in TDM networks, non wavelength conversion in DWDM networks): in this case, a label is completely available either for advance or for immediate reservations and the existence of a solution for the non-label swapping problem along the designated path could exist just in one of two disjoint solutions spaces (i.e. the advance or the immediate).

More complex partitioning schemes might include resource re-partitioning in time, e.g. triggered by the notification of a complete use of the resources in one partition or by other performance benchmarks. In this case the filled partition could include free and available resources of the other partition (or a further sub-partition of the other partition) up to a certain threshold. The definition of these schemes and procedures relies above all on the Network Operator policies for offering connection services and depends on the expected connections requests (i.e. more advance reservation – and resources for them – could occur in case of a high deployment of Grid services, more immediate reservation could occur in case of traditional end-to-end permanent connection services, etc.).

The answers by some NRENs to the questionnaire on the possible adoption of GMPLS/G^{2}MPLS Network Control Plane do not identify further requirements on resource partitioning. However, the possible coexistence of multiple technologies in the Transport Plane controlled by the same instance of Control Plane pushes for a partitioning internal to GMPLS. This could be needed to group homogeneous resources (i.e. all the L2, or TDM or the optical ones) under the same Control Plane domain and deploy the full set of network interfaces (i.e. UNI, I-NNI, E-NNI). The proposed solution is very tight to the state of the art and could be considered as a short-term solution to the multi-technology problem. It could be superseded with the definition of node-wide and network-wide procedures for a seamless control of multi-region and multi-layer environments. These solutions have been evaluated by IETF to not rely on architectural or additional protocol specific extensions in the GMPLS Control Plane, thus they are not under standardization.

5.2 Integration of G^2MPLS SCN in the existing DCN

G^2MPLS complies with the native GMPLS separation of Transport Plane and Control Plane.

G^2MPLS messages are supposed to be exchanged among G^2MPLS entities through the configured SCN, which for these purposes is supposed to export IP Layer 3 services.

The specific support of Grids does not impose specific requirements on the G^2MPLS SCN and the same means available in GMPLS may be used to establish a Control Plane adjacency.

However, a difference between GMPLS SCN and G^2MPLS SCN is exactly the existence of a Grid among the users attached to the Control Plane. Grid-specific transactions for the configuration, booking and monitoring of the jobs in the dispersed Grid performers are currently conveyed through the Internet (with secure connections). In the G^2MPLS case, part of them feed the extended Network Control Plane at the UNI reference point and are directly progressed by the G^2MPLS entities towards the far end Grid performers. Therefore, these middleware messages need to share the same SCN used for the setup and maintenance of the Transport Plane connections among the performers. Other control messages by the Grid middleware but not in the primary scope of the G^2MPLS NCP could:

- share the DCN with G^2MPLS and other Network Operator's management means
- or flow across dedicated Transport Plane connections (LSPs)
- or continue using the (secure) Internet.

5.2.1 Grid resource information dissemination and service discovery

The need for disseminating Grid resource information across the Control Plane elements for the full-fledged G^2MPLS operations and for implementing a Grid service discovery at the User-to-Network edge imposes the availability of an access to the G^2MPLS SCN from the accessing (Grid) user sites. These interfaces could be provided by the Network Operator DCN and have the same requirements of the client part in the UNI.

5.2.2 Security Issues

The Grid power-users of the G^2MPLS represent a critical vulnerability point of the Control Plane and Transport Plane because of their capability to configure both Grid and network resources with seamless and automated procedures.

The same considerations for GMPLS SCN security in the Transport and Control Plane apply also to the G^2MPLS SCN. In fact, the basic procedures for guaranteeing integrity to the protocol messages and transactions are still applicable. Moreover, G^2MPLS inherits the mechanisms for establishing trust chains among the components of a Grid, being the network just one of them. Therefore, G^2MPLS users are supposed to access network and far-end Grid resources after a successful phase of authentication and authorization against an Authorization Entity.

5.3 G^2MPLS network reference points

The deployment of the enhanced Grid-GMPLS Control Plane sets analogous reference points with respect to the ASON/GMPLS ones (ref. section 4.3). The resulting network interfaces are a Grid-aware evolution of the standard interfaces (UNI, I-NNI, E-NNI), with a set of procedures that maintains the backward compatibility with the original ASON references, but provides also the seamless and one-step control of both Grid and network resources.

Five network interfaces are identified in the G^2MPLS Network Control Plane (ref. Figure 5-2):

- **G.OUNI**, i.e. the Grid Optical User-Network Interface that supports Grid and network signalling and discovery between the Grid site and the G^2MPLS domain.
- **G.I-NNI**, i.e. the Grid Internal Node-Node Interface (G.I-NNI) that supports the routing and signalling procedures between adjacent nodes.
- **G.E-NNI**, i.e. the Grid External Network-Network Interface that propagates Grid and network topology information across different Control Plane domains and supports the inter-domain signalling mechanisms.
- **SBI**, i.e. the Southbound Interface that retrieves resource status from the specific Transport Plane and translates Control Plane actions into appropriate configurations of those resources.
- **NBI**, i.e. the Northbound Interface that groups two interfaces towards upper layers: one towards the Grid layer (G.NBI) and one towards the Network Service Plane (including NRPS, N.NBI).

These reference points are described in terms of abstract messaging by deliverable D2.1 – Section 8 and briefly summarized in the following subsection.

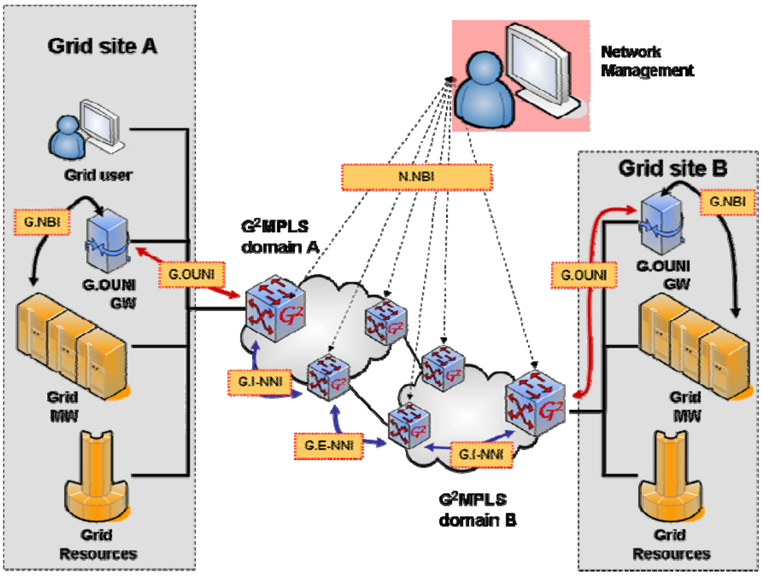

Figure 5-2: G^{2}MPLS network reference points.

5.3.1 Summary of G^{2}MPLS network interfaces

5.3.1.1 *G.UNI*

Grid Optical User Network Interface (G.OUNI) comprises a number of procedures to facilitate on demand as well as in-advance access to Grid services over G^{2}MPLS. It has knowledge of the JSDL semantic [JSDL] and implements its functionalities of

- (Grid and network) Service Discovery, in order to discover Grid (e.g. CPU types, OS types, etc.) and network capabilities of the two participating parties;
- (Grid and network) Neighbour Discovery, in order to learn resource availability in the Grid layer and from the network side;
- Grid Network Service creation/deletion integrated with AAA capabilities, in order to set-up or delete calls an connections for Grid jobs;
- Grid Network Service status inquiry, in order to retrieve the status of an established connection;
- Grid Network Service notifies service, in order to let the network indicate a change in the status of the Grid Network Service (e.g., un-restorable, connection failure, remote Grid resource failure, etc.).

Further details of G.OUNI interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.7.

5.3.1.2 *G.I-NNI*

The G.I-NNI interface has knowledge of the Grid Network Service semantic and implements the routing and signalling mechanisms for the creation, deletion, monitoring, recovery and crankback of the network connections from a node to a neighbour node.

This interface is basically the GMPLS NNI. It progresses the call setup initiated at the G.I-NNI and participates actively in the TE routing. Further details of G.OUNI interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.2.

5.3.1.3 *G.E-NNI*

The G.E-ENI interface has also knowledge of the Grid Network Service semantic and implements the inter-domain routing and signalling mechanisms for the creation, deletion, monitoring, recovery and crankback of the network connections.

This interface is based on ASON E-NNI. It progresses the call setup initiated in a domain and participates actively in the hierarchical per-domain TE routing. Further details of G.E-NNI interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.7.

5.3.1.4 *SBI*

The SBI interface allows the semantic translation of Control Plane actions into specific actions to be carried out by underlying Transport Network Equipment. Depending on the interface exported by the equipment (e.g. TL1, SNMP, CLI, etc.), the SBI syntax adapts and implements its basic actions of:

- Transport Network resource discovery, aimed to retrieve resource availability
- Transport Network resource configuration request, aimed to configure resources (e.g. cross-connections)
- Transport Network resource status, aimed to retrieve resource status (free/busy, up/down)
- Transport Network resource notify, aimed to inform the upper-lying Control Plane of change status.

The SBI interface is conceived to have a client side on the Control Plane and a server side on the management agent of the Transport Network Equipment. Most of the transactions are initiated by the client towards the server. Just the notification is asynchronously initiated by the server upon the occurrence of an event in the set of monitored resources (e.g. link down/up, etc.).

Further details of SBI interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.3.

5.3.1.5 *G.NBI*

G.NBI interface is conceived to link the Grid layer with the G^2MPLS and the G.OUNI in particular, mediated by gateway functionality (the G.OUNI gateway). Its services are under the root of Execution Management Services and comprise:

- Resources indexing (processing, storage, executables, resource management and provisioning)
- Resource selection (to decide where to execute a unit of work)
- Reservation

Further details on this interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.3, D2.7, and D2.8.

5.3.1.6 *N.NBI*

The N.NBI is used only to request/tear down network connections and to retrieve topology and connection status information. It has been defined in the PHOSPHORUS WP1 framework, in the form of a WS-based interface [PHOSPHORUS-D1.1]. This interface is based on a simple request/response model, which abstracts and generalizes transactions between NRPS/NSP and GMPLS.

Services exported by N-NBI are:

- Path Creation Service
- Path Termination Service
- Topology Discovery Service
- Path Discovery Service

Further details on this interface are available in deliverable D2.1 – Section 8 and will be defined in deliverable D2.9.

5.3.2 PHOSPHORUS G^2MPLS grid – network layering models

Two Control Plane models have been identified in the G^2MPLS architecture:

- G^2MPLS Overlay
- G^2MPLS Integrated

These models refer to different layering solutions of Grid and network resources and have a different scope with respect to the Overlay, Augmented and Peer models described in section 4.3. In fact, they refer principally to the positioning between the Grid Service Layer and the Network Control Plane and not to the relationships between different Control Plane segments/domains.

5.3.2.1 *G²MPLS Overlay*

In G²MPLS Overlay model, the Grid layer has Grid and network routing knowledge in order to provide Grid resource configuration and monitoring (as in its standard behaviour) plus network resource configuration and monitoring. G²MPLS acts as an information bearer of network and Grid resources and as a configuration "arm" just for the network service part.

This model is intended to be mainly deployed when most of the computational and service intelligence need to be maintained in the Grid layer for specific middleware design and functional behaviours. The leading role in this case is played by the Grid scheduler, which is the overall responsible for initiation and coordination of the reservation process through the participating Grid sites and the network in between.

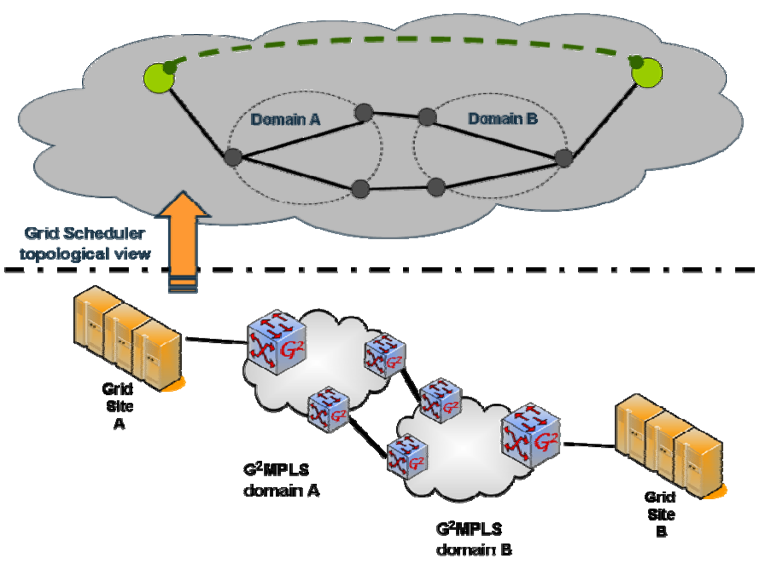

Figure 5-3: Topological view in the Grid scheduler as per G²MPLS overlay model.

It is up to the Grid scheduler to localize and select the Grid sites where to execute the requested job (immediately or later in time) and for this purpose it needs to learn about the existence of remote Grid sites and a special Grid node implementing the network connection service between them. This is sketched in Figure 5-3, where the Grid layer topological view (green lines and shapes) is augmented with the (possibly summarized) network topological view (dark grey lines and shapes). The Grid topology overlays the network topology, but the Grid scheduler needs to know both in order to send detailed connection requests towards the G²MPLS, e.g. by specifying the ingress and egress network attachment points or possibly the explicit route to follow.

G^{2}MPLS is responsible for scheduling, configuring and monitoring of the job part related to the network, thus implementing advance reservations, recovery and connection monitoring mechanisms.

This model is intended to be a short-term evolution of state-of-the-art, because it improves knowledge and capabilities of the Grid layer, relies on advanced GMPLS features (e.g. advance reservations) that are still missing in state of the art, but it does not provide the aimed one-step co-allocation of Grid and network resource yet. In fact, the two layers (i.e. Grid and network) remain asymmetrically aware of each other (the more the Grid the less the G^{2}MPLS), and the level of awareness much depends also on the capability of the Grid layer – and the Grid scheduler in particular – to compute service distributions based on the details of network topology retrievable from the G.OUNI.

5.3.2.2 *G^{2}MPLS Integrated*

In the G^{2}MPLS Integrated model, most of the functionalities for resource advance reservation and commit are moved to the G^{2}MPLS Network Control Plane. G^{2}MPLS is responsible for scheduling and configuring all the job parts, those related to the Grid sites and those related to the network.

Grid sites are modelled as special network nodes with specific additional Grid resource information (ref. green and dark grey shapes in Figure 5-4). The resulting topology is flat and integrated with respect to the positioning of the Grid layer against the network layer.

The Grid scheduler functionality is still needed to support the many user applications that rely on specific Grid infrastructure. Moreover, the Grid scheduler is needed to support workflow services. In fact, the elementary unit of service managed by the G^{2}MPLS Control Plane is the Grid job, which is a component of the workflow.

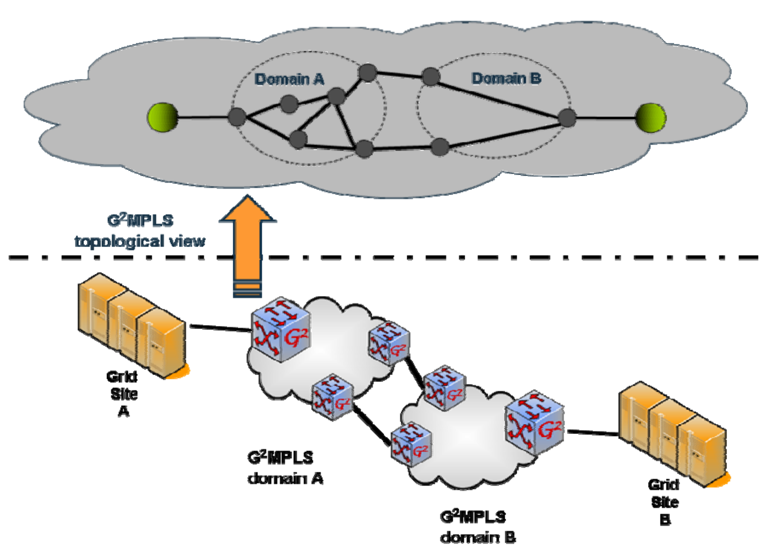

Figure 5-4: Topological view in Domain A as per G^{2}MPLS integrated model.

This model implements the one-step co-allocation of Grid and network resources, because job requests are directly managed by G^2MPLS. The full path between the selected Grid job providers is determined by the G^2MPLS. Similarly, the coordination of the signalling procedures to negotiate a timeframe and trigger execution of a Grid job is under G^2MPLS responsibility. One of the advantages of this full-fledged G^2MPLS deployment model is the availability of recovery functionalities, involving network but also Grid resources. In fact, new escalation strategies could be defined for mitigating the failure condition during the execution of a job. For example, G^2MPLS could start recovering the network connections, then (in case of failure) move to the selection of new job performers – if not strictly specified by the application/middleware – and/or move to the notification towards the Grid layer/user that could take the final decision on pausing re-scheduling the job on other resources.

Due to the impact of the model on the existing Grid middleware and the advanced functionalities to be implemented by the G^2MPLS Network Control Plane, the PHOSPHORUS integrated model could be considered a mid/long-term deployment solution for the Grid and network co-allocation.

5.4 Main G^2MPLS procedures

This chapter is aimed to summarize the procedures that can be provided by the G^2MPLS Network Control Plane. The main reference for the contents of the following subsections is deliverable D2.1 – section 6. Details on the protocol extensions and specific mechanisms of the different procedures will be provided in the further WP2 deliverables, and particularly D2.2, D2.3, D2.7, D2.8 and D2.9.

5.4.1 Recovery

An important feature provided by the use of G^2MPLS integrated model is the availability of recovery functionalities, which are the native realm of GMPLS. In general, the term recovery refers to the capability of the system to preserve or recover the integrity of an existing transport service in the event of any failures within the Transport Network. To deliver reliable services G^2MPLS, like GMPLS, requires a set of procedures to provide protection or restoration of the data traffic. In case of failure occurrence on working LSP, these recovery procedures have to be fast enough not to disrupt application's data connections. The additional advantage of G^2MPLS for this situation stems from the ability of G^2MPLS to take care both network and grid resources and therefore not only network recovery procedures can be used but also some grid recovery procedures may take place.

Depending on the definition of the job contents and objectives, a number of network calls and connections are created in the G^2MPLS network and each network connection, segment; span could be recovered in case of fault in the network, according to the procedures described in section 4.4.1. However, some critical and unrecoverable fault conditions might occur in the network, especially in case of limited connectivity or any grid node failure that can interrupt the grid workflow. Depending on the seriousness and impact of the occurring network fault, it could be impossible for the Network Control Plane to recover the service.

In such case an escalation from the network layer to the Grid layer could be needed, triggered by timely fault notifications across the G.OUNI. Solutions for grid computing recovery are based on schemes that are directly implemented by the user in the applications and also on schemes implemented in the middleware to provide automate response to failures.

The G^{2}MPLS integrated model allows the definition of new escalation strategies, which could start recovering the network connection, then (in case of failure) move to the selection of new job performers – if not strictly specified by the application/middleware – and/or move to the notification towards the Grid layer/user that could take the final decision on releasing all already used resources, re-scheduling the job and sending a new request to the G^{2}MPLS Control Plane.

5.4.2 Crankback

When a failure occurs, the crankback error message (RSVP PathErr or Notify) carries precise failure details to the ingress node, in order to let the next computed path avoid the known failures. However, in particular conditions, it may be impossible to discover a way to send a new path message to the same egress node (destination node). In this case crankback operation will be repeated a few times and when the number of retries is exceeded the path-making efforts will be ended. G^{2}MPLS in this situation could consider escalation to Grid middleware or select a new grid node if available to which path will be send. In the first scenario, the Grid middleware probably is responsible for sending a new job request to G^{2}MPLS Control Plane. The second scenario is possible if Grid middleware used the implicit service identification (i.e. without any network attachment point specification) and Traffic Engineering Database (TED) in G^{2}MPLS Control Plane nodes contains alternative grid resources.

Also crankback could be used in case of Grid resource unavailability when Path message come to Vsite. Detailed information about this unavailability can be sent back to the ingress node and alternative Grid resources can be applied for the non-explicit Grid resources job request.

In the context of G^{2}MPLS crankback can also be applied to advance reservation mechanisms to search advance paths within the network that are not fully recognized by the routing protocol. Thus PCE can provide ERO for a new Path message. This ERO may contain intermediate nodes which haven't provided future links availabilities for the PCE. As there is a higher possibility of adversity it can be normal to repeat the crankback procedure a few times. The crankback operation timeout can be greater in advance reservation that in on-fly path setup. Additionally, the crankback error message should be able to carry the full future link availabilities on taken path (time extended crankback routing). This information can be used in PCE to compute next advance paths.

5.4.3 Claiming of existing circuits

The deployment of G^{2}MPLS Control Plane in an existing infrastructure and Management Planes should cope with the possible case of maintaining the configured connections and migrating their control under G^{2}MPLS. This could occur when deploying G^{2}MPLS Control Plane in a domain previously managed by NRPS.

The requirements and proposed protocols extensions are under definition for standard GMPLS (draft-ietf-ccamp-pc-and-sc-reqs-00.txt) and completely apply to the G^2MPLS. In fact, claiming existing circuit implies using the G^2MPLS just in the network scope and specifically in the PHOSPHORUS overlay model. A specific G^2MPLS requirement with respect to this procedure is the capability to extend the circuit migration also to planned circuits, i.e. those under advance reservation.

5.4.4 Advance reservations

Traditional immediate reservations (IR) allocate the resources just after the request; in turn, AR specifies a start time and a holding period as well as source, destination and bandwidth for connections that can be requested prior in time. The most common approach for dimensioning AR is to define slot time units, thus reducing processing capacity and increasing scalability. Other possible implementations of AR allow a flexibility degree to adjust the start slot and fit a connection reservation that otherwise would be blocked.

In a heterogeneous network in which different clients may coexist, both IR and AR services may be requested. The introduction of AR in a real network lowers the quality of IR due to their disadvantage in time and the uncertainty of the holding time of IR may force to pre-empt active connections when there is a conflict of resources already reserved by AR in a future slot. The most straightforward approach consists in partitioning resources, dedicating different partitions for IR or AR (ref. section 5.1).

Regarding scheduling information for AR in G^2MPLS, two different solutions may be implemented, centralized or distributed. As outlined in D2.1 the centralized one may apply to single domain purposes and the distributed one to inter-domain.

- Centralized: Each single domain would be managed by a Path Computation Element (PCE), further detailed in the following section, which will perform the calculation of routes based on a network graph with updated information of network resource state. This is the traditional approach, and requires that the PCE has an updated picture of the network topology at all times and PCE to PCE information exchange between domains.
- Distributed: Each node of the domain implements a PCE instance and resource scheduling state information is exchanged using the OSPF-TE protocol.

For NRENs, Advance Reservation Service would be useful to guarantee resource provisioning for instance in real-time applications (e.g. live broadcast) that know the start time before it is actually needed.

5.4.5 Path Computation Issues

Path computation procedures identified in the context of PHOSPHORUS should be in compliance with the IETF Path Computation Element (PCE) architectural model [IETF-RFC4655] since most of the protocol-specific issues are defined and solved in this framework. "*The PCE is an entity (component, application or network node) that is capable of computing a network path or route based on a network graph and applying computational constraints during the computations"* [IETF-RFC4655]. The PCE can provide Traffic Engineering

(TE) functions to the G^{2}MPLS enabled network and run sophisticated constraint routing algorithms in a convenient way to address issues like: i) resource coordination (e.g. CPUs, storage) ii) advance reservation iii) physical layer impairments iii) different connection types (unicast, multicast or anycast) and iv) QoS.

For the PHOSPHORUS project, it is assumed that each domain is independent and PCE can be centralized only at a single domain level (as per chapters 5.2 in [IETF-RFC4655]). The PCE can be also distributed within each domain, so that multiple PCEs cooperate to fulfil user requests (ref. chapters 5.3 and 5.4 in [IETF-RFC4655]). Each PCE then has partial knowledge of the network, which may provide the benefit of faster computation for large, complex networks. The local PCE in the former case represents a local Autonomous Domain (AD) that acts as a protocol listener to the intra-domain routing protocols e.g. OSPF-TE, and is also responsible for inter-domain routing. PCEs peer across domains and exchange abstract or actual topology information to enable inter-domain path computation and also utilize a modified version of OSPF-TE to share a link state database between domains. It is advised that PCE has limited knowledge about all domain internals supported by the overall control plane, at least as abstracted information. This is required for multi-domain path computation. In case that the PCE has full knowledge of neighbour topologies, it is possible to perform path searching instantly. Such path is then called "explicit path", as it contains all hoops from source to destination points. If the PCE is allowed to see only abstracted topologies of neighbour domains, it is possible to find "strict/loose path". The term "strict" refers to a local domain, where all links and nodes are known, while "loose" refers to the neighbour domains, where only abstract nodes or LERs are known. The strict/loose paths must be sent to neighbour domains for further negotiation regarding domain internals.

The PCE includes advanced algorithms that integrate calendar, topology discovery and routing functions to allow path computation with multiple constraints like physical layer impairments, network resource coordination, advance reservations and different connection types. For the path computation upon a request arrival, the user specified parameters carried by the LSP request are parsed into constraints inside the PCE which takes the responsibility to provide the required end to end path if possible. The coordination of Grid applications requires the rapid discovery of appropriate connections which are the result of very complex and intensive path computations. PCE can assist to this direction through the abstracted information of the global topology stored in the TED, of each domain, especially in cases where the network management plane is not able to provide this functionality. From Grid constraints (Grid resource scheduling and coordination) the PCE may construct a reduced topology of the network based on which integrated approaches that consider multiple parameters (e.g. physical layer impairments, QoS, different connection types) can compute optimal paths maintaining high network performance and implementing optimize protection and restoration schemes.

5.5 Integration of Grid and Network AAA mechanisms

In addition to GMPLS signalling G^{2}MPLS functionality should be available as PHOSPHORUS project results. Therefore the AAI infrastructure must be extended to support Grid services. That can be done in two ways, depending on expected integration:

- Full integration of Grid and network AAI systems
- Separated services with information exchange

The first approach assumes that there is single AAA mechanism that is used for both network and Grid. This may potentially cause problems with maintenance of users and privileges database due to large number of database entries. The network AAA systems usually do not include large number of users, as it is limited to system administrators and operators. In contrary, Grid may consist of large number of users. Moreover, users may be not physical ones, but virtual rather, like projects, groups, etc. The most advantage of this approach however, is full single sign-on, where users logged into Grid systems, are automatically authenticated to access network resources.

The second approach assumes that Grid and network AAA services are separated, however there is an information exchange channel. User authenticated for Grid resources must be again authenticated to access network resources. This leads to situations where a single user must be authenticated twice. To avoid this issue, it is possible that the user is authenticated in Grid services only, and is transparent for network services, while the Grid service represents the user for network purposes. The network AAA infrastructure does not need to know all users then; however it must be aware of Grid services which must be authenticated. As this authentication is performed automatically, the user is not forced to log in twice. The advantage of this approach is that Grid and network AAA are kept separately under jurisdiction of specific administrators, who are not crossing responsibilities. The drawback is that there are two databases which must be kept consistence, so this causes a maintenance problem. Keeping two independence accounting infrastructures may also complicate charging users for services.

5.5.1 SLA operation and maintenance

The SLA management system proposed for GMPLS can be extended to service G^2MPLS architecture. In general all procedures of SLA operations and maintenance are the same but in case of G^2MPLS there are several modifications that complicate the process in some aspects.

The main extension refers to the SLA template that has to include additional parts with description of grid specific resources requirements and job specifications.

SLA maintenance and scheduling procedures are also a bit more complicated because grid and network resource access and reservations must be synchronized in advance to ensure reliable job execution. The SLA management system has to cooperate closely with grid job scheduling and queuing systems and take into account many constraints and time dependencies between submitted jobs.

The process of SLA negotiations can be split into two depended steps. In the first step network SLS is processed according to the rules defined for GMPLS architecture but grid constraints should also be met. If network negotiations are finished the grid resources reservations procedure is triggered. When required grid resources are available in local domain then grid SLS record is performed locally and if resources belong to some other domain the grid part of SLA is sent to the next peer domain toward grid domain.

Negotiations are successfully finished and SLA is accepted only if, both network and grid parts requirements, dependencies and constraints can be fulfilled.

5.6 Integration of G^{2}MPLS management in the existing NMS

The network operators, who will be deploying G^{2}MPLS Control Plane, will want to use the existing Management Plane which is composed of Network Management System. These expectations will motivate to introduce GMPLS functionalities, which are described in chapter 4.6 and will not be repeated here, and the Grid related extensions into NMS, which should give the ability at least to monitor Grid resources localization, capability, availability and usage. In this way the network/service operator should be able to view grid and network resources existence and state. This knowledge will be gathered and provided to MP by Vsite adjacent CP entities, and it will be very helpful during the failure localization and reparation processes. We can imagine also that MP will have the information about a grid and network resources interrelation which exists in relation to particular GNS request.

Some requirements to ensure management and control plane interaction are (all pure network related requirements are skipped; they can be found in chapter 4.6):

- For Grid resource discovery, the MP must be capable of receiving notifications from the CP for the discovery of the addition or removal of a Vsite (provided by G-SDA module of the Vsite adjacent CP node)
- The MP must support the capability to query the CP Grid resource localization, capabilities and availabilities
- The MP should be capable of receiving notifications from the CP for Grid and network resources allocation or release
- The MP must be able to query the information about all calls which are using particular Grid or network resource
- The MP must be able to query the information about all calls which are using particular Grid or network resource
- The MP must be able to query the particular call related Grid and network resources

5.6.1 Monitoring of G^{2}MPLS calls and connections (SCs and recovery LSPs)

Since Management Plane (MP) is not directly involved in all connection setup processes (e.g. SC) monitoring of existing connections and newly requested calls is very important to ensure efficient and reliable resource management and fault detection. Also SLA, performance, recovery and scheduling mechanisms need monitoring data to support reliable services. ASON/GMPLS network calls and connections monitoring and management are described in section 4.6.1 while this section is focused on grid parts of G^{2}MPLS calls. Of course, to support complete end-to-end G^{2}MPLS call and connection status, data from network and grid monitoring systems have to be performed jointly.

The reference architecture for grid monitoring systems was proposed by the Global Grid Forum (GGF) as the Grid Monitoring Architecture (GMA). It is defined in the GFD-I.7 document that has been submitted as an informational document by the Performance Working Group within the Information Systems and Performance Area of the GGF. The document describes the core GMA components and models for high-level communication between components of different types. The Grid Monitoring Architecture consists of three types of components, shown in Figure 5-5:

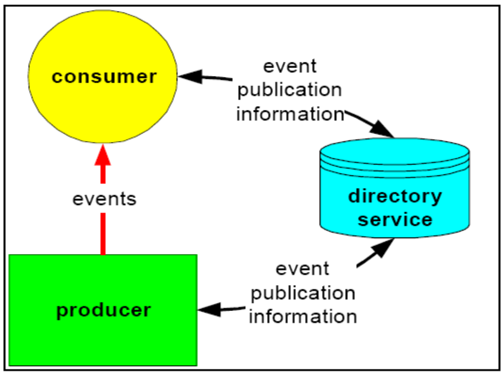

Figure 5-5: Grid Monitoring Architecture Components

Directory Service – database supports information publication and producers discovery,

Producer - producers register themselves with the directory service and describe the type and structure of information they want to make available to the other services,

Consumer – consumer queries the directory service to find out what type of information is available and locate producers that provide such information. To obtain the relevant data consumers contact the producers directly.

There are two main implementations of GMA architecture proposal:

- Relational GMA - R-GMA
  The Relational Grid Monitoring Architecture provides a service for information, monitoring and logging in a distributed computing environment. R-GMA bases on a Relational Data Model, it uses individual RDBMS and SQL statements to provide the functionality outlined in GMA.

- Monitoring and Discovery System - MDS
  The Monitoring and Directory Service (MDS) is a part of Globus Toolkit and it provides information about the status of Globus system components. There is also the Heart Beat Monitor (HBM) that allows the system administrator or user services to detect failure of application processes.

To monitor G^2MPLS calls according to the GMA model MP should communicate with a particular grid monitoring system over customer API and probably also some new instances of monitoring data producers should be provided. The G^{2}MPLS monitoring service requiring information initiates request/query via the consumer service using consumer API. The query is first passed to the Directory Service to identify which producers are able to answer that query, and a list of appropriate producers is established. In the next step the query is passed by the consumer service to each relevant producer, to obtain the requested data directly. When customer receives all required data it can pass them back to the client service.

6 Models for coexistence/interoperability of GMPLS and G²MPLS Control Planes

The G²MPLS Control Plane derives from the ASON/GMPLS architecture and has been defined as a superset of the original ASON/GMPLS functions and procedures in support of the specific functionalities required by Grids.

The study of the relationships between G²MPLS and GMPLS domains is an interesting use-case, from which an evaluation of the achievable levels of interworking in signalling and routing procedures can be derived. In this chapter the applicability of the network reference points is discussed, mainly in an inter-domain framework. Moreover, the applicability of the G²MPLS transaction model to the GMPLS call/connection concept is analyzed and some possible network scenarios of co-existence are provided.

6.1 Applicability of the network reference points

The ASON architecture is built around the identification and description of network reference points (i.e. UNI, I-NNI, E-NNI), which represent a convergence point for the main architectural features and protocols' actions. G²MPLS inherits these reference points and elaborate on them in order to provide the needed extensions in support of Grids (thus defining a G.OUNI, G.I-NNI, G.E-NNI).

Any evaluation of the coexistence issues between G²MPLS and ASON/GMPLS domains needs to start considering the compatibility and possible interworking between the two set of network interfaces. In fact, this translates in the identification of the continuity points of the different Network Service and Control Planes, extended in the PHOSPHORUS case to the Application (Grid)-Service Plane, as depicted in Figure 6-1.

This section describes the problem under two perspectives, i.e. the intra-domain and the inter-domain, being these the main frameworks for a complete analysis of the co-existence of two different Network Control Planes.

Discussion on the detailed applicability of the different procedures and protocol messages at the different network interfaces (i.e. UNI vs. G.OUNI, I-NNI vs. G.I-NNI, E-NNI vs. G.E-NNI) is out of the scope of this deliverable and will be defined in subsequent detailed specifications of all those interfaces (deliverable D2.7)

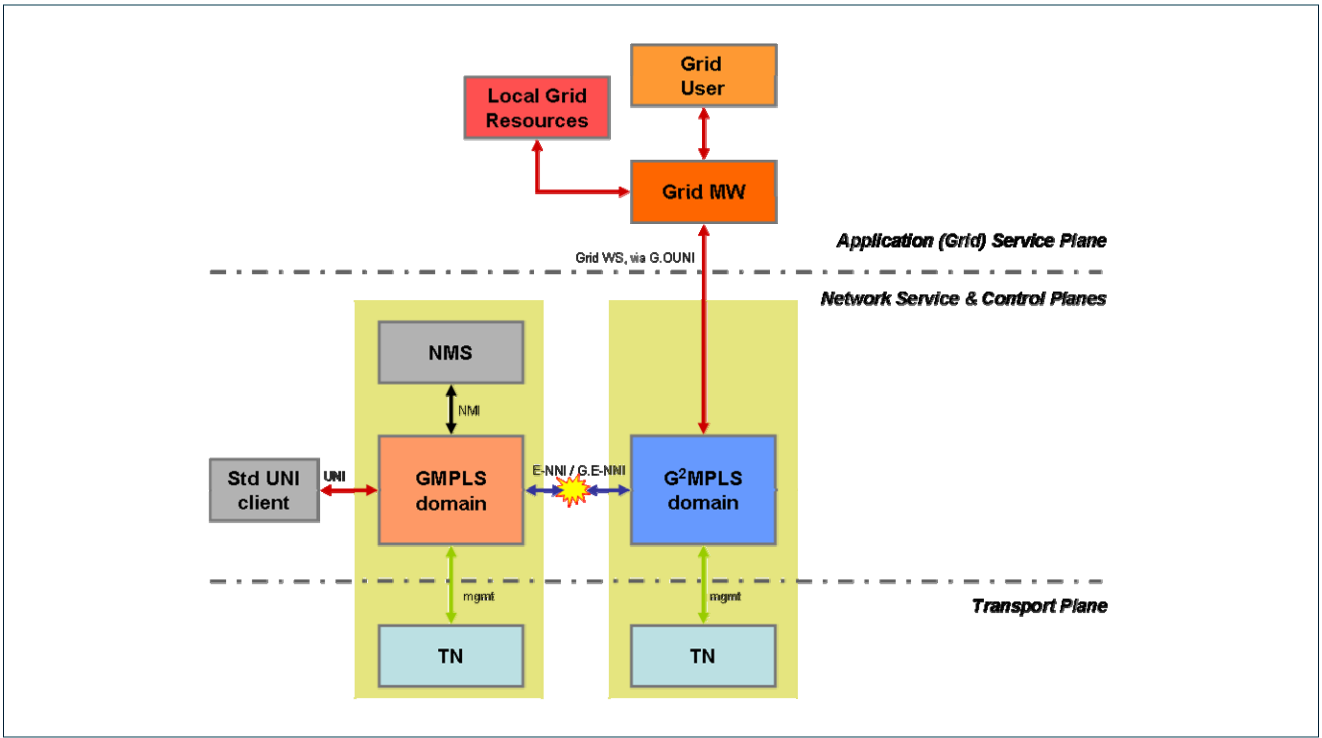

Figure 6-1: G^{2}MPLS and ASON/GMPLS positioning in the different planes.

6.1.1 Intra-domain coexistence

Intra-domain is generally related to the Internal Node-to-Node Interface between Control Plane performers. ASON/GMPLS I-NNI is a standardized network reference point in which GMPLS-controlled equipments by different vendors could establish peering relationships. It is a common practice to test and validate the standardized set of GMPLS features developed in a stack against third-party protocol controllers; this procedure assess the standard compliance of the developments and, implicitly, define the base for a possible interworking between GMPLS-empowered equipments by different vendors in the same domain.

G^{2}MPLS requires a complex set of extensions to the base GMPLS protocols for Grid purposes, both in terms if routing (e.g. advertisement of Grid capabilities and information permeability at the G.UNI) and of signalling (e.g. procedures for the setup and maintenance of Grid Network Services through GNS transactions). While routing extensions rely on OSPF Opaque LSAs and could be just forwarded – though neglected – by standard GMPLS controllers, part of the signalling extensions cannot be ignored by pure GMPLS nodes and could lead into unsupported operational conditions resulting in a final block of signalling.

For this reason, the PHOSPHORUS position is to analyze the interoperation/coexistence issues between G^{2}MPLS and GMPLS just at the domain granularity (i.e. inter-domain co-existence), in order to adapt and progress signalling and routing messages according to the declared neighbouring domain capabilities.

Use of the G²MPLS architecture and controllers in a mode reduced to standard, i.e. for the implementation of just standard connection services as per ASON and GMPLS architectures, does not raise any interworking problem between G²MPLS and ASON/GMPLS nodes – at least in terms of architectural design and protocol behaviours. In fact, G²MPLS is defined to be backward compatible and down-gradable to standard GMPLS just by cutting off new Grid-specific services (e.g. advance reservations and time limited bandwidth reservations) and resource advertisements.

The most convincing future scenario for an intra-domain coexistence of G²MPLS and ASON/GMPLS controllers seems to be that of a multi-region domain, in which PHOSPHORUS G²MPLS controllers are used to provide Control Plane capabilities to equipments currently lacking in it (e.g. carrier Ethernet switches or similar). In this context, the reduced PHOSPHORUS G²MPLS would be in line with other promoted and running research efforts in the field such as US DRAGON project and the Acreo National Broadband Test-bed (ANBT) in Sweden, but loosing most of its native benefits of bridging the Grid and the network layers towards the seamless one-step setup. The main envisaged role of the GMPLS Control Plane in this scenario is to set up and maintain Forwarding Adjacencies (FA) across the served TN section in an automatic, highly dynamic and adaptive way. On the other side, G²MPLS just uses these FAs in the technology layer/region under its control.

However, all these projects could provide narrow-scoped and preliminary short-term solutions to the multi-region support by GMPLS, above all with respect to the carrier (Transport) Ethernet technology. In fact, this is a hot and very open matter of discussion in IETF, because most of the basic architectural elements (e.g. resources/labels) are not merely related to protocol capabilities at the Control Plane (native scope of IETF WGs), but have a high technology impact possibly tangential to IETF scope. Moreover, planned features and services (e.g. asymmetrical LSPs) are not always mapped in existing commercial equipments ready to implement that kind of Transport Plane (i.e. carrier Ethernet) and its services, which are still under definition in ITU-T and IEEE.

Therefore, all these reasons could limit the possible impact of deploying open source G²MPLS Control Plane mixed with standard and commercial GMPLS implementations in an intra-domain scenario.

6.1.2 Coexistence across different domains

Inter-domain coexistence of G²MPLS and ASON/GMPLS is a more challenging goal with respect to the PHOSPHORUS objectives and target impacts. In fact, the probability of having such a deployment scenario is increased by both the interest of participating NRENs in deploying G²MPLS/GMPLS controllers delivered by WP2, and the pan-European test-bed and its extensions beyond Europe with the infrastructures by companion projects (i.e. US EnLIGHTned and JP G-Lambda) using some other standard GMPLS.

Other interworking solutions have been explored in deliverable D2.1 with components lying in the Network Service Plane, such as NRPS-es and the GN2 IDM. In those cases, G²MPLS preserved the consistency of its network reference point in the East-West direction (i.e. the G.E-NNI) by moving in an additional network element – the G.E-NNI gateway – the gluing logic between the Management Plane and the Control Plane.

Interworking across G^2MPLS and ASON/GMPLS domains can be established can be established according to the same design principle, but without the need of any gateway functionality, being the two parties natively peers.

However, in a G^2MPLS domain both Grid and network resources are under the same control, while in an ASON/GMPLS domain only network resources are controlled. This implies a reduction of the information set carried out through the G.E-NNI, both for routing purposes and, above all, for signalling.

Concerning E-NNI routing, the G^2MPLS advertisement process (TE + Grid) is designed to be based on OSPF Opaque LSAs as in standard ASON/GMPLS and, thus, any node in these domains must be an opaque-capable router that generates and floods LSAs towards its neighbour (cf. IETF RFC2370). Therefore, any G^2MPLS Grid-specific LSA must flow through the GMPLS routers, though not supported by them, and standard ASON/GMPLS LSAs are natively generated and processed (learning) by G^2MPLS, either if received from within its domain or generated by a neighbour domains. A real issue in the G.E-NNI routing applicability is the detail of network topology exported by each domain towards its peers, which impact the path computation and advance reservation capabilities. However, this issue rises also in case of homogeneous G^2MPLS vs. G^2MPLS interfacing and has different solutions with different levels of optimality and scalability. Some of these solutions are described in details in deliverable D2.1 and range in:

- Advertisement of just domain capabilities
- Advertisement of domain border nodes and inter-domain links
- Advertisement of domain border nodes, inter-domain links and a summarization of virtual TE-links among the border nodes (intra-domain links)
- Coordination with signalling means such as crankback.

Concerning E-NNI signalling, the main architectural issue which limits the applicability of the two E-NNI models relies on new G^2MPLS concept of destination end-point, which, in the anycast case, could be unspecified in terms of network attachment point but detailed in terms of Grid capabilities (cf. deliverable D2.1 section 5). In fact, the destination network attachment point is mandatory for an ASON/GMPLS call. A possible solution in this case could be to limit this G^2MPLS behaviour in a single domain, or to run the Grid resource localization process and bind to a network attachment point in the first G^2MPLS domain traversed. Concerning the Grid Network Service (GNS) transaction concept introduced by G^2MPLS, it has no impact on the G.E-NNI/E.NNI, because it is a container that provides a common root to different network call/connections traversing the same G^2MPLS User to Network reference point. The same rules and mechanism of a call traversing an ASON E-NNI applies to GNS transactions.

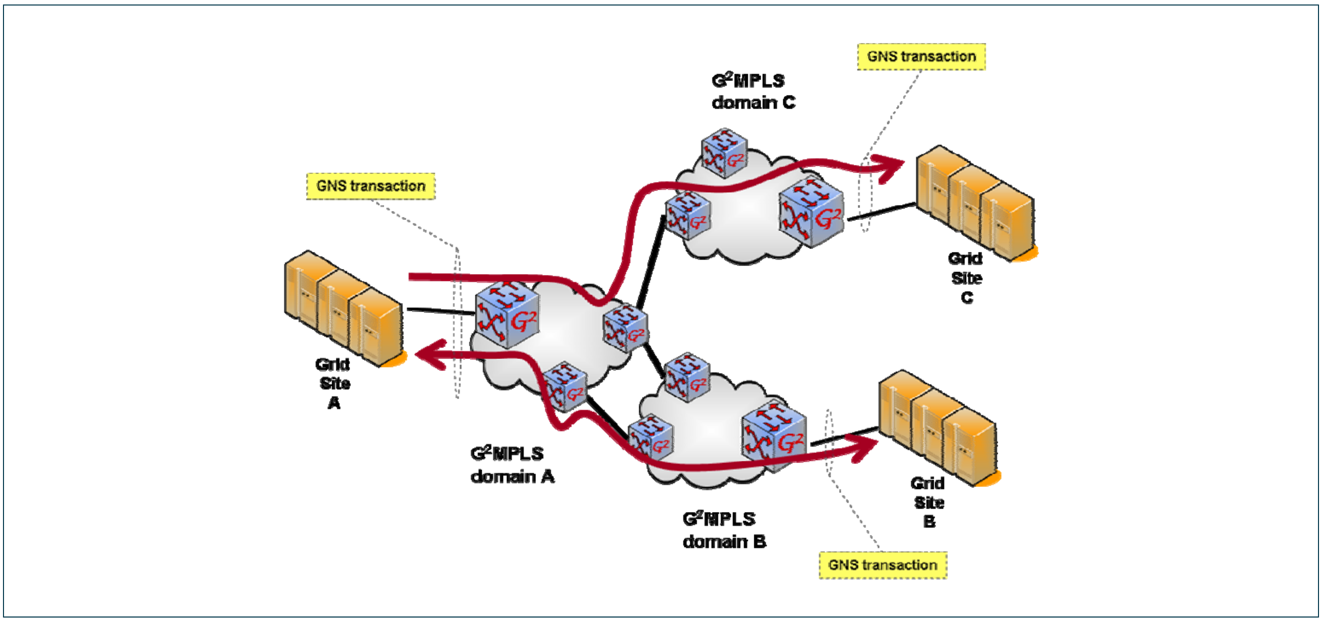

Figure 6-2: Grid job breakdown into GNS transaction and G^{2}MPLS calls.

6.2 Network scenarios

In this section two relevant scenarios are described to summarize the models for coexistence of G^{2}MPLS and ASON/GMPLS Network Control Planes. Most of them apply to the PHOSPHORUS Overlay model only, because of the network scope limitation of the ASON/GMPLS. Moreover, services initiated in a GMPLS domain and terminating or traversing a G^{2}MPLS domain cannot be modified in G^{2}MPLS Network Services as defined in deliverable D2.1 (e.g. with advance reservation), but must stick to the only service mode defined in ASON, i.e. immediate and long-living calls/connections.

6.2.1 Two-parties scenario

In this scenario one GMPLS and one G^{2}MPLS domains are directly interfaced (ref. Figure 6-3). Grid users and resources, wrapped by the Grid middleware, result localized just in the G^{2}MPLS domain. Both PHOSPHORUS Overlay and Integrated models are possible for establishing the GNS transactions.

G.OUNI interface conveys the Grid WS semantic from the middleware towards the G^{2}MPLS for establishing the GNS (Integrated case) or jus the network part of it (Overlay case). However, G.OUNI could “degrade” to a standard UNI for standard ASON call services.

G.E-NNI is not needed in this scenario and, thus, it “degrades” to standard E-NNI.

ASON call is segmented according to the ITU recommendations into connections (LSP), with the end-points briefly sketched in Figure 6-3 (reverse grey triangle).

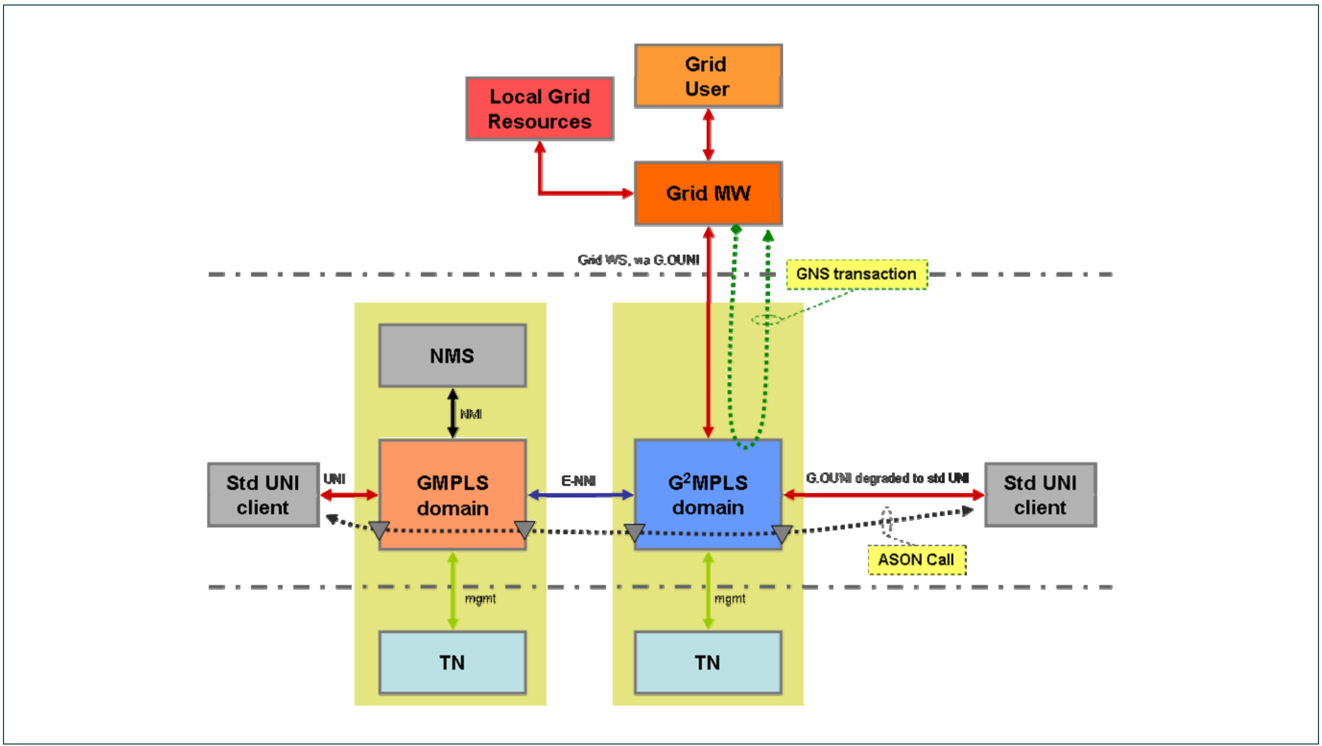

Figure 6-3: Two-parties interworking between G^{2}MPLS and ASON/GMPLS.

6.2.2 Three-parties scenario

In this scenario two G^{2}MPLS domains are “bridged” by a GMPLS domain (ref. Figure 6-4). Grid users and resources on both sides are locally managed by the Grid middleware components. Both PHOSPHORUS Overlay and Integrated models are possible for establishing the GNS transactions, with the limitation of explicit indication of the destination end points, in order to create call and connections in the GMPLS domain with standard signalling session descriptors.

G.OUNI is the full PHOSPHORUS interface for Grid and network services and the inter-domain interface is a full G.E-NNI in order to reciprocally feed the far end G^{2}MPLS instances of routing and signalling information.

G.E-NNI and G.OUNI can degrade to standard ASON behaviour in case of standard ASON call setup requests by standard users attached at the G^{2}MPLS domains.

This scenario could be possible in the PHOSPHORUS project in case the GÉANT2 network would support GMPLS and two remote local test-beds would deploy the G^{2}MPLS Network Control Plane.

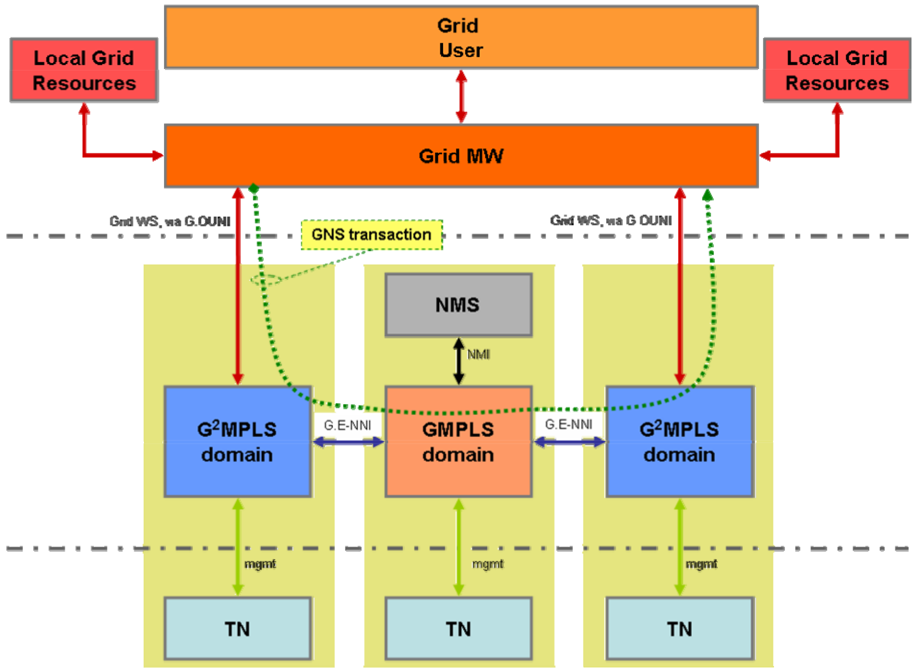

Figure 6-4: Three-parties interworking: ASON/GMPLS in the middle.

7 Conclusions

This document analyses the two proposed deployment models of G^2MPLS Control Plane (overlay, integrated) and provides a high level discussion of the identified network reference points and procedures some of which will be implemented and demonstrated by the PHOSPHORUS experimental activities on the test-bed. Specific description of current NRENs environments is provided and their willingness to adopt the G^2MPLS framework is investigated, whereas solutions to issues that arise from such adaptation are addressed.

Case studies toward the deployment of PHOSPHORUS Control Plane into specific NRENs local test-beds have been further discussed and network scenarios dealing with the coexistence of GMPLS and G^2MPLS architectures have been presented. In this document issues like AAA functionality and resource partitioning under the G^2MPLS Control Plane have been also investigated and requirements of the G^2MPLS SCN compared to the GMPLS SCN have been defined, particularly in terms of security considerations. Finally the PHOSPHORUS deployment models and the required extensions to the standard interfaces and procedures have been introduced.

This deliverable will drive the subsequent low-level designs and developments of the identified Grid-GMPLS interfaces that will be provided through D2.7. Following the G^2MPLS Control Plane architecture presented in D2.1, specific protocol enhancements will be detailed in deliverable D2.2, while the G^2MPLS high level system design will be discussed in deliverable D2.3. The overall set of these deliverables will build the structure of the G^2MPLS Network Control Plane architecture.

8 References

8.1.1 Normative references

[ASON-ARCH] ITU-T G.8080/Y.1304 Recommendations, “Architecture for the Automatic Switched Optical Network (ASON)”, 2001.

[ASON-DCN] ITU-T G. 7712/Y.1703 Recommendations, “Data Communication Network”, 2003.

[ASON-DCM] ITU-T G.7713/Y.1704 Recommendations, “Generalised Distributed Connection Management”, 2001.

[ASON-LDP] ITU-T G.7713.3/Y.1704 Recommendations, “Distributed Call and Connection Management – GMPLS CR-LDP Implementation”, 2003.

[ASON-PNNI] ITU-T G.7713.1/Y.1704 Recommendations, “Distributed Call and Connection Management – PNNI Implementation”, 2003.

[ASON-RSVP] ITU-T G.7713.2/Y.1704 Recommendations, “Distributed Call and Connection Management – GMPLS RSVP-TE Implementation”, 2003.

[ASON-RSVP-SIG] ITU-T G.7713.2/Y.1704.2 Recommendations, “Distributed Call and Connection Management : Signalling mechanism using GMPLS RSVP-TE ”, 2003

[ASON-ROUT] ITU-T G.7715/Y.1706 Recommendations, “Architecture and requirements for routing in automatically switched optical networks”, 2002.

[ASON-OSPF] ITU-T G.7715.1/Y.1706.1 Recommendations, “ASON routing architecture and requirements for link state protocols”, 2004.

[GMPLS-ARCH] E. Mannie (Ed.), “Generalized Multi-Protocol Label Switching (GMPLS) Architecture”, IETF RFC 3945, October 2004.

[GMPLS-MLN] K. Shiomoto, D. Papadimitriou, J.L. Le Roux, M. Vigoureux, D. Brungard, ”Requirements for GMPLS-based multi-region and multi-layer networks (MRN/MLN)”, IETF draft, draft-ietf-ccamp-gmpls-mln-reqs-02.txt, work in progress, October 2006.

[GMPLS-MESH] JP. Vasseur (Ed.), “Routing extensions for discovery of Multiprotocol (MPLS) Label Switch Router (LSR) Traffic Engineering (TE) mesh membership”, IETF draft, draft-ietf-ccamp-automesh-04.txt, work in progress, January 2007.

[GMPLS-PCSC] D. Caviglia, D. Bramanti, D. Li, “Requirements for the Conversion Between Permanent Connections and Switched Connections in a Generalized Multiprotocol Label Switching (GMPLS) Network”, IEFT draft, draft-ietf-ccamp-pc-and-sc-reqs-00.txt, work in progress, December 2006.

[GMPLS-SIG]	L. Berger (Ed.), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Functional Description", IETF RFC 3471, January 2003.
[G.RSVP-TE]	L. Berger (Ed.), "Generalized Multi-Protocol Label Switching (GMPLS) Signaling Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Extensions", IETF RFC 3473, January 2003.
[IETF – RFC4655]	A. Farrel J.-P. Vasseur, J.Ash, "A Path Computation Element (PCE) – Based Architecture", IETF RFC 4655, August 2006.
[IETF – RFC4139]	D. Papadimitriou (Ed.), "Requirements for Generalized MPLS (GMPLS) Signaling Usage and Extensions for Automatically Switched Optical Network (ASON)", IETF RFC 4139, July 2005.
[IETF – RFC4206]	K. Kompella(Ed.) "Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS)Traffic Engineering (TE)",IETF RFC 4206, October 2005
[IETF – RFC3209]	D. Awduche(Ed.) "RSVP-TE: Extensions to RSVP for LSP Tunnels ", IETF RFC 3209, December 2001
[IETF – RFC2750]	S. Herzog "RSVP Extensions for Policy Control ", IETF RFC 3750, January 2000
[IETF – RFC2747]	F. Baker (Ed.) "RSVP Cryptographic Authentication ", IETF RFC 2747, January 2000
[IETF – RFC4208]	G. Swallow (Ed.), "Generalize Multiprotocol Label Switching (GMPLS) User-Network Interface (UNI): Resource ReserVation Protocol-Traffic Engineering (RSVP-TE) Support for the Overlay Model", IETF RFC 4208, October 2005.
[IETF – gmpls-mln]	R. Callon (Ed.) "A Framework for Layer 3 Provider Provisioned Virtual Private Networks", IETF draft, draft-ietf-l3vpn-framework-00.txt, March 26, 2003
[OGSA EMS]	R. Subramaniam "Execution Management Services (Re-factoring)", OGFA, February 2005
[OGF-GFD.11]	M. Roehrig (Ed.)"Grid Scheduling Dictionary of Terms and Keywords", OGF GFD-I.11 November.2002
[OGF-GFD.81]	J. Treadwell, "Open Grid Services Architecture Glossary of Terms Version 1.5", OGFA, July 2006
[OIF-E-NNI-Sig-1.0]	Optical Interworking Forum, "Intra-Carrier E-NNI 1.0 Signaling Specification", http://www.oiforum.com/public/documents/OIF-E-NNI-Sig-01.0-rev1.pdf.
[OIF-E-NNI-Sig-2.0]	Optical Interworking Forum, "Intra-Carrier E-NNI 2.0 Signaling Specification", draft OIF2005.381.03, work in progress.
[OIF-E-NNI-Rtr-1.0]	Optical Interworking Forum, "External Network-Network Interface (E-NNI) OSPF-based Routing - 1.0", http://www.oiforum.com/public/documents/OIF-ENNI-OSPF-01.0.pdf.
[OIF-UNI1.0]	Optical Interworking Forum, "User Network Interface (UNI) 1.0 Signalling Specification", http://www.oiforum.com/public/documents/OIF-UNI-01.0.pdf.
[OIF-UNI1.0R2-COMM]	Optical Interworking Forum, "UNI 1.0 Signaling Specification, Release 2: Common Part", http://www.oiforum.com/public/documents/OIF-UNI-01.0-R2-Common.pdf.
[OIF-UNI1.0R2-RSVP]	Optical Interworking Forum, "RSVP Extensions for User Network Interface (UNI) 1.0 Signaling, Release 2", http://www.oiforum.com/public/documents/OIF-UNI-01.0-R2-RSVP.pdf.
[OIF-UNI2.0-COMM]	Optical Interworking Forum, "UNI 2.0 Signaling Specification: Common Part", draft OIF2005.204.02, work in progress.
[OIF-UNI2.0-RSVP]	Optical Interworking Forum, "RSVP Extensions for User Network Interface (UNI) 2.0 Signaling", draft OIF2005.205.01, work in progress.
[OSPFv2]	J. Moy, "OSPF Version 2", IETF RFC2328, April 1998.
[OSPF-OPQ]	R. Coltun, °The OSPF Opaque LSA Option", IETF RFC2370, July 1998.

[S-GMPLS] K. Kumaki (Ed.), "Interworking Requirements to Support operation of MPLS-TE over GMPLS networks", IETF draft, draft-ietf-ccamp-mpls-gmpls-interwork-reqts-00.txt, work in progress, December 2006.

[JSDL] A. Savva (Ed.), "Job Submission Description Language (JSDL) Specification, Version 1.0", GFD.056, http://forge.gridforum.org/projects/jsdl-wg, November 2005.

8.1.2 Informational references

[GEANT2] GÉANT2, The Pan-European Network for Research, http://www.geant2.net.

[G^{2}MPLS-ARCH] PHOSPHORUS WP2, "The Grid-GMPLS Control Plane architecture", deliverable D2.1.

[OGF-GNS] G. Clapp, T. Ferrari, D.B. Hoang, A. Jukan, M.J. Leese, P. Mealor, F. Travostino, "Grid Network Services", draft-ggf-ghpn-netserv-2, Grid High Performance Networking Research Group working draft, May 2005.

[QOS-SEQUIN] M. Campanella, P.Chivalier, A. Sevasti, N. Simar, "Definition of Quality of Service" SEQUIN Deliverable D2.1, May 2001.

9 Acronyms

AAA	Authentication, Authorisation, and Accounting
AAI	Authentication and Authorization Infrastructure
AMPS	Advance Multi-domain System
ANBT	Acreo National Broadband Test-bed
API	Application Programming Interface
AR	Advanced Reservation
ARGON	Allocation and Reservations in Grid-enabled Optical Networks
AS	Autonomous System
ASON	Automatic Switched Optical Network
ATM	Asynchronous Transfer Mode
BGP	Boarder Gateway Protocol
BoD	Bandwidth on Demand
CallC	Call Controller
CBR	Constant Bit Rate
CC	Connection Controller
CCAMP	Common Control and Measurement Plane
CP	Control Plane
DCC	Data Communication Channel
DCN	Data Communication Network
DFN	Deutsches Forschungsnetz
DiffServ	Differentiated Services
DRAC	Dynamic Resource Allocation Controller
DRAGON	Dynamic Resource Allocation via GMPLS Optical Networks
DWDM	Dense Wavelength Division Multiplexing
EF PHB	Expedited Forwarding Per-Hop Behaviour
EGP	Exterior Gateway Protocol
ELC	Explicit Label Control
E-NNI	External Network-Network-Interfaces
ERO	Explicit Route Object
FA	Forwarding Adjacency
FSC	Fiber Switched Capable

GE	Gigabit Ethernet
GÉANT	Pan-European Gigabit Research Network
GGF	Global Grid Forum
GLIF	Global Lambda Integrated Facility
GMA	Grid Monitoring Architecture
GMPLS	Generalized MPLS
G^2MPLS	Grid-GMPLS (enhancements to GMPLS for Grid support)
GNS	Grid Network Services
GRE	Generic Router Encapsulation
GUI	Graphical User Interface
HBM	Heart Beat Monitor
HPC	High Performance Computing
IA	Implementation Agreements
IETF	Internet Engineering Task Force
IGP	Interior Gateway Protocol
I-NNI	Internal Network-Network-Interfaces
IP	Internet Protocol
IPsec	IP Security
IR	Immediate reservations
IPv6	IP version 6
L2	Layer 2
L2SC	Layer 2 Switched Capable
LHC-OPN	Large Hadron Collider Optical Private Network
LRM	Link Resource Manager
LSC	Lambda Switched Capable
LSP	Label Switched Paths
LSR	Label Switch Router
MDS	Monitoring and Directory Service
MLN	Multi Layer Network
MP	Management Plane
MPLS	Multi Protocol Label Switching
MRN	Multi Region Network
NBI	Northbound Interface
NCC	Network Call Controller
NCP	Network Control Plane
NE	Network Elements
NMS	Network Management System
NNI	Network to Network Interface
NREN	National Research and Education Networks
NRPS	Network Resource Provisioning Systems
NSP	Network Service Plane
OAM	Operations, Administration and Maintenance
OIF	Optical Internetworking Forum
OPC	Optical Supervisory Channel

OSPF	Open Shortest Path First protocol
OXC	Optical Cross Connects
PC	Permanent Connection
PCE	Path Computation Element
PSNC	Poznan Supercomputing and Networking Center
PVC	Permanent Virtual Circuit
QoS	Quality of Service
RDBMS	Relational Database Management System
R-GMA	Relational Grid Monitoring Architecture
RRO	Record Route Object
RSVP	Resource reSerVation Protocol
RSVP-TE	RSVP with Traffic Engineering extensions
SBI	Southbound Interface
SC	Switched Connection
SCN	Signalling Control Network
SDH	Synchronous Digital Hierarchy
SEQUIN	SErvice QUality across Independently managed Networks
S-GMPLS	Segmented GMPLS
SLA	Service Level Agreement
SLS	Service Level Specification
SONET	Synchronous Optical Networking
SPC	Soft Permanent Connection
SP	Service Provider
TDM	Time Division Multiplexing
TE	Traffic Engineering
TMN	Telecommunications Management Network
TP	Transport Plane
UCLP	User-Controlled Lightpath Provisioning system
UNI	User to Network Interface
VC	Virtual Circuit
VIOLA	Vertically Integrated Optical test-bed for Large Applications
VLL	Virtual Leased Line
VPLS	Virtual Private LAN Service
WDM	Wavelength Division Multiplexing
WG	Working Group

Appendix A Case studies

A.1 PSNC

PSNC PHOSPHORUS test-bed will contain network and grid (computing and storage) resources. The network layer will be built from optical switches and Gigabit Ethernet switches, grid resources will include PC cluster nodes and storage server and disk volumes, Figure A.1-1. Most of the local connections as well as the links to the other PHOSPHORUS networks will use Gigabit Ethernet as the transmission technology. The Gigabit Ethernet transmissions will be switched in two layers – on layer 1 by optical switches and on layer 2 by Gigabit Ethernet switches.

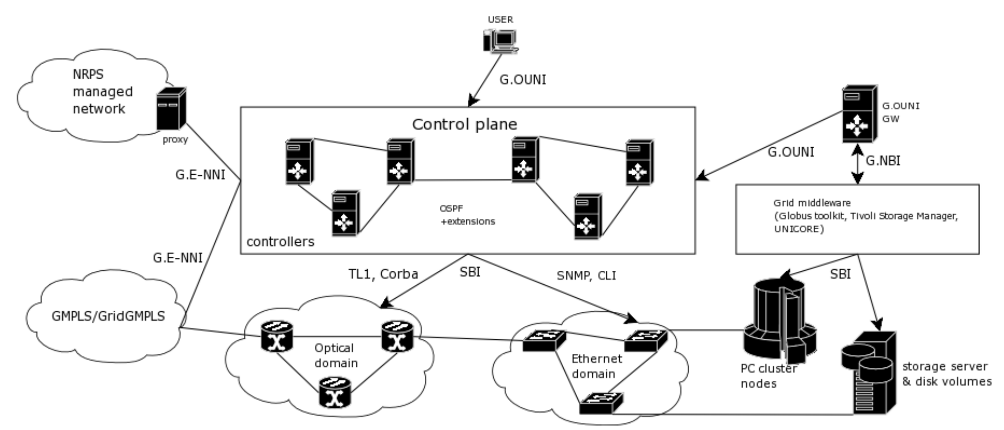

Figure A.1-1: PSNC case scenario

The core of the local test-bed will be built from three optical switches and the rest of the test-bed resources will be connected to them. The switching infrastructure will also contain three Gigabit Ethernet switches. The network core will switch lightpaths in the optical domain without conversion to electrical signals so it will be independent of the transmission technology. The optical switches will be accompanied by three Gigabit Ethernet switches which will switch Gigabit Ethernet VLANs.

The transport network and grid resources will be controlled by GMPLS (in the first phase of the project) and G^2MPLS protocol (in the second stage). Each of the switches will have an external GMPLS/G^2MPLS controller made of a PC host with a GMPLS/G^2MPLS implementation and there will be also additional controller/gateway for grid resources. The controllers will create the control-plane network and they will use the G.E-NNI interfaces to communicate with other networks and G.OUNI to receive the requests from the applications/Grid middleware/users. The controllers will also support a few specialised southbound interfaces (SBI) to communicate with the controlled transport plane network equipment: SNMP or CLI interface for Ethernet domain equipment and TL1 or Corba interface for optical domain equipment.

The PSNC scenario is supposed to use only GMPLS/G^{2}MPLS to establish lightpaths and to communicate with the control/provisioning plane of other networks over G.E-NNI interface. Direct integration of the PSNC network and the other networks using NRPSes will be possible over proxy interface between GMPLS/G^{2}MPLS protocol and NRPS managed networks.

The control-plane will be also interconnecting with GRID resources (computing and storage nodes) over the G.OUNI gateway that provides G.NBI interfaces to communicate with proper grid middleware system. The configuration allows for the exchange of control information between the GRID middleware and the control plane of the network, e.g. signalling requests from GRID to the network.

In the grid layer the Globus Toolkit version 4, IBM Tivoli Storage Manager and UNICORE middleware will be installed on PC cluster nodes in order to support test cases with DDSS Grid FTP, DDSS backup/archive, KoDaVis and WISDOM applications. When the applications are equipped with the G^{2}MPLS G.OUNI interface, they will be able to communicate with the control plane of the transport network in order to automatically provision optical paths and request required grid resources inside the local network and between other PHOSPHORUS test-beds.

A.2 CESNET

CESNET planes to deploy and verify NRPS in a testbed and later, when sufficiently stable versions of NRPS are available, in the operational network CESNET2. The test network (the local PHOSPHORUS testbed) is based on the CESNET Experimental Facility, built in previous years and including GOLE CzechLight as a part of the GLIF.

The PHOSPHORUS project has supposed some availability of GMPLS-capable software for Cisco equipment from the vendor, but a development roadmap has been changed and no implementation of GMPLS on Cisco equipment can be expected for the foreseeable future.

Working capacities (MMs) needed to overcome this obstacle has not been planned in PHOSPHORUS project and we are working on their evaluation. One feasible solution can be interfacing of Cisco equipment by an external GMPLS/G^{2}MPLS controller made of a PC host with a GMPLS/G^{2}MPLS implementation similar as in the PSNC case. Equipment already deployed supports and communicates with SNMP, CLI or TL1.

Open (programmable) photonic equipment developed outside of the PHOSPHORUS project can be used widely in different networks (not only NRENs) and it is being tested in the CESNET Experimental Facility and in the CESNET2 network. Two vendors are licensed for production of CzechLight Amplifiers. Evaluation of possibilities to use this equipment for PHOSPHORUS project is ongoing. Linux based operation allows access to all parameters of lighting devices, giving good possibilities of network management.

The transport network layer of the testbed and CESNET2 is build with the use of DWDM equipment with ROADMs, SONET/SDH cross-connects (MSPPs), Ethernet switches and routers and experimental optical equipment based on the planar lightwave circuit (PLC) technology. In addition, other equipment like optical amplifiers and tuneable dispersion compensators can be managed by GMPLS controllers (when building new circuits or lightpaths). In some cases (CESNET-developed optical switches) GMPLS capabilities (i.e. control software) can be embedded into equipment and external controllers may not be required.

Local links can use DWDM 1 Gbps or 10 Gbps channels or dark fibres if DWDM channels are not suitable for certain applications (like advanced modulation formats and/or transmission speeds above 10 Gbps). The core transport network CESNET2 is equipped with ROADMs therefore no electrical conversions are needed and it means format and protocol independency (terminology between ROADMs and OXCs is sometimes fuzzy as different people use different naming conventions). Ethernet switches can be used when VLANs have to be implemented. These switches have L3 capabilities (router/switch) and routing can be implemented for some special applications without any problems. Storage and computing resources are to be connected to the core network using 1 GE or 10 GE if needed.

One interesting resource is a direct 10 Gbps link between Praha (CESNET/CzechLight) and Chicago (StarLight): the link is a part of GLIF. The question how to utilize and distribute requested capacity among different applications is really important and it is not an easy task.

Although CESNET participation in PHOSPHORUS has been planned for testing of management of transport resources, it is possible to discuss connection of storage and computing resources in conjunction with GMPLS-enabled equipment deployed in both before mentioned environments – the CESNET Experimental Facility and the CESNET2 network. More information about these resources can be found for example on http://meta.cesnet.cz/cms/opencms/en/index.html. The use of some other resources can be discussed too (e.g. High Energy Physics cluster in Prague, PlanetLab, etc.).

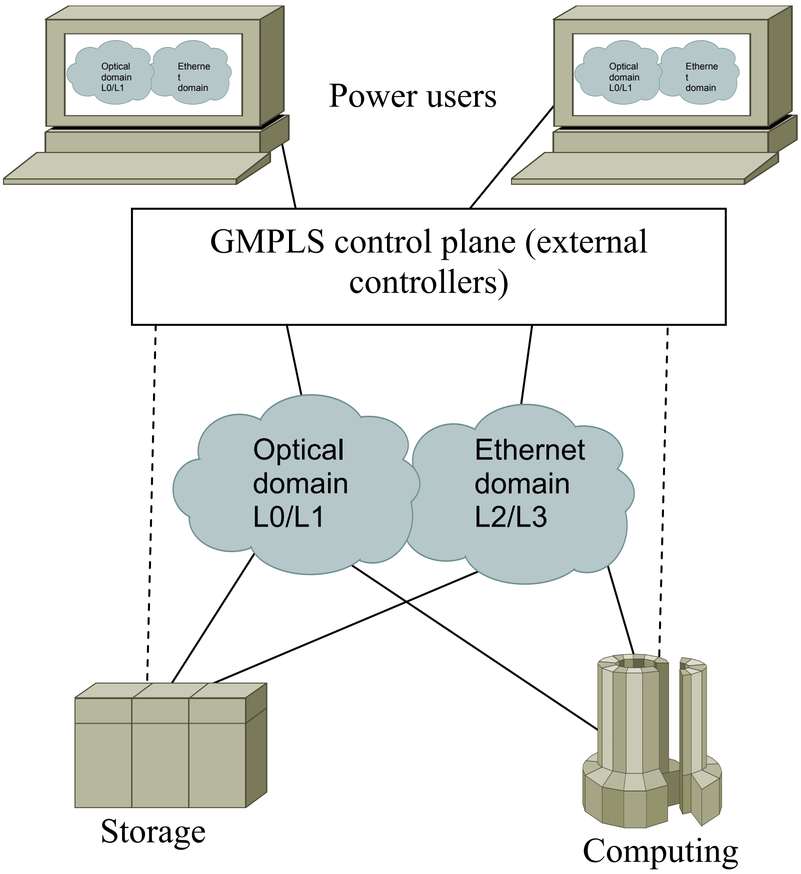

Figure A.1-2: CESNET case scenario

Appendix B NREN questionnaires

The following questionnaire has been defined in order to evaluate the requirements and possible plans of the NRENs towards GMPLS adoption in their infrastructures. In a wider perspective, questionnaire is aimed also to publicize the PHOSPHORUS project and its planned results to NREN managers and technical officers, thus preparing the framework for a possible deployment of PHOSPHORUS results in operational networks.

The questionnaire template is provided in the following sub-section. It was circulated to the wider NREN community during the TERENA European Future Networking Initiative Workshop, held on February 22nd 2007 in Amsterdam.

Some received answers are attached to this section, provided by two consortium internal partners (PSNC and CESNET) and four external partners (FCCN, DFN, GARR and HEAnet).

The low answering rate at the time of writing is assessed to be mainly due to the lack of running and public demonstrations of the PHOSPHORUS concept, which are planned for a later phase with respect to the ongoing design and implementation processes.

From an internal perspective, answers and interest towards GMPLS/G^{2}MPLS has been received by those NRENs who are currently missing any other control/management system for network and Grid resources. In fact, these NRENs are directly and actively involved in the design and development activities of WP2. On the contrary, other NRENs participating in PHOSPHORUS are more focused on the interoperability of their existing solutions with other systems (e.g. SURFnet with D-RAC) and, thus, have interest towards other PHOSPHORUS objectives, out of the WP2 scope.

Concerning the external dimension of the answers, FCCN, DFN, GARR and HEAnet official interest towards the PHOSPHORUS seamless and one-step reservation of Grid and network resources, as well as some informal declarations of interest towards the project results by other NREN managers and technical officers, supports the expectation of a wide impact of the project results on the NREN community. In this perspective two factors may be considered key for PHOSPHORUS take-up:

- The capability of PHOSPHORUS GMPLS/G^{2}MPLS to control efficiently the optical infrastructures integrated with the switched Ethernet. NRENs are planning upload towards these infrastructures through incremental steps: at first some reserved portions of their networks until a full operational deployment in service network.
- The smooth integration of PHOSPHORUS GMPLS/G^{2}MPLS with the Bandwidth on Demand service currently developed by most of the NRENs through the GÉANT2 project.

It is expected that early PHOSPHORUS dissemination of results may increase the rate of answers to the questionnaire and define a wide set of possible “customers” of the PHOSPHORUS GMPLS/G^{2}MPLS.

B.1 Template

<table>
<tr><td colspan="3">

NREN Questionnaire

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.

Contact: Artur Binczewski artur@man.poznan.pl

</td></tr>
<tr><td></td><td>Country</td><td></td></tr>
<tr><td></td><td>Name of the NREN</td><td></td></tr>
<tr><td></td><td>Contact person</td><td></td></tr>
<tr><td>1</td><td>Do you have GMPLS already deployed in your operational network?</td><td>Yes / No</td></tr>
<tr><td>2</td><td>One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.

Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).</td><td>Yes (date) / No</td></tr>
<tr><td>3</td><td>What would you use the GMPLS procedures primarily for?</td><td>Please choose from the following list:

• Transport services restoration
• Automatic and dynamic transport services provisioning to users through a UNI
• Automatic and dynamic transport services provisioning through adjacent operators</td></tr>
</table>

4	Which transport plane technologies do you want to cover with GMPLS?	*Please choose from the following list:* • IP • ETH • SDH/SONET • DWDM / ROADM / All optical switching • Other:
5	What kind of transport network devices would you like to empower with GMPLS?	*Please specify a vendor name and a type of the device:* • vendor, type, number
6	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth, advanced services, etc.)	*Please choose from the following list:* • Grids / Projects: o EGEE o ATLAS o VLBI o DEISA • Radio-astronomers • Others:
7	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G²MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Yes / No
8	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^2MPLS) in your network?	

9	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS Control Plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	Yes / No
10	Which transport plane technologies do you want to cover with an NRPS?	*Please choose from the following list:* • IP • ETH • SDH/SONET • DWDM / ROADM / All optical switching • Other:
11	Which processes would you like to automate using an NRPS?	
12	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	*Please specify a vendor name and a type of the device:* • vendor, type, number
13	Other comments and remarks	

B.2 Filled forms

B.2.1 CESNET2

NREN Questionnaire **PHOSPHORUS** **Lambda User Controlled Infrastructure for European Research** *Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.* *Contact: Artur Binczewski artur@man.poznan.pl*		
	Country	Czech Republic
	Name of the NREN	CESNET2
	Contact person	Vaclav Novak, Jan Radil
1	Do you have GMPLS already deployed in your operational network?	No
2	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	Yes, when available and secure
3	What would you use the GMPLS procedures primarily for?	*Please choose from the following list:* • Transport services restoration • Automatic and dynamic transport services provisioning to users through a UNI

4	Which transport plane technologies do you want to cover with GMPLS?	*Please choose from the following list:* • <u>IP</u> • <u>ETH</u> • <u>SDH/SONET</u> • <u>DWDM / ROADM / All optical switching</u>
5	What kind of transport network devices would you like to empower with GMPLS?	*Please specify a vendor name and a type of the device:* • vendor, type, number Cisco ONS15454 MSPP, ONS15454 MSTP
6	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth, advanced services, etc.)	*Please choose from the following list:* • Grids / Projects: o <u>EGEE</u> o ATLAS o VLBI o DEISA • Radio-astronomers • Others: <u>LHC-OPN</u>
7	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G^{2}MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	<u>Yes</u>
8	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^{2}MPLS) in your network?	<u>EGEE</u>

9	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS Control Plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	We plan to deploy in test network and later in our operational one.
10	Which transport plane technologies do you want to cover with an NRPS?	*Please choose from the following list:* • IP • ETH • SDH/SONET • DWDM / ROADM / All optical switching • Other:
11	Which processes would you like to automate using an NRPS?	BoD (GE/10GE lightpaths)
12	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	*Please specify a vendor name and a type of the device:* • vendor, type, number Cisco ONS15454 MSPP, ONS15454 MSTP
13	Other comments and remarks	No

B.2.2 PIONIER

NREN Questionnaire

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.

Contact: Artur Binczewski artur@man.poznan.pl

	Country	Poland
	Name of the NREN	PIONIER
	Contact person	Artur Binczewski
1	Do you have GMPLS already deployed in your operational network?	No
2	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	Yes (2008)
3	What would you use the GMPLS procedures primarily for?	*Please choose from the following list:* • Automatic and dynamic transport services provisioning to users through a UNI (the client application should be able to request the path via UNI) • Automatic and dynamic transport services provisioning through adjacent operators (MANs networks) • Transport services restoration

4	Which transport plane technologies do you want to cover with GMPLS?	*Please choose from the following list:*
		• <u>DWDM/ROADM</u>
		• Other:
		  • <u>Optical switches</u>
		  • <u>Ethernet switches supporting MPLS/VPLS</u>

5	What kind of transport network devices would you like to empower with GMPLS?	*Please specify a vendor name and a type of the device:* • Calient FibreConnect • Foundry XMR 8000 • Foundry XMR 16000 • Cisco 6500

6	Are there any users with non-average transport service requirements in your network?
	(e.g. users needing large bandwidth, advanced services, etc.)
		Please choose from the following list:
		• Grids / Projects:
		  o <u>EGEE</u>
		• Radio-astronomers:
		  o <u>Expres</u>
		  o <u>E-VLBI</u>
		• Others:

7	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G^2MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Yes
8	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^2MPLS) in your network?	<u>PSNC, TASK</u>

9	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS Control Plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	Yes (we would like to connect to GLIF with UCLPv2)

10	Which transport plane technologies do you want to cover with an NRPS?	• DWDM/ROADM
		• Other:
		  • Optical switches
		  • Ethernet switches supporting MPLS/VPLS

11	Which processes would you like to automate using an NRPS?	Bandwidth on Demand services
12	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	*Please specify a vendor name and a type of the device:* • Calient, FibreConnect • Foundry XMR 8000 • Foundry XMR 16000 • Cisco 6500
13	Other comments and remarks	

B.2.3 FCCN

NREN Questionnaire **PHOSPHORUS** **Lambda User Controlled Infrastructure for European Research** *Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.* *Contact: Artur Binczewski* artur@man.poznan.pl		
	Country	Portugal
	Name of the NREN	FCCN
	Contact person	João Nuno Ferreira
1	Do you have GMPLS already deployed in your operational network?	No
2	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	No. Presently we do not have requirements to run an operational layer of GMPLS. We may however deploy it on separate lambdas or fibres, for research and testing purposes, if our users request it. Please see annex.
3	What would you use the GMPLS procedures primarily for?	Taking into account that we don't have plans to have a GMPLS layer, we answer the following: 1- To have a fast and controlled way of provisioning services to advanced users. 2- For enabling networking researchers to test and "play" with it.

4	Which transport plane technologies do you want to cover with GMPLS?	Taking into account that we don't have plans to have a GMPLS layer, we answer the following: • IP (No) • ETH (Yes) • DWDM / ROADM / All optical switching (Yes)
5	What kind of transport network devices would you like to empower with GMPLS?	Taking into account that we don't have plans to have a GMPLS layer, we answer the following: • IP Routers, 12000,Cisco • IP Routers, M20, Juniper • Ethernet Switch, Foundry • Lambda Switches, not yet purchased.
6	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth, advanced services, etc.)	Presently no. Not known.
7	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G^2MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Our GRID community is in its infancy and it was decided recently that classical switched Ethernet services (1G and 10G) would be the preferred choice. There were no requirements for an extra layer of control like G/MPLS, as our network has a coherent Ethernet control plane.

8	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^2MPLS) in your network?	Portugal do not have any real HPC site yet. There are a couple of clusters being built right now. We are presently running a survey to map them. The most important locations are the following: - Minho University in Braga and in Guimarães. - Porto University - Aveiro University - Coimbra University - All Lisbon universities
9	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS Control Plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	We don't have any NRPS deployed, and don't have any plan to deploy an NRPS in our operational network in the near future.
10	Which transport plane technologies do you want to cover with an NRPS?	Taking into account that we don't have short term plans to have a NRPS layer, we answer the following: • IP (No) • <u>ETH</u> (Yes) • SDH/SONET (No) • DWDM / (Yes) ROADM / All optical switching • Other:
11	Which processes would you like to automate using an NRPS?	Taking into account that we don't have short term plans to have a NRPS layer, we answer the following: - L2-VPN provisioning - Lambda provisioning

12	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	• Not defined yet.
13	Other comments and remarks	

B.2.4 DFN

NREN Questionnaire

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.

Contact: Artur Binczewski artur@man.poznan.pl

	Country	Germany
	Name of the NREN	DFN
	Contact person	Kaufmann
1	Do you have GMPLS already deployed in your operational network?	No, only in testbed
2	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	No, only within testbed
31 4	What would you use the GMPLS procedures primarily for?	*Please choose from the following list:* • Transport services restoration • At first: Automatic and dynamic transport services provisioning to users through a UNI • Automatic and dynamic transport provisioning through adjacent operators

4	Which transport plane technologies do you want to cover with GMPLS?	*Please choose from the following list:* Probably: • ETH • SDH/SONET • DWDM / ROADM / All optical switching
5	What kind of transport network devices would you like to empower with GMPLS?	*Please specify a vendor name and a type of the device:* • Cisco router • DWDM/ROADM from various Vendors (in testbed)
6	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth , advanced services, etc.)	*Please choose from the following list:* • Grids / Projects: o EGEE o ATLAS o VLBI o DEISA • Radio-astronomers • Others: e.g. HDTV
7	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G^{2}MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Yes
8	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^{2}MPLS) in your network?	e.g. FZ Jülich

9	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS control plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	Yes, but only in testbed (ARGON)
10	Which transport plane technologies do you want to cover with an NRPS?	*Please choose from the following list:* • ETH • SDH/SONET • DWDM / ROADM / All optical switching
11	Which processes would you like to automate using an NRPS?	
12	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	*Please specify a vendor name and a type of the device:* • vendor, type, number Alcatel 1678 MCC Alcatel 1850 TSS Siemens hiT 7070 Sycamore SN 16000
13	Other comments and remarks	

B.2.5 GARR

NREN Questionnaire

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.

Contact: Artur Binczewski artur@man.poznan.pl

	Country	ITALY
	Name of the NREN	GARR
	Contact person	Mauro Campanella
15	Do you have GMPLS already deployed in your operational network?	No
16	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	Yes 4Q 2008
17	What would you use the GMPLS procedures primarily for?	• Transport services restoration • Automatic and dynamic transport services provisioning through adjacent operators
18	Which transport plane technologies do you want to cover with GMPLS?	• IP • ETH • DWDM / ROADM /
19	What kind of transport network devices would you like to empower with GMPLS?	• vendor, type, number Juniper, M320, MXseries ADVA, FS3000

20	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth , advanced services, etc.)	• Grids / Projects: o EGEE o ATLAS o VLBI o DEISA • Radio-astronomers • Others:
21	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G²MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Yes
22	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^2MPLS) in your network?	INFN-CNAF CINECA
23	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS control plane). These are the NRPS (Network Resource Provisioning Systems) considered under the Phosphorus framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	Yes, not before 2H2008
24	Which transport plane technologies do you want to cover with an NRPS?	• IP • ETH • DWDM / ROADM / All optical switching • Other: • Possibly SDH/SONET
25	Which processes would you like to automate using an NRPS?	Engineering and provisioning of circuits, resiliency

26	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	• vendor, type, number Juniper, M320, MXseries ADVA, FS3000
27	Other comments and remarks	Key issues for the use of third party software are: o short term maintenance and support policies (bugs fixing, new development to maintain compatibility with standards) o medium term support o software complexity and level of documentation available (manuals and in-line comments) o integration / interfaces with monitoring systems and control and management system/databases o proven interoperability with other existing systems

B.2.6 HEAnet

IreNREN Questionnaire

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Please fill-in the questionnaire. Don't hesitate to modify this document, e.g. by removing any options, which are not supported by your National Research and Education Network.

Contact: Artur Binczewski artur@man.poznan.pl

	Country	Ireland
	Name of the NREN	HEAnet
	Contact person	Victor Reijs
28	Do you have GMPLS already deployed in your operational network?	No
29	***One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by NREN AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.*** Would you be interested in running it in your operational network? If yes, please estimate a foreseen date (if possible).	Don't know, we spent time on related activities like GN2-JRA3 and UCLP related work on layer 0, 2 and 3. The multi-domain aspect is seen as essential. GMPLS can have a role, so who knows.
30	What would you use the GMPLS procedures primarily for?	*Please choose from the following list:* 1. Transport services restoration 2. Automatic and dynamic transport services provisioning to users through a UNI 3. Automatic and dynamic transport services provisioning through adjacent operators In some way indeed all of them, but I see that the easy of implementation starts with 1, 2 and 3. I wonder if GMPLS will work in the case 3. I see many large scale interworking protocols fail, and that seems still be open with regard to GMPLS. So if it is a real inter domain protocol, I don't know.

31	Which transport plane technologies do you want to cover with GMPLS?	*Please choose from the following list:* • IP • ETH • ROADM (using DWDM) Some equipment of our Ethernet service is not capable of GMPLS (3750) while another part is (760x). The IP is likely to be GMPLS capable (in middle of tender, so equipment not known yet). For ROADM we plan to do possibly a pilot this year (using GMPLS might be possible for this). Our present ADVA is not capable of GMPLS.
32	What kind of transport network devices would you like to empower with GMPLS?	*Please specify a vendor name and a type of the device:* See above. We have implemented UCLPv1.5 for the layer 2 network (760x and 3750), We hope to start a project MANTICORE for the IP equipment we will chose using UCLP(v2). But GMPLS is still interesting.
33	Are there any users with non-average transport service requirements in your network? (e.g. users needing large bandwidth , advanced services, etc.)	*Please choose from the following list:* • Grids / Projects: • Grid-Ireland
34	***The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID worlds. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G^{2}MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.*** Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?	Yes, that is also why we do UCLP (seeing network or virtual machine, etc systems as resources in WS environment).

35	Could you please indicate which HPC centers/GRID sites would be your potential partners to run Grid-enabled GMPLS protocol stack (G^2MPLS) in your network?	If we go for GMPLS (even in pilot, they are not informed yet even!): Grid-Ireland (or the possible continuation: e-INIS). But it might be we will do also something on UCLP with them, so there is competition.
36	***Another of the results of the PHOSPHORUS project will be the development of a set of interfaces that will allow the full interoperation between ARGON, DRAC and UCLPv2 (including interfaces to interact with the developments done within GN2-JRA3 and a standard GMPLS control plane). These are the NRPS (Network Resource Provisioning Systems) considered under the PHOSPHORUS framework. The results of the PHOSPHORUS project will allow users and applications to create end to end connections and advance reservations through different domains after their access is granted.*** Do you have an NRPS already deployed or you plan to deploy an NRPS in your operational or test network?	Yes: * In layer 2 we use our own tool BLUEnet for making p2p links over native ethernet and L2 MPLS VLL, we are implementing an interface towards Gn2-JRA3. * The above is also using UCLPv1.5 * We have worked on UCLPv1.5 for pure optical switch (OOO): Glimmerglass * We are thinking of UCLP(v2) for logical IP networks. * It might be we will make interface between UCLP(v2) and GN2-JRA3
37	Which transport plane technologies do you want to cover with an NRPS?	*Please choose from the following list:* • IP • ETH • ROADM • Virtual machines (as proposed in Federica).
38	Which processes would you like to automate using an NRPS?	Provisioning, monitoring
39	What kind of transport plane devices do you have (or are you going to have) with NRPS support?	*Please specify a vendor name and a type of the device:* Should be comparable equipment as stated above.

40	Other comments and remarks	Up to now we did not really invest much time in GMPLS (looked to be too much telco oriented). We know many NRENs already invest time in GMPLS, so we saw no need to do it and thus concentrated on other steps (UCLP: resource based networking and GN2-JRA3: allowing each NREN to keep its autonomy in protocols). It might be GMPLS is getting more mature, so perhaps it indeed some more time from HEAnet;-), to be sure we make the right decision.

Appendix C Questionnaires to super-computing centers

The following questionnaire has been defined in order to evaluate the requirements and plans of the Super-computing centers towards G^2MPLS adoption in their infrastructures.

The limited answers to the questionnaire at the time of writing are assessed to be mainly due to the lack of public demonstrations of the PHOSPHORUS concept and a late circulation of the document among the selected centers' representatives. The received answers from Barcelona Supercomputing Center and CINECA have been included in the remainder of this section.

The dissemination of the PHOSPHORUS (G^2MPLS) concepts through WP2 design deliverables, participation in public events and demonstrations in the worldwide test-bed is expected to increase the interest of the HPC community towards the project results.

C.1 Template

PHOSPHORUS

Lambda User Controlled Infrastructure for European Research

Questionnaire for Super-computing Centers

Contact: Artur Binczewski artur@man.poznan.pl

What is PHOSPHORUS?

A new generation of scientific applications is emerging that couples scientific instruments, data and high-end computing resources distributed on a global scale. Developed by collaborative, virtual communities, many of these applications have requirements such as determinism (e.g. guaranteed QoS), shared data spaces, large data transfers, that are often achievable only through dedicated optical bandwidth.
High capacity optical networking can satisfy bandwidth and latency requirements, but software tools and frameworks for end-to-end, on-demand provisioning of network services need to be developed in coordination with other resources (CPU and storage) and need to span multiple administrative and network technology domains.
In response to the above requirements, PHOSPHORUS will address some of the key technical challenges to enable on-demand e2e network services across multiple domains. The PHOSPHORUS network concept and test-bed will make 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.
PHOSPHORUS will enhance and demonstrate solutions that facilitate vertical and horizontal communication among applications middleware, existing Network Resource Provisioning Systems, and the proposed Grid-GMPLS Control Plane.

Aim of this survey

The aim of this survey is to gather information about the characteristics of the environments in which PHOSPHORUS results may be deployed. Results of this survey will help refine PHOSPHORUS software and network requirements.
We very much appreciate your answers to the questions below. If you feel you cannot or don't know how to answer some question, please leave all options blank.

PHOSPHORUS Team

Please provide the following information:

Country:
Organisation:
Contact Person:

1) In what Grid projects your organisation is involved? Please indicate the names and URLs to project web sites.

(a) Example1 http://localhost
(b)

2) Please provide a high level architectural description of mentioned projects with major emphasis on the adopted middleware, number of involved locations and network interconnections between them.

3) One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by your AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.

Would you be interested in testing PHOSPHORUS GMPLS in your Grid environment? If yes, please estimate a foreseen date (if possible).

Yes (date: ……) / No

4) What would you use the GMPLS procedures primarily for?

Please choose from the following list:

(a) Automatic and dynamic network transport services provisioning to end-users through a UNI
(b) Automatic and dynamic transport services provisioning through adjacent network operators
(c) Transport services restoration

5) Regarding the network architecture supporting your GRID, please indicate which transport technologies may be covered with GMPLS.

Please choose from the following list:

(a) IP
(b) ETH
(c) SDH/SONET
(d) DWDM / ROADM / All optical switching
(e) Other:

6) From the perspective of your Grid community, please indicate what network devices you would like to empower with GMPLS.

Please specify a vendor name and a type of the device:

(a) vendor, type, number
(b)

7) The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G2MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.

Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?

Yes / No

8) Other comments and remarks:

C.2 Filled forms

C.2.1 BSC

Please provide the following information:
Country: Spain
Organisation: Barcelona Supercomputing Center
Contact Person: Franesc Guim Bernat

1) In what Grid projects your organisation is involved? Please indicate the names and URLs to project web sites.

- http://www.hpc-europa.org/
- http://www.coregrid.net/
- http://www.xtreemos.org
- http://www.beingrid.com/
- http://www.eu-brein.com/
- http://lagrid.fiu.edu/
- http://asds.dacya.ucm.es/GridMiddleware/
- http://www.sorma-project.org/
- http://recerca.ac.upc.edu/eDragon/
- http://www.xtreemfs.org/

(These projects are related to the Computer Science group)

2) Please provide a high level architectural description of mentioned projects with major emphasis on the adopted middleware, number of involved locations and network interconnections between them.

- http://www.hpc-europa.org/ The objective of this scheme is to support the integrated provision of innovative HPC services to the research community at a European level. Middleware used are: Globus & UNICORE and GRIA. 6 different locations. The interconections are done over TCP/IP (http).
- http://lagrid.fiu.edu/ LA Grid, pronounced "lah grid," is an international multi-disciplinary research community and virtual computing grid enabling institutions and industry to extend beyond their individual reach to facilitate collaborative IT research, education and workforce development. The main goal to design infrastructures for the interoperability of different brokering architectures. The current number of different locations is: 4. Middleware used: UNICORE, Globus. The interconnections are done also over TCP/IP (http)
- http://www.xtreemos.org The XtreemOS project aims at investigating and proposing new services that should be added to current operating systems to build Grid infrastructure in a simple way. XtreemOS targets the Linux well-accepted open source operating system extending it to Grid with native support for virtual organizations. One of the most important challenges in XtreemOS is the identification of the basic functionalities which are to be embedded in the Linux kernel. 12 Academic partners and 9 industrial partners, each of them is a potential location for the computational resources of the project. (see the file

http://www.xtreemos.org/intranet/communication-tools/flyer_v5.ppt/download) . The middleware used will be the same xtreemOS.

3) One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by your AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.

Would you be interested in testing PHOSPHORUS GMPLS in your Grid environment? If yes, please estimate a foreseen date (if possible).

No (although it looks interesting it is currently is out of scope of our research interests)

4) What would you use the GMPLS procedures primarily for?

Please choose from the following list:

(a) Automatic and dynamic network transport services provisioning to end-users through a UNI
(b) Automatic and dynamic transport services provisioning through adjacent network operators
(c) Transport services restoration

5) Regarding the network architecture supporting your GRID, please indicate which transport technologies may be covered with GMPLS.

Please choose from the following list:

(a) IP
(b) ETH
(c) SDH/SONET
(d) DWDM / ROADM / All optical switching
(e) Other:

6) From the perspective of your Grid community, please indicate what network devices you would like to empower with GMPLS.

Please specify a vendor name and a type of the device:

(c) vendor, type, number
(d)

7) The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G2MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.

Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?

Yes / No

8) Other comments and remarks:

C.2.2 CINECA

Please provide the following information:
Country: ITALY
Organisation: CINECA
Contact Person: Andrea Vanni a.vanni@cineca.it

1) In what Grid projects your organisation is involved? Please indicate the names and URLs to project web sites.

- DEISA www.deisa.org
- AWARE www.a-ware.org
- BeinGrid www.beingrid.eu
- HPC-Europa http://www.hpc-europa.org

2) Please provide a high level architectural description of mentioned projects with major emphasis on the adopted middleware, number of involved locations and network interconnections between them.

a) DEISA 11 sites interconnected by high performance network 10Gbit/s with Global parallel filesystem. Middleware deployed Unicore, DESHL, gridftp et al.

b) AWARE Grid Access via EngineFrame portal and Unicore/GS 5 partners.

c) BeinGrid. Consortium of 75 partners. Implementing and deploying Grid solution in industrial key sectors.

d) HPC-Europa. The objective of this scheme is to support the integrated provision of innovative HPC services to the research community at a European level. Middleware used are: Globus & UNICORE and GRIA. 6 different locations. The interconections are done over TCP/IP (http).

3) One of the results of the PHOSPHORUS project will be a free and widely available Open Source implementation of GMPLS protocol stack. With the framework provided by PHOSPHORUS, users and client applications (previously authorized by your AAA component) will be allowed to make ad-hoc and advance reservations of the network resources.

Would you be interested in testing PHOSPHORUS GMPLS in your Grid environment? If yes, please estimate a foreseen date (if possible).

No (for the time being)

4) What would you use the GMPLS procedures primarily for?

Please choose from the following list:

5) Regarding the network architecture supporting your GRID, please indicate which transport technologies may be covered with GMPLS.

Please choose from the following list:

(a) <u>IP</u>
(b) <u>ETH</u>
(c) SDH/SONET
(d) DWDM / ROADM / All optical switching
(e) Other:

6) From the perspective of your Grid community, please indicate what network devices you would like to empower with GMPLS.

Please specify a vendor name and a type of the device:

7) The added value given by PHOSPHORUS to GMPLS protocol stack is a possibility of integration between NETWORK and GRID. The PHOSPHORUS implementation of GMPLS protocol stack will provide a seamless and efficient access to both grid and network resources, by the use of a network signaling and routing protocols. It will allow to perform discovery of the grid and network resources available in your network and make a resources booking/allocation using G2MPLS (Grid-GMPLS) procedures. By exposing the GMPLS network discovery and reservation capabilities via standard Grid protocols, it will also enable Grid middleware (resource brokers and schedulers) to use network connections as Grid resources.

Would you, together with your Grid communities, be interested in a comprehensive mechanism for seamless and single-step Grid and Network resources reservation?

Yes / <u>No</u>

8) Other comments and remarks: