

# Experimental implementation of Grid enabled ASON/GMPLS Networks

Eduard Escalona<sup>1</sup>, Jordi Perelló<sup>1</sup>, Salvatore Spadaro<sup>1</sup>,  
Jaume Comellas<sup>1</sup>, and Gabriel Junyent<sup>1</sup>

<sup>1</sup> Optical Communications Group, TSC Dept.  
Universitat Politècnica de Catalunya (UPC), Jordi Girona, 1-3  
08034 Barcelona, Spain  
{escalona, jperello, spadaro, comellas, junyent}@tsc.upc.edu

**Abstract.** Current GMPLS protocols can be used in a Grid enabled environment by properly defining a new interface between Grid and Transport layers. However, the asymmetric nature of Grid traffic makes GMPLS inefficient when establishing bidirectional connections. The User to Grid traffic is unlikely to be the same as the Grid to User traffic; whilst the location of the processing sites may be of no importance to the user, the destination address of the Grid to User downstream is the same for all the different flows akin to the same job. This may result in asymmetric resource allocation throughout the network. We propose and experimentally evaluate an implementation for a middleware interface as well as a novel scheme that reduces blocked job requests due to lack of network resources in such environment.

**Keywords:** GMPLS, Control Plane, Grid, Asymmetry, Testbed.

## 1 Introduction

As applications demand for higher processing and storage capacity, the concept of Grids become more and more useful. Single machines cannot cope with the amount of capacity required for next generation applications and most small to medium enterprises (SME) cannot afford specialized hardware to provide such services. Moreover, supercomputers are a very expensive solution and are not scalable. Grid architectures allow the division of huge tasks into smaller jobs to process them separately in remote locations and then reassemble them to obtain the resulting information.

The main problem arises when the amount of data to be transmitted is large. Optical networks have made large bandwidths available. Nevertheless, if Grids networks are to better utilize this capacity, mechanisms to efficiently manage the optical resources are required.

The emerging Grid scenario requires a closer cooperation of the service and transport layers as bandwidth requirements of applications increase. Lately, several studies have been done to improve the performance of optical networks and considerable effort made to define optimal ways to manage and control network

resources. To this effect, the Internet Engineering Task Force (IETF) has developed a set of protocols defined as the Generalized Multi Protocol Label Switching (GMPLS [1]) paradigm. GMPLS, being an extension of MPLS, supports different types of switching; it has mechanisms that allow the establishment of optical channels in a fast, effective and dynamic way, acting as the control plane of the automatically switched optical networks (ASON [2]). Grid networks can take advantage of GMPLS capabilities as it fits the requirements detailed by the standardization organisms [3].

To implement the required procedures that enable the establishment and deletion of connections across heterogeneous networks, the Optical Internetworking Forum (OIF) defines a User to Network Interface (UNI). The UNI describes the service control interface between user and transport network addressing the procedures required for its interoperation based on the adaptations of two GMPLS signaling protocols, Label Distribution Protocol (LDP) and Resource ReserVation Protocol with Traffic Engineering extensions (RSVP-TE).

In order to extend the services offered by the UNI to facilitate on-demand access to Grid services, the Open Grid Forum (OGF) settled the concept of the Grid UNI (G.UNI) abstractly defining procedures and functionality considerations [3]. As the implementation of the G.UNI is open, we have designed an overlay architecture based on Web Services that will be further described in Section 2.

However, the deployment of Grid Services in a GMPLS-based transport network discovers that the asymmetric nature of the Grid traffic requires a more efficient management of the resources. When a job request is split into different connection requests the data to be processed is forwarded to diverse locations. In turn, the processed information should be sent back to the originating user or to another specific address. In such environment, unidirectional, symmetric and asymmetric bidirectional connections may be required when handling distributed processing. Within this scenario, the resource allocation becomes asymmetric, which means that, in a network link, available bandwidth may not be found for both directions.

Current GMPLS protocols support the establishment of unidirectional and bidirectional connections. However, when creating bidirectional connections, GMPLS is constrained to use the same path for downstream and upstream. Furthermore, the upstream label is chosen using local information. In Lambda Switched Capable networks (LSC), this results in a high blocking probability when wavelength conversion is not available as the same wavelength has to be maintained along the whole path.

In this paper, a novel connection scheme for GMPLS networks is proposed in order to reduce the rate of blocked job requests due to the lack of available network resources. It is based on calculating routes for both downstream and upstream connections separately, allowing the diversification of routes for bidirectional connections. It is beneficial when no upstream resources are available along the same path of the downstream connection. The proposed scheme has been experimentally tested over a real ASON/GMPLS network testbed, CARISMA [4].

This paper is organized as follows. Section 2 defines the interface between Grid and the optical control plane as well as our specific implementation in the CARISMA testbed. Section 3 addresses the asymmetry issue in GMPLS networks and Section 3.1 describes a novel connection scheme to improve its efficiency. Some experimental

results are shown in Section 3.2 for performance evaluation. Finally, Section 4 concludes the paper.

## **2 Interfacing Grid and GMPLS**

The Open Grid Forum (OGF) defines the requirements for an optical infrastructure for grid and describes the characteristics of a Grid User network Interface (G.UNI) [3]. The G.UNI is the interface between the Grid users and the network to dynamically allocate and provision bandwidth and Grid resources on request.

The OGF groups G.UNI functionalities in signaling and transport. Signaling deals with bandwidth allocation, scheduling, AAA (Authentication, Authorization and Accounting) information, resilience and propagation of service events. Transport deals with traffic classification, traffic shaping, grooming and security.

The first step in the introduction of the GUNI is the integration of the optical control plane following an overlay model in which Grid user applications will not have knowledge of the network topology. The next step is a peer model in which users and grid clusters are also elements of the network. This paper deals with the implementation, deployment and evaluation of the overlay model at our research premises in the CARISMA testbed.

### **2.1 The CARISMA Testbed**

The CARISMA project was initiated in 2002 as an initiative to build a high performance Wavelength Division Multiplexing (WDM) based network to be used as a field-trial for the integration and evaluation of the current emerging innovative technologies. It is intended to provision bandwidth on demand while ensuring Quality of Service (QoS) between IP networks. The CARISMA testbed implements the ASON architecture based on a GMPLS control plane.

The control plane has been built associated to the transport plane, thus having a Signaling Communication Network (SCN) formed by Ethernet point-to-point links, following the same architecture as the transport plane. In such point-to-point architecture, only one control channel between Optical Connection Controllers (OCCs) is supported over each interface. Particularly, the protocols implemented are RSVP-TE [5] for signaling, OSPF-TE [6] for IP routing and information dissemination and Link Management Protocol (LMP) for discovery and control channel maintenance. These protocols have been implemented using Linux-based routers (Pentium 4 – 2GHz) running the GMPLS control plane and forming the OCCs.

CARISMA supports both soft-permanent and switched connections as defined by ASON. The management plane, responsible of the establishment of soft-permanent connections and the monitoring of the network topology and equipment status, uses the Simple Network Management Protocol (SNMP) for the communication with the transport and control planes. For the establishment of switched connections, OIF UNI 1.0 has been implemented.

## 2.2 G.UNI Implementation

The design of the G.UNI in the CARISMA network allows the Grid application to decide itself the type of connection required for a better performance. For the implementation we have used Java and Globus Toolkit 4 (GT4) to build the G.UNI as a WSRF (Web-Service Resource Framework) Grid Service. The result is a Web Service capable of accepting requests for optical channel establishments. Initially just three basic services have been defined to be supported by the G.UNI: Connection Setup, Connection Tear-Down and Connection Status Query. The functional block of the G.UNI that handles the requests is called Lightpath Provisioning Server (LP Server). The overlay G.UNI implementation should support different service invocation reference configurations likewise to the OIF UNI, namely Direct and Indirect Invocation Models. Under the direct model, the Grid User (or Grid Broker) invokes the transport services directly over the G.UNI (Fig. 1).

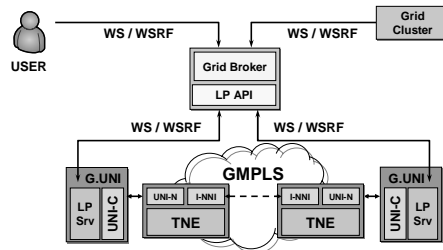


Fig. 1: G.UNI Direct Invocation Model

In the Indirect Invocation model a G.UNI proxy acts on behalf of one or more clients (Fig. 2). The interface between the G.UNI and the GMPLS control plane is open to any implementation. Specifically, our design accepts SNMP (soft-permanent connections) and OIF UNI 1.0 (switched connections) requests and uses the Indirect Invocation model.

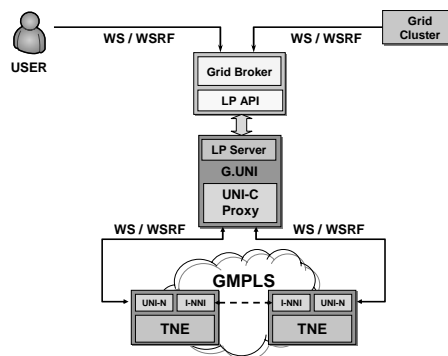


Fig. 2: G.UNI Indirect Invocation Model

### 2.3 Preliminary Results

In order to evaluate the performance of the designed G.UNI a Grid client has been implemented to emulate a Grid application requesting for optical connections. Results regarding setup latency prove the effectiveness of our implementation (Table 1).

**Table 1:** Connection Setup Time Flow using G.UNI

Action	Signaling Protocol	Delay
Client → G.UNI (LP Server)	SOAP (WS)	32 ms
G.UNI (LP Server) → UNI-C Proxy	UNIX Socket	13 ms
UNI-C Proxy → UNI-N (Source)	RSVP UNI 1.0r2	8 ms
UNI-N (Source) ↔ UNI-N (Dest)	RSVP-TE	27 ms
UNI-N (Dest) → UNI-C Proxy	RSVP UNI 1.0r2	8 ms
TOTAL		88 ms

## 3 Grid asymmetry impact over GMPLS

If we think about typical Grid applications like video streaming, Grid computing or storage we can easily see that the traffic generated is mainly asymmetric. In a general network with Grid applications over an optical control plane, in which unidirectional, symmetric and asymmetric bidirectional connections have to coexist, the distribution of network resources becomes also asymmetric (i.e. available bandwidth is not the same for the upstream and downstream on the same link).

When using current GMPLS standards, bidirectional connections are established adding the RSVP Upstream Label (UL) object to the Path messages so upstream resources are reserved hop by hop using local wavelength information [5]. With sparse (or lacking) wavelength conversion capabilities it results in very high blocking rates. Previous studies propose methods to select, if available, any free upstream resources (e.g. wavelengths) for the chosen path using an Upstream Label Set [7], or Global Wavelength information [8] (GW). Despite providing better performance than the UL scheme, they are constrained to use the same path for upstream and downstream connections. In case of asymmetric allocation of the network resources, connection requests are then blocked if not enough bandwidth is available on the upstream direction of the connection. As a further contribution, in this paper we propose a novel connection scheme for GMPLS networks to face the impact of the Grid traffic asymmetry.

### 3.1 UR Connection Establishment Scheme

The scheme proposed (UR) consists in calculating a new upstream route taking into account the actual available resources also for the upstream connection. Our scheme allows choosing a disjoint route when a bidirectional path can not be established because of the lack of resources in one of the directions. The signaling schemes are the same used in [9]. The upstream tunnel signaling is automatically triggered after the egress node has calculated an available upstream route. Although this kind of

signaling doubles the setup time, this is not the predominant issue in a circuit-switched environment as flooding resource information updates take substantially longer.

In addition, we have also implemented a routing algorithm that takes into account the actual link usage, routing the traffic through the less congested links. Flooding capabilities of the OSPF-TE protocol are used to propagate the link usage information by means of Link State Advertisements (LSAs) assigning a metric field to the data links (wavelengths). This field has value zero when all data links on a TE link are free. When a connection is established and some wavelengths are reserved, the metric value is increased for all the data links belonging to the same TE link than the reserved one. This information, after disseminated, is used to calculate less congested routes using (1) and (2) for every wavelength following the wavelength continuity constraint and applying the Dijkstra algorithm. Note that, in (1)  $C_j$  is the cost for the route  $j$ ,  $H_j$  is the number of hops for the route  $j$ ,  $DL_i$  is the number of data links in the link  $i$  (total or free) and  $n$  is the weight given to the new optical metric value added. The higher value of  $n$ , the more importance is given to the optical metric. The resulting route is the one with minimum cost of all the computed routes. In our studies we have used a value of  $n$  equal to 3 in order to assign more weight to the optical metric than to the distance without making the number of hops negligible.

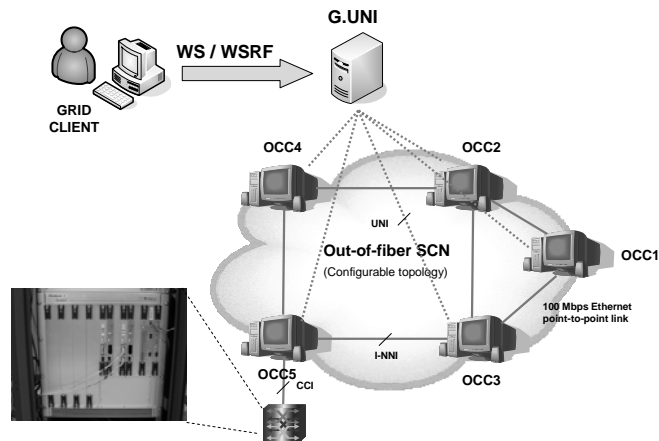
$$C_j = H_j + \sum_{i=link_1}^{i=link_{H_j}} (DL_i^{TOT} - DL_i^{FREE}) \cdot n \quad (1)$$

$$C_F = \min \{C_j\} \quad (2)$$

Studies with different weight values should be done to observe the effects over the network performance. It is important to note that no extra packet or byte overhead is added as the metric is included in the metric field of the OSPF opaque LSAs and routes are updated upon route changes. We call this algorithm the Optical Metric algorithm (OM).

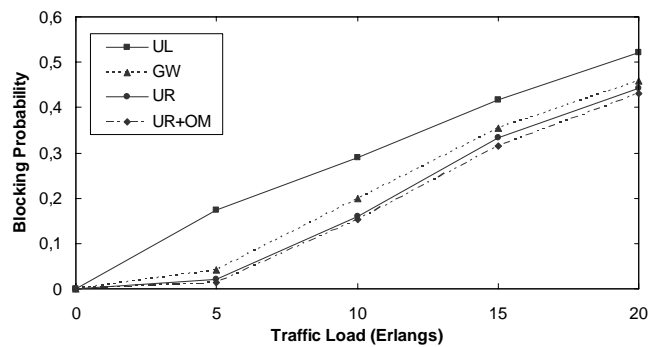
### 3.2 Performance Evaluation

We have carried out some experiments to evaluate the performance of the proposed scheme for Grid networks. The meshed-based network scenario shown on Fig. 3 composed by 5 nodes and 4 bidirectional wavelengths per link (8 data links) has been considered. To emulate the asymmetric allocation of the network resources (typical in a Grid network context), we assume that three permanent unidirectional connections using one wavelength each (OCC2-OCC1, OCC3-OCC2 and OCC4-OCC5) are initially established.



**Fig. 3.** CARISMA testbed configuration

Three different schemes to establish bidirectional connections have been compared experimentally as a function of the traffic load, considering blocking probability as the figure of merit: the standard GMPLS scheme with a First Fit wavelength assignment (UL), the GW scheme and the UR scheme. We have also evaluated the improvement of the performance of the UR scheme with the OM algorithm (UR+OM). Grid connections requests arrive with an exponentially distributed inter arrival time (IAT) of mean 20s. The connection holding time (HT), also exponentially distributed, is varied to generate different traffic load. The requests are set up between pairs of nodes chosen randomly so the traffic load is distributed uniformly among the nodes.



**Fig. 4.** Blocking probability

As a sample of the obtained results, Fig. 4 shows the resulting curves for the four compared schemes (UL, GW, UR, UR+OM). These results have been obtained using

a statistically relevant number of connections (10000). GW improves the results of the standard GMPLS implementation (UL) as expected, as it uses global wavelength information to maintain the wavelength constraint.

Nonetheless, in the considered network scenario, UR shows better performance than the GW scheme, since resource allocation is not symmetric. In fact, when no upstream resources are available on the same route, the UR scheme is able to select a diverse one. The improvement of the performance on the average in the tested scenario is about 4% which should increase in conditions of more traffic asymmetry.

It can be observed, on the other hand, that the OM algorithm allows to further improve performance of the UR scheme in about 4% with high traffic loads. Please note that UR and UR+OM have the same behavior for low traffic loads. This is due to the fact that UR+OM may use longer paths than UR to avoid links that are already carrying traffic for a better load balance. However, in low loaded networks these links are not as prone to be saturated as with high traffic loads. This effect compensates the gain obtained with the wavelength occupancy distribution.

Results regarding average lambda occupancy per link have also been taken. Lower differences between maximum and minimum occupancy values for the UR+OM scheme can be seen, which means that this scheme offers a more compensated traffic distribution among links than GW. Whereas the GW scheme makes some links more loaded than others, OM tries to compensate these values, shortening the difference and as a consequence, providing a better load balance.

## 4 Conclusions

The amount of bandwidth required for the emerging intensive grid applications makes necessary to build a service-aware optical transport network. As GMPLS seems to fulfill the requirements needed for a Grid applications correct performance, a Grid middleware has to be defined and implemented to merge service and optical transport environments.

Experimental results show that the obtained setup delays are suitable for the requirements of Grid applications.

Moreover, a novel connection scheme for bidirectional connections in multi-service networks has been presented to address the asymmetric nature of Grid networks. It is based on calculating a separate route for the upstream connection allowing the selection of a diverse path when no upstream resources are available along the same path used by the downstream connection. The results obtained show that the UR scheme outperforms the already proposed GW and UL schemes and it is further improved if a routing algorithm that takes into account the global wavelength availability is also considered for optimizing the load balance. Ongoing further studies will consider larger networks and varied topologies.



## Acknowledgements

The work reported in this paper has been partially supported by the Spanish Science Ministry through Project "Red inteligente GMPLS/ASON con integración de nodos reconfigurables (RINGING)", (TEC2005-08051-C03-02), and by the i2CAT Foundation ([www.i2cat.cat](http://www.i2cat.cat)).

## References

1. E. Mannie: Generalized Multi-Protocol Label Switching (GMPLS) Architecture, RFC 3945, Oct. 2004.
2. ITU-T Rec. G.8080/Y.1304, Architecture for the Automatically Switched Optical Network (ASON), November 2001
3. D. Simeonidou et al.: Optical Network Infrastructure for Grid, GFD-I.036, March 2004
4. J. Perelló et al.: Control Plane protection using Link Management Protocol (LMP) in the ASON/GMPLS CARISMA network, IFIP Networking 2006, May 2006.
5. L. Berger: GMPLS Signaling RSVP-TE Extensions, RFC3473, January 2003.
6. D. Katz et al.: Traffic Engineering (TE) Extensions to OSPF Version 2, RFC3630, September 2003
7. E. Oki et al.: Bidirectional Path Setup Scheme Using an Upstream Label Set in Optical GMPLS Networks, IEICE Trans. Communic., vol E87-B.
8. R. Martinez et al.: Experimental GMPLS-based Dynamic Routing in All-Optical Wavelength-Routed Networks, Proc. ICTON 2005, Barcelona.
9. E. Escalona et al.: Establishing Source-Routed Bidirectional Connections over the Unidirectional ASON/GMPLS CARISMA Testbed, Proc. ICTON 2005, Barcelona.