Grid Optical Burst Switched Networks (GOBS)

Grid High Performance Networking Research Group	Dimitra Simeonidou & Reza Nejabati (Editors) University of Essex
GRID WORKING DRAFT	Lina Battestilli North Carolina State University
draft-ggf-ghpn-GOBS-0	Piet Demeester Marc De Leenheer Bart Dhoedt University of Gent
Category: Informational Track	Gigi Karmous-Edwards MCNC, Research & Development Institute
http://forge.gridforum.org/projects/ghpn-rg/	Ken-ichi Kitayama OSAKA University
	Masafumi KOGA NTT lab
	Hiroshi Onaka Fujitsu Japan
	Yongmei Sun Tokyo Univeristy
	Anna Tzanakaki, Ioannis Tomkos AIT
	Emmanuel (Manos) Varvarigos University of Patras
	Luca Valcarenghi Scuola Superiore Sant'Anna

Status of this Memo
This memo provides information to the Grid community in the area of high performance networking. It does not define any standards or technical recommendations. Distribution is unlimited.

Comments: Comments should be sent to the GHPN mailing list (ghpn-wg@Gridforum.org).

1. INTRODUCTION 3

2. OPTICAL BURST SWITCHING, A REALISTIC TECHNOLOGY FOR GRID NETWORKING3

2.1 SPECIFIC IMPLEMENTATIONS 4

2.1.1 OBS for consumer Grid applications 4

3. CONTROL PLANE AND SIGNALLING FOR GRID-OBS 6

3.1 CONNECTION SETUP MECHANISMS........ 7

3.2 QOS PROVISIONING IN GRID-OBS NETWORKS 10

3.3 CONSTRAINED BASED ROUTING INCORPORATED IN THE OBS CONTROL 13

3.4 SPECIFIC IMPLEMENTATION OF CONTROL PLANE 13

3.4.1 Control plane issues for consumer Grid application 13

3.4.2 Reliable transport control technology for Grid-OBS........ 14

3.4.3 Ultra fast optical path setup protocol and its optical implementations 14

4. GRID-OBS NETWORK ELEMENTS 15

4.1 CORE OBS ROUTER 15

4.2 EDGE OBS ROUTER 16

5. ADVANCED NETWORK CONCEPTS AND SOLUTIONS (PROGRAMMABLE, ACTIVE, SELF ORGANISED GRID OBS NETWORK) 18

5.1 SELF-ORGANISED OBS NETWORK FOR CONSUMER GRIDS 18

5.2 PROGRAMMABLE OPTICAL BURST SWITCHED NETWORK 20

6. SECURITY ISSUES IN GRID-OBS NETWORKS 23

7. REFERENCES 23

1. Introduction

Optical networking for the Grid computing is an attractive proposition offering huge amount of affordable bandwidth and global reach of resources [1]. Currently, Grid computing using optical network infrastructure is dedicated to a small number of well known organizations with extremely large jobs (e.g. large data file transfers between known users or destinations [1]. Due to the static or semi-static nature of this type of Grids, long-lived wavelength paths between clients and Grid resources with centralized job management strategies are usually deployed (Lambda Grids). This type of Grid networking relies on carrier provision of optical network resources while the Grid users have no visibility of the lambda infrastructure. In other words, the Grid user is not able to setup paths over the optical Grid network.

As Grid applications evolve, the need for user controlled network infrastructure is apparent in order to support emerging dynamic and interactive services. Examples of such applications may be high resolution home video editing, real-time rendering, high-definition interactive TV, e-health and immersive interactive learning environments. These applications need infrastructures that makes vast amount of storage and computation resources potentially available to a large number of users. Key for the future evolution of such networks is to determine early on the technologies, protocols, and network architecture that would enable solutions to these requirements.

In an attempt to address this problem, in this draft novel network paradigms and solutions based on the optical burst switching are discussed.

2. Optical burst switching, a realistic technology for Grid networking

Optical burst switching (OBS) is a promising technology for the future networks where the bandwidth needs to be accessible to users with different traffic profiles. The OBS technology combines the advantages of optical circuit switching and optical packet switching [2]. An optical burst is usually defined as a number of continuous packets destined for a common egress point. The burst size can vary from a single IP packet to a large data set at milliseconds time scale. This allows for fine-grain multiplexing of data over a single wavelength and therefore efficient use of the optical bandwidth through sharing of resources (i.e. light-paths) among a number of users. The fundamental premise of OBS technology is the separation of the control and data planes, and the segregation of functionality within the appropriate domain (electronic or optical). Prior to data burst transmission a Burst Control Packet (BCP) is created and sent towards the destination by an OBS ingress node (edge router). The BCP is typically sent out of band over a separate signalling wavelength and processed at intermediate OBS routers. It informs each node of the impending data burst and setup an optical path for its corresponding data burst. Data bursts remain in the optical plane end-to-end, and are typically not buffered as they transit the network core. The bursts' content, protocol, bit rate, modulation format, encoding are completely transparent to the intermediate routers. The main advantages of the OBS in comparison to the other optical networking schemes are that: a) unlike the optical wavelength switched networks the optical bandwidth is reserved only for the duration of the burst; b) unlike the optical packet switched network it can be bufferless.

The OBS technology has the potential to bring several advantages for Grid networking:

- Native mapping between bursts and Grid jobs: the bandwidth granularity offered by the OBS networks allows efficient transmission of the user's jobs with different traffic profiles
- Separation of control and data plan: this allows all-optical data transmission with ultra-fast user/application-initiated light-path setup
- Electronic processing of the burst control packet at each node: this feature can enable the network infrastructure to offer Grid protocol layer functionalities (e.g. intelligent resource discovery and security)

2.1 Specific implementations

2.1.1 OBS for consumer Grid applications

For the average home user today, the network cannot sustain Grid computing. With a home access bandwidth of only a few Mbps, to at most 100 Mbps download speeds, and an order of magnitude smaller upload speeds, transmission of jobs would simply take too long. However, if the current trend holds and bandwidth availability (doubling each year) keeps growing faster than the computing power (at most doubling every 18 months) of an average end user, tapping into the Grid at home becomes viable.

Let us assume that in such a future Grid, home users are connected through a symmetrical access link offering a bandwidth of about 2.5 Gbps (in the optical range). While this kind of bandwidth is certainly not readily available to end-users at the time of writing, extrapolation of past trends shows that within 15 years such an evolution can be expected. Indeed, a typical broadband connection offers around 4 Mbps download speeds and 512 Kbps upload speeds. This means download bandwidth will have reached 2.5 Gbps within the next 10 years, and the same upload bandwidth will be available within 15 years. In analogy, if computational capacity doubles every 18 months, an increase in high-end desktop PC performance with a factor in the order of magnitude 210 should be envisaged. The resulting processing power will offer the possibility to process extremely demanding applications (by today's standards) on an ordinary desktop PC. However, as we will show, it is reasonable to assume that application demands will experience a similar increase in their requirements, making it unfeasible to execute them locally. The needed aggregate power for these applications is drawn from the Grid, where end users share their otherwise idle resources (most desktop computers have a low average processing load) and commercial providers offer dedicated computing farms (with a processing power comparable to that of hundreds or thousands of desktop PCs). This means that in this future Grid, a large user base will have direct access to a vast pool of shared resources as access bandwidth catches up with processing power.

In what follows we present some typical application requirements and their impact on the underlying Grid system, indicating that existing Grid infrastructures will fail to cater for their needs. A first application example comes from the area of multimedia editing; video editing applications are widely adopted, and allow users to manipulate video clips, add effects, restore films etc. Advances in recording, visualization and video effects technology will demand more computational and storage capacity, especially if the editing is to be performed within a reasonable time frame (e.g. allowing user feedback).

More specifically, 1080p High Definition Television (HDTV) [3] offers a resolution of 1920x1080, which amounts to about 2 MPixel per frame. Suppose now that a user would like to evaluate an effect for 10 different options, where applying an effect requires 10 floating-point operations per pixel per frame. It follows then that processing a 10 second clip (25 fps) already requires over 50 GFlop. This will take about 0.5 s to complete locally (we assume local processing power is 100 GFlops), while execution on a provider's resource should only take 5 ms (assuming the capacity of providers is a factor 100 higher). Transmission time of 10 s of compressed HDTV video (bitrate 20 Mbit/s or a 25 MB filesize) on a 2.5 Gbit/s access link is 80ms. While the 2.5 Gbps is likely to be realized through optical technologies, it is unfeasible to assume that each end user is allowed to set up end-to-end wavelength paths for each multimedia editing operation. Indeed, unless wavelength path set-up times were to decrease sharply (currently in the range of 100 ms), the use of optical circuit switching (OCS) would waste a considerable amount of network resources and one would have to devise a mechanism able to handle path set-up and tear-down requests from vast amounts of users.

A second application example is the online visualization of (and interaction with) a virtual environment. Virtual environments are typically made up of various objects, described by their shape, size, location, etc. Also, different textures are applied on these objects. A user should not only be able to visualize selected scenes in the environment by adjusting his viewing angle, but should also be able to interact with the rendered objects. Usually the description of a scene can be realized in limited storage space, the size of a texture being limited to a few kilobytes. Thus a scene can be stored in a rather small storage space, typically around a few Megabytes. However, rendering the scene is a different problem altogether; if we demand a performance of 300 million polygons per second, computational capacities as large as 10000 GFlops are required [4]. Clearly, the rendering of these scenes, preferably in real-time, is unfeasible using only local resources. Suppose a user has at its disposal an archive of different scene descriptions, with a requested frame rate of 25 frames per second. This amounts to a latency smaller than 40 ms between the submission of the scene description, and the actual displaying of the scene. Assuming a scene is 2.5 MB in size, we obtain a transmission time of only 8 ms per frame (excluding overhead); this leaves us with about 30 ms for processing and retransmission of the final rendering, which should be possible with the given capacities of the (local) resource providers. Considering the delay associated with setting up an optical circuit, OCS can only be used when a user employs the same resource, hereby severely limiting the flexibility of the Grid concept. On the other hand, the lack of adequate QoS in the standard IP protocol makes it near-impossible to meet the strict real-time constraints.

When looking at the requirements of Grid technology for consumer applications, the following conclusions can be drawn:

The current network solutions (OCS for computational and through the current internet for peer-to-peer) unsuited for providing Grid access to everyone. The dedicated infrastructure will be too wasteful and inflexible, while the request/grant based architecture with electronic management of Grid resources will be too complex.

To overcome these problems, a new infrastructure will be required. There is no doubt that these will be based on optics, in particular OBS based architecture will be well suited to the task: low processing, with high resource utilization and simple control.

2.1.2 Contribution from (Fujitsu Laboratories, Japan) to be added

3. Control plane and signalling for Grid-OBS

The utilization and improvement of the GMPLS control plane (i.e., routing and signalling protocols) allows Grid-OBS to provision Grid application with the required QoS. The GMPLS control plane would contribute not only on improving Grid-OBS resilience but it will indeed impact Grid-OBS ability of providing QoS connectivity. Currently deployed optical networks are still based on permanent and semi-permanent optical connections terminated at each network node by optoelectronic transponders. Because of their high cost, fixed bit data rate, and fixed protocol data format, optoelectronic transponders limit the network evolution. Novel emerging technologies, such as Optical Burst Switching (OBS), can boost the network evolution from the technological viewpoint by allowing the introduction of all-optical sub-networks at whose edges optical data signals undergo optoelectronic conversion.

High performance applications, such as several Grid applications, may significant benefit from the introduction of advanced network features provided by OBS networks, e.g. data transparency at extremely high bandwidth. For some Grid applications however, there is the need for all bursts to travel the same route through the network. These applications are particularly sensitive to jitter and out-of-order delivery of packets. In these cases the setup of persistent routes can guarantee the required level of Quality of Service (QoS). Persistent OBS connections require a session declaration separated from the cross-connect setup phase and the data burst transmission phase.

During the session declaration phase the routing decision is taken for the burst data flow and an identifier (or label) is associated to the flow in such a way that every burst belonging to that flow is treated in the same way from source to destination.

The cross-connect setup phase refers to the signaling messages that travel out-of-band ahead of the data burst. These messages notify how to configure the switch for the incoming burst (explicit or estimated setup/release).

Data burst transmission phase refers to the transparent flow of optical data bursts.

The management of persistent connection in OBS networks however seems to have many similarities to connection setup and data forwarding in Generalized Multi-Protocol Label Switching (GMPLS) networks, where every data packet is characterized by a label defined during the initial path setup phase. Because of the flexible structure that characterizes the GMPLS protocol suite, GMPLS seems to be a qualified candidate to incorporate the aforementioned OBS session declaration phase.

3.1 Connection Setup Mechanisms

- **Signaling**

In most OBS variants, the signaling of connections is accomplished using a one-way signaling scheme whereby the burst is transmitted after an offset without any knowledge of whether the optical path has been successfully established end-to-end. Therefore it is possible that a burst may be lost if the control packet is not able to reserve resources at any of the OBS nodes along the burst's path. The OBS network, however, does not retransmit lost bursts as this is left to the upper network layers. Note also that it is very important that the offset is calculated correctly. If the offset is too short then the burst may arrive at a node prior to the control packet and thus be lost. On the other hand, offsets that are too long reduce the throughput of the end-device

- **Routing**

An OBS network needs an effective routing algorithm. One approach is to route the control packets on a hop-by-hop basis, as in an IP network, using a fast table look-up algorithm to determine the next hop. The second approach is to use the multi-protocol label switching (MPLS) techniques. In MPLS, a packet is marked with a label, which is used to route the packet through the network. At each node, the label of an incoming packet is looked up in a table in order to obtain the destination output port and a new label valid on the next hop. A third routing approach is to use the constrained-routing version of MPLS, which can be used to explicitly setup routes. This explicit routing is very useful in a constrained-based routed OBS network, where the traffic routes have to meet certain Quality of Service (QoS) metrics such as delay, hop-count, BER or bandwidth.

- **Wavelength Allocation**

As in any other type of optical network, each OBS network has to assign wavelengths at the different WDM fibers along the burst route. This wavelength allocation in OBS depends on whether or not the network is equipped with wavelength converters, devices that can optically convert signals from one wavelength to another. In an OBS network with no wavelength converters, the entire path from the source to the destination is constrained to use the same wavelength. In an OBS network with a wavelength conversion capability at each OBS node, if two bursts contend for the same wavelength on the same output port, then the OBS node may optically convert one of the signals from an incoming wavelength to a different outgoing wavelength. Wavelength conversion is a desirable characteristic in an OBS network as it reduces the burst loss probability, however it is still an expensive technology. An OBS network will most likely be sparsely equipped with wavelength converters, i.e., only certain critical nodes will have that ability.

- **Pre-transmission Offset Time**

An OBS user first transmits a control packet and after an offset time it transmits the burst. This offset allows the control packet to reserve the needed resources along the transmission path before the burst arrives. Furthermore, the OBS nodes need this offset time to set up their switching fabrics so that the data burst can ``cut-through" without the need for any buffers. Ideally, the offset estimation should be based on the number of hops between the source and the destination and the current level of congestion in the network. Obviously, an incorrect offset estimation would result into data loss because the burst may arrive at an OBS node before the optical cross-connect has been completely set up. Therefore, determining this offset is a key design feature of all OBS networks and its effectiveness is measured in terms of the burst loss probability. There are variations in the OBS literature on how exactly to determine the pre-transmission offset time and how to reserve the needed resources at the core OBS nodes. Despite their differences, however, all of the proposed OBS architectures have a dynamic operation, which results in high resource utilization and adaptability.

- **Scheduling of Resources: Reservation and Release**

Upon receipt of a control packet, an OBS node processes the included burst information and allocates resources in its switch fabric that will permit the incoming burst to be switched out on an output port toward its destination. The resource reservation and release schemes in OBS are based on the amount of time a burst occupies a path inside the switching fabric of an OBS node.

There are two OBS resource reservation schemes, namely, immediate reservation and delayed reservation. In the immediate reservation scheme, the control unit configures the switch fabric to switch the burst to the correct output port immediately after it has processed the control packet. In the delayed reservation scheme, the control unit calculates the time of arrival tb of the burst at the node, and it configures the switch fabric at tb.

There are also two different resource release schemes, namely, timed release and explicit release. In the timed release scheme, the control unit calculates when the burst will completely go through the switch fabric, and when this time occurs it instructs the switch fabric to release the allocated resources. This requires knowledge of the burst duration. An alternative scheme is the explicit release scheme, where the transmitting end-device sends a release message to inform the OBS nodes along the path of the burst that it has finished its transmission. The control unit instructs the switch fabric to release the connection when it receives this message.

Combining the two reservation schemes with the two release schemes results in the following four possibilities: immediate reservation/explicit release, immediate reservation/timed release, delayed reservation/explicit release and delayed reservation/timed release.

- **Limited Buffering Using Fiber Delay Lines**

One of the main design objectives for OBS is to build a bufferless network, where the user data travels transparently as an optical signal and ``cuts-through" the switches at very high rates. Bufferless transmission is important to OBS because electronic buffers require optical-to-electronic-to-optical conversion, which slows down the transmission, and optical buffers are still quite impractical. In fact, as of today, there is no way to store light and so the only possible optical buffering is to delay the signal through very long fiber lines. Fiber delay lines (FDLs) can potentially improve the network throughput and reduce the burst loss probability. In the presence of FDL buffers, the OBS reservation and release schemes have to be revised. In addition to scheduling the wavelengths at the output ports, the OBS nodes also have to manage the reservation of their available FDL buffers.

- **Variations on Burst Dropping**

Most of the OBS literature specifies that if all the resources are occupied at the moment of the burst arrival then the entire data burst is lost. An interesting OBS variation, is to divide each burst into multiple segments and in the case of resource contention, instead of dropping the entire burst, either the head or the tail segment is deflected to an alternative route to the destination.

- **Classes of Traffic**

In an OBS network, the filtering of upper layer data and the assignment of classes to bursts will occur at the edge of the network during the burst assembly process. In order to minimize the end-to-end delay of the high priority traffic, the burst assembly algorithm can vary parameters such as the pre-set timers or the maximum/minimum burst sizes. However, selecting the values for these parameters is a difficult task because of the throughput interdependency between the different classes of traffic. Here are some of the proposed solutions:

a) Classes Based On Extended Offsets: The higher priority traffic is assigned a longer offset between the transmission of its control packet and its corresponding data burst. The burst blocking probability decreases as the offset time increases. One of the main constraints of this scheme is the maximum acceptable upper layer delay, i.e., certain high priority applications cannot tolerate long pre-transmission offsets.
b) Classes based on the Optical Signal Properties and Preemption: This scheme is based on the physical quality of the optical signal such as the maximum bandwidth, the error rates, the signal to noise ratio and the spacing between the different wavelengths. These parameters are included in the control packets. A connection is established only if all of these requirements can be met, possibly using a constrained-based routing algorithm. In addition to the intrinsic physical quality, it is possible to implement priorities based on a preemption mechanism, where a lower priority burst, which is in the process of being transmitted, can be preempted by a higher priority one.

- **Multicast**

In OBS, as in wavelength-routed networks, the multicasting is achieved through light splitting, which inherently results in signal losses. Therefore, there is a limit on the number of times the signal can be split and the number of hops it can traverse. In addition, the multicasting in all WDM network is tightly coupled with wavelength allocation and is greatly dependent on the availability of wavelength converters. It is important to note, however, that the dynamic nature of OBS makes it suitable for optical multicasting because the resources of the multicast tree are reserved on a per-burst-basis.

3.2 QoS provisioning in Grid-OBS networks

The aim of this section is to evaluate benefit and limits of the OBS session declaration phase managed using GMPLS and to investigate the requirements and the extensions that should be introduced into the GMPLS protocol suite. In particular ReSerVation Protocol with Traffic Engineering extensions (RSVP-TE), Link Management Protocol (LMP) and Open Shortest Path First with TE extensions (OSPF-TE) protocol require new objects and procedures, such as new properly formatted label, new interface switching capability descriptors and proper routing and signaling procedures to allow Grid applications to exploit the benefit of the emerging powerful OBS technology.

Optical networks have been identified as the network infrastructure that would enable the widespread development of Grid computing, i.e. global Grid computing. However just offering large bandwidth connections is not sufficient for the requirements of Grid computing applications. Thus not Optical Networks but Intelligent Optical Networks must be considered as the suitable network infrastructure for global Grid computing. Intelligent Optical Networks, i.e. optical networks equipped with the Generalized Multiprotocol Label Switching (GMPLS) protocol suite, are able to dynamically adapt to both network and applications changes to satisfy the Grid computing application requirements. Intelligent optical networks are also able to offer different optical bandwidth granularities. Indeed while wavelength routed optical network research is already tackling its advanced issues, Optical Burst Switching (OBS) is gaining momentum in the optical network research field. OBS is able to offer finer optical granularity connectivity service to Grid computing applications than wavelength-routed networks. This would allow users to pay just for what they need for running their applications. Indeed, while applications that need to move large amount of data, e.g. data Grids, might require the entire bandwidth offered by all-optical connections, i.e. light paths, other applications would just require fraction of the bandwidth. OBS represents the solution for providing Grid computing applications with the fraction of bandwidth they need while maintaining the protocol transparency advantages of wavelength routed networks. Thus by offering both wavelength routed, OBS, connectivity services the optical network infrastructure would allow not only users to pay what they asked for but also optical network service connectivity providers to better optimize their network utilization.

However different bandwidth granularities cannot be the only service offered by Intelligent Optical Networks. In particular Grid computing applications pose strict constraints on delay and delay jitter. Thus, at each granularity (i.e., wavelength routed, OBS), Intelligent Optical Network connectivity services must guarantee the suitable quality of service (QoS) considering also delay and delay jitter constraints. In addition, connectivity service differentiation must be guaranteed within each connection granularity. On the one hand guaranteeing connectivity service differentiation at the lightpath granularity appears to be achievable through the utilization of GMPLS protocol extensions for traffic engineering . On the other hand guaranteeing QoS of service at the OBS granularity is still matter of thorough research. In particular the synergy between GMPLS with traffic engineering extension control plane with OBS protocols appears to be necessary.

Finally another important issue to be addressed is the matchmaking of the application requirements to the connectivity services. For example, applications requiring a fraction of lightpath bandwidth, thus suitable for OBS, but requiring stringent constraints on delay and delay jitter might be better served by over provisioning them with a lightpath than utilizing for them an OBS connection.

Quality of Service (QoS) support for GRID Applications requires several characteristics referring to different elements such as networks, CPUs and storage devices. Typical network requirements are: end-to-end delay the traveling packet time from the sender to the receiver, delay jitter the variation in the end-to-end delay of packets between the same node pair, throughput (i.e., bandwidth) the rate at which the packets go through the network and packet loss rate the rate at which the packets are blocked, loss or corrupted [5,6]. Optical Burst Switching (OBS) networks will be able to satisfy GRID Applications high bandwidth requirements combining the strengths of both Wavelength Routed (WR) and Optical Packet Switching (OPS) networks, moreover several approaches for QoS provisioning in OBS networks have been proposed in the literature. The main aim here, is to provide relative service differentiation with regards to packet loss probability, nevertheless they are based on relative QoS model in which the service requirements for a given class of traffic are defined relatively to the service requirements of another class. It is possible to distinguish in:

- *Offset-based schemes* [7,8] that introduce an extra-offset time between control burst (CB) and data burst (DB) to differentiate bursts in several service classes. These technique have been proposed utilizing Just-Enough-Time (JET) protocol in buffer-less OBS networks, and it has been proved that, opportunely setting the extra-offset time (the higher priority, the higher extra-offset time), high class bursts loss rate can be independent from lower classes traffic. The main drawback of these schemes is represented by the aware increase of end-to-end delay for high priority burst.
- *Strict priority schemes* [9], minimize high priority bursts loss rate allowing them to preempt reservations of lower priority bursts. Therefore a specific burst can be only blocked by reservations of higher class bursts or in-going transmission of lower priority bursts, in this case the end-to-end delay is proved to be less with respect to offset-based schemes, but the lower class burst loss rate is still strongly dependant on the higher priority traffic as in offset-based schemes.
- *Segmentation-based schemes* [10 , 11] avoids bursts collisions in core nodes providing preemptive high class bursts combined with low class bursts segmentation and deflection. In particular when a contention occurs, lower class contending burst is divided into multiple segments and only overlapping segments

are dropped or deflected. This approach can decrease low priority bursts loss rate but it significantly increases the physical layer architecture.

Other schemes propose to differentiate bursts classes allowing each class to utilized different network functionalities (e.g, extra-offset, wavelength conversion, deflection routing) considering class specific QoS requirements [12]. The usefulness of end-to-end re-routing with respect to deflection routing is investigated in [13], it improves network throughput reducing nodes congestion and decreases delay jitter avoiding unpredictable delays typically introduced by deflection routing; moreover end-to-end re-routing is able to more efficiently provide network resilience in case of node or link failures.
Other proposal for OBS networks [14,15,16,17], aim to provide quantitative QoS guarantees with regard to packet loss rate, worst case end-to-end delay and throughput. These kind of QoS schemes seems to be more suitable to be applied in a Grid environment where each application needs specific QoS requirements. Proportional QoS schemes are proposed in [14,15], to adjust the service differentiation of a particular QoS metric to be proportional to particular weights that a network service provider can set; these schemes feature in advance discard of lower class optical bursts. In [16] an early dropping mechanism, which probabilistically drops lower class bursts, and a wavelength grouping mechanism, which provisions necessary wavelengths for high class busts are proposed. In [17] a possible architecture to provide quantitative QoS guarantees with respect to worst case end-to-end delay, throughput, and packet loss probability in buffer-less Labeled OBS networks is proposed. In particular [17] shows that deploying fair scheduling algorithms in both the data plane of the edge nodes and the control plane of core nodes it is possible to support a wide range of service guarantees with regards to throughput, end-to-end delay and packet loss probability.

In conclusion there are different ways to provide QoS in OBS networks, the key issues in providing QoS for Grid applications is to understand the requirements for each specific application and find out the right strategy to quantitatively provide them.

Providing Grid computing applications with resilient connectivity appears one of the QoS requirements of increasing importance. In addition maintaining, even upon failure occurrence, QoS differentiation among the connections utilized by the applications, i.e. differentiated resilience (reliability), is required. Resilience in OBS network has just started to be addressed by the optical network community [18,19]. In general OBS dynamic routing, on which hop-by-hop OBS routing is based, is able to overcome “by nature” network failure. However because of the high recovery time [18,20], mainly due to the routing table updates [21], dynamic OBS rerouting is not able to guaranteed the required QoS.

Already proposed pre-planned global rerouting based on Labeled Optical Burst Switching has shown to be promising for balancing the network load and recovery bursts after a physical network kink failure [22]. However resilient schemes based on deflection routing have shown the ability of improving the performance, in terms of burst blocking probability, of resilient schemes based on global routing updates during the failure recovery phase. In both cases the utilization of schemes based on traffic engineering extensions to GMPLS already developed for wavelength routed network might help in improving OBS network performance, in terms of burst loss probability, upon failure occurrence[19]. Previously proposed schemes are based on proposed extensions to routing and signaling protocols of the GMPLS protocol suite. Therefore routing and signaling protocols are also important for Grid-OBS resilience.

For example a better choice for the deflection path taken by the bursts involved in the failure can be obtained by utilizing a weighted stochastic approach, such as the one utilized in [19]. The approach proposed in[19] represents a scheme fairly simple to be implemented applicable to both local and global rerouting. In addition failure notification based on RSVP-TE signaling might improve failure notification time.

The main issue in utilizing resilient schemes already proposed for wavelength routed network consists in the different dynamic characteristics of OBS and wavelength routed networks. Indeed OBS network parameters, such as load, change much more quickly that the correspondent ones in wavelength routed network. A possible solutions therefore would be to apply schemes typical of OBS in the short time scale and periodically improving their performance by changing their behavior through the feedback obtained by wavelength routed alike resilient schemes.

3.3 Constrained based routing incorporated in the OBS control

The OBS routing protocols offer the opportunity to take into consideration the physical layer characteristics of the network infrastructure as part of the routing algorithm and the Grid service offering. In addition to the information relating to the traditional Grid resource characteristics, physical layer characteristics (i.e. chromatic dispersion, polarization mode dispersion, amplifier gains, amplifier noise, launch power level, span length, loss of a pan and node, crosstalk levels) will be considered.

Based on these parameters information, carried by the burst control packet, a set of available Grid and network resources can be identified by the OBS routers. These costs will be taken into account when finding the possible paths to establish the Grid services as and when required across the network. The Grid service will be established across the path that satisfies the service policy requirements in terms of all critical parameters specific to the requested service. This is a novel way of implementing user controlled constrained based routing across the two network domains (Grid and optical).

3.4 specific implementation of control plane

3.4.1 Control plane issues for consumer Grid application

- When looking at the requirements of Grid technology for consumer applications, the following conclusions can be drawn:
- It is economically unsound to build a dedicated network for each application. Although there exist several high bandwidth and computationally intensive applications, constructing a seperate network to which individual users connect,

seems unrealistic. The current convergence of phone, television and data networks ("triple play") clearly proves this point.

- Grid service requests will be, in most cases, highly unpredictable, implying a dedicated, static infrastructure is not the most efficient solution.
- The sheer potential volume of requests makes electronic processing highly complex. In other words, we need to simplify intelligence in the network as much as possible, as well as use optics wherever appropiate to deal with the huge bandwidth requirements.
- In many cases, the transmission times (job sizes) will be rather short (few 100 μs to tens of ms). This means that using end-to-end circuit switched connections will prove to be too wasteful, as the holding time of a wavelength path will be too small compared to its setup time. Real time applications place even further importance on this point.

We can easily deduce several essential requirements which the control and signaling plane should be able to satisfy:

- The ability for new application types to be deployed quickly and efficiently, which implies a flexible control plane is required. Indeed, as mentioned before, it is infeasible to build separate networks for each application type. As such, the basic infrastructure offered by the OBS network should be able to support all types of applications, each with its own typical resource usage patterns.
- Flexibility also indicates that the features offered by the control plane should be of relative simplicity. Features which are usable by only one application group introduce complexity in the signaling protocols and can usually be assembled from simpler, generally deployable components.
- Support for a huge number of users implies scalability of the control plane is essential. In light of this, research should focus on minimizing the control and signaling traffic. This point becomes even more important when users have a highly unpredictable traffic pattern.
- Support for highly dynamic user access patterns means the control plane should be adaptable to the Grid's status, e.g. by reducing signaling data in favor of more actual data transfers.
- Sufficient levels of speed and flexibility in the control plane are imposed by real time applications. As we mentioned repeatedly, the main disadvantage of traditional circuit switching is its inability to react quickly to dynamic traffic demands. Adding real time constraints to this setting is only possible with networks which have a minimal latency imposed by the control plane, thus leaving more time for the actual data transfers.

3.4.2 Reliable transport control technology for Grid-OBS

Contribution from (University of Tokyo, Japan) **to be added**

3.4.3 Ultra fast optical path setup protocol and its optical implementations

Contribution from (Osaka University, Japan) **to be added**

4. Grid-OBS network elements

4.1 Core OBS router

As future optical technology moves to 40Gb/s and beyond, networking solutions must be designed to be compatible these bit rates, in order to reduce the cost per bit [23]. OBS has been introduced as a switching technology relaxed on fast switching requirements, as the relatively slow switch set-up times (milliseconds rather than nanoseconds) are small compared to the payload duration (usually hundreds of milliseconds or seconds) and therefore throughput is almost unaffected [24]. However, the introduction of Grid services over OBS implies new constrains for the switching speed requirements, which become particularly important when high speed transmission is considered.

A flexible Grid network will require also the support of users with small job requests. For example, a relatively small burst, 300ms, transmitted at 10Gb/s can be switched by a MEMS based switch typically within 20ms. Considering only the switching time, the throughput of the system is 93.3%. If the same burst is transmitted at 160Gb/s then its duration is 18.75ms and routing through the same switch would decrease the system's throughput to less than 50%. This becomes more severe when users with even smaller job requests are treated. These small jobs are implied by the small bursts and may be with short offset time. These types of bursts with small length (typical 100 to 1000 bytes), requires ultra-fast switching in nanoseconds. Additionally, the support of multicasting is particularly advantageous, in order to enable parallel Grid processing services latency [25] as well as resource discovery. For these reasons the deployment of fast switching technology is essential for future high speed OBS networks that can support Grid applications. It should be noted though, that the core OBS for the Grid computing may require intensive and intelligent processing of control information and BCP (i.e. performing some Grid network functionality, e.g.: taking part in resource discovery), which can only be performed by specially designed fast electronic circuits. Recent advances in the technology of integrated circuits allow complicated processing of bursty data directly up to 10Gb/s [26]. This sets the upper limit in the transmission speed of the control information and BCP. On the other hand the much longer transparently switched optical bursts (i.e. no conversion to electronic domain) are those that determine the capacity utilisation of the network. The optical bursts can be transmitted at ultra-high bit rates (40 or 160Gb/s), providing that the switching elements can support these bit rates. Faster bursts indicate higher capacity utilisation of the existing fibre infrastructure and significantly improved network economics.

The fast switching solutions that have been proposed are based on the use of fast active components, like Semiconductor Optical Amplifiers (SOAs). Switching is achieved either by broadcasting the signal (passive splitting) and selecting the appropriate routes using fast gating [27,28] or by converting the signal's wavelength and routing it to an output port of a passive routing device (AWG) [29,30,31]. The gating solution is independent of the signal's bit rate and also supports multicasting but scales poorly to a large port-count switch. The wavelength conversion and selection solution is scalable but bit-rate dependent on the utilised conversion technique.

The deployment of fast switching assists the efficient bandwidth utilisation but provides an expensive solution when it scales to many input port. On the other hand, there is no additional benefit for long bursts of data (e.g. originated from large GRID users) if fast switching is utilised. Therefore, a proper OBS networking solution needs to consider a combination of fast (e.g. SOA-based) and slow (e.g. MEMS-based) switches.

One solution can be based on the use of OXCs that has a number only of output ports connected to a fast optical switch that follows. Several OXCs and fast switched can be placed in parallel in a scalable wavelength modular architecture. At the switch input the wavelength channels per input fibre are separated. When a BCP appears the control mechanism must first recognise if the BCP belongs to a long, a short burst. In the first case the OXC is reconfigured so that when the long burst arrives it automatically routed to the appropriate output port. In the other two cases the short and the active bursts are routed directly to the fast switch (through pre-defined paths) and switched immediately to the next node. This set-up requires all the switching paths inside the OXC to be initially connected to the fast switch ports and special design constrains must be considered to avoid collision. The benefit of the proposed scheme is that it reduces the requirements on fast switching and therefore smaller and cost efficient matrices are only required

4.2 Edge OBS router

In OBS networks, a data burst and its burst control header are transmitted separately on different wavelength channels and switched respectively in optical and electronic domains. Thus, in an OBS network an ingress edge router able to initiate a burst control header and also map user traffics traffic into the optical domain in the form of variable length optical bursts is mandatory.
An edge OBS router in a Grid enabled OBS network must be able to perform the follow functionalities:

a) Traffic aggregation and optical burst assembly
b) Optical burst transmission
c) Grid user to network as well as Grid resource to network signaling

- **Burst aggregation**

The burst aggregation algorithm at the edge router can greatly impact the overall OBS network operation because it sets the burst characteristics and therefore shapes the burst arrival traffic. The algorithm has to consider the following parameters: a pre-set timer, a maximum burst length, and a minimum burst length. The timer determines when the end-device is to assemble its collected traffic into a new burst. The maximum and the minimum burst length parameters shape the size of the bursts. It is necessary to set a maximum burst length since very long bursts hold on to the resources of the network for a long time and, thus, they cause the unfair loss of other bursts. On the other hand, the minimum burst length is necessary because very short bursts may give rise to too many control packets, which can overload the control unit of the OBS node. The burst aggregation algorithm may use bit-padding if there is not enough data to assemble a minimum size burst.

- **User and resource network interface functionality:**

To facilitate on demand access to Grid services, interoperable procedures between Grid users and optical network for agreement negotiation and Grid service activation have to be developed. These procedures constitute the Grid User Optical Network Interface (G-OUNI). The G-OUNI functionalities and implementation will be influenced by number of parameters as follows:

- Service invocation scenarios
- Control plane architecture

The GUNI in a grid enabled OBS network needs to provide the following main functionalities:

- Flexible bandwidth allocation
- Support for claiming existing agreements
- Automatic and timely light-path setup
- Traffic classification, grooming, shaping and transmission entity construction

On the other hand, geographically distributed processing and storage resources across the network constitute fundamental elements of the large scale Grid network. In such network scenario the Grid resources (i.e. storage and processing) can dynamically enter and leave the OBS network based on pre-established agreements. This fact imposes the necessity of a dedicated signalling and control interface between such resources and the Grid network. Like the GUNI, the Grid resource network interface (GRNI) must perform interoperable procedures between external network elements and the OBS network. But unlike the GUNI, the interface will be between resources-end elements (processing and/or storage distributed across network) and the optical network. The similarity between GUNI and the GRNI makes it possible to extend the GUNI model to provide required functionalities for the resource network interface. Main functionalities of such an interface can be:

- Support for existing agreements
- Job submission to local Gird resources
- Support for advance resource reservation schemes
- Propagation state of the local resources (available storage/ processing resources)
- Propagation of service related events
- Sending back results to source or multiple alternative destinations

AS both GUNI and GRNI with aforementioned functionalities are related either to the Grid users or Grid resources (i.e. Grid network end elements), thus their functionalities must be integrated into an edge OBS router device. Such edge router must be an agile and user-controlled interface able to map user traffic into optical domain at sub-wavelength granularity (i.e. in the form of optical bursts).

5. Advanced network concepts and solutions

5.1 Self-organised OBS network for consumer Grids

It is usually assumed that OBS networks employ shortest path routing, seeking to minimize the end-to-end delay. It is however well known that this approach may lead to inefficient usage of network resources; certain links are hardly used, while others can become severely congested, which of course leads to sub-optimal network performance. This is especially true when the burst dropping probability is the main metric of interest, as is usually the case in OBS networks. Several approaches have been proposed to overcome this problem, such as deflection routing and multi-path routing. In any case, both the sender and the receiver are usually known in traditional data transfers. This differs from a Grid OBS network where the destination is not always known, as we'll show in the next section.

- **Anycast Routing in Grid OBS networks**

In the consumer Grid scenario [32,33,34], it doesn't matter where exactly the job is processed. Instead, the user is only interested in the fact that his job is processed within certain predetermined requirements. In general, there will exist multiple locations where a job can be executed, and the selection of a suitable resource is left to the routing protocol. This represents a shift in the nature of the employed routing algorithm; whereas previously bursts had an exact destination, now we only require the burst to be sent to any end node capable of processing the burst. The former approach is called unicast routing, while the latter is usually denoted by anycast routing. [35,36,37]

- **From User to Resource**

The basic operation of the Grid network is now as follows. First, the user realizes that a computing task cannot be completed within a reasonable timeframe on the local system, and decides to post it on the Grid to accelerate processing. The job is then transformed in an optical burst (containing code and data), accompanied by a header indicating various parameters (e.g. processing, storage and policy requirements). Note that a very important design decision has been made, i.e. the mapping of one job onto one optical burst. As discussed earlier, no destination address is needed, and thus the burst is simply handed over to the OBS network. The burst travels along a link, while the intermediate routers are not notified in advance of its arrival, much like JIT or JET based schemes. On arrival of the burst, an intermediate router decides on the fly where to forward the burst, based on information contained in the preceding header and on network and resource status information. Examples of such information are link load and blocking probability, delay requirements, estimated free computing or storage capacity which can be reached through a certain interface, and estimated computing and storage requirements of the burst. Since the end user doesn't specify the network location where the burst will be processed, the job is scheduled implicitly through its progress in the network. This makes the Grid architecture completely distributed, which naturally implies better scalability and robustness. Note that an intermediate router does not need a detailed view of where the resources are located and how much (free) capacity they have. As long as there is enough information to push the burst closer to a suitable destination, a good decision can be made. This means that the aggregation of status information can be used to reduce control traffic.

- **Processing a Job**

Each intermediate router in the network goes through the same process, and eventually the burst arrives at a Grid resource. If this resource is able to handle the job contained within the burst, it will process it. If this is not the case, a deflection mechanism can be used to repost the job in the OBS network. It is also possible to drop a burst which cannot be timely processed.

- **From Resource to User**

Once the job is completed, its results must be delivered back to the user (most likely where the burst originated). Here the asymmetry of the Grid OBS network becomes clear; although posting a job uses the anycast paradigm, sending results back most likely will not. There is a distinct return address, and more traditional forwarding solutions have to be used. A variety of options and choices can be made, depending on such parameters as the processing time, storage availability, size of results, etc. For instance, a real time application requires its results to be transmitted as fast as possible, while for an offline calculation the results can be stored on the processing node until network availability improves. Also, we can consider a returning burst to be "more valuable" than one which hasn't been processed yet. Naturally, this notion gives rise to the introduction of different QoS classes in the network traffic.

- **Burst Correlation**

Up until now, we have assumed that all bursts are sent completely independent of each other in the network. However, we will show that it can be advantageous to dispose of a method to send consecutive bursts to the same resource.

a) The proposal of mapping one job onto one optical burst is mainly inspired by the simplicity and general application of this approach. However, this technique will prove insufficient whenever jobs are generated which are too large to fit into one optical burst. In this case, the original job has to be segmented into smaller sub-bursts, which are sent individually in the network. The routing algorithm must be adapted to make sure these sub-bursts arrive at the same resource. Also, resources must contain the functionality to reassemble the individual segments into the original job request.

b) A second scenario where burst correlation can be useful is for specific applications which can reuse input and output data of preceding bursts. For instance, in a virtual reality application, there is no need to re-render the complete scene when the user changes his viewing angle of the scene. Instead, it is better to make use of the rendering results of a previous burst, and incorporate only the changes generated by the user's actions. Note though that specific support for this feature will have to be built into the application logic.

Because of the architectural requirement to scale to large numbers of users, it is impossible to maintain the forwarding decision of each burst in all routers. An alternative approach is to let the user wait for the results of the first burst, extract the address of the employed resource, and send all following bursts to the same destination address. Yet another possibility is that the first burst sets up a path which is followed by all later bursts, similar to the label switching technique. Aggregation techniques may be applicable too, such as merging common portions of serveral OBS paths, very like merging and stacking in label switching. As a logical extension this may result in OCS-like operation (wavelength switching), supporting the more static portions of the network.

- **Robustness**

Robustness of a network is typically evaluated based on the number of requests (jobs in our case) that cannot be handled whenever resources are failing. The heterogeneous nature of the Grid implies two types of resources can fail; the network resources (links and routers) and the server resources (the processing elements). We describe two methods to introduce robustness against failing resources of both types.

a) Spare capacity

Before deployment, a network is usually dimensioned based on load estimates or experienced job request rates. In case more network or server resources are introduced in the network than are strictly necessary, this remainder of capacity can be used in case certain Grid components fail. Research needs to be done on different restoration strategies, focusing on how and when this spare capacity will be utilized.

b) Duplicate Submission

If the same job is sent into the network more than once, the possibility that this job reaches a different server resource, or reaches the same server but arrived there over a different path, is non-negligable. Thus, this method can also introduce robustness in the Grid OBS network. Observe though that more capacity is used than strictly necessary.

5.2 Programmable Optical Burst Switched Network

In this section a novel solution towards ubiquitous photonic Grid networking is proposed. This solution utilizes optical burst switching and active router technologies. It aims to provide a physical infrastructure able to fulfill both existing data-intensive and future Grid application requirements and make efficient use of network resources. The solution is based on programmable network architecture, in which the optical network topology can be programmed by Grid users and services.

The architecture is based on the novel concept of using active OBS routers for resource discovery and routing of the Grid jobs to the appropriate resources across the network. The network comprises active and non-active OBS routers. A non-active OBS router is a conventional OBS router and performs the burst forwarding functionality. The router is informed in advance about the data burst characteristics (duration, type, class of service, etc.) by the Burst Control Packet (BCP). Upon the data burst arrival the router, forwards the data to the appropriate output port. An active OBS router, in addition to the burst forwarding, can intercept with data carried by some optical bursts (active bursts) and perform dedicated Grid networking functionality. The proposed active OBS networking scheme has the potential to offer global reach of computing and storage resources to a large number of anonymous users with different traffic profiles. In such a network, OBS offers efficient network resource utilisation while the active networking offers intelligent Grid functionality. One of the main advantages of the proposed scenario is that both traditional data traffic and Grid traffic can be supported by a common infrastructure. All OBS routers perform burst forwarding when normal traffic transits across the network while in addition some OBS routers (active routers) support transport of Grid traffic over the network.

- **Description of Transport format**

There are several major OBS variants differing in bandwidth reservation schemes [38]. Among all of them, the just-enough-time (JET) is the most appropriate protocol for the proposed Grid network architecture [39]. The JET protocol employs a delayed reservation scheme which operates as follows: an output wavelength is reserved for a burst just before the arrival of the first bit of the burst; if, upon arrival of the BCP, it is determined that no wavelength can be reserved at the appropriate time, then the BCP is rejected and the corresponding data burst dropped. The proposed network concept utilizes the JET scheme and extends it to support both active and non-active network operations. Non-Grid traffic is injected into the network in the form of a normal, non-active burst and active routers do not intercept the traffic. In this mode, once data is ready to be transmitted, a BCP is sent from the edge router into the optical network and the required resources are reserved for the duration of the burst. For efficient transmission of Grid traffic, we have developed a two-stage OBS networking scheme including an active stage and a non-active stage. Grid traffic is transmitted in two stages as follows: job specification is transmitted in the form of an active burst prior to the actual job (user data) which is transmitted in the form of a non-active burst. The user with a Grid job sends a request to the edge router informing about the job specification and resource requirements. The edge router then constructs and transmits the active optical burst for which the BCP only informs intermediate active routers that the incoming optical burst is active. After an offset time, the active burst is transmitted carrying information about the Grid job characteristics (i.e. processing and storage requirements). With this mechanism active routers prior to arrival of the job specification have been informed about the arrival of an active burst. Upon arrival of a job specification burst, an active router performs a resource discovery algorithm to find out whether there are enough Grid resources available within its Grid resource domain to perform the job. In addition, each active router multicasts both the BCP and data burst of an active burst towards the other active routers in the network. The user is informed about the result of resource discovery by each active router through acknowledgment or not-acknowledgment messages (optical burst). In case of resource availability the user transmits the actual job in the form of a non-active burst through the edge router.

In order to accommodate the requirements of this active Grid network scenario the JET scheme is modified. The job submission is divided into two steps:

1. The BCP of an active burst is sent to all active routers through intermediate nodes (active or non-active). After an offset time the active data burst is sent to the network. The result of the resource discovery algorithm in each active router produces an acknowledgment (Ack) or a notacknowledgment message (Nack). These messages are transmitted back to the user through an optical burst (non-active burst). In case of acknowledgement, the active OBS router also informs the corresponding resource manager. At that point the resource manager reserves the local resources for a predefined and limited duration of time.

2. Receiving all ACK and NACK messages, the user can choose one or multiple appropriate destinations among all available resources across the network. The actual job is now sent within the reservation period to the appropriate destination in normal (non-active) optical burst format.

In summary, the proposed programmable OBS concept is a two mode networking scheme:

• It is an active network when the Grid job specification is routed through the network to discover the suitable Grid resources

• It is non-active when Grid jobs or normal data traffic are routed across the network

This combination provides bandwidth efficiency especially when a large data set needs to be transferred because the actual job is submitted to the network only when both the Grid resources and the network resources have been reserved. In addition it provides a secure and policy based Grid environment where the users have the ability to choose among the available resources in different Grid domains across the network. Furthermore, active routers in each domain can respond positively only to the requests that match with the applied policy in their corresponding domain.

- Grid enabled active OBS routers

Central to the programmable OBS network architecture is the possibility of using network processors (NPs) in active OBS routers, capable of analyzing data traveling through the network at wire speed. In the proposed network architecture active OBS routers utilize high-performance network processors (NPs) for routing the active jobs. The NPs are capable of executing specific processing functions on data contained within an active burst at line rates (e.g. Grid resource discovery algorithm). Active OBS routers are key enablers for the support of user-controlled networking functionalities: 1) quality of service (QoS) provisioning 2) reliable multicasting and 3) constrained base routing.

It has been shown in [40] that services and applications are concerned about QoS based on network, bandwidth and delay. In the proposed network architecture, a combination of the control protocol and active routers' processing power can be used to deploy an advanced burst-scheduling algorithm. This algorithm is able to reduce delay whilst maintaining high bandwidth efficiency and low burst loss rate.

In the active Grid network environment, multicasting performs an important role, where interactive and distributed applications are deployed. A reliable multicast protocol framework is deployed, in order to minimize the traffic load across the network and also reduce the recovery latency [41].

6. Security issues in Grid-OBS networks

7. References

[1] D. Simeonidou et al., “Optical Network Infrastructure for Grid”, Grid Forum Draft , GFD-I.036, Oct 2004

[2] C. Qiao, M. Yoo, “Optical Burst Switching - A new Paradigm for an Optical Internet”, Journal of High Speed Networks, Spec. Iss. On Optical Networking, vol. 8, no. 1, Jan. 2000, pp. 36-44

[3] L. Zong and N. G. Bourbakis, “Digital Video and Digital TV: A Comparison and the Future Directions”, Proc. of the International Conference on Information Intelligence and Systems, Mar 1999.

[4] T. Ikedo, “Creating Realistic Scenes in Future Multimedia Systems”, IEEE MultiMedia, 9(4):56-72, Oct 2002.

[5] R. Al-Ali, G. von Laszewski, K. Amin, M. Hategan, O. Rana, D. Walker, N. Zaluzec, “QoS Support for High-Performance Scientific Grid Applications,” in *IEEE International Symposium on Cluster Computing and the Grid CCGrid,* Chicago, 19-22 Apr. 2004, pp.134 – 143.

[6] D. Simeonidou et al., “Optical Network Infrastructure for Grid,” GFD-I.036 draft, Informational GHPN-RG *http://forge.Gridforum.org/projects/ghpn-rg*

[7] M. Yoo, C. Qiao, and S. Dixit, “Optical Burst Switching for Service Differentiation in the Next Generation Optical Internet,” *IEEE Communications Magazine*, pp. 98–104, Feb. 2001

[8] M. Yoo, C. Qiao and S. Dixit, “ QoS Performance of Optical Burst Switching in IP-Over-WDM Netwoks,” *IEEE Journal on Selected Areas in Communications,* pp. 2062-2071, Oct. 2000.

[9] A. Kaheel and H. Alnuweiri, “A Strict Priority Scheme for Quality-of-Service Provisioning in Optical Burst Switching Networks,” in *Proceedings of IEEE Symposium on Computers and Communications-ISCC’03*, Turkey, Jun. 2003, pp. 16–21.

[10] V. M. Vokkarane and J. Jue, “Prioritized Routing and Burst Segmentation for QoS in Optical Burst Switched Networks,” in *Proceedings of Optical Fiber Communication Conference- OFC’02*, Anaheim, March 2002, pp. 221–222.

[11] V. M. Vokkarane, Jason P. Jue, “Prioritized Burst Segmentation and Composite Burst-Assembly Techniques for QoS Support in Optical Burst-Switched Networks,” *IEEE Journal on Selected Areas in Communications,* pp. 1198-1209, Sep. 2003.

[12] M. Casoni, M.L. Merani, “Resource Management in Optical Burst Switched Networks: Performance Evaluation of a European Network,” in *Proceedings of 1st International Workshop on Optical Burst Switching*, 16 Oct. 2003, Dallas.

[13]M. Casoni, M.L. Merani, A. Giorgetti, L. Valcarenghi, P. Castoldi, “Guaranteeing Seamless end-to-end QoS in OBS Networks,” in *Proceedings of OpNeTec,* Oct. 2004, Pisa.

[14]Y. Chen, M. Hamdi, and D. Tsang, “Proportional QoS over OBS Networks,” in *Proceedings of IEEE Global Telecommunications Conference- Globecom’ 01*, San Antonio, Nov. 2001, pp. 1510–1514.

[15] Y. Chen, M. Hamdi, D. H. K. Tsang, and C. Qiao, “Proportional differentiation – A scalable QoS approach,” *IEEE Commun. Mag.*, vol. 41, pp. 52–58, Jun. 2003.

[16] Q. Zang, V. M. Vokkarane, J. P. Jue, B. Chen, “Absolute QoS Differentiation in Optical Burst-Switched Networks,” *IEEE Journal on Selected Areas in Communications,* pp. 1781-1795, Nov. 2004.

[17] A. Kaheel, H. Alnuweiri, “Quantitative QoS Guarantees in Labeled Optical Burst Switching Networks”, in in *Proceedings of IEEE Global Telecommunications Conference, Globecom’01*, Dallas, Nov. 2004.

[18]Yufeng Xin *et al.*, “Fault Management with Fast Restoration for Optical Burst Switched Networks”, IEEEXplore

[19] L. Valcarenghi and A. Fumagalli, “Implementing Stochastic Preplanned Restoration with Proportional Weighted Path Choice in IP/GMPLS/WDM Networks", in Photonic Network Communications, Special Issue on “Routing, Protection, and Restoration Strategies and Algorithms for WDM Optical Networks", Kluwer Academic Publishers, vol. 4, no. 3/4, July/December 2002

[20] Yufeng Xin *et al.*, “A Novel Fast Restoration Mechanism for Optical Burst Switched Networks”, IEEEXplore

[21]S. Darisala, A. Fumagalli, P. Kothandaraman, M. Tacca, L. Valcarenghi, M. Ali, and D. Elie-Dit-Cosaque "On the Convergence of the Link-State Advertisement Protocol

in Survivable WDM Mesh Networks", in Proceedings of ONDM 2003, Budapest,

[22] Jing Zhang et al., "Pre-planned Global Rerouting for Fault Management in Labeles Optical Burst Switched WDM Networks", ICC 2004

[23] Mintera Optical Networks, "Myths and realities about 40G optical technology", white paper, www.mintera.com/WhitepaperWEBnew.pdf, 2002.`

[24] C. Qiao, M. Yoo, "Optical Burst Switching - A new Paradigm for an Optical Internet", Journal of High Speed Networks, Spec. Iss. On Optical Networking, vol. 8, no. 1, Jan. 2000, pp. 36-44

[25] M. Maimour and C. Pham, "Dynamic Replier Active Reliable Multicast (DyRAM)", Proceedings of the 7th IEEE Symposium on Computers and Communications (ISCC 2002)

[26]J. Gaither, "300-Pin MSA Bit-Error Rate Tester for the ML10G Board and RocketPHY Transceiver", XILINX, Application note: XAPP677 Virtex-II Pro family, Jan. 2004.

[27]C. Guillemot et. al. "Transparent optical packet switching: The European ACTS KEOPS project approach", IEEE Journal of Lightwave Technology, vol.16, no.12, pp. 2117-2134, Dec. 1998.

[28]A. Stavdas, S. Sygletos, M. O'Mahony, H.L. Lee, C. Matrakidis, A. Dupas, "IST-DAVID: concept presentation and physical layer modelling of the metropolitan area network", IEEE Journal of Lightwave Technology, vol.21, no.2, pp. 372-383, Feb. 2003.

[29]D.K. Hunter, M.H.M Nizam, M.C. Chia, K.M. Guild, A. Tzanakaki, M.J. O'Mahony, J.D. Bainbridge, M.S.C. Stephens, R.V. Penty, I.H. White, "WASPNET: A wavelength switched packet network", IEEE Communications Magazine, vol.37, pp.120-129, Mar. 1999.

[30]F. Xue, Z. Pan, H. Yang, J. Yang, J. Cao, K. Okamoto, S. Kamei, V. Akella, S.J.B. Yoo, "Design and Experimental Demonstration of a Variable-Length Optical Packet Routing System With Unified Contention Resolution", IEEE Journal of Lightwave Technology, vol.22, no.11, pp.2570-2581, Nov. 2004.

[31]D. Klonidis, R. Nejabati, C.(T.) Politi, M.J. O'Mahony, D. Simeonidou, "Demonstration of a fully functional and controlled optical packet switch at 40Gb/s", in proc. 30th European Conf. on Optical Comm., Stockholm, Sweden, PD Th4.4.5, Sep. 2004.

[32] P. Thysebaert, B. Volckaert, et al. "Towards Consumer-Oriented Photonic Grids", Workshop on Optical Networking for Grid Services (ECOC 2004), Sep 2004.

[33]] E. Van Breusegem, M. De Leenheer, et al., "An OBS Architecture for Pervasive Grid Computing", Workshop on Optical Burst Switching (Broadnets 2004), Oct 2004.

[34] M. De Leenheer, E. Van Breusegem, et al., "An OBS-based Grid Architecture", Workshop on High Performance Global Grid Networks (Globecom 2004), Nov 2004.

[35] C. Partridge, T. Mendez, W. Milliken, "Host Anycasting Service", IETF RFC 1546, Nov 1993.

[36] S. Bhattacharjee, M.H. Ammar, et al., "Application-Layer Anycasting", Proc. of IEEE Infocom'97, Apr 1997.

[37] E. Basturk, R. Engel, et al., "Using Network Layer Anycast for Load Distribution in the Internet", IBM Research Report (RC20938), Jul 1997.
[38] Rosberg, Z., et al.," Performance analyses of optical burst-switching networks", Selected Areas in Communications, IEEE Journal on, Volume: 21, Issue: 7, Sept.2003 Pages:1187 - 1197

[39] M. Yoo and C. Qiao, "Just-enough-time (JET): A high speed protocol for bursty traffic in optical networks", In IEEE/LEOS Technol. Global Information Infrastructure, pages 26{27, August 1997.

[40] Foster, I., Kesselman, C. and Tuecke, S., " The Anatomy of the Grid: Enabling Scalable Virtual Organizations", International Journal of High Performance Computing Applications,15 (3). 200-222. 2001.

[41] M. Maimour and C. Pham, "Dynamic Replier Active Reliable Multicast (DyRAM)", Proceedings of the 7th IEEE Symposium on Computers and Communications (ISCC 2002)