OBS/GMPLS Interworking Network with Scalable Resource Discovery for Global Grid Computing
(Invited Paper)

J. Wu, L. Liu, X.B. Hong, and J.T. Lin

P.O. Box 55#, Key Laboratory of Optical Communication and Lightwave Technologies, Ministry of Education, Beijing University of Posts and Telecommunications, Beijing, China
{jianwu,xbhong}@bupt.edu.cn, liulei.bupt@gmail.com

Abstract. In recent years, Grid computing is more common in the industry and research community and will open to the consumer market in the future. The final objective is the achievement of global Grid computing, which means that the computing and networks are flexibly integrated across the world and a scalable resource discovery scheme is implemented. In this paper, a promising architecture, i.e., optical burst switching (OBS)/generalized multi-protocol label switching (GMPLS) interworking network with Peer-to-Peer (P2P)-based scheme for resource discovery is investigated to realize a highly scalable and flexible platform for Grids. Experimental results show that this architecture is suitable and efficient for future global Grid computing.

Keywords: optical Grid, OBS, GMPLS, resource discovery, P2P.

1 Introduction

Optical networks have been identified as the network infrastructure that would enable the widespread development of Grid computing, i.e. global Grid computing. However, the large bandwidth connections are not sufficient for the requirements of global Grid computing. In order to achieve this goal, a scalable and efficient architecture must be considered.

In this paper, a promising architecture, i.e., OBS/GMPLS interworking network with P2P-based scheme for resource discovery is investigated through experiments. The results show that this architecture is suitable and efficient for future global Grid computing.

The rest of this paper is organized as follows. Section 2 proposes the OBS/GMPLS interworking network. Section 3 investigates the P2P-based resource discovery scheme. Section 4 concludes this paper.

2 Network Infrastructure

The choice of the network infrastructure for global Grid computing is mainly driven by the fulfillment of a number of requirements, such as high bandwidth utilization (user's

P. Primet et al. (Eds.): GridNets 2008, LNICST 2, pp. 243 – 250, 2009.

requirement), a controllable and manageable infrastructure (network manager's requirement), and efficient job transmission in a large-sized network, etc.

Clearly, OBS can certainly fulfill the user requirement of efficient bandwidth utilization for its statistic multiplexing [1]. But due to the one-tier signaling protocols (e.g. JET, JIT), OBS is not suitable for large-sized networks. On the contrary, GMPLS utilizes a unified control plane to manage heterogeneous network elements and is expected to be gradually deployed to provide flexible control and management for carrier's optical transport networks in the near future. So the best solution for future global Grid computing is OBS over GMPLS network, which limits the number of nodes involved in the OBS signaling transactions, and rely on GMPLS control plane to handle recovery and network resource optimization issues.

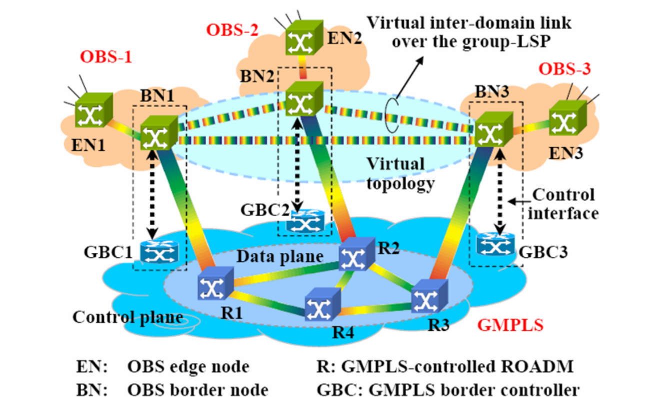

Fig. 1. OBS over GMPLS multi-layer architecture

This architecture is firstly proposed by Open Grid Forum (OGF) and investigated in detail in [2]. In this paper, such network architecture is demonstrated through experiments. We investigated an overlay model, as shown in Fig. 1, for the integrated OBS over GMPLS network. In this overlay network, the OBS networks are located at the edge and are capable of accessing various bursty traffic, while the GMPLS core network is responsible for providing layer-1 interconnections for the attached OBS client networks. An OBS Border Controller (OBC) and a GMPLS Border Controller (GBC) are introduced at the border between OBS client network and GMPLS core network. Each pair of OBC and GBC has a dedicated control interface between them in order to enable the GBC to dynamically establish or release a light path in GMPLS network for Data Burst (DB) transmission upon the request from OBC. Moreover, a simple client/server model with request/reply transactions can be applied between OBC and GBC, so no inter-layer routing or signaling between GMPLS and OBS networks is required, which is suitable for the network migration.

This overlay model is simple for the integrated OBS over GMPLS network. However, since the Burst Header Packet (BHP) in the OBS network needs to precede its corresponding DB (i.e. Grid job) by a particular offset time in order to configure the optical switch in advance, the BHP must traverse the same route and undergo the

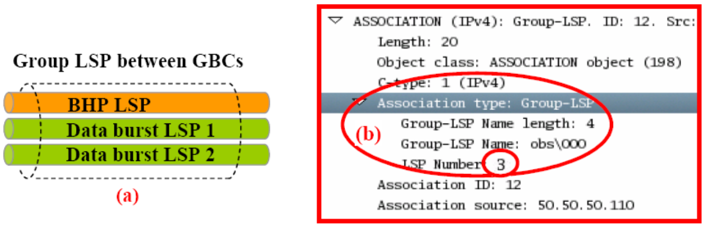

Fig. 2. Group LSP in GMPLS network for BHP/DB transmissions

same transmission latency with the DB. This requirement cannot be guaranteed if the BHP is transported over a conventional GMPLS control plane. Therefore, a light path extendedly provisioned by a so-called group-LSP as shown in Fig. 2, which consists of multiple LSPs with one dedicated LSP for BHPs, is introduced here for BHP/DB transmissions in the GMPLS data plane.

Fig. 3 illustrates the detailed routing and signaling procedures on how to conduct the inter-domain Grid job transmission in this overlay-based network. The OSPF protocol can still be used for intra-domain routing in each independent OBS network. In addition, considering only one control channel is required during the period of inter-domain routing, the group LSP established beforehand may contain only one single BHP LSP, while other DB LSPs will be dynamically created upon the request of traffic. Once the inter-domain reachability information is distributed to intra-domain nodes, the OBS edge node will be able to access the traffic destined to

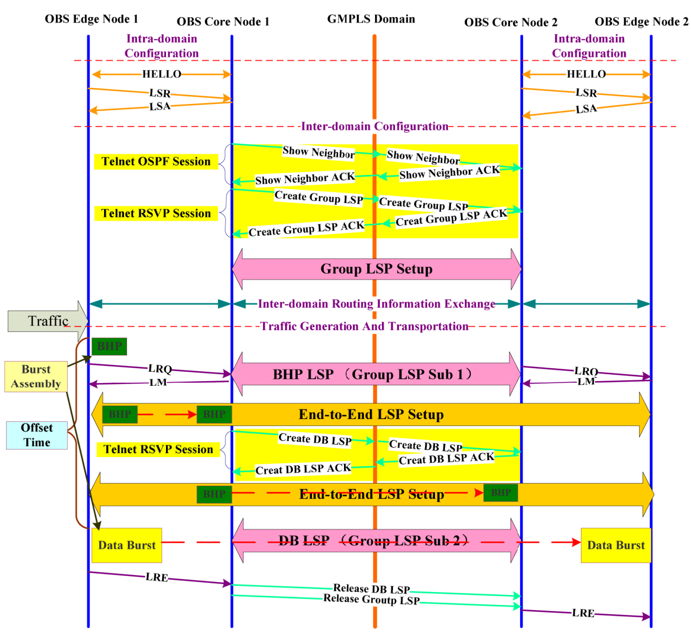

Fig. 3. Routing and Signaling procedures in the OBS over GMPLS network

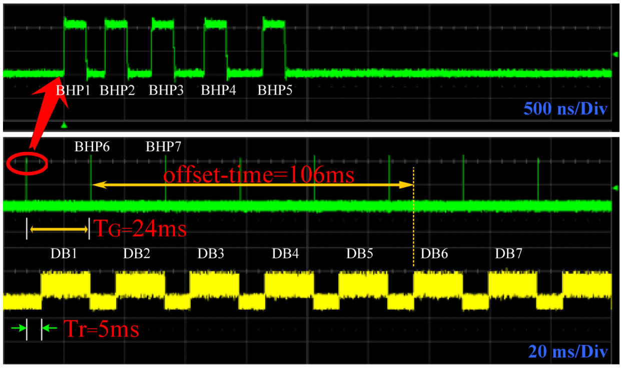

Fig. 4. Experimental results for an e2e LSP setup for Grid job transmission

external domains. Meanwhile, an e2e LSP via the BHP LSP can also be established across the GMPLS network in order to forward the inter-domain BHP/DB pair. Finally, when the BHP destined to peered OBS domain arrives at the exit of OBS border node, the group LSP containing BHP LSP and DB LSPs will be selected to forward the BHP and its corresponding DB. However, if the DB LSP on specified wavelength does not exist, OBC will send a request to the attached GBC to create a new one.

The results in Fig.4 verified the end-to-end Grid job transmissions. In the experiment, the BHP length is about 200ns, the assembled burst length and burst interval are configured as 16ms and 25ms respectively. The offset time between data burst and its BHP is set to 106ms taking into account the LSP setup latency of about 101ms. It can be seen from the Fig.4 that after the first data burst (DB1) destined to OBS Edge Node 2 (as shown in Fig.3) is generated, a new E-E LSP setup is triggered. Before the end of this procedure, there are total 5 data bursts and BHPs are generated and buffered at OBS Edge Node 1 (as shown in Fig.3). As a new LSP is available, the 5 BHPs buffered inside the edge node 1 are sent out in a very short time, which make them seems to be only one BHP in the large time scale as shown in the middle part of Fig.4. But with small time scale of 500ns/Div, they could be clearly distinguished as shown in the top part of Fig.4. Because the first BHP (BHP1) is buffered at edge node 1 for about 101ms, the offset time remained between DB1 and BHP1 is only T_r=106ms-101ms=5ms, though the initial offset time is 106ms. For the successive 4 data burst, the remained offset time can be calculated as 30ms, 55ms, 80ms and 105ms, respectively. When the sixth data burst and its BHP are generated, the LSP is already established. The BHP will be sent out immediately and the T_r remains 106ms as it is configured, as shown in Fig.4. At this moment, a small offset time could be set instead in order to reduce the end-to-end latency in case of the delay-sensitive traffic.

3 Scalable Resource Discovery Scheme

Currently, all the available resource discovery and management schemes are conventional Client/Server-based (centralized schemes) [3-7]. In these solutions, several centralized servers, which stores real time information about the available grid resources, resides in the network to deal with all the resource discovery requests and make decisions to these requests. The decision will be sent back to tell the users where to send out the job. However, this C/S mechanism confines the scalability of the whole Grid network. These solutions are poor scalability, fault-tolerance and low efficiency with the increasing of Grid users and job frequency.

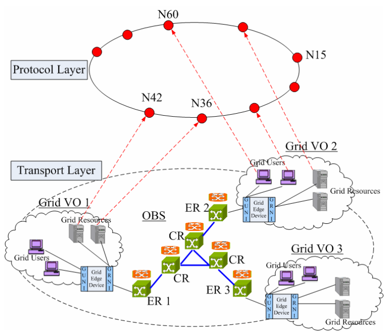

Fig. 5. Optical Grid architecture with P2P-based resource discovery capability

In order to solve these disadvantages, a P2P-based resource discovery scheme (decentralized scheme) is proposed in [8], which distributed the resource information on many nodes using distributed hashing table [9-10]. Fig.5 shows the Grid over OBS architecture, in which a virtual protocol layer is introduced to implement the peer-to-peer resource discovery scheme. It is composed of the virtual nodes mapping from the actual Grid users/resources. The consistent hash function assigns each node in the protocol layer an m-bit identifier using SHA-1[9] as a base hash function. The signaling process of the P2P-based resource discovery is introduced in detail in [8].

As same as the experimental setup in [3] and [8] on our OBS testbed [11], the performance between C/S-based resource discovery and P2P-based resource discovery can be evaluated. As shown in Fig.6, when job request frequency increasing, the performance of P2P model is not degraded, which is contrary to the results of C/S-based scheme. It is because in the C/S-based solution, the server stores so many items about the resource information, while in P2P solution, these items are distributed on many nodes due to distributed hashing table. A job request needs to check these items to find out a most appropriate resource. So in the C/S solution, when the request frequency is high, the server cannot handle the resource query and update timely. It will result in sharp increased average response time (Fig.6(a)) and the out-of-date resource information is often queried by the requests, which will causes a lower resource discovery successful rate (Fig.6(b)).

Clearly, by employing this P2P-based resource discovery mechanism, it is not necessary to keep the powerful centralized server and distribute real-time information to it, which makes up a more feasible and scalable Grid network. The experimental results show that P2P-based scheme outperforms C/S schemes in terms of discovery successful rate and discovery time in a large-scale Grid environment with high job request frequency. Thus it is suitable for future global Grid computing.

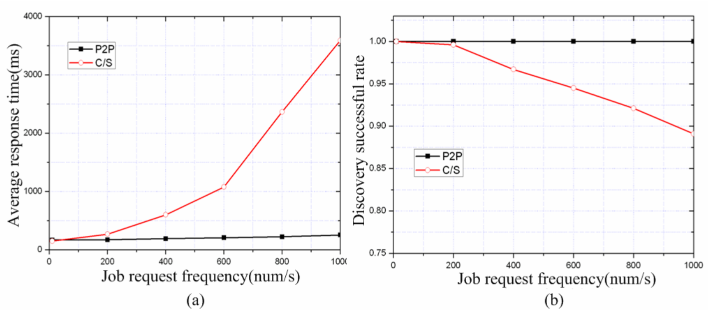

Fig. 6. Comparison of C/S-based and P2P-based resource discovery schemes (a) average response time v.s. job request frequency (b) discovery successful rate v.s. job request frequency

Although the P2P-based resource discovery schemes can improve the scalability of the Grid networks, the shortcoming of the P2P scheme is that only non-network resource is taken into consideration for resource discovery.

In order to solve this problem, a novel Self-organized Resources Discovery and Management (SRDM) scheme based on the P2P scheme is investigated in [12]. In the SRDM scheme, each user runs the P2P protocol proposed in [8], and two new tables: Latency Table (LT) and Blocking Table (BT) are added, which are used for storing the end-to-end latency and blocking probability from the current user to different resources. Such information is obtained by self-learning mechanism. Once a resource is discovered, the IP address of this resource will be saved in LT and BT. For each resource, user periodically sends “Hello” signaling to it to get the end-to-end latency and save it in LT. Meanwhile, user records the job history (success or failure), calculates job blocking probability for each resource and saves it in BT. BT is cleared in every T minutes to eliminate the out-of-data information. Note that although the exact network information may be obtained from the network manager, in many operational cases, it will increase the burden of manager and confine the scalability. Moreover, carriers may not want to expose too much detailed internal network information to customers due to the security concern.

The process of SRDM is described as follows: firstly, user can specify the job requirements and job characteristic (i.e. loss-sensitivity or delay-sensitivity) through a web portal in which dynamic Web Service technology is implemented. After that, P2P-based resource discovery scheme will be utilized to obtain a list of candidate resources satisfying the specified requirement. Then the end-to-end blocking probability and latency from user to these resources will be compared in order to choose a best one (least blocking, least latency) according to the job characteristic. Random or first-fit mechanism can be used if there is no relevant record in LT and BT. After a resource is chosen, non-network resource can be reserved by using the same method in [8]. Together with OBS-JET and GMPLS-RSVP signaling, SRDM enables a flexible end-to-end reservation of both network and non-network resources in a fully decentralized manner.

The experimental setup is shown in Fig.7. Various latency and blocking probability from users to resource 1 and 2 was introduced by injecting background traffic. The SRDM signaling (e.g. Hello) were encapsulated into bursts for transmission for

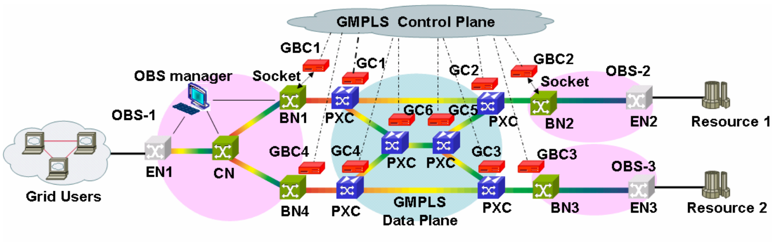

Fig. 7. Experimental setup

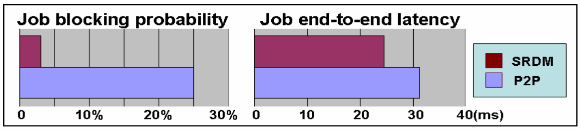

Fig. 8. Comparison between P2P [8] and SRDM

avoiding the O/E/O conversion and message processing delay. About 1000 jobs were randomly generated with random resource requirement, job characteristic and start time. The results in Fig.8 show that SRDM outperforms P2P scheme in terms of job blocking and end-to-end latency since the resource discovery in SRDM is capable of consideration of both network and non-network resources. It can be seen that in SRDM, each user has its own intelligence to manage resource discovery requests and make proper decision based on its own information about the whole Grid network. Clearly, by employing this distributed mechanism, it is not necessary to deploy powerful centralized servers for storing Grid resource information, which enables to construct a more scalable and fault-tolerant network for future global Grid computing.

4 Conclusions

In this paper, a promising architecture, OBS over GMPLS network infrastructure with P2P-based resource discovery scheme is proposed for future global Grid computing. Experimental results show that this architecture is efficient and scalable. Due to the wide using of GMPLS and P2P technology in the world now, we believe the work in this paper can facilitate the achievement of global Grid computing.

Acknowledgments. This work was supported by 863 Program (2007AA01Z248), MOST Program (No.2006DFA11040), PCSIRT (No.IRT0609) and 111 Project (B07005).

References

1. Qiao, C., Yoo, M.: Optical Burst Switching (OBS)—a New Paradigm for an Optical Internet. J. High Speed Netw. 8(1), 69–84 (1999)
2. Nejabati, R. (ed.): Grid Optical Burst Switched Networks (GOBS). Technical report, Open Grid Forum (OGF), GFD.128 (2007)

3. Zervas, G., Nejabati, R., Wang, Z., Simeonidou, D., Yu, S., O'Mahony, M.: A Fully Functional Application-aware Optical Burst Switched Network Test-bed. In: Optical Fiber Communication Conference (OFC), Anaheim, California, USA (2007)
4. Vokkarane, V.M., Zhang, Q.: Reliable Optical Burst Switching for Next-generation Grid Networks. In: IEEE/CreateNet GridNets, Boston, USA, pp. 505–514 (2005)
5. De Leenheer, M., et al.: An OBS-based Grid Architecture. In: Global Telecommunications Conference (GLOBECOM) workshop, Dallas, USA, pp. 390–394 (2004)
6. Farahmand, F., De Leenheer, M., Thysebaert, P., Volckaert, B., De Turck, F., Dhoedt, B., Demeestert, P., Jue, J.P.: A Multi-layered Approach to Optical Burst-switched Based Grids. In: International Conference on Broadband Networks (BroadNets), Boston, USA, vol. 2, pp. 1050–1057 (2005)
7. Zervas, G., Nejabati, R., Simeonidou, D., Campi, A., Cerroni, W., Callegati, F.: SIP Based OBS Networks for Grid Computing. In: Tomkos, I., Neri, F., Solé Pareta, J., Masip Bruin, X., Sánchez Lopez, S. (eds.) ONDM 2007. LNCS, vol. 4534, pp. 117–126. Springer, Heidelberg (2007)
8. Liu, L., Hong, X.B., Wu, J., Lin, J.T.: Experimental Demonstration of P2P-based Optical Grid on LOBS Testbed. In: Optical Fiber Communication Conference (OFC), San Diego, USA (2008)
9. FIPS 180-1, Secure Hash Standard. U.S. Department of Commerce/NIST, National Technical Information Service, Springfield, VA (1995)
10. Stoica, I., Morris, R., Karger, D., Kaashoek, F., Balakrishnan, H.: Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. In: ACM SIGCOMM, San Diego, USA, pp. 149–160 (2001)
11. Guo, H., Lan, Z., Wu, J., Gao, Z., Li, X., Lin, J., Ji, Y., Chen, J., Li, X.: A Testbed for Optical Burst Switching Network. In: Optical Fiber Communication Conf (OFC), Anaheim, California, USA (2005)
12. Liu, L., Guo, H., et al.: Demonstration of a Self-organized Consumer Grid Architecture. In: European Conference on Optical Communications (ECOC), Brussels, Belgium (accepted) (2008)