Erlang Reduced Load Model for Optical Burst Switched Grids

Marc De Leenheer, Chris Develder, Filip De Turck, Bart Dhoedt, Piet Demeester
Dept. of Information Technology - IBBT
Ghent University
Gaston Crommenlaan 8 bus 201, 9050 Gent, Belgium
Email: marc.deleenheer@intec.ugent.be

***Abstract*—This paper presents an Erlang reduced load model to analyze Optical Burst Switched Grid networks. The model allows the evaluation of job blocking probabilities, in which blocking occurs in both the transport network and the resources where jobs are processed. Additionally, a novel routing strategy is introduced to improve the usage efficiency of the existing infrastructure. Simulation analysis is used to confirm the validity and accuracy of the model, and the effectiveness of the novel routing strategy is shown for Grids over OBS.**

I. Introduction

Grid computing aims to offer a unified interface to access various resources such as computational clusters, data storage sites and scientific instruments. In general, these resources are heterogeneous in nature, are distributed on a global scale and have differing access policies. The main driver to realize this Grid technology are the highly challenging applications which emerge mainly from large-scale, collaborative experiments and the eScience field. Since the datasets involved in such applications pose a challenge to the transport network, photonic networks appear to be the most suitable solution. In particular, Wavelength Division Multiplexing (WDM) allows simultaneous access to multiple wavelengths on a single strand of fiber, and each wavelength offers data rates of 40 Gbps and more. Additionally, optical cross connects (OXC) make it possible to switch these wavelengths over multiple network hops (generally referred to as a lightpath), without costly O/E/O conversions. In this way, an Optical Circuit Switched (OCS) network is created, which can make efficient use of available bandwidth as long as data traffic between end nodes remains high.

However, efficiency drops rapidly in case bandwidth requirements of individual end users decrease [1]. This is frequently the case for applications geared towards enterprise and consumer markets [2]. A possible solution is Optical Burst Switching [3], since it allows access to bandwidth on a sub-wavelength scale, and as such statistical multiplexing of several data transfers (called bursts) is possible on a single wavelength. This approach could prove essential in the realization of true global-scale Grid computing, where a very diverse set of applications is supported on a single, common data plane.

The initial proposal for OBS has quickly gained attention in research communities and has delivered a number of theoretical studies on performance evaluation and fairness [4]–[6]. The first accurate model to evaluate the blocking behaviour of OBS networks, appeared in [7]. This was based on a reduced load model such as the one that is presented in this paper.

These works focused almost exclusively on the OBS technology in itself, without incorporating Grid-related concepts. In contrast, several papers have addressed the role of introducing network awareness in Grid scheduling algorithms [8], [9]. These clearly demonstrated the need for an integrated approach (i.e. network and end resources) for optimal job scheduling in Grid networks, even though no specific attention was given to Grids based on an OBS network.

This paper presents a reduced load model for an Optical Burst Switched Grid network, supporting various job scheduling and routing approaches. The model explicitely incorporates elements from both OBS networks and Grids. Additionally, a novel routing strategy is proposed, which tries to optimize the efficiency of OBS-based Grid networks. The main idea of the strategy is the realization that jobs can be processed at multiple resources in the network, and as such we should abandon the idea of routing towards fixed destinations. A similar approach has been pursued in e.g. [10]. In this paper, we establish the validity and accuracy of our model through comparison with simulation analysis, and show the increased efficiency of existing infrastructure by our novel routing strategy.

The remainder of this paper is structured as follows. In Section II, the general operation of an OBS-based Grid is presented, and job routing and scheduling strategies are discussed. We then proceed to present the reduced load model, which is followed by a validation and analysis of results in Section IV. Consequently, future work is discussed and our conclusions are summarized in Section VI.

II. Job Routing and Scheduling Strategies

In general, an OBS-based Grid is composed of an OBS network, which has clients and resources attached. Clients can generate jobs (possibly composed of executables and data), and assemble them into one or more optical bursts. After transferring these bursts over the optical network, they will be processed in a resource. Afterwards, it is possible that job results need to be sent back to the client or to some other specified destination.

The actual decision of where to process the burst and how to reach that destination, is traditionally made in scheduling entities. This decision is based on the current Grid state, the specific job requirements and various pre-determined optimization criteria. This approach has proven sufficient for most scenarios, but is not well adapted to the possible highly dynamic nature of a Grid environment. Indeed, in case large users groups are to be supported (e.g. consumer grids), the unpredictable and highly dynamic behaviour of user requests (and correspondingly, the resource and network states) can result in non-optimal use of existing infrastructure. A possible solution lies in the realization that there usually exist multiple, feasible resources for the execution of a specific job. As such, the assignation of a fixed, *hard destination* for a job should be abandoned in favour of a *soft destination*. Even though a job should still try to reach that soft destination, any suitable resource which is passed during the transfer, should be considered as a possible location for processing that job. As such, soft destination assignment can be regarded as an approach to schedule a job at multiple resources at once, whereby the availability of each resource is checked in a sequential manner. The soft destination approach can also be viewed as a form of *anycast routing* [11], since clients are not aware of which resource will do the actual servicing of the job. Finally, note that this mode of operation requires explicit support of the network's control plane; in the soft destination approach, the network router will be made aware of the resource availability. In this way, the router can quickly decide whether a specific job can be executed on the locally attached resource, instead of offering the job to the local scheduler and await its scheduling decision.

III. Reduced Load Approximation

This section starts with an overview of the various parameters and notations necessary to describe a Grid based on an OBS network. We then present the various steps required to obtain the global job blocking probability.

A. Grid Network Model

As shown in Figure 1, consider a network composed of a set L of directed links[1], each link l having W_l wavelengths with transmission rate α_l (expressed in jobs per time unit). Each link is terminated at both ends by a router, which are all capable of full wavelength conversion. The network also contains a set of sources S, with each source s generating jobs according to a Poisson arrival process with mean job arrival rate λ_s. Jobs are executed on a set of resources R, each resource r composed of C_r CPUs which have a mean processing rate β_r (jobs per time unit). Finally, each source and resource are connected to a single network router and their access link is neglected in this model (i.e. no blocking occurs on the access links).

Scheduling and routing policies are incorporated as follows. Let d_{sr} be the probability that a job which originated at

[1]shown undirected in the figure for clarity

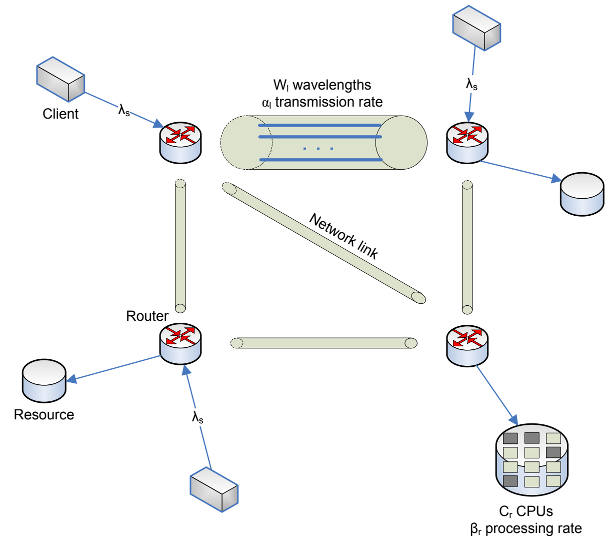

Fig. 1. Overview of Grid network model

source s is sent to resource r. This probability represents the scheduling policy of a source, and obviously, for each source s it holds that $\sum_r d_{sr} = 1$. The single routing path between each (source, resource) pair is represented by $P(s,r)$, which equals an ordered set of links. For notational convenience, we also introduce the following sets:

- $P^{\text{net}}_{sr}(l)$ is the set of links which come before link l on path $P(s,r)$, empty if $l \notin P(s,r)$
- $Q^{\text{net}}_{sr}(u)$ is the set of links which come before resource u on path $P(s,r)$, empty if $\forall v \in R : (u,v) \notin P(s,r) \wedge (v,u) \notin P(s,r)$
- $P^{\text{res}}_{sr}(l)$ is the set of resources which come before link l on path $P(s,r)$, empty if $l \notin P(s,r)$
- $Q^{\text{res}}_{sr}(u)$ is the set of resources which come before resource u on path $P(s,r)$, empty if $\forall v \in R : (u,v) \notin P(s,r) \wedge (v,u) \notin P(s,r)$

The mutual effects of soft destination assignment and the deployed routing protocol could have dramatic effects on the network's performance, but this study is postponed for future work. Instead, we assume for the remainder of this paper that destination assignment (both soft and hard) follows a uniform distribution, i.e. each source sends an equal fraction of jobs to all resources. Additionally, shortest path routing is used.

B. Model Overview

Figure 2 shows an overview of our model and the different calculation steps. In general, we start from a given topology, the location and properties of clients and resources, and the implemented scheduling and routing policy. Based on this information, we want to obtain an estimate for the blocking probability of jobs in the Grid network. It is important to note that blocking can occur at two distinct locations in the network:

- in network links, due to network congestion, or
- in resources, due to overloaded resources.

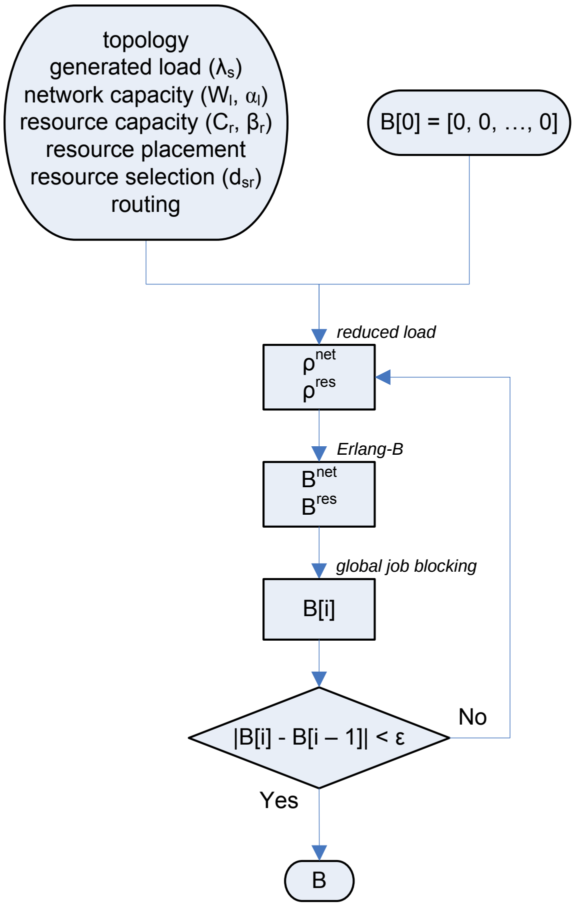

Fig. 2. Overview of reduced load approximation

To incorporate these two different causes for blocking, the algorithm starts by estimating the load on individual network links (ρ_l^{net}) and resources (ρ_r^{res}), based on a reduced load approach. This implies that the load on a network link or resource is reduced because of blocking events on other network links and/or resources. Consequently, we can calculate the individual blocking probabilities (B_l^{net} and B_r^{res}) by using the Erlang-B formula. This is based on the assumption that jobs are generated following a Poisson process, but this can evidently be replaced by other distributions if sufficient evidence can be gathered. Previous work [12] has shown that, in a large scale Grid, jobs do arrive (irrespective of the job creation process) at resources according to a Poisson process. Based on the blocking probabilities of individual network links and resources, we can obtain an estimate for the global job blocking probability ($B[i]$). This process is repeated until two successive iterations obtain an estimate for the global blocking which are sufficiently close to each other. The accuracy of the approximation can, as such, be varied by determining an appropriate value of this parameter ϵ. This technique is generally referred to as fixed point approximation. For more details on the convergence of the fixed point technique in this modelling approach, the reader is referred to [7].

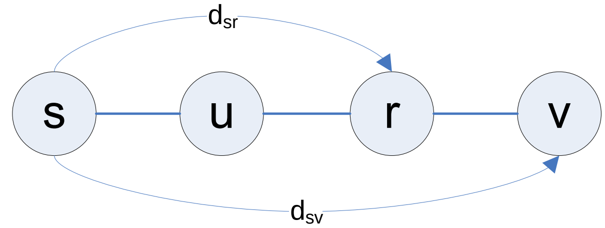

Fig. 3. Soft destination assignment: source s contributes to the blocking in resource r in two distinct ways

C. Load

In this section, we develop the relevant expressions to model the load experienced by individual links (ρ_l^{net}) and resources (ρ_r^{res}). This will be done for both traditional routing policies (i.e. hard destination assignment) and our novel routing policy (i.e. soft destination assignment). First note that we can write the individual loads in function of the arrival rate of all jobs that are offered to a network link λ_l^{net} (resp. resource λ_r^{res}):

$$\rho_l^{\text{net}} = \frac{\lambda_l^{\text{net}}}{\alpha_l}$$

$$\rho_r^{\text{res}} = \frac{\lambda_r^{\text{res}}}{\beta_r}$$

1) Hard Destination Assignment:

$$\lambda_l^{\text{net}} = \sum_{s\in S} \sum_{\substack{r\in R \\ l\in P(s,r)}} \left(\lambda_s d_{sr} \prod_{k\in P_{sr}^{\text{net}}(l)} (1 - B_k^{\text{net}}) \right)$$

$$\lambda_r^{\text{res}} = \sum_{s\in S} \left(\lambda_s d_{sr} \prod_{l\in P(s,r)} (1 - B_k^{\text{net}}) \right)$$

The previous expressions show the essence of the reduced load approach; the load experienced by a network link (resp. resource) is reduced because of the blocking which occurs on the preceding links of the path. Also note the assumption that blocking events occur independently of each other.

2) Soft Destination Assignment:

$$\lambda_l^{\text{net}} = \sum_{s\in S} \sum_{\substack{r\in R \\ l\in P(s,r)}} \left(\lambda_s d_{sr} \prod_{k\in P_{sr}^{\text{net}}(l)} (1 - B_k^{\text{net}}) \prod_{u\in P_{sr}^{\text{res}}(l)} B_u^{\text{res}} \right)$$

$$\lambda_r^{\text{res}} = \sum_{s\in S} \left(\lambda_s d_{sr} \prod_{k\in P(s,r)} (1 - B_k^{\text{net}}) \prod_{u\in Q_{sr}^{\text{res}}(r)} B_u^{\text{res}} \right) + \sum_{s\in S} \sum_{\substack{v\in R \\ r\in Q_{sr}^{\text{res}}(v)}} \left(\lambda_s d_{sv} \prod_{k\in P_{sv}^{\text{net}}(r)} (1 - B_k^{\text{net}}) \prod_{u\in Q_{sv}^{\text{res}}(r)} B_u^{\text{res}} \right)$$

Recall that soft destination assignment means that jobs are given a resource as destination, but can be executed on an intermediate resource which has free capacity. This implies that the load on a network link is not only dependent on the preceding links, but also on blocking of preceding resources. Also note the two distinct cases which contribute to the load on a resource r (Figure 3). The first term represents the load of jobs which were originally meant to be processed on that resource r. The second term is load generated by the fraction of jobs which were originally destined for resource v, but resource r (located on the path between source s and resource v) still has available capacity.

D. Blocking Probabilities

Based on the reduced loads offered to individual network links and resources, we can deduce their respective blocking probabilities, by using the Erlang-B formula:

$$
\begin{aligned}
B_l^{\text{net}} &= Erl(\rho_l^{\text{net}}, W_l) = \frac{\frac{(\rho_l^{\text{net}})^{W_l}}{W_l!}}{\sum_{i=0}^{W_l} \frac{(\rho_l^{\text{net}})^i}{i!}} \\
B_r^{\text{res}} &= Erl(\rho_r^{\text{res}}, C_r) = \frac{\frac{(\rho_l^{\text{res}})^{C_r}}{C_r!}}{\sum_{i=0}^{C_r} \frac{(\rho_r^{\text{res}})^i}{i!}}
\end{aligned}
$$

Once the blocking probabilities of the individual links and resources have been established, it follows that the global job blocking probability is given by:

$$
\begin{aligned}
B &= 1 - \frac{\sum_{r\in R} \lambda_r^{\text{res,eff}}}{\sum_{s\in S} \lambda_s} \\
&= 1 - \frac{\sum_{r\in R} \lambda_r^{\text{res}}(1 - B_r^{\text{res}})}{\sum_{s\in S} \lambda_s} \\
&= 1 - \frac{\sum_{r\in R} \beta_r \rho_r^{\text{res}}(1 - B_r^{\text{res}})}{\sum_{s\in S} \lambda_s}
\end{aligned}
$$

In the previous expression, $\lambda_r^{\text{res,eff}}$ is the effective arrival rate of jobs to resource r, and thus represents the arrival rate of jobs which will effectively be processed on that resource. Likewise, λ_r^{res} represents the arrival rate of all jobs that are offered to resource r, as discussed in Section III-C.

IV. Validation and Analysis

The basic European topology, depicted in Figure 4, was used for our validation. This network is composed of 28 network routers and 41 bidirectional links. Each router has a client attached with a fixed job arrival rate ($\lambda_s = \lambda = 1000$ jobs per second). Six resources are installed at a fixed location (routers: Amsterdam, Paris, Berlin, Budapest, Rome and Madrid), and have a fixed processing rate (i.e. $\beta_r = \beta$) depending on the load scenario. Each resource r contains $C_r = 20$ CPUs, while each network link l has $W_l = 20$ wavelengths and a fixed transmission rate ($\alpha_l = \alpha$), also depending on the specific load scenario. As mentioned previously, we implemented a uniform scheduling policy, i.e. $d_{sr} = \frac{1}{|R|}$, and shortest path routing was used for all results.

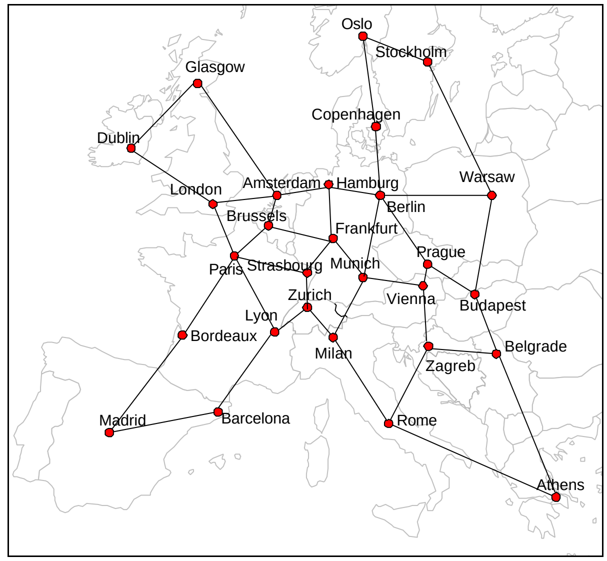

Fig. 4. Simulated topology: basic European network (28 nodes, 41 bidirectional links)

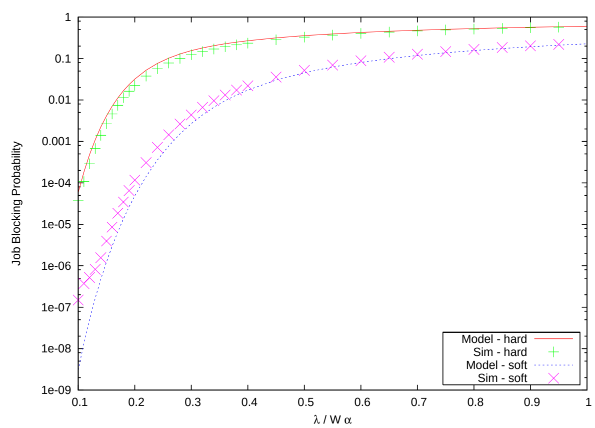

Fig. 5. Job blocking probability for varying generated network load and fixed mean generated resource load ($\frac{\lambda}{C\beta} = .01$)

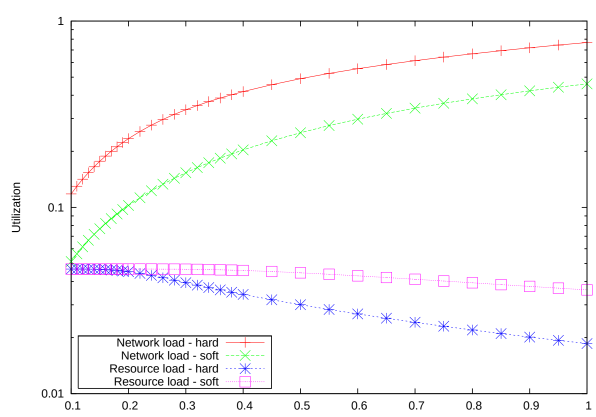

Fig. 6. Network and resource utilization for varying generated network load and fixed mean generated resource load ($\frac{\lambda}{C\beta} = .01$)

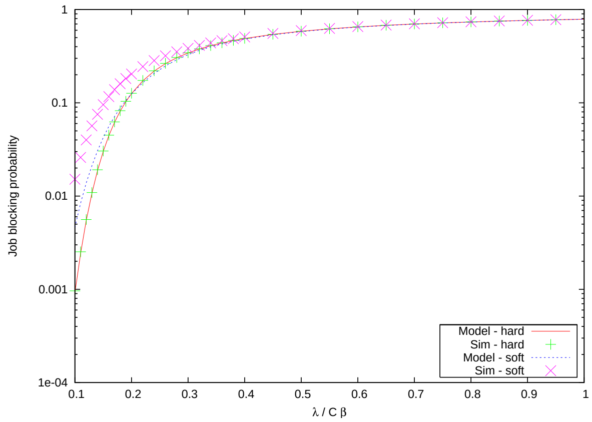

Fig. 7. Job blocking probability for varying generated resource load and fixed mean generated network load ($\frac{\lambda}{W\alpha} = .01$)

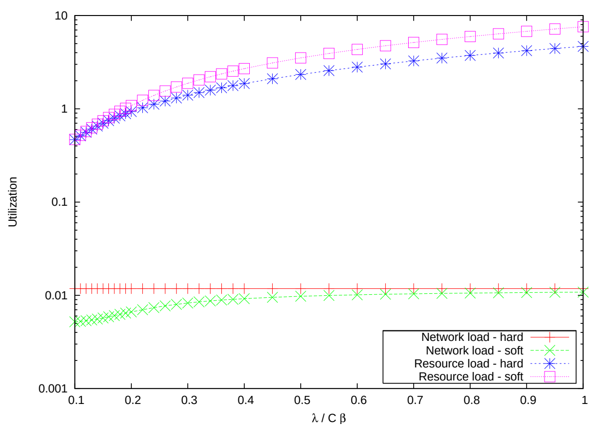

Fig. 8. Network and resource utilization for varying generated resource load and fixed mean generated network load ($\frac{\lambda}{W\alpha} = .01$)

Figure 5 shows the job blocking probability for varying generated network loads and a fixed mean generated resource load ($\frac{\lambda}{C\beta} = .01$ which implies resource blocking should be negligable). This varying load $\frac{\lambda}{W\alpha}$ can also be interpreted as a varying link dimensioning, i.e. $\frac{\lambda}{W\alpha} \in [0, 1]$ is equivalent to $\alpha \in [\frac{W}{\lambda}, 0]$ for fixed values of W and λ. An immediate observation is the accuracy of the proposed model in comparison to the simulation results. Another important conclusion is that the soft destination approach clearly outperforms hard destination assignment. The difference in blocking behaviour can clearly be attributed to network blocking events (see Figure 6). Soft destination assignment makes use of resource capacity as soon as possible, and as such generates a lower utilization of the transport network. In summary, soft destination improves the blocking behaviour whenever network capacity is the limiting factor.

In Figure 7, the job blocking probability is shown for varying generated resource loads and a fixed mean generated network load ($\frac{\lambda}{C\beta} = .01$). Similarly to the previous discussion, we can conlude the validity and accuracy of the reduced load model. The soft destination approach initially shows worse blocking behaviour than hard destination assignment, which is a consequence of the limited availability of resource capacity. Indeed, jobs with intermediate resources on their path toward their soft destination, will almost always be processed on that resource. However, jobs without intermediate resources on their path will arrive at their soft destination which is experiencing an increased utilization, and thus a higher job blocking probability. This is confirmed by Figure 8, which shows the increased resource utilization for soft destination assignment, although network utilization is decreased. This shortcoming of the soft destination approach is likely to be resolved by incorporating algorithms for advanced routing and intelligent resource dimensioning.

V. Future Work

Future work will continue to analyze various scheduling and routing strategies and their effect on the network's blocking behaviour. Another focus is the design of dimensioning and network planning algorithms; the placement of resources and determination of capacities and service rates of both resources and network links is essential to optimize the Grid network operation. In particular, the mutual effects of intelligent dimensioning before operation of the network, and routing algorithms in an online setting, are of great interest and will likely realize the best optimization. Finally, the relation between these algorithms and practical implementation issues, such as protocols for the control and signaling plane, will be investigated.

VI. Conclusion

In this paper, we presented an Erlang reduced load approximation to analyze the blocking behaviour of Optical Burst Switched Grid networks. Since job blocking in such networks can occur in both network links as well as in resources, we proposed a novel routing strategy (referred to as soft destination assignment) to improve the network's efficiency. Through simulation analysis, we showed the validity and accuracy of our model. Additionally, we demonstrated the improved blocking behaviour of the Grid network for our novel routing strategy, when compared to the traditional routing method. In particular, soft destination assignment is able to improve the Grid's blocking behaviour whenever network blocking is the main restriction. In contrast, when resource capacity is the limiting factor, soft destination performed slightly worse than hard destination. Results suggest that a combination of intelligent routing and resource dimensioning are able to remove this disadvantage.

Acknowledgment

M. De Leenheer thanks the IWT for their financial support through his Ph.D. scholarship. C. Develder thanks the FWO for his post-doctoral grant.

REFERENCES

[1] F. Xue, S.J.B. Yoo, H. Yokoyama, and Y. Hoiuchi, *Performance Comparison of Optical Burst and Circuit Switched Networks*, Proc. of the Optical Fiber Communication Conference (OFC), Anaheim, CA, USA, Mar 2005.

[2] M. De Leenheer, P. Thysebaert, B. Volckaert, F. De Turck, B. Dhoedt, P. Demeester, D. Simeonidou, R. Nejabati, G. Zervas, D. Klonidis, and M.J. O'Mahony, *A View on Enabling Consumer Oriented Grids through Optical Burst Switching*, IEEE Communications Magazine, 44(3):124-131, Mar 2006.

[3] C. Qiao, and M. Yoo, *Optical Burst Switching: A New Paradigm for an Optical Internet*, Journal of High Speed Networks, 8(1):69-84, Mar 1999.

[4] Q. Zhang, V.M. Vokkarane, J.P. Jue, and B. Chen, *Absolute QoS Differentiation in Optical Burst-Switched Networks*, IEEE Journal on Selected Areas in Communications, 22(9):1781-1795, Nov 2004.

[5] M. Düser, and P. Bayvel, *Analysis of a Dynamically Wavelength-Routed Optical Burst Switched Network Architecture*, Journal of Lightwave Technology, 20(4):574-585, Apr 2002.

[6] N. Barakat, and E.H. Sargent, *Separating Resource Reservations from Service Requests to Improve the Performance of Optical Burst-Switching Networks*, IEEE Journal on Selected Areas in Communications, 24(4):95-107, Apr 2006.

[7] Z. Rosberg, H.L. Vu, M. Zukerman, and J. White, *Blocking Probabilities of Optical Burst Switching Networks Based on Reduced Load Fixed Point Approximations*, Proc. of IEEE Infocom, Vol. 3, pp. 2008-2018, Mar 2003.

[8] P. Thysebaert, B. Volckaert, F. De Turck, B. Dhoedt, and P. Demeester, *Network Aspects of Grid Scheduling Algorithms*, Proc. of the ISCA 17th Int. Conference on Parallel and Distributed Computing Systems (PDCS), San Francisco, CA, USA, Sep 2004.

[9] K. Ranganathan, and I. Foster, *Simulation Studies of Computation and Data Scheduling Algorithms for Data Grids*, Journal of Grid Computing, 1(1):53-62, Mar 2003.

[10] J. Teng, and G.N. Rouskas, *A Traffic Engineering Approach to Path Selection in OBS Networks*, Journal of Optical Networking, 4(11):759-777, Nov 2005.

[11] T. Stevens, M. De Leenheer, F. De Turck, B. Dhoedt, and P. Demeester, *Anycast Routing Algorithms for Effective Job Scheduling in Optical Grids*, Proc. of the European Conference on Optical Communication (ECOC), Cannes, France, Sep 2006.

[12] K. Christodoulopoulos, M. Varvarigos, C. Develder, M. De Leenheer, and B. Dhoedt, *Job Demand Models for Optical Grid Research*, Submitted to the 11th Conference on Optical Network Design and Modelling (ONDM), Athens, Greece, May 2007.