Joint Path and Resource Selection for OBS Grids with Adaptive Offset Based QoS Mechanism

Mehmet Koseoglu
kmehmet@ee.bilkent.edu.tr

Ezhan Karasan
ezhan@ee.bilkent.edu.tr

Department of Electrical and Electronics Engineering
Bilkent University
06800 Ankara, Turkey

ABSTRACT

In this paper, we first propose a joint resource and path selection algorithm for optical burst switched grids. We show that path switching and network-aware resource selection can reduce burst loss probability and average completion time of grid jobs compared to the algorithms that are separately selecting paths and grid resources. In addition to joint resource and path selection, we present an adaptive offset algorithm for grid bursts which minimizes the average completion time. We show that the adaptive offset based QoS mechanism significantly reduces the job completion times by exploiting the trade-off between decreasing loss probability and increasing delay as a result of the extra offset time.

Keywords

Optical Burst Switching, Grid Networking

1. INTRODUCTION

As the bandwidth requirements of grid applications increase, optical networks are now being considered for the network infrastructure of grids. A well known example for high bandwidth grid applications is the particle physics experiments performed at CERN[1]. These experiments generate terabytes of data which cannot be processed by using only local resources. For that reason, data has to be carried over long distances to be processed by distributed resources. For this type of high bandwidth applications, deployment of optical networks becomes a necessity [11].

In addition to e-science applications, it is predicted that more interactive applications such as remote rendering and interactive TV will run on the grid in the future. In these applications, local computational resources are not sufficient to process jobs interactively, so consumers get service from a remote computational facility. With further deployment of optical networks and improvement of QoS guarantees, it is possible to get seamless service for this type of applications. An experiment of remote rendering over an intercontinental optical network is described in [10] which shows that remote rendering is possible between continents.

GridNets 2007, October 17-19, 2007, Lyon, France

DOI 10.4108/gridnets.2007.2195

For interactive high bandwidth grid applications, optical burst switching (OBS) is seen as a more suitable technology than circuit switching [13]. The bandwidth granularity of OBS allows efficient for transmission of relatively small jobs in the grid and also the separation of the control and data planes in OBS can provide consumer initiated light-path setup. Moreover, it can be possible to map grid jobs to grid bursts one-to-one, so the grid jobs can be effectively transmitted using the bandwidth granularity of OBS.

Despite these advantages, there are several problems with OBS for grid computing. One of them is the high loss rates associated with one-way reservation. Although OBS allows fast transport of bursts, burst contentions in the core network occurs when the reservation attempt by the burst control packet is not successful if the capacity of a link is fully occupied by other bursts.

There are many studies in the literature to reduce burst contention. These studies can be divided into two main approaches: Edge assisted contention avoidance techniques and contention resolution techniques at the core routers. Contention resolution techniques requires incorporation of fiber delay lines, deflection routing and burst segmentation necessitating expensive hardware and complex software at the core routers. On the other hand, the edge assisted contention avoidance techniques are more practical for the near future since buffering and processing and processing at the edge nodes is feasible at relatively low costs.

One of the edge assisted contention avoidance techniques is path switching. Path switching is alternating transmission paths to a destination depending on the congestion in the network. It is shown that path switching can reduce loss rates especially when some of the nodes in the network become highly congested [14, 15].

There are several studies in the literature to support different classes of service over OBS networks. One of these methods is to use a larger offset time for high priority bursts [16]. The relationship between the offset duration and the loss rate is modeled analytically in [2, 7]. The drawback of this QoS mechanism is that the delay increases as the extra offset increases.

In this paper, joint path and resource switching is used to reduce loss rates in a grid network. Since the consumers in a grid can request service from several providers, it is possible to select resources and paths considering the congestion in the network. This joint selection mechanism outperforms algorithms which perform resource selection and path selection separately especially for high levels of traffic load. This scheme is then extended using an adaptive offset based QoS algorithm which computes offset value for grid bursts minimizing average completion time. Although applying an extra offset to grid bursts increases transmission delay, average completion time can be reduced by decreasing loss probability. Numerical results show that the proposed algorithm achieves smaller job completion times compared with using fixed offset especially under non-static traffic conditions.

This paper is organized as follows. Section 2 explains the OBS grid architectures proposed in the literature. The consumer and resource-side optimizations performed to minimize the average completion time are described in Section 3 and Section 4, respectively. Simulation results are presented in Section 5.

2. OBS GRID ARCHITECTURE

An OBS grid architecture is first proposed by [4]. In this architecture, grid jobs are mapped into OBS bursts and information about the grid job is embedded to the burst header. The bursts are sent to the network without a specific destination address and they are deflected to suitable resources by the intelligent OBS routers. Anycasting is used when a request can be performed by more than one servers. Several anycasting algorithms for grid OBS architecture is proposed and analyzed in [5].

A detailed OBS grid architecture is described in [9, 12] in which active networking is used for job specification dissemination. Difference of this architecture from the previous one is that anycasting is not used and the intelligent routers are sparsely placed. The job specification is sent to the nearest active router as an optical burst and it is multicasted to the other active routers by this router. These active routers send an acknowledgment (ACK) or a negative acknowledgment (NACK) burst to the consumer about the situation of the resources and they reserve the resource for a limited time. After receiving all ACK and NACK messages, the consumer selects a resource and transmits the job data using an OBS burst.

A modified version of this architecture is proposed in [6]. In this architecture, job specification is transmitted using the control plane instead of active bursts and all routers in the network are intelligent routers. There are two reservation mechanisms presented in this paper: In implicit discovery and reservation, the control packet of grid bursts are anycasted to a suitable resource reserving both the grid resources and network resources. In the explicit discovery mechanism, job specification is disseminated using the control plane and routers in the network return an acknowledgment to the consumer. Then, the consumer selects the resource and sends the job burst.

Since computational complexity should be placed at the edge routers instead of core routers, a consumer controlled version of these architectures is studied in this paper. The phases of grid job completion are explained as follows.

- Job Specification Dissemination: Instead of using a fully intelligent network, using sparse intelligent routers reduces hardware costs. For that reason, we study a partially intelligent network in which multicasting is used for job specification dissemination. The consumer sends the job specification to the nearest intelligent router and the specification is multicasted to other intelligent routers using the control plane.
- Grid Resource Reservation: When the intelligent routers receive the job specification, they query the resources. In [6, 12], the intelligent routers send an ACK or NACK to the consumer about the availability of resources. However, binary signaling is not sufficient for resource selection when there are more than one available resource. For that reason, we study an architecture where the intelligent routers send processing time estimations to the consumer and consumers use this information to perform resource selection. There are also other metrics that can be transferred to the consumer such as processing cost but for simplicity we use completion time as the single metric. The intelligent routers reserve the grid resources for a limited time in order to guarantee processing time offers as in [12].
- Resource selection: In contrast to [6], the resource selection is solely performed by the consumer in the proposed architecture not by the intelligent routers.
- Path Selection: Path selection is performed by consumers and core routers does not perform anycasting. A list of two link-disjoint paths between each consumer-grid resource pair is computed and one of these paths to a resource or consumer is adaptively chosen for sending a burst. These link-disjoint paths are computed in advance using an edge-disjoint path pair algorithm [3].
- Network Resource Reservation: In [6], wavelength reservation can be performed at the same time with the resource reservation. Since the resource is selected on-the-fly by intelligent routers, it is possible to make wavelength reservations at the same time. In contrast to this, we study a consumer controlled architecture so the consumer sends the job burst after performing resource and path selection as in [12].
- Feedback Collection: To perform congestion-based path and resource selection, feedback messages which carry information from the core routers to the edge routers must be employed. In the architecture we study, we use acknowledgment messages and probe packets to transfer feedback to consumers and resources, respectively. Core routers write their congestion information to the ACK packets and consumers receive the congestion information when they receive the resource situation. Also, resources send probe packets to consumers just before the completion of the grid job. Core routers write their congestion information to these probe packets and consumers send these packets back to the resource. On their way back, probe packets collect congestion information from the core routers and this information is used by the resource for choosing the path. The acknowledgment messages and probe messages are sent over both link-disjoint paths between the resource and the consumer in order to collect congestion information along both paths.
- Notification Burst Losses: Consumers and resources send a burst acknowledgment message to the source node when they receive a burst. Without this acknowledgment, the loss of a grid burst cannot be understood

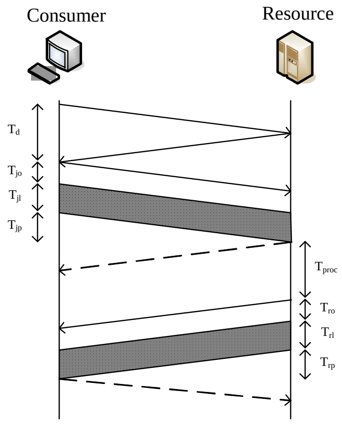

Figure 1: Timeline of a successfully transmitted grid job.

until a timeout duration large enough to include completion time and this results in increased delay.

In the next section, we analyze the phases of grid job execution and present the completion time optimization strategies from the consumer's point of view.

3. CONSUMER-SIDE OPTIMIZATION

In this section, we describe how the grid consumer chooses the grid resource, path and offset to minimize the grid completion time.

3.1 Completion Time and Retransmission Cost

The timeline of a successfully completed OBS grid job can be seen in Figure 1. The components of the lifetime of an OBS grid job are the following:

- T_d: Resource discovery delay
- T_{jo}: Offset time of the job burst
- T_{jl}: Transmission time of the job burst
- T_{jp}: Propagation delay of the job burst
- T_{proc}: Job processing time
- T_{ro}: Offset time of the job result burst
- T_{rl}: Transmission time of the job result burst
- T_{rp}: Propagation delay of the job result burst

From Figure 1 it can be seen that the minimum required time to complete a job is

$$T_{min} = T_d + T_{jo} + T_{jl} + T_{jp} + T_{proc} + T_{ro} + T_{rl} + T_{rp}$$

For simplicity, we assume that the job result burst size is equal to the job burst size, i.e., $T_{jl} = T_{rl} = T_l$ and the propagation delay of the job burst is equal to the propagation delay of the job result burst, i.e., $T_{jp} = T_{rp} + T_p$. We also assume that the required transmission offset, which is equal to the product of the number of hops on the path and the per-hop processing delay, is negligible with respect to other components. Under these assumptions the required time to transmit the job becomes

$$T_{min} = T_d + T_{jo} + 2T_l + 2T_p + T_{proc} + T_{ro} \qquad (1)$$

However, if the job burst is lost, the time needed to complete the job increases. The timeline of a grid job when the job burst is lost once can be seen in Figure 2.

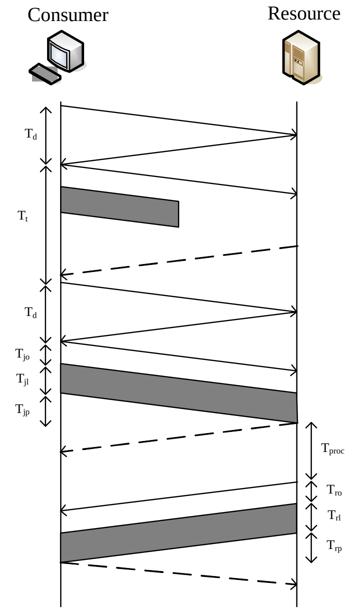

Figure 2: Timeline of a grid job when the job burst is lost.

The consumer detects the loss after a timeout since it does not receive the burst acknowledgment sent by the resource. When a job burst is lost, time required to detect the loss of a burst is denoted as T_t. This timeout duration consists of job burst transmission delay, job burst propagation delay, job burst QoS offset and propagation delay of the burst acknowledgment. Timeout duration should also include a guard band, T_g, for unpredictable delays.

$$T_t = T_l + T_p + T_{jo} + T_p + T_g$$

In addition to the timeout duration, resource discovery phase has to be performed again because the computational resources reserve their processors for a limited time. Consequently, the retransmission cost, i.e., the difference between job completion time and T_{min}, is given by

$$\begin{aligned} T_{rt} &= T_t + T_d \\ &= T_l + 2T_p + T_{jo} + T_g + T_d \qquad (2) \end{aligned}$$

Next, we use (1) and (2) to minimize the expected completion time of a grid job.

3.2 Expected Completion Time

Let $P_b^{(n)}$ be the loss probability of the grid job burst and $T_{rt}^{(n)}$ be the retransmission cost in the n^{th} transmission attempt, and, T_{min} is given by (1). Then the expected completion time can be written as

$$\overline{T} = T_{min} + \sum_{i=1}^{\infty}(\prod_{j=1}^{i} P_l^{(j)})T_{rt}^{(i)} \qquad (3)$$

Assuming that the network and computational resource conditions does not change between transmission attempts, we have $P_l^{(n)} = P_l$ and $T_{rt}^{(n)} = T_{rt}$. Then, the expected completion time of a grid job can be expressed as

$$\overline{T} = T_{min} + T_{rt}\frac{P_l}{1-P_l}$$

Next, we discuss how this expected retransmission cost can be used in resource and path selection.

3.3 Joint Resource and Path Selection

Each core router keeps a record of grid traffic and background traffic loads on its outgoing links. The length of bursts corresponding to each class is added to find T_{of}^G and T_{of}^B which are the total length of bursts offered to a link for grid traffic and background traffic, respectively. These values are set to zero periodically at the end of a predetermined time window in order to dynamically record traffic load changes over a link. The duration of this time window should be small enough to reflect short-term changes in the network and large enough to collect enough data about the traffic. At the end of a time window, the load on link l for each traffic class is computed using

$$A_l^G = \frac{T_{of}^G}{WT_{win}}, \quad A_l^B = \frac{T_{of}^B}{WT_{win}} \tag{4}$$

where T_{win} is the length of the time window and W is the number of wavelengths.

This load levels are transferred to the edge routers using acknowledgment and probe packets as described previously. Consumers can use this feedback to compute path loss probability of each disjoint path to a resource. When using this feedback from the core routers, consumer should also consider the traffic generated by itself during the previous time window because most of the traffic load on a link might be generated by the consumer itself.

Let us denote the overall traffic load on link l which is received from the corresponding core router as A_l in Erlangs. Load level estimation when the traffic will be routed over this link can be expressed as

$$A_l' = A_l + \Delta$$

where Δ is the difference between the traffic offered by the consumer on link l between the next and previous time windows.

If the burst arrival distribution is Poisson, then the loss rate of link l when there are W wavelengths can be computed using the Erlang B formula.

$$\pi_l = \frac{\frac{A_l^W}{W!}}{\sum_{i=0}^{W}\frac{A_l^i}{i!}} \tag{5}$$

Using the link independence assumption, the loss probability over path p can be written as

$$P_l^p = 1 - \prod_{l\in p}(1-\pi_l)$$

For each resource-path pair, the consumer computes an expected completion time as follows.

$$\overline{T}^{r,p} = T_{min}^{r,p} + T_{rt}^{r,p}\frac{P_l^p}{1-P_l^p} \tag{6}$$

where

$$T_{min}^{r,p} = T_d + 2T_l + 2T_p^{r,p} + T_{proc}^r$$

and

$$T_{rt} = T_t^{r,p} + T_d$$

assuming that no extra offset is used for job and job result bursts. After computing expected completion time for each resource and path pair, the consumer selects the pair (r,p) which minimizes $\overline{T}^{r,p}$, i.e., $(r,p) = \arg\min \overline{T}^{r,p}$.

In the next section, we present an adaptive offset based QoS mechanism for grid bursts which operates jointly with the resource-path selection mechanism.

3.4 Effect of extra offset for job bursts on completion time

Extra offset based QoS mechanism is used to guarantee a minimum burst loss rate for high priority bursts in the literature. However, the effect of the delay caused by the extra offset is application dependent and may be very significant for some time sensitive applications. For an OBS grid application, the extra offset can also be used to reduce the burst loss probability for high-priority grid bursts. However, the increase in the offset time will increase the minimum required completion time so the trade-off between delay increase and loss reduction needs to be addressed.

The minimum required completion time increases linearly in response to T_{jo} as it can be observed from (1). Similarly, the retransmission cost increases linearly with T_{jo}. However, it is possible to reduce the expected completion time function if loss probability can be reduced sufficiently.

In order to analyze the effect of offset time on completion time, we used the mathematical loss probability analysis given in [7]. In this model, there are two classes of traffic. We assume that grid bursts(job and result) constitute the high-priority traffic whereas all other bursts, called background, constitute the low priority traffic. Bursts belonging to both classes arrive according to processes. The overall loss probability of OBS traffic can be computed using the Erlang B formula for an offered load A_l and W wavelengths as given by (5).

To find the loss probability of the high priority traffic, the affect of the low priority traffic on the grid traffic must be considered. It is possible to write the loss probability of grid traffic as

$$\pi_l^G = B(A_l^G + Y_B(\delta_G), W) \tag{7}$$

where $Y_B(\delta_G)$ is the low priority background traffic which is seen by the grid traffic with a QoS offset of δ_G. Then, the loss probability of the background traffic can be approximated using the conservation law as

$$A_l\pi_l = A_l^G\pi_l^G + A_l^B\pi_l^B \tag{8}$$

where A_l^B is the offered load of the background traffic. The background traffic affecting the grid traffic, $Y_B(\delta_G)$, can be computed using

$$Y_B(\delta_G) = A_l^B(1-\pi_l^B)(1-F_B^f(\delta_G)) \tag{9}$$

where $A_l^B(1-\pi_l^B)$ is the background traffic which is not lost and $F_B^f(\delta_G)$ is the distribution function of residual life of background burst length. Since there is a mutual dependency between π_l^G and π_l^B, these equations has to be solved iteratively as described in [7].

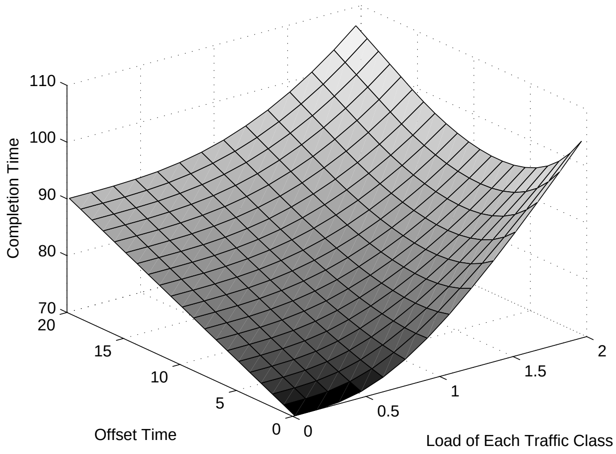

Figure 3: Graph of completion time vs. QoS offset and traffic load

To understand the effect of offset on the completion time, we computed the average completion time with respect to traffic load and extra offset. In this scenario, we assume a 3 hop path between the consumer and resource and 4 wavelengths at each link. Using the analytical model in [7], we estimate the loss probability P_l^p over the path with respect to different load levels and offset times, and used this loss probability value to compute the estimated completion time using (6). It is assumed that the burst size for each traffic class is uniformly distributed between 0.5 and 15 ms. The change of estimated completion time with respect to offset and load can be seen in Figure 3 for $T_{min} = 70ms$ and $T_{rt} = 30ms$. From the figure, it can be deduced that applying an extra offset can reduce completion time especially when the traffic load is high.

3.5 Computing the optimum extra offset for a path

The feedback received from the core routers include the total traffic load generated by grid and background bursts. Using the same method explained in the previous section, the consumer separately estimates the grid load levels at each link along a path.

Then, the consumer performs an iterative procedure to compute the value that minimizes the completion time given by (6). At each iteration, the consumer computes the loss probability using the analytical model given by (7),(8) and (9) for the given offset value and evaluates the completion time function using (6).

4. RESOURCE-SIDE OPTIMIZATION

In contrast to the consumer side optimization where the consumer can choose any resource to send the job, the only problem of the resource is to choose the path to send the result burst since the destination of the job result is readily known. Similar to the consumer, the resource also knows two disjoint shortest paths to consumer and it uses a similar approach to select the path to send the job result. The timeline of a job result burst which is successfully transmitted can be seen in Figure 4.

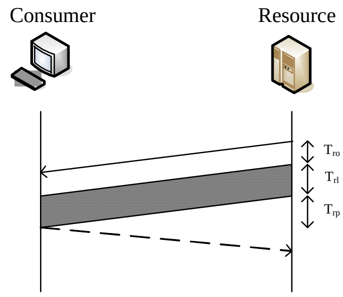

Figure 4: Timeline of a successfully transmitted grid job result.

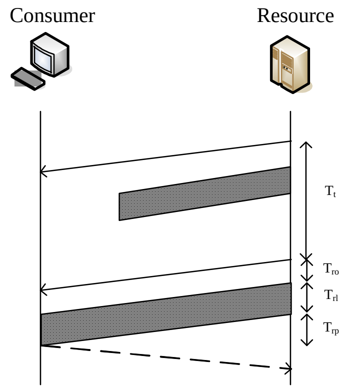

Figure 5: Timeline of a grid job when the job result burst is lost.

It can be seen that the minimum required transmission time for job result is

$$T_{min} = T_{ro} + T_{rl} + T_{rp}$$

If the job burst is lost, the retransmission cost is the timeout duration, which is required to notice the loss of the burst in addition to a guard band. The timeline of this situation can be seen in Figure 5.

$$T_{rt} = T_{ro} + T_{rl} + T_{rp} + T_g$$

The difference of the extra offset mechanism for job result bursts from the one for job bursts is that the minimum required time and the retransmission cost functions changes. The retransmission cost of a job result burst is smaller than a job burst so it is expected that the optimum offset computed for job result bursts is smaller.

5. SIMULATION AND RESULTS

In this section, we present the simulation framework and performance evaluation of the proposed algorithms.

5.1 Grid network model

The OBS grid network shown in Figure 6 is used in simulations where the length of each core link is indicated. In this topology, there are 11 customers and 3 resources. Each customer and resource is connected to the core network through an edge router. Also, for each resource, there is an intelligent router adjacent to the resource which performs resource

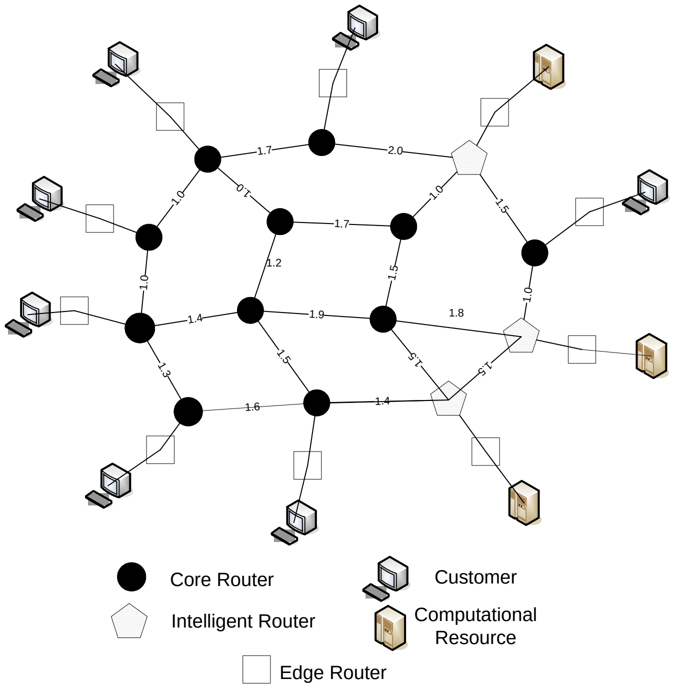

Figure 6: The simulated OBS grid topology. The numbers show the propagation delay of each link in ms.

querying and sends acknowledgments to consumer regarding that resource.

The length of the background bursts and grid bursts is distributed uniformly between 0.5 ms and 15 ms. Each optical burst carries a single grid job or grid job result. We assume that the result of a job has the same data size with the job itself. The switching time for the core switches is 0.1 ms and control packet processing time is negligible. There are $W = 5$ wavelengths per fiber at each link and one of them is reserved for the control plane. Also, we assume that there are five links between edge routers and core network in order to prevent congestion at the edges of the network. The core routers take their load measurements using $T_{win} = 5s$.

In order to evaluate the performance of the proposed congestion avoidance mechanism, we simulate a background burst traffic independent of the grid traffic. In the simulations, the background traffic is generated at the edge routers and sent to a randomly selected other edge router. The simulations are performed for 50,000 jobs.

5.2 Resource and Job Model

In order to perform a realistic simulation of a burst switched consumer grid, we use a parallel workload model [8]. This model is used to generate grid job parameters, to schedule jobs at the grid resources and to estimate execution times of jobs in our simulations.

In a grid environment, computational resources have multiple processors. Parts of the submitted jobs can be executed in parallel on these multiple processors. However, depending on the characteristics of the job, the number of processors that will be used in execution may be fixed or variable. The speedup obtained by executing a job on multiple processors does not change linearly as the number of processors increase. This affects the scheduling decisions made by the resource. Downey's speedup model estimates the speedup of a job using its average parallelism, A, and its variance in parallelism, V which is defined as $V = \sigma(A-1)^2$ where σ is the coefficient of variance in parallelism.

$$S(n) = \begin{cases} \frac{An}{A+\sigma(n-1)/2} & \sigma < 1,\ 1 \leq n \leq A \\ \frac{An}{\sigma(A-1/2)+n(1-\sigma/2)} & \sigma < 1,\ A \leq n \leq 2A-1 \\ A & \sigma < 1,\ n \geq 2A-1 \\ \frac{nA(\sigma+1)}{A+A\sigma-\sigma+n\sigma} & \sigma \geq 1,\ 1 \leq n \leq A+A\sigma-\sigma \\ A & \sigma \geq 1,\ n \geq A+A\sigma-\sigma \end{cases}$$

Using this speedup estimation, resource can estimate the execution time of a job and schedule submitted jobs over multiple processors. There are several scheduling strategies in [8]. In our simulations, we use a simple scheduling strategy which allocates a number of processors equal to the average parallelism of the job, A. If A processors are not available at time of the job request, the resource postpones the execution of this job until A processors become available.

The processing characteristics of jobs are determined by three parameters: Job instruction count in Million Instructions (MI), average parallelism and variance in parallelism. We chose the job instruction count to be distributed uniformly between 200 and 6,000 MI and average parallelism distribution between 1 and 20. We take parallelism variance distribution between 0 and 2. Resources are characterized with the number of processors and the processing speed of each processor in terms of million instructions per second. In simulations, each computational resource has 2000 processors and each processor has a processing power of 20,000 Million Instructions per Second (MIPS).

5.3 Results

The proposed joint path/resource selection algorithm is compared with other path switching algorithms. For the resource selection, an algorithm which selects the resource offering minimum computation time (MCR) is used in order to make comparisons. If the resources offer the same computation time, it selects the nearest resource.

In combination with this resource selection algorithm, following routing algorithms are used for comparison.

- Shortest Path Algorithm (SP): The shortest path between a consumer and resource is always used.
- Weighted Link Congestion Strategy (WLCS): This path switching strategy is proposed in [15]. It computes the successful transmission probability of a path using the loss reports of each core router and divides this value to the number of hops and selects the path which gives the larger value.
- Weighted Bootleneck Link Utilization Strategy (WBLU): This is also proposed in [15]. It uses the utilization of the most congested link along a path weighted by the hop length and select the path accordingly.

The algorithms proposed in this paper are JR-NO, which corresponds to the joint resource and path selection algorithm with no offset, and JR-AO, which corresponds to the joint resource and path selection algorithm with adaptive offset.

For a static background load of 2 Erlangs on each edge router, the average completion time, job burst ratio and job

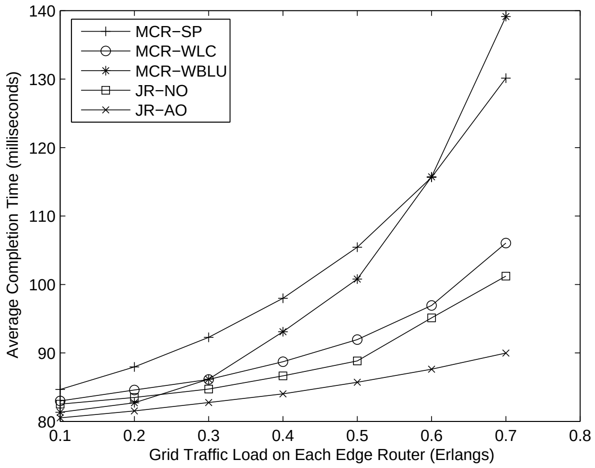

Figure 7: Average completion time for different resource and path selection strategies for a background load of 2 Erlangs.

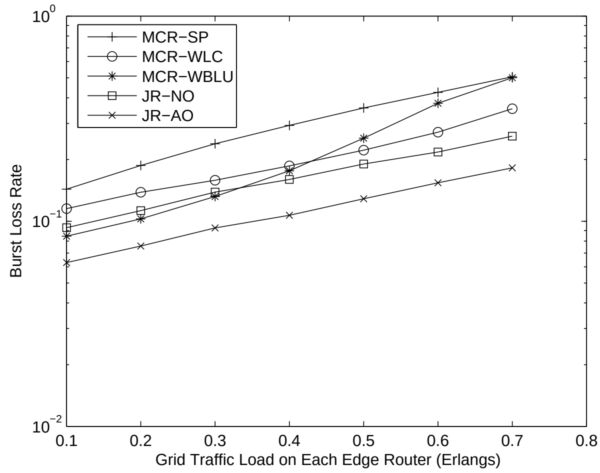

Figure 8: Job burst loss probability for different resource and path selection strategies for a background load of 2 Erlangs.

result burst ratio for different resource and path selection algorithms are depicted in Figures 7, 8 and 9, respectively. The average completion time is a more suitable metric to compare algorithms because the ultimate goal of contention avoidance is to minimize the average job completion time. Burst loss probability of grid bursts can be reduced to low levels by using a large offset, but it leads to longer completion times.

From these figures, it can be observed that JR-NO and JR-AO algorithms show better performance in terms of both average completion time and burst loss rate in comparison to other path switching algorithms. For low load levels, the routing algorithms does not have an important effect on the completion time but as the load increases the routing algorithm becomes critical.

For a non-static traffic load, the behavior of the JR-AO is compared with JR-NO and JR-FO in Figure 10. In JR-FO, the fixed offset is equal to 0.6 ms. In this scenario, the background traffic load is 2 Erlangs at the beginning until t=75 s, it is increased to 5 Erlangs and kept at that level until t=225 s, after which the background load is reduced to 2 Erlangs again. The average offset generated by the JR-AO algorithm with respect to time can be seen in the first plot. The second plot gives the loss rate achieved with each algorithm and the third plot gives the average completion time for each algorithm. It can be seen that, although three algorithms show similar performance for the low load region, JR-AO achieves better performance than JR-NO and JR-FO in the high load region. The problem with the fixed offset scheme is the determination of the fixed offset. The optimum offset value for grid bursts are strongly dependent on the background traffic load, so it is not possible to find a fixed offset scheme for every traffic condition.

6. CONCLUSIONS

In this paper, we show that combining resource selection and path selection reduces congestion in the OBS grid network. This reduction affect the completion times of grid jobs directly showing that network aware resource selection is important for OBS grids especially when load levels are high.

Also, an adaptive optimum extra offset decision algorithm for OBS grids is proposed. This algorithm balances the drawbacks of extra offset based QoS system (increased delay) with its advantages (reduced loss rates). The algorithm can adapt to the load changes in the network, applying larger offset when the congestion is high and applying a smaller offset when the congestion is low, and it reduces completion times significantly.

7. ACKNOWLEDGMENTS

This work is supported in part by the Scientific and Technological Research Council of Turkey (TUBITAK) under project EEEAG-104E047, and by the European Commission through the Network of Excellence e-Photon/ONe+.

8. REFERENCES

[1] Lhc computing project (lcg) : http://lcg.web.cern.ch/lcg/.

[2] N. Barakat and E. Sargent. Analytical Modeling of Offset-Induced Priority in Multiclass OBS Networks. *Communications, IEEE Transactions on*, 53(8):1343–1352, 2005.

[3] R. Bhandari. *Survivable Networks: Algorithms for Diverse Routing*. Kluwer Academic Publishers, 2001.

[4] M. De Leenheer et al. An OBS-based grid architecture. *Global Telecommunications Conference Workshops, 2004*, pages 390–394.

[5] M. De Leenheer et al. Anycast algorithms supporting optical burst switched grid networks. *Networking and Services, 2006. ICNS '06. International conference on.*

[6] M. De Leenheer et al. A view on enabling-consumer oriented grids through optical burst switching. *IEEE Communications Magazine*, 44(3):124–131, 2006.

[7] K. Dolzer et al. Evaluation of reservation mechanisms for optical burst switching. *AEU International*

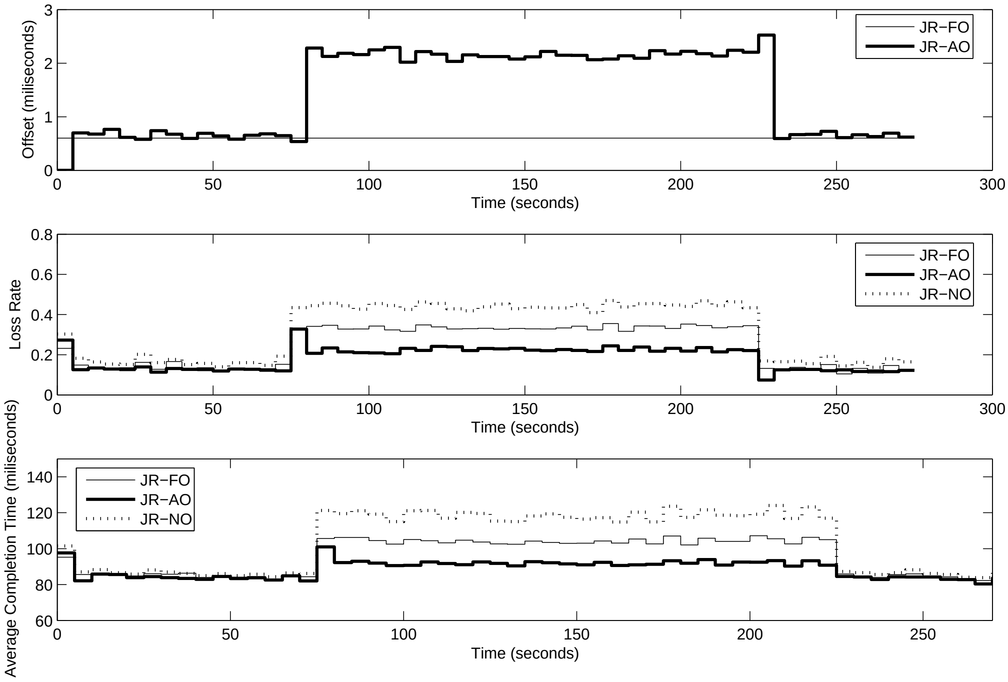

Figure 10: The graph of average offset time, loss rate and average completion time with respect to time for a non-stationary background traffic.

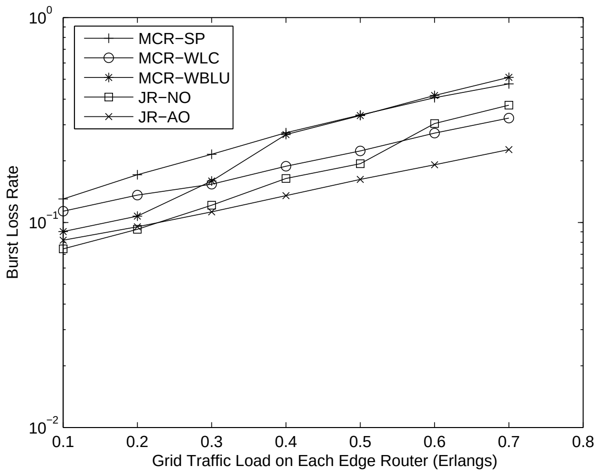

Figure 9: Job result burst loss probability for different resource and path selection strategies for a background load of 2 Erlangs.

Journal of Electronics and Communications, 55(1):18–26, 2001.

[8] A. B. Downey. A parallel workload model and its implications for processor allocation. *Cluster Computing*, 1(1):133–145, 1998.

[9] R. Nejabati et al. Programmable optical burst switched network: A novel infrastructure for grid. pages 993–999. Cluster Computing and the Grid, 2005. CCGrid 2005. IEEE International Symposium on, May 2005.

[10] L. Renambot et al. Griz: experience with remote visualization over an optical grid. *Future Generation Computer Systems*, 19(6):871–882, 2003.

[11] D. Simeonidou et al. Optical Network Infrastructure for Grid. *Grid High Performance Networking Research Group*, 2003.

[12] D. Simeonidou et al. Dynamic optical network architectures and technologies for existing and emerging grid services. *IEEE Journal of Lightwave Technology*, 23(10):3347–3357, October 2005.

[13] D. Simeonidou et al. Grid Optical Burst Switched Networks (GOBS). *Global Grid Forum Draft, May*, 2005.

[14] G. Thodime, V. Vokkarane, and J. Jue. Dynamic congestion-based load balanced routing in optical burst-switched networks. In *Global Telecommunications Conference, 2003. GLOBECOM '03. IEEE*, pages 2628–2632, December 2003.

[15] L. Yang and G. Rouskas. Adaptive Path Selection in OBS Networks. *Journal of Lightwave Technology*, 24(8):3002, 2006.

[16] M. Yoo, C. Qiao, and S. Dixit. QoS performance of optical burst switching in IP-over-WDM networks. *Selected Areas in Communications, IEEE Journal on*, 18(10):2062–2071, 2000.