Does Header Length Affect Performance in Optical Burst Switched Networks?

Felisa J. Vázquez-Abad*,+, Jolyon A. White+,
Lachlan L. H. Andrew+ and Rodney S. Tucker+

*Département d'informatique et recherche opérationnelle
Université de Montréal, Canada

+ARC Special Research Centre for Ultra-Broadband Information Neworks
Department of Electrical and Electronic Engineering
The University of Melbourne, Australia

This paper investigates the impact of non-negligible header length (HL) in optical burst switching on blocking probability. The header length is the total delay of a control packet at the controller. We first develop a model that explicitly presents the distribution of offset times as a function of the header length. Next we argue that the variance of this distribution (and not the mean) affects the blocking probability. In particular, the total blocking probability of a burst is dominated by the blocking on its last link, where its offset is shortest. We derive a lower bound for a HL threshold value below which blocking is not sensitive to the reservation algorithm. This threshold depends on network connectivity, number of channels per fiber and burst length. The blocking probabilities of both the Just Enough Time (JET) and horizon reservation algorithms were empirically found not to be very sensitive to the distribution of burst sizes.

1. Introduction

Optical Burst Switching (OBS) [1–4] has been proposed as a possible future transport network, in the hope that it will be able to provide more efficient multiplexing of the available optical bandwidth. Many recent publications focus on a key feature of OBS networks: the choice of reservation scheme [5–11]. In OBS networks, a control packet (or "header") is sent in advance of each burst to reserve an output channel on the outgoing link at each switch on its path. Upon its arrival each header is converted to electronic form and a *reservation algorithm* is executed in order to allocate an output channel (wavelength) to the burst (if one is available) for the given time period when the burst requires it. The chosen output channel then carries the burst towards the next switch in the network. Since the channel is allocated based on the known start and end times of the burst as well as the availability of resources, the algorithm in effect makes a *reservation* of the resource for a specific time in the future. If the reservation algorithm cannot allocate a channel, then the burst is lost. The blocking probability of the system is the expected fraction of lost bursts.

Since optical networks lack optical buffering, if bursts are to travel from origin to destination without conversion to the electronic domain, the cross-connect switch fabrics must be already configured to route them from input port to output port when they arrive. Because switch fabric reconfiguration takes non-negligible time, it is necessary to assemble bursts at the entry points to the network (ingress nodes) which contain a "large" amount of multiplexed information so that the cross-connect can be reconfigured less frequently. These set up and release times must be taken into account in the reservation algorithms.

This study focuses on the impact on blocking probabilities of non-negligible delay at the controller of an optical switch. In this paper, the term *header length* (HL) of a control packet will refer to the total delay at a controller. This includes the time required to read the header, as well as the execution time of the control algorithms. The time delay between a header and its companion burst is known as the "offset time", or more simply the "offset", and it is defined at each visited OBS switch along the path of the burst. Because the control packet has to be processed at each of the controllers, the offset after processing will be decreased by HL. We compare the performance of two popular reservation schemes: the Horizon [1] and Just Enough Time (JET) [5,8] algorithms, explicitly taking into account their execution times.

The details of the model are in Section 2. Section 3 establishes the model for the offset distribution at each OXC, as a function of the processing time. The main results of this paper are in Section 4 and Section 5, where we present a comprehensive study comparing the performance JET and Horizon as a function of processing times. We conclude the paper with a series of design recommendations, based on our simulation experiments.

2. Model Formulation

We consider a network with several access networks where packets are directed to the *edge routers* (entry points to the OBS network). There, packets are assembled into bursts. The core network consists of optical cross connects (OXCs) and each link joining OXCs represents a fiber with C channels or wavelengths.

When a burst has been assembled, a *control packet* is sent via dedicated control channels to the OXC's along its route for processing. The main function of the processor is to make the reservation of channels for the upcoming burst. Following [13], suppose that the route $p_1, \ldots, p_H$ is assigned at the edge router, and consider any intermediate OXC $p_i, i \in \{1, \ldots, H\}$, where H is the number of hops in the route. The control packet arrives at time $t(p_i)$ and contains the information $(b, (p_{i+1}, \ldots p_H), \delta(p_i))$, where b is the burst size in units of time and $\delta(p_i)$ is the *offset* representing the delay between the arrival of the control packet at the OXC and the arrival of the corresponding burst.

The controller keeps a list of reservations per output fiber, each of which has C optical carriers or channels. This list enables the controller to establish appropriate connections across the switch fabric just in time for each burst arrival. Upon arrival of a control packet at a cross-connect p, it is processed by the (electronic) controller, which retrieves the information $(b, p', \delta(p))$. This control packet will request a reservation of (any) channel at the output port towards node p' for the period of time $[t(p) + \delta(p), t(p) + \delta(p) + b]$, assuming that the burst size b already contains enough slack time for the switching to be performed before the actual burst arrival.

The Horizon Algorithm

The horizon algorithm searches through the C wavelengths to find those that are currently free of reservations for the required time period, and then it chooses the one which minimizes the void which would be introduced by accepting the new reservation (Figure 1). The time when the reservation is requested is represented as the origin in the x-axis, and the state of the list is represented in the picture by indicating the time sub-intervals that have been reserved for each wavelength. Channel 3 is chosen in this instance, the only one eligible for the algorithm. If every wavelength has reservations after time $t(p) + \delta(p) + b$ the burst will be blocked upon arrival.

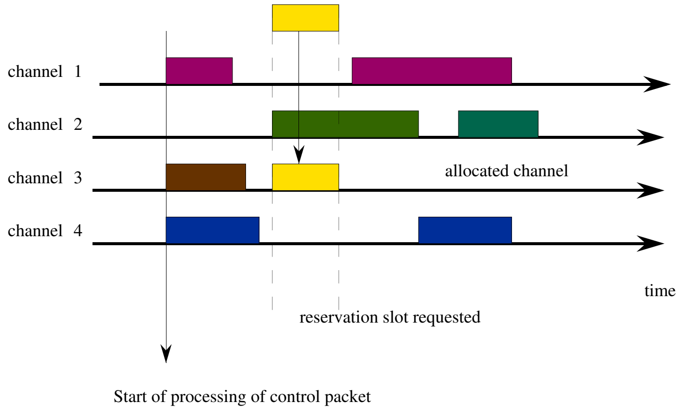


Fig. 1. Example of burst allocation by the Horizon algorithm

The Just Enough Time (JET) Algorithm

In this method, the algorithm looks for voids to fit the requested time interval reservation [5, 8]. If several candidates exist, then it chooses the one with minimal introduced void between the preceding burst and the new arrival. In Figure 2 we show an example of such an allocation, where now channel 4 is chosen. In [12] we propose an implementation, called "slotted JET" that requires a constant execution time per channel in order to find the eligible channels for the requested reservation.

The execution time of the reservation algorithms is therefore assumed to be a constant (not random) that may depend on the number of channels, the

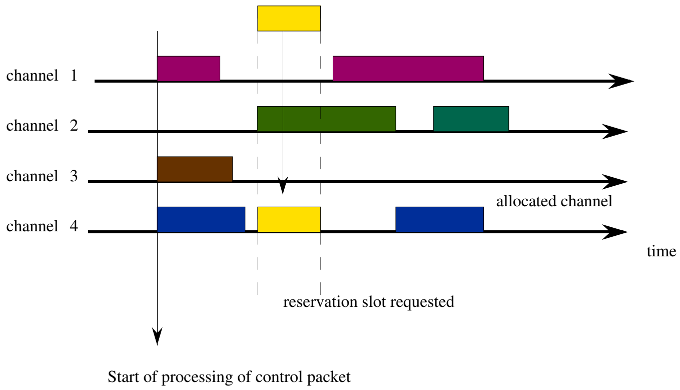


Fig. 2. The Just Enough Time algorithm

processor speed and the implementation of the algorithms.

3. Offset Distribution and Header Length

It takes κ units of time to read the information of the header. Call $\tau(p)$ the execution time of the reservation algorithm. Therefore the header length is defined to be $\tau = \tau(p) + \kappa$. In the event that no reservation is made because all the channels will be busy at that time slot, the header and corresponding burst must be discarded. Otherwise, the controller updates the offset for the next node (see Figure 3):

$$\delta(p_{i+1}) = \delta(p_i) - \tau(p_i) - \kappa, \tag{1}$$

because both control packet and burst travel at the speed of light, assuming zero dispersion (constant group delay).

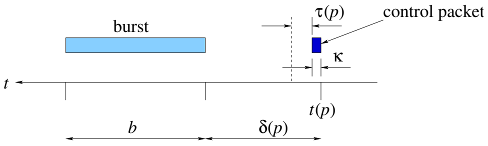


Fig. 3. Arrival epochs of control packets and bursts.

From this basic dynamic equation on the evolution of offsets along the control packet's route, it follows that the execution time of the scheduling algorithm has an effect on the offsets. In what follows, we show that it is through this dependency that the performance of the network (in terms of blocking probabilities) can be greatly affected.

Because the route $p_1, p_2, \ldots, p_H$ of any burst is known at the time of burst creation at the edge router, it is possible to set the initial offset time as:

$$\delta(p_1) = \sum_{i=1}^{H} \tau(p_i) + H\kappa. \tag{2}$$

Using (1), this assignment gives just enough delay between the header and the burst arrival times upon burst assembly, in order to avoid early arrivals during flight.

Lemma 1 *Suppose that all OXC's controllers have the same processor speed, and number of channels, so that the header length* $\tau(p) + \kappa = \tau$ *is constant for all* p*. Then the offset distribution at any OXC* p *is a discrete distribution on* $\{\tau, 2\tau, \ldots, H\tau\}$ *and*

$$\mathsf{P}\{\delta(p) = h\tau\} = \frac{\lambda_h(p)}{\sum_{i=1}^{H} \lambda_i(p)}, \tag{3}$$

where $\lambda_h(p)$ *is the effective arrival rate at the OXC* p *of bursts on paths whose destination lies exactly* h *hops ahead of* p*.*

Proof: Assume that (2) is established to ensure zero offset upon exiting the final hop. When a burst arrives at p, if its route still has h more hops, then the offset is necessarily $h\tau$. Because the arrival processes at p have different sources, in stationary operation the fraction of arrivals coming from such a path is proportional to the effective arrival rate. □

At the OXC, the header arrival process is a superposition of arrival processes per route. Under the assumption that there is a large number of routes passing through any OXC such that they have h hops to reach their destination, it is reasonable to model the consecutive scheduled arrival times of headers as a Poisson processes with aggregate arrival rates. Under this assumption, the effective arrival process of these packets is also Poisson.

4. Impact of Header Length

Consider a single OXC with a superposition of arrival processes, separated by their *class*. All arrivals from class $h \in \{1, \ldots, H\}$ have the same offset $\delta_h \equiv h\tau$, as in Lemma 1. Control packets of class h belong to routes that need h more hops to reach their destination. Therefore, small values of the class represent bursts that are close to destination, and they have smaller offsets. Because reservations made well in advance have more chances to find free resources, it turns out that for both JET and horizon there is a bias favoring the customers of higher class h [14].This establishes a *priority system*.

4.A. Comparison of Scheduling Algorithms

Intuitively, JET (and other void-filling algorithms) would be expected to have a lower blocking probability than horizon. This is not always the case; there are many realistic network scenarios when it is never possible to fill voids. For example, if all offsets are the same ($H = 1$), then the system becomes a simple $M/G/C/C$ system and void filling is impossible. In this scenario, JET and horizon yield identical performance. This result is generalized by the following lemma.

Lemma 2 *Let $\bar{b}$ be the minimal burst size. For the model of offset distribution given in Lemma 1, if the processing times for both Horizon and JET satisfy $\tau < \bar{b}/(H-1)$, then both reservation schemes yield the same allocation of channels.*

Proof: There cannot be any reservations starting after $H\tau$ units of time from the time a reservation is requested. The earliest request possible is for the time interval $[\tau, \tau + b]$. Because $b > (H-1)\tau$, it is impossible that such a request be allocated to a channel unless the channel is indeed free of future reservations, so JET and horizon yield the same allocation. □

From the above analysis, it seems natural to expect that the advantages of the void filling algorithms *will only be noticeable when the offset distribution has a large variance*. In our model, this happens when the HL $\tau = \kappa + \tau(p)$ is relatively large. As our simulation experiments confirm, the blocking probabilities are the same up to a threshold, after which the blocking increases dramatically for the Horizon algorithm, while staying relatively stable for the JET algorithm. However, the processing times of the reservation schemes may be different and this may negatively impact the performance of JET.

4.B. Number of Channels

It has been shown [15] that as the number of channels grows, the voids vanish under both algorithms. Thus blocking attains the same value for both algorithms, even using a scaling of the traffic intensity to keep the same utilization. However both algorithms have a processing time that is dependent on the number of channels. In this paper we show by simulation experiments that the upper bound $\bar{b}/(H-1)$ is conservative as the number of channels grow. The point where both allocations start to differ increases with the number of channels.

4.C. Burst Size Distribution

Some researchers have assumed that burst sizes are exponentially distributed [5, 9, 11, 16]. In this work we will depart from this common model, and we will use simulations to assess the performance of various scheduling schemes. While independence between burst sizes may be justified in principle under the assumption that the bursts are formed with a large number of multiplexed packets from the access network, we find it very difficult to motivate the exponential distribution of the burst sizes. Because the bursts are formed by aggregating information at the access networks, if the sources have a stationary behavior, then the burst size must necessarily be positively correlated to the inter-arrival times: the longer it takes to send the next burst, the larger its size.

At the edge router, bursts will be assembled by aggregating traffic that shares the same route. Suppose that the traffic demand for route $R = (p_1, \ldots, p_H)$ multiplexed from the access networks is a Poisson process with rate λ_R. We mention here two likely scenarios for burst assembly that will characterize the burst size distribution. In our *first scenario*, the information is collected until the burst reaches a target size with $N(\bar{b})$ bytes, which require $\bar{b}$ units of time for transmission. In this case, the inter-arrival times between bursts constitute a sequence of iid random variables with an Erlang distribution of mean $N(b)\lambda_R$, being the sum of $N(b)$ iid exponential random variables. The *second scenario* assumes that bursts along each route are sent periodically, every $\bar{T}$ units of time. Under this model, the consecutive burst sizes are iid random variables with Poisson distribution of mean $\lambda_R\bar{T}$. For large enough $\bar{T}$ the burst size can be approximated by a Gaussian random variable, using the Central Limit Theorem. In Section 5.B we evaluate the impact of burst size distribution on performance via simulations using exponential, constant and Gaussian distributions.

5. Simulation Experiments

We consider performance in regions of relatively low blocking probability ($\sim 10^{-3}$), which allows us to approximate the blocking probability of any route as the blocking probability it experiences on its last hop, neglecting blocking at upstream nodes, as a result of the priority effect. We call this performance indicator the *lowest priority blocking*.

Following Lemma 1, we simulated a single link with control packets arriving as a Poisson process with rate λ, each with an offset time drawn from a uniform distribution over $\{\tau \ldots, H\tau\}$. This model is representative of a number of network topologies, such as the linear networks depicted in Figure 4. Other discrete distributions will yield analogous results.

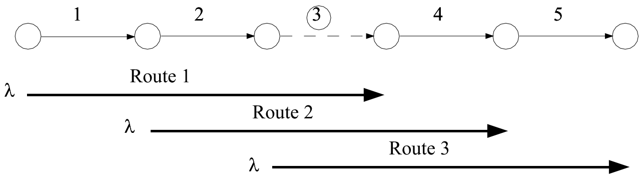
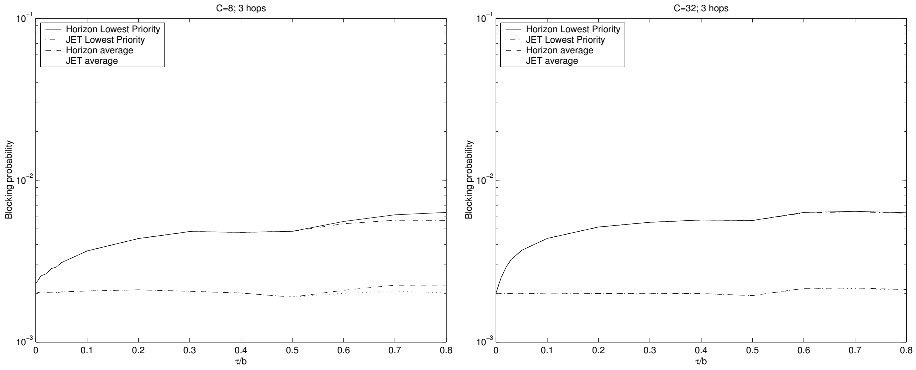


Fig. 4. Network modeled in simulation study. Blocking probability at the third link (dashed) was the performance measure studied.

We report on the characteristics of the blocking probability of the last link under both Horizon and JET as the HL τ varies. When $\tau = 0$ both Horizon and JET operate like a $M/G/C/C$ system and their blocking probabilities can be calculated by Erlang's B formula. As a baseline therefore, we chose to set the traffic intensity in each simulation scenario to yield a blocking probability of about 0.002 when $\tau = 0$.



(a)

(b)

Fig. 5. Blocking probability vs. $\tau/\bar{b}$ for constant burst size, 3 hops, (a) 8 wavelengths, (b) 32 wavelengths, 3 hops.

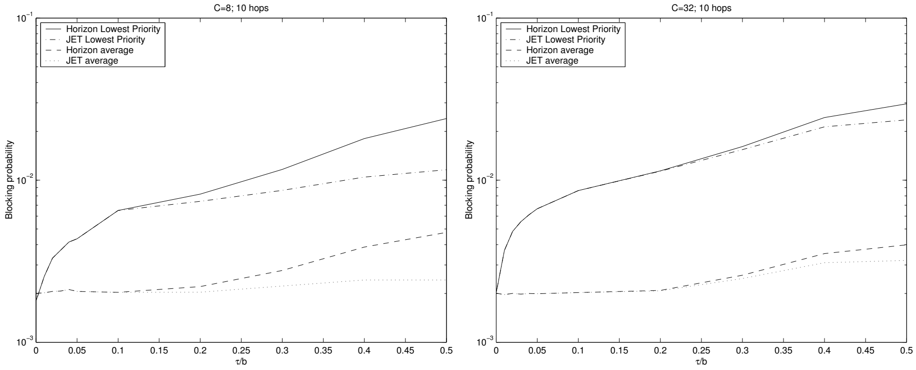
5.A. Constant Burst Sizes

The burst sizes were set to be constant at $\bar{b} = 100.0\,\mu$s. Figures 5(a)–5(b) show the average and lowest priority blocking probabilities as a function of $\tau/\bar{b}$, for 8 and 32 wavelengths respectively, when $H = 3$. Note that comparing the results for a particular value of $\tau/\bar{b}$ will be misleading since Horizon and JET will have different processing times τ. The graphs should therefore be read by identifying the two points of interest, one for JET and one for Horizon, and comparing the blocking probability at those two points. For the system with 10 hops the results are shown in Figures 6(a)– 6(b). In both cases, the lower bound of $\bar{b}/(H-1)$ given in Lemma 2 can be verified, but the threshold increases with the number of wavelengths, a result that is consistent with previous findings that the performance gap between Horizon and JET is reduced as the number of channels grows to infinity.

The significant feature of these results is that the worst case blocking *increases sharply as τ increases from zero*. This is important since any burst's probability of being blocked is dominated by this quantity.

5.B. Burst Size Distribution

To investigate the impact of the burst size distribution we used Gaussian and exponential distributions. These were simulated for a single link with 32 wavelengths and $H = 10$. We compared the results to the corresponding ones when the burst size was constant and equal to the mean of the Gaussian and exponential distributions, $\bar{b} = 100\,\mu$s. For the Gaussian, we chose two values of the standard deviation parameter: $\bar{b}/5$ and $\bar{b}/2$. The results, in Figures 7(a)–7(b), show that the higher variance of the burst duration causes higher blocking, which is an expected result. The dependence on the distribution is fairly weak, which is intuitively pleasing given the



(a) (b)

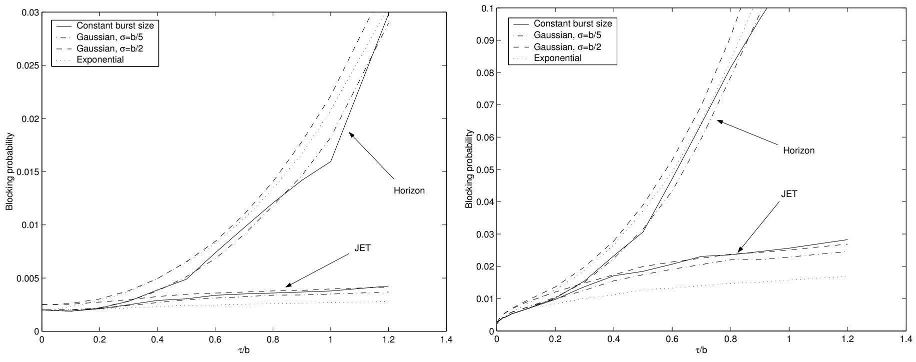
Fig. 6. Blocking probability vs. $\tau/\bar{b}$ for constant burst size, 10 hops, (a) 8 wavelengths, (b) 32 wavelengths

well known insensitivity of the $M/G/C/C$ system to the holding time distribution. These plots show the blocking probabilities in a linear scale (not logarithmic) to distinguish the small differences in values. It is noticeable in this scale how JET becomes almost insensitive to processing times after the threshold is attained.

5.C. Header length values

The foregoing results show the impact of HL on blocking probability for the JET and horizon algorithms. However, the execution time may not be equal for the two algorithms. If the control packets are assumed to be 1000 bits long, the transmission time will be $\kappa \approx 1\,\mu$s at 1 Gbps. For both algorithms, a basic magnitude comparator implemented in parallel hardware would require at most a few clock cycles, yielding a conservative estimate of $\tau(p) \approx 0.1\,\mu$s for the Horizon scheme with current technology. Most implementations of JET would require sequential operations to find the candidate gaps and may require five times as much time as horizon. This is the motivation for our current research [12], which proposes a novel implementation of JET that can reduce $\tau(p))$ also to a few clock cycles, reasonably between $0.1\,\mu$s and $0.5\,\mu$s.

We can see from the from Figure 6(a) that the point where JET and horizon diverge in blocking performance depends on the design parameters τ and $\bar{b}$, through their ratio. For example, for the network with $C = 8$, $H = 10$, if $\tau/\bar{b} = 0.2$ then the blocking probabilities have already diverged. At that point, the ratio of the average link blocking probabilities of JET and horizon is 1.08, assuming equal execution times. Thus, supposing that burst lengths are $\bar{b} = 100\,\mu$s, the break point occurs at an unrealistically high value of $\tau = 2\,\mu$s. However, if burst lengths are designed to be $10\,\mu$s then the break



(a) (b)

Fig. 7. Blocking probability vs. $\tau/\bar{b}$ for different burst distributions, 32 wavelengths, 10 hops. (a) Average link blocking, (b) Lowest priority blocking.

point will occur when $\tau = 0.2\,\mu$s which falls within realistic values of the HL. This places an upper bound on the HL of 200 bits (25 octets) at 1 Gbps. Outside of these limits the incentive to implement efficient versions of JET becomes significant.

Consider now situations when JET and horizon have the same performance. If $\tau/\bar{b} = 0.05$ for the case $C = 32, H = 10$, then the ratio of lowest priority to average blocking is 3.34, whereas when $\tau/\bar{b} = 0.01$, the ratio is 1.87. Thus $\tau/\bar{b} = 0.05$ creates more unfair differentiation among routes compared to 0.01. However, If $\bar{b} = 100\,\mu$s then this means $\tau = 1\,\mu$s is in this "fairer" regime, whereas if $\bar{b} = 10\,\mu$s then $\tau = 0.5\,\mu$s is already in the "unfair" regime. This again provides guidance to engineers, on bounds for τ and $\bar{b}$.

6. Concluding Remarks and Design Recommendations

This paper focussed on systems with small blocking probabilities. This regime is important because client protocols which are envisioned to be transported over OBS networks (such as TCP/IP networks) cannot tolerate excessive loss and still deliver acceptable throughput performance to the end user. The processing time, τ, of an OBS switch determines the minimum offset time that should be set at the source, if the control packet is to arrive at its destination before its burst does (assuming no delay lines are used). The *variation* of the offset time causes a priority system whereby those bursts which are closest to their destinations (with the least offset time) are disadvantaged significantly. However, for our model, every burst becomes a lowest priority burst at its final hop, and intermediate hop blocking is insignificant compared to final hop blocking, so final hop blocking probability is a critical performance measure.

Through simulation we have shown that as τ increases from zero, the final hop blocking probability of the OBS node increases dramatically. Furthermore, JET and horizon experience identical blocking probability until τ reaches a threshold value. The threshold depends on the maximum number of hops in the routes as well as burst size. For small number of hops (say $H \approx 5$), the scenario of τ being comparable to $\bar{b}$ is highly unrealistic, because bursts comprise large amounts of data. In this case it may be better to use the Horizon algorithm. However for networks with high connectivity and a moderate number of channels and small burst sizes there may be a benefit in using JET, provided that an efficient implementation is used with constant and small execution time, making it comparable with that of the Horizon algorithm. This is the subject of a parallel research project [12].

There is significant value in designing the OBS network to ensure that blocking probabilities are equalized among classes such that final hop blocking is no longer dominant, using either delay lines [17] or software based preemption [13]. Along with more realistic traffic models, these two aspects of OBS will be a major focus for our continuing study.

Acknowledgements

This work was supported by the ARC, NSERC Strategic Grant STPG 246159, and Redfern Broadband Networks.

References and Links

[1] J. Turner, "Terabit burst switching," *J. High Speed Networks*, vol. 8, no. 1, pp. 3–16, Jan. 1999.

[2] C. Qiao and M. Yoo, "Optical burst switching (OBS) – a new paradigm for an optical internet," *J. High Speed Networks*, vol. 8, no. 1, pp. 69–84, Jan. 1999.

[3] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-Over-WDM networks," *IEEE J. Selected Areas in Communications*, vol. 18, no. 10, pp. 2062–2071, Oct. 2000.

[4] L. Xu, H. Perros, and G. Rouskas, "Techniques for Optical Packet Switching and Optical Burst Switching," *IEEE Commun. Mag.*, pp. 136–142, Jan. 2001.

[5] M. Yoo and C. Qiao, "Just-Enough-Time (JET): A high speed protocol for bursty traffic in optical networks," in *IEEE/LEOS Technologies for a Global Information Infrastructure*, Aug. 1997, pp. 26–27.

[6] ——, "A new optical burst switching protocol for supporting quality of service," in *Proc. SPIE'98 Conf. All Optical Commun. Syst.: Architecture, Control, Network Issues*, vol. 3531, Boston, Nov. 1998, pp. 396–405.

[7] J. Wei, J. Pastor, R. Ramamurthy, and Y. Tsai, "Just-in-time optical burst switching for multiwavelength networks," in *IFIP Broadband Commun.*, Hong Kong, Nov. 1999, pp. 339–352.

[8] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Control architecture in optical burst switched WDM networks," *IEEE J. Selected Areas in Communications*, vol. 18, no. 10, pp. 1838–1851, Oct. 2000.

[9] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-Over-WDM networks," *IEEE J. Selected Areas in Communications*, vol. 18, no. 10, pp. 2062–2071, Oct. 2000.

[10] ——, "Optical burst switching for service differentiation in the next-generation optical internet," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 98–104, Feb. 2001.

[11] J.-B. Chang and C.-S. Park, "Efficient channel-scheduling algorithm in optical burst switching architecture," in *Proceedings of IEEE Workshop on High Performance Switching and Routing (HPSR 2002).* Kobe, Japan: IEICE, May 2002, pp. 194–198.

[12] B. Sansò, E. Gutiérrez, and F. J. Vázquez-Abad, "Efficient scheduling algorithms for OBS switches," *in preparation*, 2003.

[13] J. A. White, R. S. Tucker, and K. Long, "Merit-based scheduling algorithm for optical burst switching," in *Proceedings of the International Conference on Optical Internet.* Seoul, Korea: Korean Institute of Communication Sciences, July 2002.

[14] K. Dolzer and C. Gauger, "On burst assembly in optical burst switching networks – a performance evaluation of just-enough-time," in *Proceedings of the 17th International Teletraffic Congress*, vol. 4, Salvador da Bahia, Brazil, Sept. 2001, pp. 149–160.

[15] L. L. H. Andrew, Y. Baryshnikov, E. G. Coffman, S. V. Hanly, and J. A. White, "Asymptotically optimal greedy algorithms for large optical burst switching systems," *Perf. Eval. Review*, vol. 31, no. 2, pp. 14–16, Sept. 2003.

[16] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *Int. J. Electron. Commun.*, vol. 55, no. 1, pp. 18–26, 2001.

[17] B.-C. Kim, J.-H. Lee, Y.-Z. Cho, and D. Montgomery, "A novel hop-by-hop priority increasing scheme for multi-hop optical burst switching networks," in *Proceedings of the International Conference on Optical Internet.* Seoul, Korea: Korean Institute of Communication Sciences, July 2002.