Separating Resource Reservations from Service Requests to Improve the Performance of Optical Burst-Switching Networks

Neil Barakat, *Student Member, IEEE*, and Edward H. Sargent, *Senior Member, IEEE*

***Abstract*— In this paper, we introduce a new signalling architecture called Dual-header Optical Burst Switching (DOBS) for next generation burst-switching optical networks. DOBS decouples the resource reservation process from the service request process in core nodes and allows for delayed scheduling to be implemented. This relaxes the constraints on burst scheduling operations and allows the offset sizes of bursts to be precisely controlled in core nodes without the use of fiber delay line buffers. This allows for increased flexibility, control, and performance.**

To demonstrate the benefit of delayed scheduling and core-node offset control, we examine the performance of a DOBS system in which the offset size of every burst on a core link is set to a constant value. Using simulation and analysis, we show that the resulting constant-scheduling-offset (CSO) system realizes lower ingress delay, higher throughput, and better fairness than conventional single-header OBS systems, while simultaneously requiring only $O(1)$ burst scheduling complexity.

In a 16-channel system with full wavelength conversion and no fiber delay line buffers, the CSO DOBS system achieved a blocking probability 50% lower than that of a similar LAUC-VF JET OBS system. The CSO DOBS system also achieved perfect fairness, both with respect to burst length and with respect to the residual path length of bursts.

***Index Terms*— Delayed scheduling, dual-header optical burst switching, optical burst switching, optical networks, signalling architecture.**

I. Introduction

IN next-generation optical networks, WDM technology will be employed to provide aggregate data rates on the order of terabits per second. In order to efficiently handle this huge amount of data in a cost-effective manner, new network architectures based on all-optical switching technologies are required. One such architecture that has garnered much attention is Optical burst switching (OBS). In OBS, IP packets with the same destination are assembled electronically at the edge of the network into long bursts, which are transmitted through the network core entirely within the optical domain. By separating the transmission of the burst header (or control packet) and the burst payload by a time-offset, OBS eliminates the need for buffering of the burst during header processing and switch configuration [1]. Further, because control packets are transmitted on a separate control channel and are processed electronically in each node, there is no need for complex optical processing.

Although the operation of OBS is relatively simple for a single-hop network, the operation of OBS can become significantly more complex in networks where bursts travel over multiple links between their source and destination. Because control packets incur delay while they are processed in each node, the size of a burst's offset will shrink as the burst travels through the network [2], [3]. This results in a system with *variable offset sizes*. The presence of this variability complicates the task of burst scheduling and can compromise the throughput performance, and it has been found that there is an inherent tradeoff between throughput efficiency and scheduling complexity in these systems [4], [5], [6]. Offsets also affect the fairness of OBS systems because bursts with longer offsets implicitly receive higher priority in burst schedulers than bursts with smaller offsets [7], [8], [9].

The complexity-efficiency tradeoff and the unfairness inherent in OBS systems is an artifact of the variable offsets that exist in each node. If one could control the size of offsets in the core of the network, not only would the throughput and fairness of these systems be improved, but the task of burst scheduling would also become much simpler.

In this paper we present a new optical burst switching architecture called Dual-header Optical Burst Switching (DOBS) that decouples the reservation request operation and scheduling operation in each node of an optical burst switched network. This decoupling makes it possible to delay scheduling operations in core nodes and individually select the offset size of each burst at each node in its path *without the use of FDL buffers*.

To illustrate the utility of DOBS signalling, we consider a system in which the offset size of every burst on a given core link is set to a constant value. We show that the resulting DOBS system realizes better performance than classical single-header OBS systems in terms of throughput, delay, and fairness, while simultaneously decreasing the required burst-scheduling complexity.

The outline of the paper is as follows. In Section II, we briefly describe past OBS signalling and scheduling architectures and provide motivation for DOBS. Using notation given in Section III, we then describe the operation of our new DOBS architecture in Section IV. We present the constant-scheduling-offset DOBS architecture in Section V and com-

Manuscript received September 13, 2004; revised January 13, 2006. This paper was presented in part at IEEE INFOCOM 2005: N. Barakat and E. H. Sargent, "Dual-header optical burst switching: A new architecture for WDM burst-switched networks," in *Proc. IEEE INFOCOM '05*, Mar. 2005.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, Ontario, Canada M5S 3G4 (e-mail: {neil.barakat, ted.sargent@utoronto.ca}@utoronto.ca).

Digital Object Identifier 10.1109/JSAC-OCN.2006.04008.

pare its performance to that of classical OBS systems using analysis in Section VI and simulation in Section VII. Areas for future work and conclusions are discussed in Section VIII and Section IX respectively.

II. Background

In this section, we describe the operation of classical, single-header OBS systems. We discuss the most commonly studied OBS signalling architectures. We also describe the problem of variable-sized offsets and describe the different scheduling strategies that have been proposed to deal with this problem.

A. Just-in-Time OBS

One of the simplest signalling architectures for OBS systems with one-way reservations is called Just-in-Time (JIT) signalling [10], [11]. In JIT systems, *immediate reservation* is used, so the optical switch fabric is configured immediately after processing the control packet. As a result, burst reservations are scheduled in first-come-first-serve (FCFS) order, so very simple scheduling algorithms can be implemented. However, this simplicity comes at the cost of wasted bandwidth because switches are configured not only for the duration of bursts, but also for the time between control packet and burst arrivals.

JIT is very well suited to systems in which the control-packet processing time is much smaller than the burst duration. In these systems, the bandwidth overhead associated with immediate reservation is very small, so JIT can provide throughput levels that are comparable to those of much more complex signalling and scheduling architectures [12].

Although current commercial optical switches based on MEMS technology have switching speeds on the order of a few milliseconds, in next generation optical networks, it is expected that optical switching speeds will increase by orders of magnitude. For example, the OBS testbed described in [13] and [14] used wavelength-selective optical switches that could be reconfigured in less than 100 ns. Faster optical switches make it possible to efficiently transmit much shorter bursts, so one can envision future network scenarios in which a burst's control-packet delay and corresponding control offset may be comparable to the burst's duration. At the same time, the number of wavelengths per channel will also increase as WDM technology improves. This will result in an increase in the load on the control channel. This increase in control channel traffic will tend to increase the level of control-packet queueing that occurs in the control-packet processors of OBS nodes, which necessitates the use of larger offsets. For these systems, the bandwidth waste associated with immediate reservation would be prohibitively large or would limit the system to transmitting only very large bursts. This motivates alternate OBS architectures that can make use of the channel bandwidth between the time a control packet arrives and the time its burst arrives.

B. Just-Enough-Time OBS

In order to improve on the efficiency of immediate reservation and allow for more efficient transmission of short bursts, Just-Enough-Time (JET) signalling was proposed [1]. In JET OBS, each control packet contains information about the offset size, duration, and channel of its burst. By storing the offset time in the control packet, JET makes it possible to implement *delayed reservation*, so that the switch fabric is only configured for the duration of the burst. Thus, JET does not waste the bandwidth that lies between each burst and its control packet. This allows JET to realize improved throughput compared to *immediate reservation* architectures such as JIT, and removes the requirement for very long bursts. However, as described below, the use of delayed reservation introduces a number of challenges and unique tradeoffs that must be considered when designing JET OBS systems.

Since control packets are processed electronically in core nodes and bursts pass through all-optically without any processing delay, the offset size of a given burst will shrink at each intermediate node by a value equal to the control-packet processing delay. Thus, the offset sizes of arriving bursts at each node in the network will vary as a function of their residual path lengths. This offset variability can profoundly affect both the scheduling complexity and throughput performance of JET OBS systems.

At a given node in a JET OBS network, the variability in the offset size of arriving bursts means that bursts may not be serviced in the same order that they arrive to the system. As a result, useable gaps between successive reservations are created on the outgoing channels. In JET OBS systems, the throughput performance generally depends strongly on the way that these gaps or *voids* are treated by the burst scheduler [4].

In *Horizon* scheduling, a single horizon value that corresponds to the end of the most recently scheduled burst is stored for each channel [15]. Bursts are scheduled on the channel that minimizes the size of the void between its horizon time and that burst's start time, so that bandwidth waste is minimized. Because of the decision criteria employed in channel selection, Horizon scheduling is also referred to as *latest available unscheduled channel* (LAUC) scheduling [16].

An extension of Horizon called LAUC with void-filling (LAUC-VF) improves the throughput of JET OBS systems by allowing bursts to be scheduled in the voids between existing reservations [16]. However, because void-filling requires that detailed channel-state information be stored for every wavelength, the increased performance of LAUC-VF comes at the cost of significantly increased complexity and memory requirements in the burst scheduler.

A void-filling scheduling algorithm that requires lower complexity than LAUC-VF was presented in [5], and a void-filling algorithm that uses a more complex channel-selection criteria was presented in [17]. Although a number of different criteria can be used to select the optimal void in which to schedule bursts, it has been found that all void-filling scheduling algorithms for JET systems result in very similar blocking performance [5].

C. Motivation: Controlling Offsets to Improve Performance

Many of the complexity and performance issues in OBS networks could be completely avoided if one could precisely control the size of offsets in core nodes. While it is relatively simple to shrink offsets in core nodes, increasing the offset size of a burst is quite difficult and usually requires delaying bursts in fibre delay line (FDL) buffers [16], [18], [19]. However, since FDL buffers introduce complexity to the system and offer only course control over offset sizes, a simpler and more flexible approach is desirable.

In the next section, we introduce a completely novel optical burst switching architecture called Dual-header Optical Burst switching which allows for precise control of offsets in core nodes of the network *without* the use of FDL buffers. This gives DOBS core nodes much more control over the order and manner in which bursts are scheduled. This added flexibility can be used in a number of ways to improve the performance and simplify the operation of optical burst switched networks.

In Section V, we describe one realization of DOBS in which the offset size of every burst on each link is selected to be equal. This allows for both delayed reservation and FCFS scheduling. As a result, it simultaneously achieves the simple scheduling complexity of JIT OBS, and the high throughput efficiency of JET OBS.

III. Notation

In this section, we define notation that will be used in the remainder of the paper. A burst-switched path with h-hops is a sequence of h consecutive links between an ingress node and an egress node in the network. We use the index $i \in \{0, 1, \ldots, h\}$ to denote the path's nodes in order from the source to the destination inclusively. Similarly, the links are indexed by $i \in \{1, 2, \ldots, h\}$. Thus, node-$i$ lies immediately downstream from link-i.

We use the following notation to describe the relevant time instants and time intervals involved in forwarding a given burst through the i^{th} node of a burst-switched path:

- SRP is the service request packet.
- RAP is the resources allocated packet.
- l_b is the duration of the burst.
- $t^i_{SRP\_in}$ is the instant at which the SRP arrives at node-i.
- $t^i_{SRP\_out}$ is the instant at which the SRP leaves node-i.
- t^i_{BS} is the instant at which the burst is scheduled in node-i onto an outgoing wavelength of link-$(i+1)$.
- $t^i_{RAP\_in}$ is the instant at which the upstream RAP arrives at node-i.
- $t^i_{RAP\_out}$ is the instant at which the newly generated RAP is forwarded downstream from node-i onto link-$(i+1)$.
- t^i_{SW} is the instant at which the optical switch fabric is reconfigured in node-i.
- t^i_b is the instant at which the burst arrives at node-i.
- $\Omega^i_{ph} = t^i_b - t^i_{SRP\_in}$ is the physical offset size at node-i.
- $\Omega^i_{BS} = t^i_b - t^i_{BS}$ is the scheduling offset size at node-i.
- Δ^i_{SW} is time required to configure node-i's optical switch fabric.
- Δ^i_{BS} is time required to execute the burst scheduling operation in node-i, including any time spent waiting in the scheduling queue.
- Δ^i_{SRP} is the time required to process the SRP in node-i.
- Δ^{max}_{BS} is the network-wide maximum value of Δ^i_{BS}.
- Δ^{max}_{SW} is the network-wide maximum value of Δ^i_{SW}.
- Δ^{max}_{SRP} is the network-wide maximum value of Δ^i_{SRP}.
- h^{max} is the maximum number of hops that the burst can traverse between its source and destination.

Similarly, for JET and JIT OBS systems, we also define the following notation:

- $\Delta^i_{CP\_X}$ is the control-packet processing delay at node-i, including any time spent waiting in the control-packet queue, and where $X \in \{JIT, JET{-}HR, JET{-}VF\}$.
- Ω^i_X is the offset size between the control packet and the burst upon arrival to node-i, where $X \in \{JIT, JET{-}HR, JET{-}VF\}$.

Lastly, when discussing the effect of over-provisioning on the performance of DOBS systems, we also make use of the following notation

- W is the number of channels on each link.
- ρ is the offered traffic load in the system.
- B is the blocking probability of the system.
- $\Delta\rho$ is the increase in the effective offered load due to over-provisioning.
- ΔB is the increase in blocking probability due to over-provisioning.

IV. Dual-Header Optical Burst Switching (DOBS)

In this section, we describe the operation of a new optical burst signalling protocol called (DOBS).

A. Dividing Control Information

DOBS signalling is characterized by the use of two different types of control packets for each burst. The *service request packet* (SRP) contains information about the service requirements of the burst. A single persistent SRP precedes the burst and communicates the burst's service requirements to each node in the path. The *resources allocated packet* (RAP) contains the burst's physical information, which is used when configuring the optical cross connect in each node. At a minimum, the SRP contains the routing and temporal information of the burst, and the RAP contains the incoming channel index of the burst. In general, however, both the SRP and RAP could also contain additional relevant information, such as signal quality, or class-of-service information.

Although DOBS uses two control packets for each burst, the combined information contained in the SRP and RAP is equivalent to the information carried in a single JET OBS control packet. As such, the traffic load on the control channel in a DOBS network is only slightly higher than that in a JET or JIT OBS network.

Since control channel bandwidth can generally be increased by using additional control channels, the practical bottleneck of most OBS systems is not control-channel bandwidth, but rather the time required to process control packets in core nodes [20]. Although DOBS uses two control packets, the amount of control processing per burst is not larger than for JET or JIT OBS. Thus, the maximum control-channel throughput of a DOBS network is as high as that of a single-header OBS network.

The architecture of a core node in a DOBS network is illustrated in Fig. 1. As in most optical burst switching architectures, a separate control channel undergoes optical-to-electronic conversion to allow for electronic processing

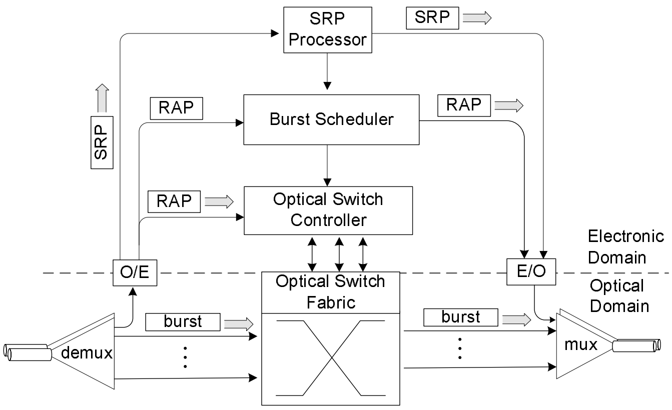

Fig. 1. Functional diagram of the different modules in a core node of a DOBS network.

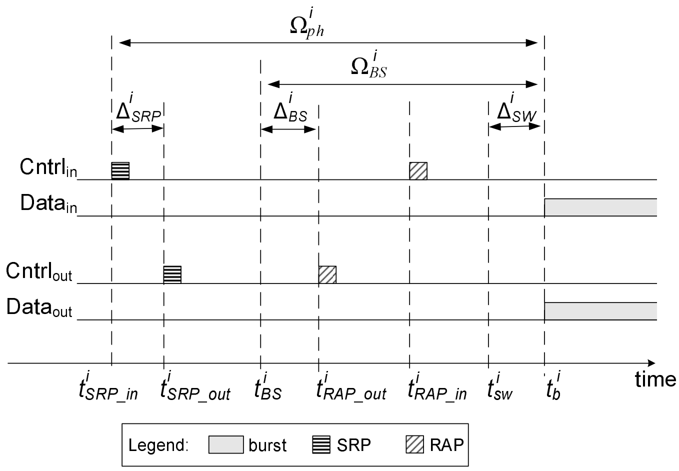

Fig. 2. Relevant timing instances involved in the forwarding procedure of a single burst through a DOBS node. The use of two control packets allows the burst reservation time t^i_{BS} and the scheduling offset Ω^i_{ph} to be chosen flexibly at each core node.

of control packets. The SRP processor, burst scheduler, and optical switch controller use information contained in the SRPs and RAPs to configure the optical switch fabric in advance of burst arrivals so that bursts can be forwarded all-optically to the appropriate output link. A detailed description of the operation of each of these modules is given in the timing and signalling discussion below.

B. Delayed Scheduling

In OBS systems, control packets carry advanced information about burst reservation requests. One might, therefore, expect that this advanced knowledge could be leveraged by the switch controller to improve throughput performance. However, in current OBS architectures, since all bursts must be scheduled as soon as their control packets arrive, this is not the case. For example, in a JET OBS system, if the offset size of every burst in the system is increased by a constant value, the throughput performance generally remains unchanged.

By separating the control information of each burst into two control packets and decoupling the resource request and resource reservation operations in each node, DOBS removes the requirement that bursts be scheduled immediately after their reservation requests are received. This makes it possible to implement *delayed scheduling* in the core nodes of DOBS networks.

Delayed scheduling allows DOBS schedulers to make use of the effective *head-start* provided by the control-packet offset, which can lead to decreased scheduling complexity, increased flexibility, and improved throughput performance. In a delayed scheduling system, for example, if the physical offset size of every burst is increased, it becomes possible to accumulate reservations over a longer duration before scheduling them. This allows for simpler and more optimal scheduling decisions to be made.

In Section V we describe how delayed scheduling can be used to implement a system that behaves as if the offset size of every burst were identical. The resulting system realizes the benefits of both JIT and JET OBS – benefiting from the reduced complexity of FCFS scheduling, while still retaining all of the efficiency advantages of delayed reservation.

C. Timing and Operation of Core Nodes

Fig. 2 depicts the relevant time instances for the forwarding procedure of a single burst through a DOBS node. The input and output control channels and the input and output data channels used by the burst are labeled as Cntrl$_{\text{in}}$, Cntrl$_{\text{out}}$, Data$_{\text{in}}$, and Data$_{\text{out}}$ respectively. The time required for the SRP forwarding, the burst scheduling, and the optical switch configuration are denoted as Δ^i_{SRP}, Δ^i_{BS}, and Δ^i_{SW} respectively.

The burst's SRP arrives at the node at time $t^i_{SRP\_in}$ and is processed immediately by the SRP processor. The SRP processing step includes a routing table lookup or a label-swap to determine the appropriate outgoing link, and the burst's reservation request is communicated to the burst scheduler. The SRP is then immediately forwarded to the next downstream node at time $t^i_{SRP\_out}$ without waiting for the burst-scheduling operation.

Sometime after $t^i_{SRP\_out}$, the burst scheduler executes the burst-scheduling algorithm at time t^i_{BS}. Once completed, an RAP is then transmitted at time $t^i_{RAP\_out}$ to inform the downstream node of the burst's new channel information.

The upstream RAP carrying the burst's incoming channel information arrives at time $t^i_{RAP\_in}$. Based on this information and the result from the burst-scheduling algorithm, the optical switch is configured at time t^i_{SW} just before the burst's arrival at time t^i_b. Since the only delay incurred by a burst at node-i is the propagation delay through the switch fabric, the difference between the time that the burst appears on Data$_{\text{in}}$ and Data$_{\text{out}}$ is negligible.

As shown in Fig. 2, we define the *physical offset* Ω^i_{ph} as the duration between the arrival time of the SRP and the arrival time of the burst. In this way, the physical offset is analogous to the control-packet offset in JIT or JET OBS. The physical offset shrinks by the SRP processing duration at each node in the burst-switched path.

We define the *scheduling offset* Ω^i_{BS} as the time between the burst reservation t^i_{BS} and the burst arrival time t^i_b. An important feature of DOBS is that t^i_{BS} can be chosen arbitrarily at each node within an allowable range. Since the order in which

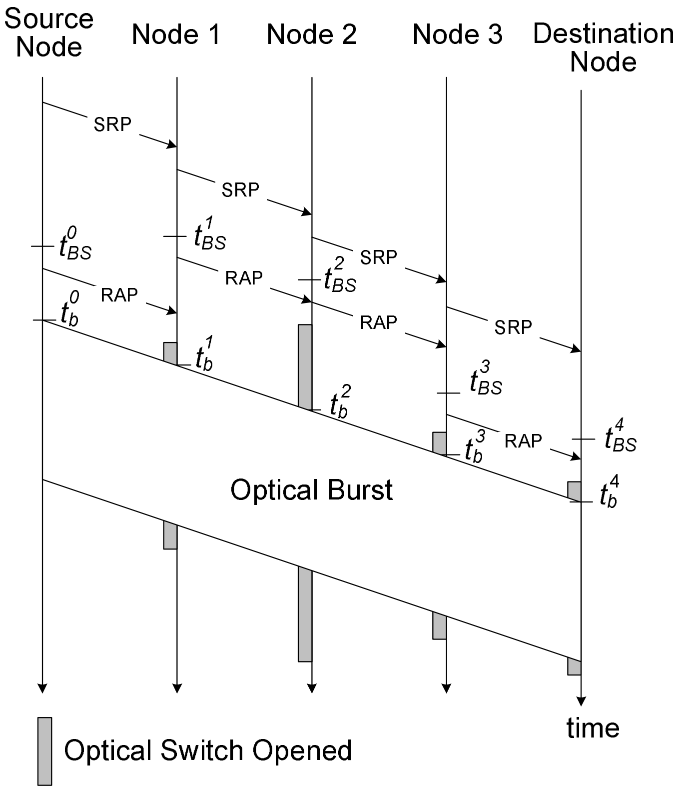

Fig. 3. The end-to-end signalling for a burst that travels through a four-hop path in a DOBS network. By separating the burst's resource request and the burst's channel information into two control packets, the burst-reservation and delayed scheduling can be implemented. This allows the burst scheduling time t^i_{BS} and the scheduling offset size to be selected independently at each node, allowing for increased flexibility and performance.

bursts are serviced is a function of their scheduling offset sizes, and since $\Omega^i_{BS} = t^i_b - t^i_{BS}$, DOBS provides complete control over the order in which arriving bursts are processed.

In Fig. 2, the burst scheduling operation is performed before the incoming RAP arrives (i.e., $t^i_{RAP\_in} > t^i_{BS}$), so the outgoing channel selection is made independently of the incoming channel. However, depending on the timing of adjacent nodes in the burst-switched path, the incoming RAP may arrive before the burst is scheduled (i.e., $t^i_{RAP\_in} < t^i_{BS}$). In these instances, the upstream channel information may also be used by the burst scheduler when scheduling the burst. In a system in which the outgoing channel decision depends on the incoming channel (e.g., systems with partial wavelength conversion), one would want to select the functional offset size on each link in the network such that upstream RAPs generally arrive before their bursts' scheduling times.

D. End-to-End Signalling

Fig. 3 illustrates the end-to-end signalling that takes place as a single burst is forwarded over a four-hop path through a DOBS network. After an SRP arrives at node-i, its contents are stored until t^i_{BS}, at which time the burst is scheduled. After the burst-scheduling is complete, an RAP containing the outgoing channel of the burst is transmitted downstream. As in classical OBS, the size of the physical offset time between the SRP and the burst shrinks as the burst travels through the network. However, since the burst scheduling time t^i_{BS} at each node can be chosen independently from the SRP arrival time, the scheduling offset size ($\Omega^i_{BS} = t^i_b - t^i_{BS}$) can also be selected independently at each node.

In general, multiple RAPs, each with a lifetime of one hop, can be used to communicate a burst's channel information downstream as it travels through the network. For example, in Fig. 3, one persistent SRP precedes the burst, but four separate RAPs are used. The pipelined transmission of RAPs is advantageous, as it not only simplifies timing in the resulting DOBS system, but also decreases the minimum required ingress physical offset size since burst-scheduling operations can be performed in parallel.

The ability to independently select offsets at each node can be leveraged in a number of ways to improve system performance. For example, in the system depicted in Fig. 3, the scheduling offsets on the second and third link in the path are selected to be significantly larger than that of the other links in order to accommodate the slow switching speed of node-2. In another scenario, making use of the fact that a burst's offset size is related to its priority [8], one could use dynamic offset provisioning in each core node to provide controllable hop-by-hop quality of service to each burst. In Section V, we describe a DOBS system with constant scheduling offsets that simultaneously allows for high throughput and ultra-simple burst scheduling.

E. Core Node Timing Constraints

Although DOBS affords a great deal of flexibility in selecting the scheduling time and the resulting scheduling offset, the size of the physical offset size imposes certain timing constraints at each node. We now derive expressions for the timing constraints in the core nodes of a DOBS network.

Consider the operation of the i^{th} node in a burst-switched path with h-hops. For the burst to be correctly forwarded through this node, the burst scheduling operation can only be performed after the SRP has been processed, and it must be completed in time for the optical switch fabric to be configured. Additionally, the downstream RAP must be transmitted early enough so that it is received by the switch reconfiguration time of the $(i+1)^{st}$ node. Using these three constraints, the allowable range of t^i_{BS} at node-i is (see Appendix for derivation)

$$t^i_{BS} \in (t^i_{SRP\_in} + \Delta^i_{SRP}, t^i_b - \Delta^i_{BS} - \max\{\Delta^i_{SW}, \Delta^{i+1}_{SW}\}), \quad \forall i \in \{0, 1, \ldots, h\} \qquad (1)$$

where for notational convenience, we have defined $\Delta^{h+1}_{SW} \equiv 0$.

Using (1), and from the definitions of Ω^i_{BS} and Ω^i_{ph}, we can obtain the following constraint on the size of the scheduling offset in terms of the physical offset (see Appendix for derivation).

$$\Omega^i_{BS} \in (\Delta^i_{BS} + \max\{\Delta^i_{SW}, \Delta^{i+1}_{SW}\}, \Omega^i_{ph} - \Delta^i_{SRP}), \quad \forall i \in \{0, 1, \ldots, h\} \qquad (2)$$

F. Ingress Offset Provisioning

The physical offset size provisioned to a burst at the ingress of the network ultimately limits the maximum size of the scheduling offset at each node in its path. While provisioning larger ingress offsets allows for more flexibility in core node scheduling operations, it is also advantageous from the perspective of delay to keep the ingress physical offset size as small as possible.

Consider an h-hop burst-switched path in a DOBS network. The physical offset shrinks at the i^{th} node by a value equal to Δ^i_{SRP}, so we have

$$\Omega^i_{ph} = \Omega^0_{ph} - \sum_{k=1}^{i} \Delta^k_{SRP}, \qquad \forall i = \{0, 1, \ldots, h\}. \tag{3}$$

Assume that a scheduling offset size of Ω^i_{BS} is desired at the i^{th} node of a burst-switched path. From (2), we require that $\Omega^i_{ph} > \Omega^i_{BS} + \Delta^i_{SRP}$. Combining this with (3), we have the following expression for the minimum required ingress offset.

$$\Omega^0_{ph} > \max_{i \in \{0,1,\ldots,h\}} \{\Omega^i_{BS} + \sum_{k=1}^{i} \Delta^k_{SRP}\} \tag{4}$$

In networks that carry delay sensitive traffic, one can minimize the end-to-end delay experienced by bursts by selecting the scheduling offset size for each link to be as small as possible. This minimizes the required ingress offset delay for each burst, resulting in reduced end-to-end delay. The minimum possible scheduling offset at node-i is given in (2). Substituting this minimum value into (4) yields

$$\Omega^0_{ph} > \max_{i \in \{0,1,\ldots,h\}} \{\Delta^i_{BS} + \max(\Delta^i_{SW}, \Delta^{i+1}_{SW}) + \sum_{k=1}^{i} \Delta^k_{SRP}\}. \tag{5}$$

The above expression specifies the absolute minimum physical offset size that must be provisioned at the ingress of the network in order to correctly forward bursts at every hop in the burst-switched path. However, the application of (5) requires accurate *a priori* knowledge of parameters which may not be available at the ingress node. For example, if hop-by-hop routing is employed, h may not be known at the source. Further, due to control-packet queueing, Δ^i_{BS} will generally vary as a function of link congestion. Thus, we propose the following slightly more conservative and much more practical ingress offset provisioning formula.

$$\Omega^0_{ph} > \Delta^{max}_{BS} + \Delta^{max}_{SW} + (h^{max}) \cdot \Delta^{max}_{SRP} \tag{6}$$

where Δ^{max}_{BS}, Δ^{max}_{SW}, and Δ^{max}_{SRP} are the network-wide maximum values of the burst scheduling, switch reconfiguration, and SRP processing durations respectively in the network, and h^{max} is the maximum number of hops that the burst can traverse between its source and destination. Each of these parameters can be estimated offline based on the topology, hardware, and expected worst-case traffic loads of the network.

G. Fantom Burst Reservations

In DOBS signalling, SRPs are forwarded before the outcome of the burst-scheduling algorithm is known. If a burst is blocked at an intermediate node in its path, bandwidth in downstream nodes may be wasted because the SRPs continue to make *fantom* reservations for bursts that will never arrive. This bandwidth waste can be reduced by transmitting a *teardown packet* (TDP) downstream to the destination node as soon as a burst is blocked. Upon receiving a TDP, each core node forwards it immediately downstream and then removes the corresponding burst reservation from the scheduler. Additionally, it might also be possible to completely eliminate fantom reservations in DOBS by including a look-ahead admission control step in the SRP processing operation.

There may also be a number of other ways to reduce the bandwidth waste associated with fantom reservations in DOBS. As shown below, however, the increase in blocking that results from fantom burst reservations is negligible in OBS systems with low or moderate blocking rates. Thus, any mechanism used to minimize the effect of fantom burst reservations will have little impact on the overall throughput performance of most DOBS systems.

Consider a single DOBS link with W channels, a fantom-free offered load of ρ, and a corresponding blocking probability of B. One can model the effect of fantom bursts as an increase in the offered load on the link, which we denote by $\Delta\rho$. In general, the blocking probability of every burst on a DOBS link need not depend on its length or its location in its path. Thus, the fraction of reservations on the link that are due to fantom bursts can be approximated (to first order) by the system blocking probability B, so we have $\Delta\rho \approxeq \rho \cdot B$.

For a system with a small value of B, assuming that the blocking probability varies smoothly as a function of the offered load, we can approximate the increase in blocking probability due to fantom bursts ΔB as

$$\Delta B \approxeq \frac{\partial B}{\partial \rho} \cdot \Delta\rho \approxeq \frac{\partial B}{\partial \rho} \cdot \rho \cdot B. \tag{7}$$

The form of the expression in (7) implies that the increase in blocking due to fantom reservations is a second order effect.

For example, if we assume that bursts arrive according to a Poisson process, B can be computed using the Erlang-B formula.

$$B = \frac{\rho^W / W!}{\sum_{i=0}^{W} \rho^i / i!} \tag{8}$$

By differentiating (8) with respect to ρ and substituting into (7), the following expression can be obtained for ΔB.

$$\Delta B \approxeq \rho \cdot \left[B^2(\frac{W}{\rho} - 1) + B^3 \right] \tag{9}$$

Thus, the improvement that can be realized by eliminating fantom bursts is on the order of B^2. Since modern communication networks are characterized by low blocking probabilities, one expects that the effect of over-provisioning due to fantom bursts would have negligible impact on the overall performance of most DOBS system. For example, in a 16 wavelength system with a blocking probability of 10^{-4}, (9) implies that the blocking probability increase due to fantom bursts would be approximately $2 \cdot 10^{-8}$. We examine the effect of fantom reservations on blocking performance using simulation in Section VII.

Interestingly, the first-order analysis above implies not only that the effect of fantom reservations is negligible, but also that the effects of the wasted *upstream* reservations of blocked bursts are also negligible if the network blocking rate is low. This is a very significant result for any one-way reservation architecture, including JIT and JET OBS, since any blocked burst in these networks leads to wasted upstream bandwidth reservations.

Because fantom reservation requests may be scheduled in downstream nodes, they may also lead to an increase in the overall processing load in the schedulers of a DOBS network. However, using analysis very similar to the analysis above, it can be shown that this increase in processing will only be significant in networks in which the number of blocked bursts is comparable to the number of unblocked bursts in the system.

V. Constant-Scheduling-Offset DOBS

The ability to control the scheduling offset size of bursts can be used in a number of ways to improve performance in DOBS systems. In this section, we describe one possible DOBS variant called *Constant-Scheduling-Offset* (CSO) DOBS. In CSO DOBS, each link in the network is associated with a single (possibly unique) scheduling offset value. By setting the offset of every burst on a given link to a single value,[1] CSO DOBS ensures FCFS operation in burst schedulers. The resulting DOBS system realizes better performance than classical single-header OBS systems in terms of throughput, delay, and fairness, while simultaneously decreasing the required burst-scheduling complexity.

We consider CSO DOBS systems with full wavelength conversion and no FDL buffers. Since bursts are serviced in the order of their arrival times in a CSO DOBS system, each outgoing link behaves like a classical FCFS loss system, except that the scheduler operates ahead of the switch by a time equal to the scheduling offset.

The fact the bursts are serviced in FCFS order is of key importance when designing CSO DOBS scheduling algorithms because it implies that *all best-effort, non-preemptive wavelength selection algorithms will result in identical blocking performance.* This is in sharp contrast with traditional single-header OBS systems [6], and it implies that one should employ the least-complex scheduling algorithm possible in CSO DOBS systems. We now present two variations of an O(1)-complexity scheduling algorithm for CSO DOBS systems called *free-channel queue* (FCQ) scheduling.

A. Free-Channel Queue Scheduling

To implement FCQ scheduling, each node stores a list of all the channels that are available for burst reservations in a *free-channel queue* (FCQ). Assume that a burst with scheduling time t^i_{BS} and length l_b arrives at node-i. If the FCQ is empty at time t^i_{BS}, it means that no channels are available at the burst arrival time, so the burst must be blocked. If the FCQ is not empty, the burst is scheduled onto the channel at the head of the FCQ at time t^i_{BS}, and that channel is removed from the FCQ. The channel is placed back into the FCQ at time $t^i_{BS} + l_b$.

At all times, the length of the FCQ will be the complement of the number of bursts in the system, except that it will lead

[1] To eliminate offset variability in burst schedulers, CSO DOBS selects the scheduling offset of all bursts on a given link to be identical. However, CSO DOBS does not require that every link in a burst switched path have the same scheduling offset. For example, the timing diagram in Fig. 3 could represent a burst-switched path in a CSO DOBS network.

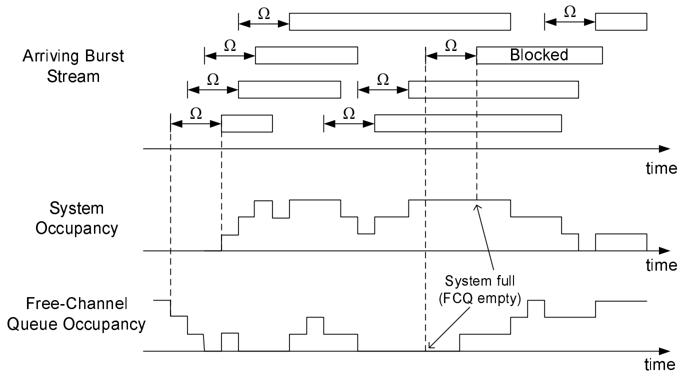

Fig. 4. Burst scheduling in a CSO DOBS system with three output wavelengths. The occupancy of the FCQ is the complement the system occupancy except that the former leads the latter by the scheduling offset time of the system Ω. The use of the FCQ allows for accurate scheduling decisions to be made Ω in advance of their corresponding burst arrivals and requires only $O(1)$ complexity.

the system state by the scheduling offset time Ω^i_{BS}. This is shown clearly in Fig 4, which depicts the time evolution of both the system occupancy and the FCQ occupancy for a given set of arriving bursts.[2]

Since the performance of a CSO system is not dependent on which free-channel is selected for an arriving burst, any data structure could be used to store the available wavelengths. However, by using a queue, it is ensured that all operations in an FCQ scheduler require only O(1) time. Thus FCQ scheduling is far simpler than previously proposed scheduling algorithms for JET OBS systems, and it is as simple as JIT scheduling. This advantage is discussed further in Section VI.

The FCQ algorithm can also be used in systems in which TDPs are used to communicate blocking events downstream in order to minimize the bandwidth waste associated with fantom burst reservations. The only modifications required to the FCQ algorithm are as follows. If the TDP arrives before t^i_{BS}, the burst reservation is canceled immediately, and the FCQ algorithm is not executed for the corresponding burst. If the TDP arrives after t^i_{BS} but before $t^i_{BS}+l_b$, then the channel that was assigned to the corresponding burst is returned to the FCQ as soon as the TDP arrives (i.e., instead of waiting for time $t^i_{BS} + l_b$, as described above). If the TDP arrives after $t^i_{BS} + l_b$, the no action is taken.

B. Scheduling Backlogs

If a number of bursts destined for the same output link arrive nearly simultaneously at a core node, the scheduler may still be busy scheduling one burst when another burst's scheduling time elapses. In such instances, it may be necessary to queue up the new reservation request before scheduling it. However, the queueing of reservation requests can affect the operation of the FCQ scheduling algorithm and can contribute to additional burst loss in two ways.

If the reservation backlog is large enough that the scheduler fails to schedule a burst before its requested arrival time, the burst will have to be blocked regardless of whether bandwidth

[2] For clarity, the superscript and subscript from Ω^i_{BS} have been dropped in the figure.

is available or not. Such losses can be avoided by selecting the scheduling offset size (and the ingress physical offset size) to be large enough to accommodate the extra delay caused by reservation queueing.

However, even with sufficiently large offsets, the reservation backlog may cause additional loss if not taken into consideration by the burst-scheduling algorithm because the queueing delay introduces timing differences in the setup and tear-down of reservations. In FCQ scheduling, if channels are immediately placed back into the FCQ but reservation requests are queued up, the FCQ occupancy may temporarily overstate the number of free channels in the system during times of congestion. This could lead to over-provisioning of the outgoing link. Similarly, if a number of burst reservations end simultaneously, and their channels cannot be instantaneously returned to the FCQ, it may be necessary to queue up tear-down requests, which could lead to under-provisioning on the output link and unnecessary burst rejections.

TABLE I
PERFORMANCE COMPARISON OF JIT OBS, JET OBS, AND CSO DOBS

	Offset Delay	Scheduling Complexity	Throughput Efficiency	Fairness
Constant Offset DOBS	$\Delta_{SW}+h\cdot\Delta_{SRP}+\Delta_{BS}$ (Lowest)	$O(1)$	Highest	No unfairness
JIT OBS	$\Delta_{SW}+h\cdot\Delta_{CP\_JIT}$ (Low)	$O(1)$	Lowest	No unfairness
JET OBS w/ Void-Filling	$\Delta_{SW}+h\cdot\Delta_{CP\_JET\text{-}VF}$ (High)	$O(\log Wm)$	High	Burst-length & path-length unfairness
JET OBS w/o Void-Filling	$\Delta_{SW}+h\cdot\Delta_{CP\_JET\text{-}HR}$ (Highest)	$O(\log W)$	Low	Path-length unfairness

W = number of wavelengths in the system
m = average number of bursts scheduled on each wavelength
h = number of hops in path of burst
Δ_{SW} = optical switch reconfiguration time
Δ_{CP_X} = control packet processing time of JET or JIT OBS systems
Δ_{SRP} = SRP processing time in DOBS system
Δ_{BS} = burst scheduling time in DOBS system

C. Input-Queued FCQ Scheduling

The problem of backlogs in a CSO DOBS scheduler can be effectively dealt with by modifying the FCQ algorithm to ensure that all reservation requests and tear-down operations are performed in the proper order regardless of congestion in the scheduler. This can be achieved by queueing up reservation requests and tear-down requests into a *reservation request queue* (RRQ) and a *tear-down request queue* (TRQ). The algorithm services the RRQ and the TRQ in parallel according to the event times at the head of each queue. This ensures that all operations on the FCQ are performed in the correct order so that over-provisioning or under-provisioning due to reservation and tear-down backlogs is avoided, while preserving the O(1) complexity of FCQ scheduling.

The occupancy of the FCQ no longer leads the occupancy of the channel by a constant time offset in the input-queued FCQ scheduler. For any input burst stream, however, the sequence of FCQ operations is unchanged. Thus, the input-queued FCQ algorithm results in scheduling decisions that are identical to those that would result if all scheduling operations could be performed instantaneously.

VI. PERFORMANCE COMPARISON OF CSO DOBS, JIT OBS, AND JET OBS

In this section, we compare the performance of CSO DOBS to that of JIT and JET OBS. We show that CSO DOBS simultaneously allows for increased throughput, increased fairness, decreased scheduling complexity, and decreased delay compared to the single-header OBS architectures. The results of the comparisons are summarized in Table I.

A. Burst-Scheduling Complexity

The amount of processing required to schedule each burst in OBS or DOBS depends directly on the amount of state information that must be searched through when selecting an output wavelength.

1) JET with Void-Filling: In JET with void-filling, the scheduler's state information consists of all future scheduled bursts in the system. By storing these burst intervals in a balanced tree structure, it has been shown that burst scheduling can be performed in $O(\log(mW))$ time, where m is the average number of void intervals on each wavelength ($m \geqslant 1$), and W is the number of wavelengths in the system [5].

2) JET without Void-Filling: In JET OBS without void-filling, a single horizon time is stored for each output channel. Each horizon time corresponds to the end of the latest scheduled burst in the channel. In LAUC scheduling, the scheduler must find the channel that minimizes the gap between the horizon time and the start time of the arriving burst. It has been reported that this takes $O(W)$ time, where W is the number of channels in the system [5]. However, by storing the horizon times in a balanced tree structure, Horizon scheduling could be implemented in $O(\log W)$ time.

3) JIT OBS: JIT trades-off throughput efficiency for scheduling simplicity. By sacrificing the bandwidth between the arrival of control packets and bursts, immediate scheduling allows for simple FCFS scheduling to be implemented in JIT networks. Thus, O(1) scheduling complexity can be realized in JIT OBS [12].

4) CSO DOBS: Since the scheduling offset of every burst is equal, the maximum amount of state information required in a CSO DOBS scheduler is a single binary value for each channel. As described in Section IV, by storing the list of available channels in a free-channel queue, all scheduling operations can be performed in O(1) time. Thus, CSO DOBS can realize the same ultra-low scheduling complexity as JIT OBS, without requiring immediate scheduling or its associated bandwidth waste.

B. Burst Blocking and Throughput

Except for systems in which the durations of bursts are very long compared to the duration of their offsets, the throughput of JIT OBS will be significantly compromised because of the bandwidth waste associated with immediate reservation. Thus, JET OBS and CSO DOBS can usually provide significantly higher throughput than JIT OBS.

Among JET OBS systems, void-filling algorithms such as LAUC-VF generally achieve higher throughput than non-void-filling algorithms by dealing more efficiently with the bandwidth fragmentation on outgoing channels caused by the presence of variable offsets and voids.

In void-filling JET OBS systems, however, some bursts may still be blocked even if all channels are never simultaneously occupied. This additional source of blocking has been described in a previous study on the throughput limitations of these systems [18].

By contrast, because offset-size variations are eliminated in CSO DOBS and bursts are scheduled in the order in which they arrive to the system, a burst in CSO DOBS will only be blocked if all channels are busy at the start of its reservation interval. Thus, CSO DOBS generally achieves as high or higher throughput levels than void-filling JET OBS.

In networks with extremely high blocking probabilities, the bandwidth wasted by fantom reservations could outweigh the benefits of void-free scheduling. In these high-loss systems, JET OBS with void-filling could achieve higher throughput than CSO DOBS. This is examined further using simulation in Section VII.

C. Ingress Offset Delay

Larger offsets result in higher delay at the ingress of the network, so it is advantageous to minimize the size of the ingress offset in burst-switched networks. We present expressions for the minimum required ingress offset delay for equivalent burst-switched paths in JIT, JET, and CSO DOBS networks. For simplicity of comparison, we assume that the control processing delay and switch reconfiguration delay do not vary from node to node (so we drop the superscript i in the notation) in this section.

For JIT and JET OBS, at each node in the burst-switched path, the residual offset must be larger than the optical switch reconfiguration time. Thus, the minimum ingress offset sizes required for transmission over h-hop paths in JIT OBS, JET OBS with void-filling scheduling, and JET OBS with Horizon scheduling are

$$\Omega^0_{JIT} = \Delta_{SW} + h \cdot \Delta_{CP\_JIT} \quad (10)$$

$$\Omega^0_{JET-VF} = \Delta_{SW} + h \cdot \Delta_{CP\_JET-VF} \quad (11)$$

$$\Omega^0_{JET-HR} = \Delta_{SW} + h \cdot \Delta_{CP\_JET-HR} \quad (12)$$

where Ω^i_X and $\Delta^i_{CP\_X}$ are the offset size and control packet processing delay at node-i, including any time spent waiting in the control packet queue ($X \in \{JIT, JET-HR, JET-VF\}$).

Since JIT scheduling is generally much simpler and faster than JET scheduling, (Δ_{CP_JIT}) is significantly smaller than Δ_{CP_JET-HR} and Δ_{CP_JET-VF}. Similarly, Δ_{CP_JET-VF} will generally be larger than Δ_{CP_JET-HR}, so we have that $\Omega^0_{JIT} \leqslant \Omega^0_{JET-HR} \leqslant \Omega^0_{JET-VF}$.

From (6), the ingress offset required on an h-hop path in a DOBS network is

$$\Omega^0_{DOBS} = \Delta_{SW} + \Delta_{SRP} + h \cdot \Delta_{BS}. \quad (13)$$

Since the burst scheduling complexities of JIT OBS and CSO DOBS are very similar, one expects that $\Delta_{CP\_JIT} \approx \Delta_{SRP} + \Delta_{BS}$, which implies that $\Omega^0_{DOBS} \leqslant \Omega^0_{JIT}$. Therefore, DOBS can achieve lower ingress offset delay than both JIT and JET OBS.

D. Fairness

There are two sources of unfairness in JET OBS systems. Firstly, in all JET OBS systems, bursts with longer offsets will tend to have higher priority and will generally experience less blocking than bursts with shorter offsets [7], [8], [9]. Since offsets shrink in a JET OBS system as bursts travel through the network, the priority of a burst is a function of its residual path length. Specifically, bursts that are close to their destinations have lower priority than bursts that are far from their destinations. This unfairness is undesirable because bursts that are close to their destination have already consumed a large amount of resources when they are dropped.

Secondly, in JET OBS systems that employ void-filling algorithms, blocking probability is generally an increasing function of burst length because shorter bursts are more likely to successfully fit into the gaps between previous burst reservations [7]. This unfairness is also disadvantageous in terms of overall system throughput, since long bursts can contain much more information and are, therefore, more valuable than short bursts.

Since immediate scheduling is employed in JIT OBS, the void between a burst and its control packet is not free for other burst reservations. In sacrificing this bandwidth, JIT OBS avoids the burst-length and path-length unfairness associated with variable offsets and void-filling algorithms.

In CSO DOBS, useful voids are not created during burst scheduling, so void-filling is unnecessary and bursts of all lengths on a given link will experience the same probability of being blocked. Further, since the scheduling offset size of a burst is not generally a function of the burst's location in its path, residual-path-length priority is also avoided in DOBS.

Burst-length and residual-path-length unfairness are examined further using simulation in Section VII.

VII. Simulation Results

In this section, we use simulation to compare the blocking performance of CSO DOBS to that of single-header OBS systems. Since JIT is intended for systems in which the burst length is much longer than the control-packet processing time, it is quite inefficient at the values of burst length and processing time considered here. Thus, we only include JIT performance curves when we examine the effect of control-packet processing time on throughput.

The simulation was written in C++, and between ten million and one billion bursts were simulated for each run. For all simulations, we assumed that core nodes were equipped with full wavelength conversion, and no FDL buffers. Each wavelength had a data transmission rate of 10 Gbps. Bursts arrived according to a Poisson process, and burst lengths followed an exponential distribution with a mean burst length of 100 kb. Unless otherwise specified, we assumed that the control-packet processing time (including control-packet queueing delay) for

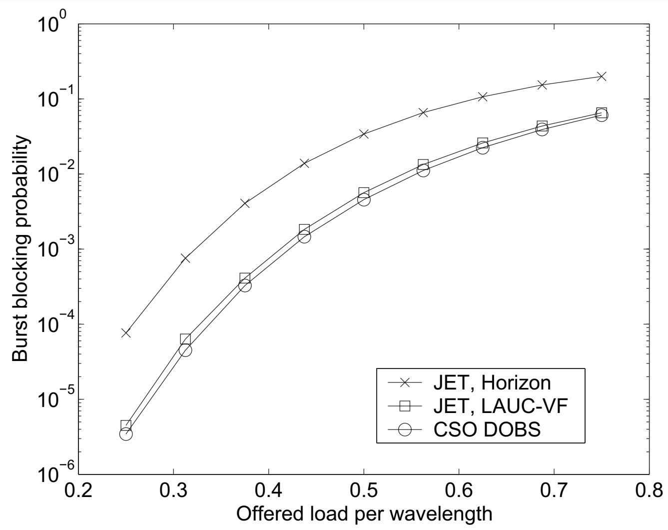

Fig. 5. Burst blocking probability as a function of offered load per wavelength for a system with 16 wavelengths. CSO DOBS achieves the lowest blocking probability, despite having lower burst scheduling complexity than the JET OBS systems.

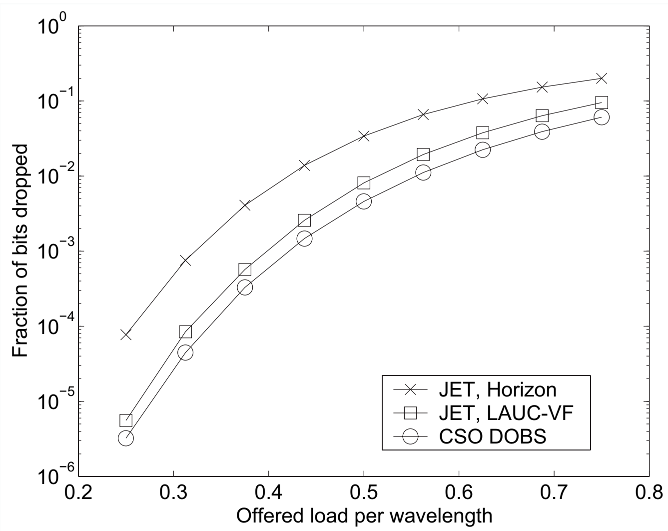

Fig. 6. Fraction of bits that are blocked versus offered load per wavelength for a system with 16 wavelengths. In the LAUC-VF system, overall throughput performance is compromised by the fact that long bursts are more likely to be blocked than short bursts.

the JET systems was 20 μs. The optical switch reconfiguration time was assumed to be 100 ns.

First, we examined the efficiency of each scheduling algorithm by simulating a single core-node of a burst-switched network. The single-node simulation had the advantage of allowing us to examine each system under the exact same input traffic conditions. At the ingress of the core node, we assumed that the residual path length of arriving bursts was uniformly distributed between one and five hops.

In addition to the single-node simulation, we also simulated a multi-hop network to determine the effect of overprovisioning in DOBS due to fantom burst reservations.

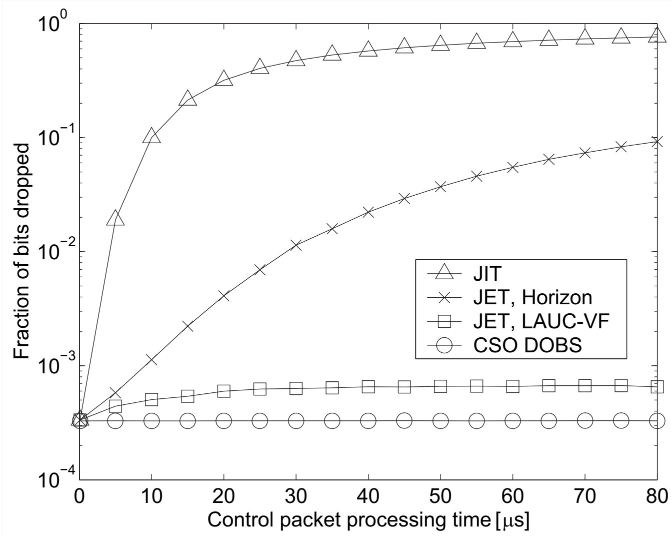

Fig. 7. Fraction of bits that are blocked as a function of processing time in a 16-wavelength system with an offered load per wavelength of 0.375. For the DOBS system, the horizontal axis represents the sum of the SRP processing time and the burst-scheduling time. For the JET OBS systems, the horizontal axis represents the control-packet processing time. The performance of the CSO DOBS system is unaffected by its control packet processing time.

A. Blocking Probability

Fig. 5 plots the burst blocking probability as a function of the offered load for a CSO DOBS system, a JET system using LAUC-VF scheduling, and a JET system using Horizon scheduling.

Void filling allows the LAUC-VF system to realize blocking probabilities of roughly an order of magnitude lower than those of the Horizon JET system. By selecting the size of the scheduling offset of each burst to be constant, the CSO DOBS system completely avoids the problem of voids, which results in a burst blocking probability that is between 25% and 50% lower than that of the LAUC-VF system.

Although the burst blocking probability is a useful measure of performance, one may also be interested in the total amount of data that is blocked in the system. Fig. 6 plots the fraction of bits blocked as a function of the offered load for CSO DOBS, Horizon JET, and LAUC-VF JET. The bit-blocking probability for the LAUC-VF system is roughly 100% higher than that of the DOBS system. The disparity between the bit-blocking performance and burst-blocking performance of LAUC-VF JET is due to the fact that LAUC-VF inherently favors short bursts over long bursts. This burst-length unfairness is further quantified in the fairness study in Section VII-C.

B. Effect of Processing Time

Fig. 7 plots the bit-blocking probability as a function of the control processing time for OBS and DOBS systems with an offered load per wavelength of 0.375.

The JIT system achieves a blocking probability very close to that of the other systems for very small values of control-packet processing time, but its performance degrades rapidly as the processing time increases. This is as expected since JIT OBS is only intended for systems in which the control-packet processing time is negligible compared to the burst duration.

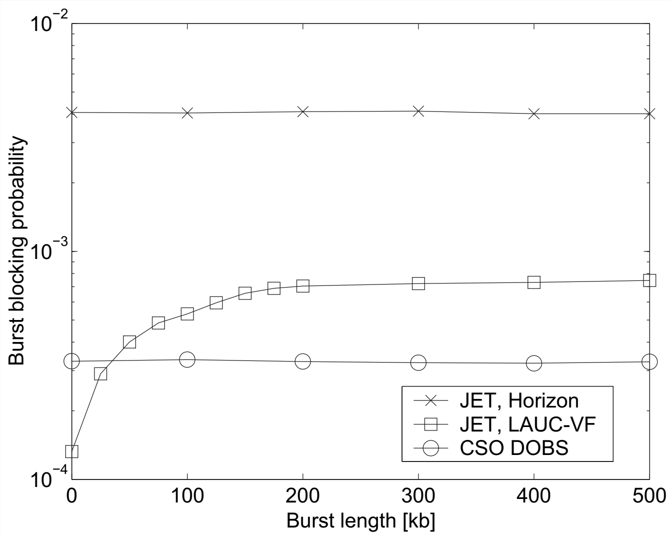

Fig. 8. Burst blocking probability as a function of burst length in a 16-wavelength system with an offered load per wavelength of 0.375. Blocking in the CSO DOBS system is completely fair with respect to burst length.

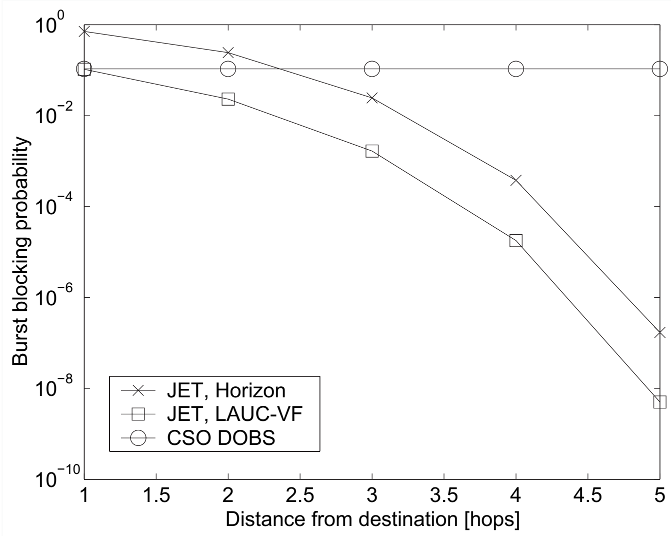

Fig. 9. Burst blocking probability as a function of the residual path length in a 16-wavelength system with an offered load per wavelength of 0.625. The probability of a CSO DOBS burst being blocked is independent of its location in its path.

In a JET OBS system, the presence of longer control processing times leads to larger voids among burst reservations. Since the bandwidth occupied by voids is wasted in Horizon systems, the blocking probability of the Horizon JET curve in Fig. 7 increases rapidly as the processing time increases. For the LAUC-VF JET system, as the processing time increases, the blocking probability initially increases to a point, after which it remains constant. This can be explained as follows. Initially, when the processing time is small compared to the length of bursts, most bursts are too large to fit into the gaps between previous reservations. In this regime, the LAUC-VF JET system only realizes a small benefit from void-filling, and larger processing times result in higher blocking probabilities. For larger processing times, the average void size is larger. Thus, LAUC-VF is able to efficiently fill these larger voids with bursts, so the blocking probability no longer increases with the processing time.

Since the scheduling offset size in DOBS is not a function of the physical offset size of incoming bursts, the performance of DOBS is unaffected by the processing time (i.e., the SRP processing time and burst-scheduling times), so the CSO DOBS curve in Fig. 7 is constant and lies below the JET and JIT curves for all points simulated.

C. Fairness

In this section, we use simulation to compare the fairness of CSO DOBS to JET OBS. Fig. 8 plots the blocking probability versus the burst length for similar CSO DOBS, LAUC-VF JET, and Horizon JET systems in which the offered load per wavelength is 0.375. Since both Horizon and CSO DOBS do not perform void-filling, their blocking probabilities are a constant function of burst length. By contrast, since short bursts are more easily scheduled into voids than long bursts, the blocking probability of LAUC-VF is an increasing function of burst length, varying from $1.3 \cdot 10^{-4}$ for very short bursts, to $7.5 \cdot 10^{-4}$ for very long bursts. Although this tendency to block long bursts over short bursts reduces the average burst-blocking probability of LAUC-VF systems, it compromises the overall throughput of the system, as was illustrated in Fig. 6.

Because the offset sizes of bursts in a JET OBS network is a function of their residual path length, bursts that are farther from their destination have higher priority and are less likely to be blocked than bursts that are closer to their destination. This is shown in Fig. 9 which plots blocking probability as a function of the physical offset size of arriving bursts for systems with an offered load per wavelength of 0.625. For both the Horizon and LAUC-VF systems, the blocking probability is a sharply decreasing function of the residual path length, varying from 0.1 to $5 \cdot 10^{-9}$ for the LAUC-VF system and from 0.7 to $1.7 \cdot 10^{-7}$ for the Horizon system. This implies that a burst that is one hop away from its destination is several million times more likely to be blocked than a burst that is five hops away from its destination in these systems. Since bursts close to their destination may have already consumed a large amount of network resources, this type of residual-path-length priority is undesirable.

The scheduling offset of every burst in the CSO DOBS system is set to a constant value, so each burst on a given link will experience the same probability of blocking, regardless of its location in its path and its corresponding physical offset size. This is verified by the constant-valued DOBS curve in Fig. 9. Thus, DOBS allows for perfect fairness with respect to the length of bursts and their distance from their destinations.

D. Effect of Fantom Burst Reservations

The analysis in Section IV-G implies that the over-provisioning caused by fantom reservations in DOBS only affects system performance in networks with very high blocking rates. In this section, we verify this conclusion by examining the overall blocking performance of CSO DOBS and JET-OBS with void-filling in a multi-hop network.

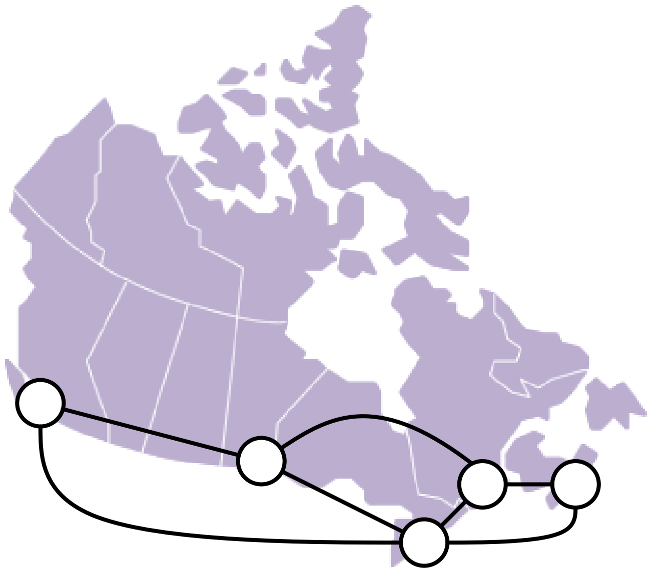

Fig. 10. The Canadian national optical network (CA*Net4) topology.

We considered a topology that is based on the core topology of the Canadian research and education optical network (CA*Net4) [21]. The topology is illustrated in Fig. 10. For the simulation, each link had 16 wavelengths, and the line-rate of each wavelength was 10Gbps. An equal volume of traffic was transmitted between each source-destination pair, and shortest-hop routing was used.

Fig. 11a) plots the burst-blocking probability versus the normalized offered load per wavelength per link (i.e., versus the average offered channel fill-rate) for both LAUC-VF JET and CSO DOBS. The blocking performance of the two systems is almost identical for all loads simulated. To examine the performance in the very-high blocking regime, Fig. 11b) plots the results over a wider range of offered loads using a linear scale. When the normalized offered load is greater than 2, the over-provisioning caused by fantom bursts causes the DOBS system to experience slightly higher blocking probabilities than the JET system. However, the maximum magnitude of the difference between the two curves is only 5%, and the effect of fantom bursts is only evident when the system blocking probability is greater than 0.5 (i.e., when over half of the bursts offered to the system are blocked).

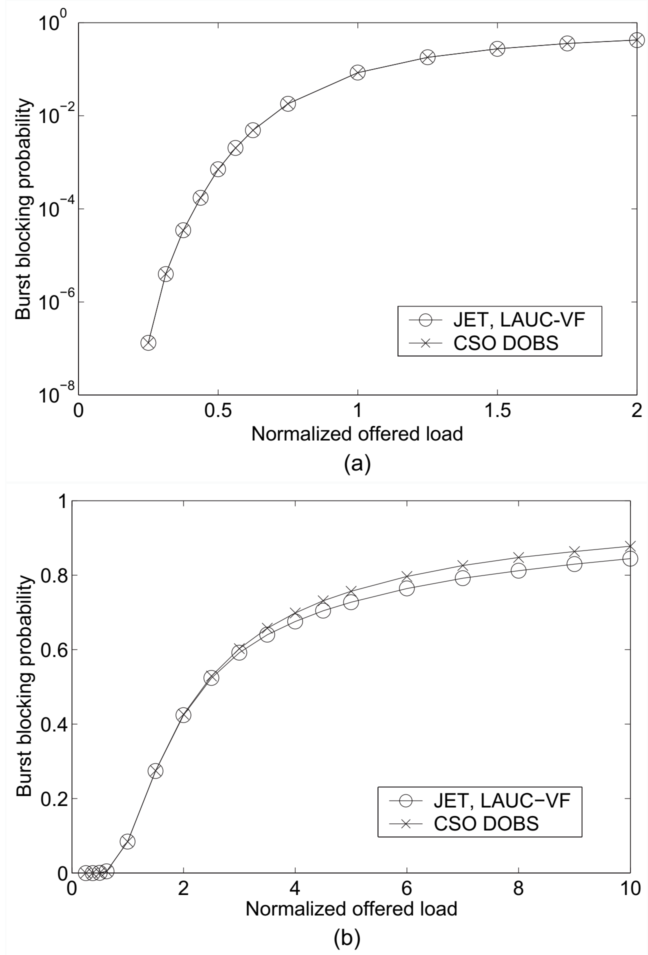

Fig. 11. Burst blocking probability as a function of normalized offered load for LAUC-VF JET-OBS and for CSO DOBS. The effect of over-provisioning due to fantom reservations in the DOBS network is negligible for blocking probabilities below 0.5. Even at extremely high blocking probabilities, the increase in blocking due to fantom reservations is less than 5%.

VIII. Future Work

In this paper, we examined a relatively simple DOBS system in order to describe the benefits of delayed scheduling and offset control. In future studies, we plan to extend the study to include DOBS systems that employ FDL buffers for contention resolution and DOBS systems with partial wavelength conversion capability. Further, we also plan to examine how an additional admission control step could be added to the SRP processing operation. This would not only allow improved throughput in networks with high blocking, but would also allow for deflection routing algorithms to be efficiently implemented in times of congestion.

In this study, we examined the benefit of setting the scheduling offset size of all bursts on each link to be constant. However, there are a number of other ways in which delayed scheduling can be used to improve the performance of burst-switched networks. For example, by sufficiently delaying the scheduling of arriving bursts, burst reservation requests could be accumulated. This would allow for batch scheduling of bursts, which could significantly increase the throughput of the DOBS system. Further, by including a class-of-service field in each SRP, a priority-based batch scheduling algorithm could be used to provide controllable hop-by-hop quality of service to high-class bursts. The design and performance evaluation of such schemes will be the focus of future studies.

IX. Conclusions

We have introduced a new signalling architecture for optical burst switching networks called DOBS that makes it possible to precisely control the scheduling offset of each burst at every node in its burst-switched path without the use of FDL buffers. We described how this added flexibility can be used to realize a system in which the scheduling offset size of every burst on a given link was set to a constant value. This not only simplified burst-scheduling but also allowed for higher throughput levels than even the most complex void-filling JET OBS systems.

We described an O(1) scheduling algorithm for the resulting CSO DOBS system called FCQ scheduling. Using simulation, we compared the performance of CSO DOBS to that of JIT and JET OBS. We found that the overall blocking probability achieved by the DOBS system was 50% lower than that of a similarly loaded LAUC-VF system. We also showed that the DOBS system was fairer than JET OBS with respect to burst length and with respect to the residual path length of bursts.

The superior throughput performance of DOBS compared to JET OBS, coupled with DOBS's dramatically reduced scheduling complexity, makes DOBS a very promising architecture for next-generation WDM burst-switched networks. Future work include extending our study to systems that employ FDL buffers or limited wavelength conversion for contention resolution, and examining other methods by which delayed scheduling in DOBS can be used to improve the performance of burst-switched WDM networks.

Appendix

Here we derive the burst scheduling and offset provisioning constraint equations that were given in (1) and (2).

For a given burst that arrives at node-i, the burst scheduling operation cannot be performed until after the SRP has been processed, and it must be completed before the switch-configuration time of the burst. These requirements impose the following two constraints on the timing of the node's operations.

$$t^i_{BS} > t^i_{SRP\_in} + \Delta^i_{SRP}, \qquad \forall i \in \{0, 1, \ldots, h\} \tag{14}$$

$$t^i_{BS} < t^i_b - \Delta^i_{BS} - \Delta^i_{SW}, \qquad \forall i \in \{0, 1, \ldots, h\} \tag{15}$$

Since the burst scheduling operation must be completed before the RAP can be transmitted downstream, we have

$$t^i_{BS} + \Delta^i_{BS} < t^i_{RAP\_out}, \quad \forall i \in \{0, 1, \ldots, h-1\}. \tag{16}$$

At each downstream node in the path, the switch cannot be configured until the upstream RAP arrives. Thus, the transmission of the downstream RAP from node-i is constrained by the timing in node-$(i+1)$. So, we require

$$t^i_{RAP\_in} < t^i_b - \Delta^i_{SW}, \qquad \forall i \in \{1, 2, \ldots, h\}. \tag{17}$$

If we denote the propagation time of link-$(i+1)$ as t^{i+1}_{prop}, we can relate the timing of the RAP and burst transmission in node-i and node-$(i+1)$ by

$$t^i_{RAP\_out} = t^{i+1}_{RAP\_in} - t^{i+1}_{prop}, \quad \forall i \in \{0, 1, \ldots, h-1\} \tag{18}$$

$$t^i_b = t^{i+1}_b - t^{i+1}_{prop}, \qquad \forall i \in \{0, 1, \ldots, h-1\}. \tag{19}$$

Starting with (16), and then substituting (18), (17), and (19), we can derive a third and final constraint for the burst scheduling time at node-i as follows:

$$t^i_{BS} + \Delta^i_{BS} < t^i_{RAP\_out}, \qquad \forall i \in \{0, 1, \ldots, h-1\} \tag{20}$$

$$= t^{i+1}_{RAP\_in} - t^{i+1}_{prop} \tag{21}$$

$$< t^{i+1}_b - \Delta^{i+1}_{SW} - t^{i+1}_{prop} \tag{22}$$

$$= t^i_b - \Delta^{i+1}_{SW}$$

$$t^i_{BS} < t^i_b - \Delta^i_{BS} - \Delta^{i+1}_{SW}. \tag{23}$$

Combining (15), (14), and (23) yields the single constraint equation for t^i_{BS} that is given in (1):

$$t^i_{BS} \in (t^i_{SRP\_in} + \Delta^i_{SRP}, t^i_b - \Delta^i_{BS} - \max\{\Delta^i_{SW}, \Delta^{i+1}_{SW}\}), \quad \forall i \in \{0, 1, \ldots, h\}.$$

Subtracting t^i_b from both sides of the above equation yields

Since $\Omega^i_{ph} \equiv t^i_b - t^i_{SRP\_in}$ and $\Omega^i_{BS} \equiv t^i_b - t^i_{BS}$, (24) can be rewritten as

$$\Omega^i_{BS} \in (\Delta^i_{BS} + \max\{\Delta^i_{SW}, \Delta^{i+1}_{SW}\}, \Omega^i_{ph} - \Delta^i_{SRP}), \quad \forall i \in \{0, 1, \ldots, h\}$$

which is the constraint equation given in (2).

References

[1] C. Qiao and M. Yoo, "Optical burst switching (OBS): a new paradigm for an optical Internet," *J. High Speed Networks*, vol. 8, pp. 69–84, Mar. 1999.

[2] T. Battestilli and H. Perros, "An introduction to optical burst switching," *IEEE Commun. Mag.*, vol. 41, pp. S10–S15, Aug. 2003.

[3] C. Qiao, "Labeled optical burst switching for IP-over-WDM integration," *IEEE Commun. Mag.*, vol. 38, pp. 104–114, Sept. 2000.

[4] J. Li, C. Qiao, and Y. Chen, "Recent progress in the scheduling algorithms in optical-burst-switched networks," *J. Optical Networking*, vol. 3, pp. 229–241, Apr. 2004.

[5] J. Xu, C. Qiao, J. Li, and G. Xu, "Efficient channel scheduling algorithms in optical burst switched networks," in *Proc. IEEE INFOCOM 2003*, pp. 2268–2278.

[6] T. Chen, C. Qiao, and X. Yu, "Optical burst switching (OBS): A new area in optical networking research," *IEEE Network*, vol. 18, pp. 16–23, May 2004.

[7] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *AEÜ International J. Electron. and Commun.*, vol. 55, pp. 18–26, Jan. 2001.

[8] N. Barakat and E. H. Sargent, "Analytical modelling of offset induced priority in multi-class OBS networks," *IEEE Trans. Commun.*, vol. 53, pp. 1343–1352, Aug. 2005.

[9] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.

[10] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *J. Lightwave Technol.*, vol. 18, pp. 2019–2037, Dec. 2000.

[11] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "Jumpstart: a just-in-time signaling architecture for WDM burst-switched networks," *IEEE Commun. Mag.*, vol. 40, pp. 82–89, Feb. 2002.

[12] J. Teng and G. N. Rouskas, "A detailed analysis and performance comparison of wavelength reservation schemes for optical burst switched networks," *Photonic Network Communications*, vol. 9, May. 2005.

[13] L. Xinwan, C. Jianping, W. Guiling, W. Hui, and Y. Ailun, "An experimental study of an optical burst switching network based wavelength-selective optical switches," *IEEE Commun. Mag.*, vol. 43, no. 5, pp. S3–S10, May 2005.

[14] H. Guo, Z. Lan, J. Wu, Z. Gao, X. Li, J. Lin, and Y. Ji, "A testbed for optical burst switching network," to appear in *Proc. OFC 2005*.

[15] J. Turner, "Terabit burst switching," *J. High Speed Networks*, vol. 8, pp. 3–16, Jan. 1999.

[16] Y. Xiong, M. Vandenhoute, and H. Cankaya, "Control architecture in optical burst-switched WDM networks," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1838–1851, Oct. 2000.

[17] M. Iizuka, M. Sakuta, Y. Nishino, and I. Sasase, "A scheduling algorithm minimizing voids generated by arriving bursts in optical burst switched WDM network," in *Proc. IEEE GLOBECOM'02*, vol. 3, pp. 2736–2740, Nov. 2002.

[18] J. Li, C. Qiao, J. Xu, and D. Xu, "Maximizing throughput for optical burst switching networks," in *Proc. IEEE INFOCOM'04*, pp. 1853–1863, Mar. 2004.

[19] L. Xu, H. Perros, and G. Rouskas, "A simulation study of access protocols for optical burst-switched ring networks," *J. Computer Networks*, vol. 41, no. 2, pp. 143–160, 2003.

[20] J. White, M. Zukerman, and H. L. Vu, "A framework for optical burst switching network design," *IEEE Commun. Lett.*, vol. 6, pp. 268–270,