An Optical Hybrid Switch With Circuit Queueing for Burst Clearing

Eric W. M. Wong, *Senior Member, IEEE*, and Moshe Zukerman, *Fellow, IEEE*

***Abstract*—We consider an optical hybrid switch that can function as an optical burst switch and/or optical circuit switch. We propose and describe in detail a new implementation whereby circuits have nonpreemptive priority over bursts. To achieve nonpreemptive priority, during circuit setup time, if there exist bursts that use wavelength channels (also called links) required by the circuit, the circuit is allowed to queue for a relatively short period of time until these bursts are cleared. We present an analysis based on a 3-D Markov chain that provides exact results for the blocking probabilities of bursts and circuits, the proportion of circuits that are delayed and the mean delay of the circuits that are delayed. Because it is difficult to exactly compute the blocking probability in realistic scenarios with a large number of wavelengths, we derive computationally scalable and accurate approximations based on reducing the 3-D state space into a single dimension. These scalable approximations that can produce performance results in few seconds can readily enable switch dimensioning. Extensive numerical results are presented to demonstrate the accuracy and the use of the new approximations.**

***Index Terms*—Blocking probability, optical burst switching (OBS), optical circuit switching (OCS), optical hybrid switching (OHS), queueing delay.**

I. Introduction

OPTICAL hybrid (circuit/burst) switching (OHS) [1]–[6] that supports both optical circuit switching (OCS) [9]–[13] and optical burst switching (OBS) [14]–[32] has been proposed as a solution that combines the benefits of OBS and OCS and could potentially provide an evolutionary path beyond OCS. It is envisaged that connections using OCS enjoy predictable and reliable performance while optical bursts benefit from improved efficiency in resource utilization and more flexible connectivity. A further step towards efficient optical network operation has been taken by the PATON architecture [33] that provides an integrated signaling and control framework that facilitates implementation of OCS and OBS and, potentially, other technologies on the same network.

Manuscript received June 11, 2007; revised December 4, 2007. Current version published January 28, 2009. This work was supported in part by a Grant from the City University of Hong Kong (Project No. 7002016) and in part by the Australian Research Council. This work was performed when M. Zukerman was with the Australian Research Council Special Research Center for Ultra-Broadband Information Networks (CUBIN), EEE Department, The University of Melbourne, Victoria, Australia.

The authors are with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China (e-mail: ewong@ee.cityu.edu.hk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2008.917320

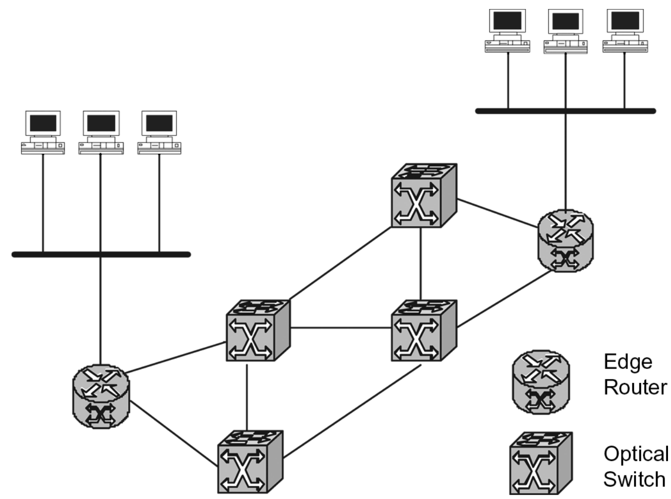


Fig. 1. Optical hybrid switching transport network architecture.

We consider a network architecture involving electronic edge and optical core as shown in Fig. 1. Data is aggregated electronically in edge routers and transmitted optically over the core. Notice the generality of this network architecture definition. We do not say anything about the size of the electronic edge or the optical core. For example, there is nothing in our definition that excludes the case where someone's PC acts as an edge router. Data aggregated at the edge router are classified into data transmitted via OBS or OCS. Examples for data that use OCS could include data associated with real-time services (streaming applications) requiring Quality of Service (QoS) guarantee, or a burst of data requiring predictable performance. Typical examples for OBS data traffic could include data associated with services that are not too QoS demanding such as email and web browsing. OBS could include real time connections to remote locations where OCS connection will be too costly to set up.

By OCS, we include all methods of switching involving exclusive connection between edge routers. It includes all cases where an edge-to-edge OCS connection is set up and the capacity is exclusively available for that connection for its entire duration. These include scenarios where capacity is permanently or semi-permanently available—the so-called Static OCS. By OCS, we also include cases where connections are set up and taken down frequently (Dynamic OCS) including the so-called OBS with acknowledgement [28], or the similar earlier proposal called Wavelength Routed OBS (WR-OBS) [29]. On the other hand, by OBS we include all methods of optically transmitting and routing bursts of data using one-way reservation, but without the ability to buffer the bursts in the optical domain. Typically, after data is aggregated into bursts in an edge

0733-8724/$25.00 © 2008 IEEE

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

router, a burst header is transmitted on a separate control channel ahead of the transmission of the burst payload to reserve capacity for the burst payload in each of the optical switches along the burst route. This way, the burst header sets up a lightpath for the burst, allowing it to remain in the optical domain through its entire route in the optical core without the need for O/E/O conversions until it reaches the edge router at the other end of its route.

In other words, in both cases of OCS and OBS, data packets are aggregated in edge routers in large buffers from where they are transmitted optically to other edge routers through the core network without being buffered on their way in the core network. While data is normally not lost inside the core optical network under OCS as capacity is exclusively guaranteed for the connection, under OBS, loss may occur if too many bursts arrive from many input wavelength channels (or input links) to be forwarded to the same output trunk, but there are not enough suitable output wavelength channels, henceforth called *output links*, there to accommodate them. In such a case, a burst is dumped.

In an OHS, users or network operators will be able to choose between OCS or OBS for each flow or stream of packets. The choice will be made based on QoS requirements, traffic behavior, and pricing. At this stage of the research it is difficult to predict how such choices are best made. It is hard to predict, for example, the average holding time of circuits, as some circuits may have very long holding time, e.g., those used for services equivalent to permanent leased lines whiles others could be smaller than some bursts.

As in many situations of designing future technology, we have a chicken-and-egg problem. The design of OHS needs to be optimized based on knowing the OCS/OBS choices namely the traffic characteristics, and the optimizing the OBS/OCS choices depends on the OHS design. Avoiding this chicken-and-egg problem, we define, in Section II, sensible rules of engagement between circuits and bursts and a circuit set-up protocol consistent with these rules of engagement.

Our rules of engagement are based on the principle that during circuit set-up time, we allow bursts that have already made a reservation to clear before the circuit set-up is complete. As we discuss in Section II, this approach is a midway between the two extreme approaches discussed in [3], namely, the preemptive priority on one hand and the no-priority on the other. In other words, on the one hand, our approach here is more favorable to bursts than the one that gives preemptive priority to circuits over bursts at the expense of increasing somewhat circuit set-up time, and on the other hand, it is more favorable to circuits than the no-priority approach where circuits are rejected if bursts have already made reservations for resources in any part of their path. It is achieved by allowing the circuit to queue for a relatively short period of time until these bursts are cleared. Although [3] provides a Markov chain analysis of models of these two extreme approaches, it does not provide protocols for their implementations. Such protocols are described in Section II.

Then, after describing the model in Section III, we provide in Sections IV and V, a Markov chain-based queueing analysis of a single strictly nonblocking OHS node in accordance with the rules of engagements and the circuit set-up protocol. Such a single node analysis can serve as the first step towards a network analysis. We focus there on an output trunk with L output links, and M input links that feed the L output links with traffic (OBS and/or OCS). The analysis leads to numerical procedures that provides performance results for the bursts and the circuits. Note that when we say M wavelength channels, or links, we mean to indicate that the same wavelength on two different fibers are two different wavelength channels (or links). Hereinafter, the concept of wavelength channel is interchangeable with link.

While Section IV provides exact results based on analysis of a 3-D Markov chain which are limited to small size problems, in Section V, we provide scalable and accurate approximations for the blocking probability of circuits and bursts, proportion of circuits that are delayed and the mean delay of the circuits that are delayed. The approximation is based on simplifying a 3-D Markov chain into a single-dimensional Markov chain. As demonstrated in Section VII, the approximation characterizes sufficiently accurately the many kinds of traffic we consider: transmitted bursts, dumped bursts, circuits in service and successful and failed circuit reservations. Because it is difficult to directly characterize all these interacting traffic types, so that they are merged into a single measure in one step, the approximation is based on fixed-point solution involving a two-module iterative procedure. A proof of existence and uniqueness of the fixed-point solution is provided in Section V. We also provide in the Appendix a binary search algorithm and prove its convergence to the fixed-point solution. As noted in [8], the equivalent proof in [3] is incomplete, so the new algorithm presented here is the first OHS algorithm that its convergence to a unique solution is proven.

Section VI provides a solution for an Erlang B type model with non preemptive priorities that serves as a bound for the case "$M = \infty$" and further provides a confirmation for the accuracy of our approximation. As we also demonstrate in Section VII, our scalable approximations provided in this paper can readily enable buffer and trunk dimensioning of any realistic size, meaning that for any given traffic loading, we can compute quickly the performance measures so we can choose the right buffer and trunk capacity such that the required quality of service is obtained.

II. Rules of Engagement

Having two types of traffic in OHS, the circuits and the bursts, there is a need to manage efficiently the interaction between them. Let us now define and try to justify our proposed rules of engagement between circuits and bursts.

First, in OHS, in general, circuits should have some priority over bursts. It does not seem reasonable to reject an entire circuit only because there are bursts in its way.

Second, we would like to avoid a situation whereby a burst that has already been partially transmitted through a switch is rejected because this can, in many cases, cause also the loss of the successfully transmitted part. This means that we will disallow circuit priority preemption [3], [4] whereby during circuit set-up time, bursts using output links may be rejected to allow the completion of the circuit set-up. This way, after a burst has been scheduled, it will not be preempted, and its header continues to visit other optical switches along its path and reserve

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

capacity for its burst "knowing" that the reservations it already made will, in fact, be allocated for the transmission of its burst. Clearly, it is not efficient to keep making false reservations for a burst that has been preempted. However, as we acknowledge the need to give priority to circuits over bursts, we propose here to make such priority "nonpreemptive". The latter is achieved by allowing a circuit to queue for a short while until the bursts that use its intended path clear out so that its end-to-end path can be set up. This is equivalent to a situation where an advanced reservation is made for a circuit when it cannot be admitted right away. This way circuits have the advantage/priority over bursts because bursts cannot be queued, and as long as there are circuits queued, newly arriving bursts are dumped.

Third, while we allow circuits to wait for bursts to clear out, we do not allow new circuit requests to queue if all output links, in any of the switches along the path, are used by circuits. Moreover, we do not allow new circuit requests to queue if a circuit will have to wait for another circuit to complete its service before the new circuit can be set up. Circuits can only wait for bursts to clear out and not for other circuits. In other words, we do not allow a new circuit request to queue if at any switch along the path of the new request, the total number of circuits—those queued to use the relevant output trunk plus those using that output trunk—is equal to the total number of output links in that trunk. As mentioned above, L designates the number of output links on a given output trunk in a given switch, so L is also the upper limit of the number of circuits served or queued for the given output trunk. Then, if there are i circuits and j bursts in progress and $i + j = L$, in any of the relevant switches, no more than j circuits are allowed to be queued. Unlike our previous OHS models with circuit queued reservations [4], [5], [6] where queued circuits wait for both circuits and bursts to clear, here we do not queue circuits to wait for other circuits to clear. This is consistent with normal circuit switching practice. The philosophy we adopt here is that it is realistic to expect that the designer will dimension the trunks so that the probability that a circuit request arrives when there are L circuits in the system is sufficiently small to meet the required quality of service (QoS), and if it happens that the circuit is blocked, it makes more sense to choose an alternative route, for example, than to wait for another (possibly long) circuit to complete its service. Of course, if a circuit holding time is completed and there happens to be other circuits waiting for burst clearing, such circuits can commence service.

Fourth, in OHS as in OCS, and in line with telephony circuit switching, when a circuit set-up is complete, the circuit has exclusive right to use the edge-to-edge lightpath already assigned to it.

Although we expect that circuits are longer than bursts, the length of a burst is not negligible, because for correct performance evaluation, we should consider its length to be associated with the time period that starts at the moment the header arrives and is processed until the moment that the burst transmission is terminated. This consideration which adds to the burst transmission time the time from the header arrival until the burst actual arrival is justified by the following: 1) recall that in OBS, a burst is scheduled or rejected at the time that its header arrives and is processed [14], [19], [21], [30], and that 2) under a practical scheduling such as the so-called *horizon scheduling* [19], no other bursts can be scheduled to be transmitted between the arrival of the header and the arrival of the burst on the same wavelength channel.

Given our model, the OBS traffic views the network operationally as if it is a pure OBS network. There is nothing in our model that excludes the use of deflection routing [31], [32], for example. From a performance point of view, on the other hand, the OBS traffic may experience increased loss during periods of increasing circuit loading.

Having adopted the above rules of engagement and established that the OBS traffic views (operationally) the network as a pure OBS network, it is important to clarify the condition for circuit blocking and circuit queueing during set-up. During a set-up of an edge-to-edge circuit for a given route, one of the following three outcomes may occur.

1) The circuit will be immediately admitted if there is a free link for its transmission in each of the optical switches along the given edge-to-edge route.
2) The circuit will be blocked from the given route if and only if it is blocked from at least one of the optical switches along the given edge-to-edge route.
3) The circuit will be queued if there is no free link for its transmission in at least one of the optical switches along the route and the total number of circuits—queued and in progress—in each of the optical switches along its route is less than the total number of output links. This implies that there is at least on optical switch along the route where there are some bursts in progress using output links and the circuit can be queued to wait until the bursts clear out and an output link is available for it.

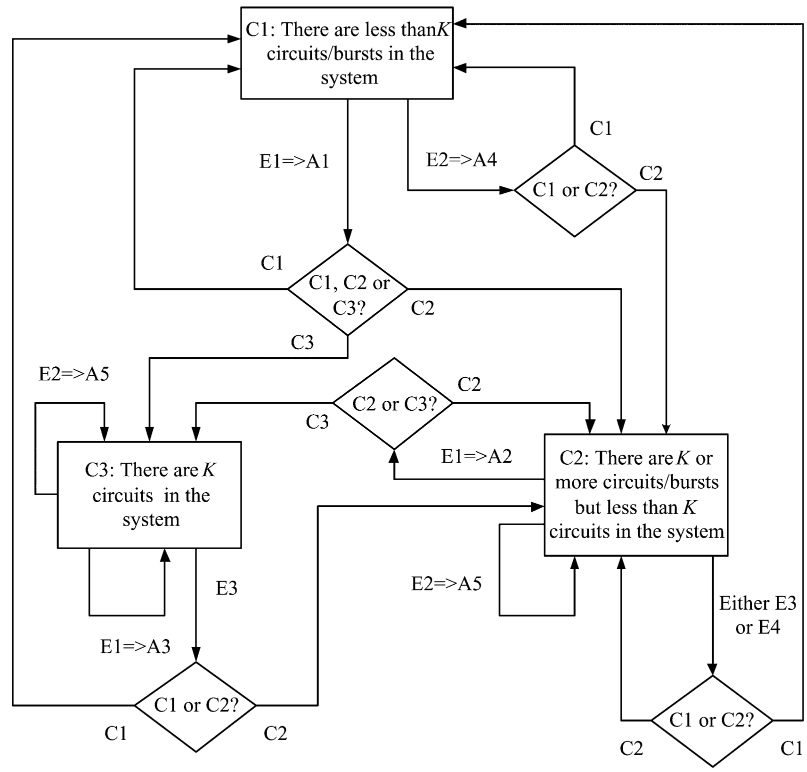
We will now outline an implementation of our new two-way circuit set-up protocol which is consistent with the above-described rules of engagements and possible outcomes.

To be consistent with our second principle not to schedule bursts that are later rejected, a circuit-set-up signalling-packet will visit each optical switch in the forward direction, and reserve an output links for the connection. To be specific, in the forward direction, the signal "instructs" each of the switches to exclusively reserve a particular output link if such is available. If non is available due to bursts in progress, the instruction is to reserve an output link in the first opportunity when such link become available. Such output link can become available when a burst/circuit in progress completes its transmission and no other circuit requests are queued ahead of the considered circuit request. Finally, we note that if we do not assume horizon scheduling, the offset period of an already scheduled burst is not covered by a circuit reservation and can be reserved and used by a new burst.

Therefore, there are three possibilities that correspond to the above three general outcomes.

1) If a link is available for a circuit request in each optical switch, then a signalling-packet will visit them all again on the backward route to confirm the successful circuit set-up.
2) If the signalling packet "realizes" that circuit request is blocked from one of the switches (i.e., it has already L circuits in progress or reserved there), then a signalling-packet will visit them all again on the backward route

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.



Events:
E1: Arrival of a forward successful-so-far circuit request
E2: Burst arrival
E3: Circuit departure due to either the end of holding time by the user or a cancellation of circuit request reservation or queued because a backward unsuccessful signal arrives
E4: Burst departure

Actions:
A1: A wavelength is assigned to this request
A2: The request is queued and if the request is at the head of the queue, the next relevant wavelength will be assigned to this request
A3: The circuit request is blocked
A4: Assign a wavelength for the burst
A5: Reject the burst

Comment: The three states in this diagram only represent the occupancy state and as such they do not reflect the waiting of the backward signal for a queued request because such waiting does not affect the occupancy state.

Fig. 2. State machine of circuit reservation operation at an optical switch.

to cancel the circuit reservations and the circuit set-up is failed.

3) If a circuit reservation is made in each optical switch in the forward direction, either by reserving a link, or by queueing a circuit request, the backward signalling-packet will visit each of the optical switches again and if it finds that a free link is reserved for the circuit it will confirm the reservation, otherwise, it finds the circuit request still in the queue, in which case it will wait until the queued circuit (request) actually obtains a link, then it will confirm the reservation and continue to move on in the backward direction. When it completes all the confirmations in the backward route, the circuit will successfully be set-up.

Notice that there is efficiency gain in this scheme. Consider a reservation made on the forward route by queuing a circuit request which progresses in the queue while the signal travels to the destination and back. This circuit request could potentially be successful in reserving a free link when the signalling packet returns on the backward route. This is especially efficient when the route is long allowing bursts to clear while the signalling packet is on its way. In Figs. 2–4, we present detailed state machine diagrams of these operations.

In [3], we analyzed two Markov chain models for two respective OHS schemes. In one, preemptive priority is given to the circuits, and in the other, no priority is given to the circuits. If we also wish to think about the best way to design a two-way circuit-set-up protocol for each of these two schemes, then for the case of the preemptive priority for circuits, it will make sense to reserve capacity for the circuit with the forward signal in each of the optical switches on the path, but while these reservations are made, to allow bursts in progress to attempt to complete their

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

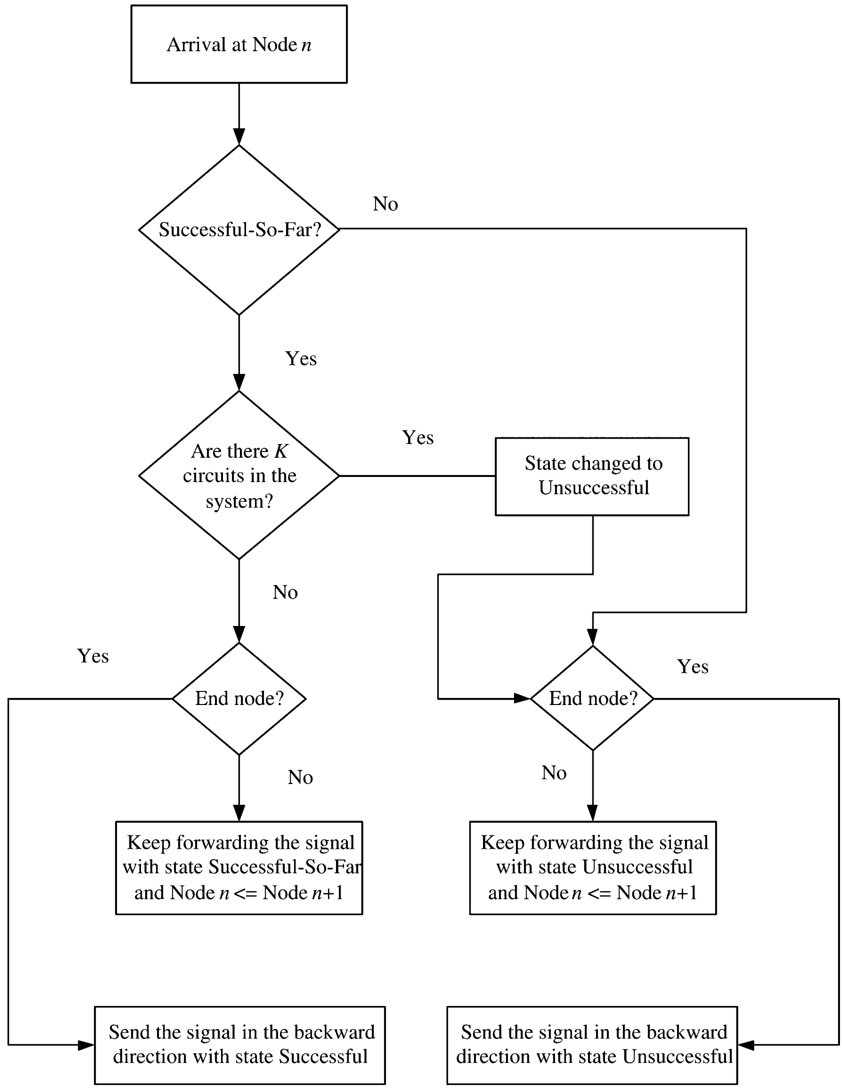


Fig. 3. Forward signal state machine of circuit reservation operation at an optical switch.

transmissions. Preferably, only on the way back, the backward signal will block all relevant burst transmissions and confirm the reservation. This way, some bursts may complete their transmission between the reservation of the forward signal and the confirmation of the backward signal.

On the other hand, the idea that no priority is given to circuits, is against our first principle of giving priority to circuits. It does not seem reasonable to reject a circuit only because there are bursts on its way. And even if we would like to implement such a scheme, there are other implementation concerns. Notice that we will need to reject a circuit if the first forward signal encounters a burst on its intended path (assuming no other links are available), because if we defer it to the backward path and some confirmations already made, a 3-way or 4-way circuit-set-up protocol will be required to cancel these reservations. Therefore, we will not consider here the case of nonpriority for circuits. Accordingly, we will only consider the preemptive priority protocol as a viable alternative to our new protocol that gives nonpreemptive priority to circuits. Henceforth, we will call the former the *preemptive protocol*, and the latter will be called the *nonpreemptive protocol*.

Notice another alternative for the preemptive protocol, where the backward signal does not reject bursts in progress, and the burst is only rejected when the circuit transmission actually commences. This can be done using a header that is transmitted ahead of the circuit data, and it gives the burst additional time to clear out and slightly reduces burst blocking.

In comparison to the preemptive protocol, where the backward signal would cancel burst transmissions and reservations, under the new nonpreemptive protocol, at the cost of small delay in circuit set-up, we avoid the burst losses due to such burst cancelations as well as the potentially false reservations made by burst headers unaware of the loss of their bursts.

There are many factors that can affect the comparison between the preemptive and the nonpreemptive protocols. These factors include the ratio of OCS connections to OBS bursts, the overall traffic load and distribution relative to available capacity and the network topology. Many of these factors are hard to predict. When the number of bursts (relatively to the number of circuits) is not too large, the nonpreemptive approach will have considerable advantage over the preemptive approach, and that the penalty of increased circuit-setup delay

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

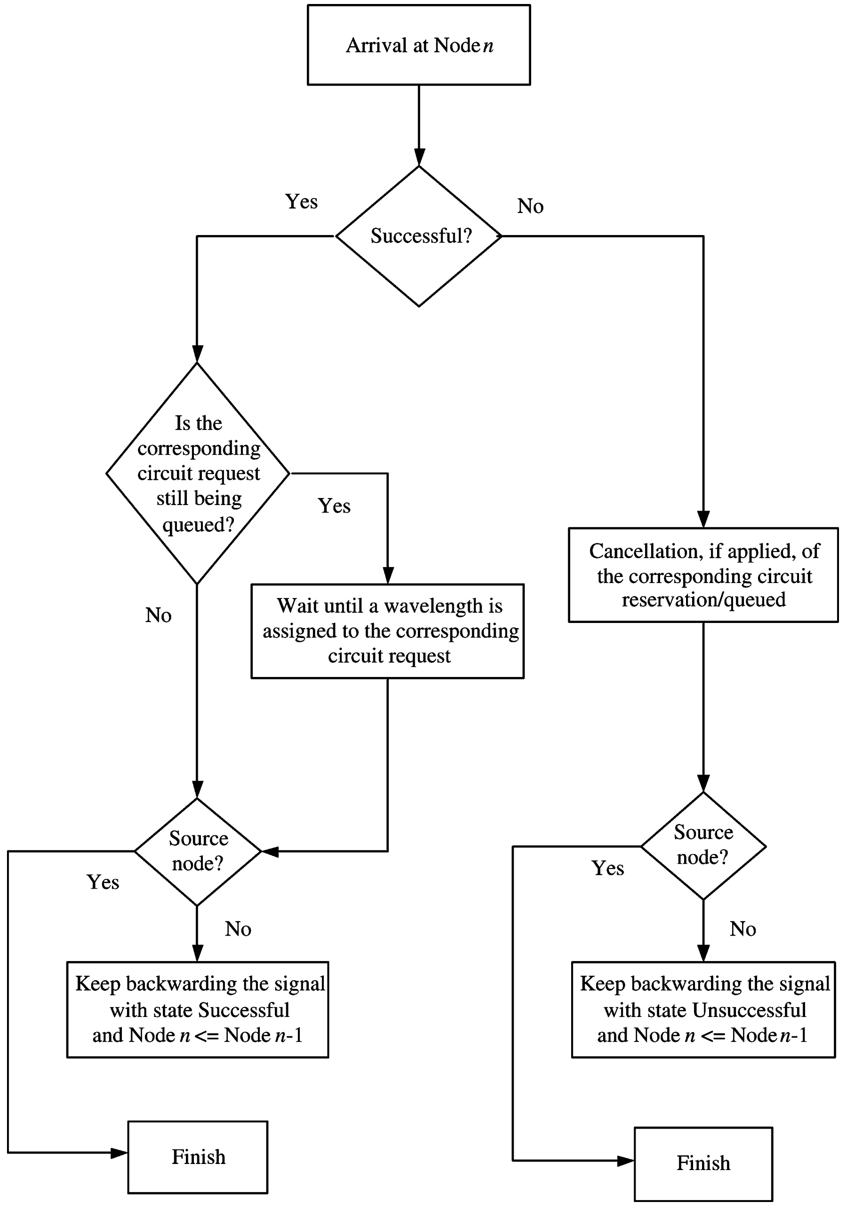


Fig. 4. Backward signal state machine of circuit reservation operation at an optical switch.

induced by the nonpreemptive approach will be acceptable if bursts are reasonably short (compared to circuit holding times). On the other hand, if the average ratio of bursts to circuits is large and if we do not consider a congested network, the benefit of our new nonpreemptive approach is negligible. Note here that the expected number of bursts per circuit and the expected length of bursts versus that of circuits can significantly vary based on applications. If, for example, circuits are used for OBS with acknowledgement [28], or WR-OBS [29], which are basically bursts of data transmitted using two-way reservations, although they may be larger than bursts, they are still considered significantly shorter than other circuits, especially to circuits used for permanent or semi-permanent leased lines which are very long and their number may be very small relative to the number of bursts. It is, therefore, important to have a scalable tool to evaluate the benefit for a wide range of parameters.

Having commented on the benefit, let us consider the cost of our approach. Consider 200 output links and a situation where all bursts are of the same size. Also assume that the burst traffic (in Erlangs) is approximately equal to the circuit traffic. This means that if a circuit is 1000 times longer than a burst, then there are going to be 1000 bursts per circuit generated. Because the traffic volume in Erlangs is assumed to be equal between the bursts and the circuits, then if a circuit set-up signal arrives and finds all output links busy, it is likely that it finds more or less 100 links busy with circuits and 100 with bursts, it will then have to wait until the first burst out of the 100 clears out minus the time the request signal propagates to the destination and back to the congested switch if this difference is positive. Considering that the time until the first burst clears out is approximated at (1/100) times the burst transmission time, and the burst is 1/1000 of a circuit. If we assume for simplicity that the switch we consider is the only bottleneck and there are no other bottlenecks,

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

then the delay due to our method is 0.00001 of the circuit time, which is normally negligible. Notice that the assumption of no other bottlenecks exist does not make a significant difference because while the circuit set-up signal waits in one bottleneck, bursts in other bottlenecks clear out.

III. The Model

In the field of performance evaluation of telecommunications networks and systems, it is common to focus the model on a single bottleneck trunk and study the performance experienced by data packets in that bottleneck trunk. Such studies can provide useful insights on their own (e.g., [34]), or they can serve as performance modules of a full network performance study, possibly based on the Erlang Fixed Point Approximation (e.g., [35] and [36]). Generally speaking, focusing on a single bottleneck trunk is a conservative approach to resource provisioning. If each output trunk in the network is separately and independently dimensioned assuming that the trunk is the bottleneck of all the traffic that may pass through it, we will have a conservatively dimensioned network.

Moreover, the single bottleneck case that we study here represents to certain extent a worst case scenario for the nonpreemptive protocol versus the preemptive protocol because we do not consider multiple nodes per path in which case under the preemptive protocol one circuit can destroy multiple bursts, and their number increases linearly in the number of hops. On the other hand, for the case of the nonpreemptive protocol, a multiple hop path does not increase the mean queueing delay of the reverse signal linearly (in the number of hops) because while the reverse signal waits for one burst to clear out, other bursts may clear out in other nodes.

Here we focus on an output trunk of a bufferless optical switch and also consider all the input links where traffic is transmitted towards that output trunk. Let F be the number of optical fibers on our output trunk, and let W be the number of wavelengths in each optical fiber. In the case of full wavelength conversion, the output trunk is considered to have $C = FW$ links. If wavelength conversion is not available, a burst/circuit that arrives on a given input wavelength must use the same wavelength at the output. In this case, the output trunk is considered to have only F links. If we let L be the number of relevant wavelength channels in our output trunk, then L takes the values of C or F for the cases of with or without wavelength conversion, respectively. Henceforth, when we say *output links* we mean to infer relevant wavelength channels in our output trunk.

Let M be the number of input links from where traffic flows are arriving towards our output trunk. As in [3] and [37], we assume that the switching fabric is strictly nonblocking. Under this assumption, if $M \leq L$ then no loss will occur. Therefore, computing blocking probability is only relevant for the case $L < M$. Fig. 5 shows the optical hybrid switch architecture.

Recall also that since we do not allow the total number of circuit in the system to be more than L, the total number of queued circuits cannot exceed L.

As in [4] and [6], we assume that a burst is transmitted on a wavelength for an exponentially distributed period of time with means $1/\mu_b$. Regarding the circuit holding time, to be precise,

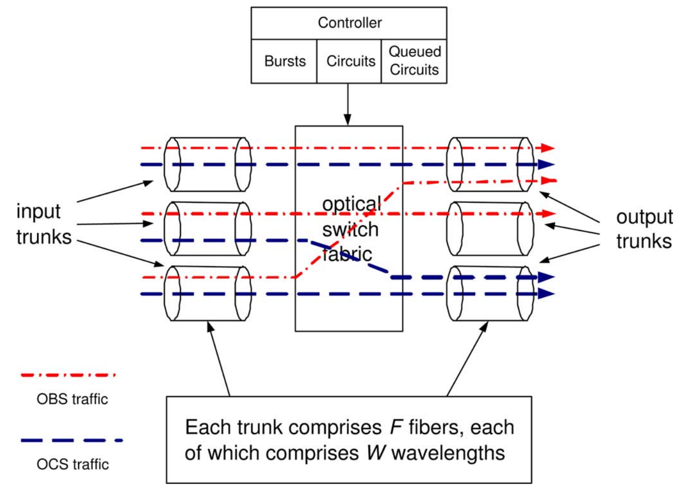


Fig. 5. Optical hybrid switch architecture.

there are two types of circuits—long and short. A *long circuit* corresponds to a successful set-up. Its holding time includes reservation time during set up plus the use of the circuit until its completion. On the other hand, a *short circuit* in a given switch corresponds to a reservation made by a signalling packet on the forward route which has been canceled on the backward route. Accordingly, although both long and short circuits are equally created—by a forward reservation signal, they do not end in the same way. The transmission of a long circuit ends when the holding time is completed, while a short circuit ends when the backward signal cancels its reservation.

Assume that circuit holding times are exponentially distributed with means $1/\mu_c$. In general, the mean length of a circuit or a burst can be adjusted to include any relevant overhead. This way, our model has a sufficient degree of flexibility to apply to various designs (e.g., PATON).

For each input link, we assume that the traffic behaves as an on-off process. This on-off process can be viewed as a process with two alternating states—on and off. A period of time used to transmit a single burst or allocated for a circuit on an input link is called an *on* period, and the time between consecutive on-periods on that input link is called an *off* period. The off-period on an input link is assumed to be exponentially distributed with mean $1/\lambda$. As was argued in [4] and [6], this assumption of exponential times is not as limiting as it may seem. It is well known that Engset formula [39] is insensitive to the on- and the off-distribution [40]. The model we consider here has exponential times, but it has been shown that the burst blocking probability is not too sensitive to the distribution of the on- and the off-periods [3], [17], [41], where there are no circuit queued reservations. We, however, note that the sensitivity of the blocking probability to the on- and the off-distribution is likely to increase with the maximal number of queued circuits.

Upon termination of an off-period on an input link, an on-period associated with a burst transmission will commence with probability p_b, or a request for circuit transmission is generated with probability $p_c = 1 - p_b$. Define $\lambda_b = p_b\lambda$ and $\lambda_c = p_c\lambda$. If a circuit is requested then with probability p_c^s it is a short circuit and with probability $p_c^l = 1 - p_c^s$, it is a long circuit. Here

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

we assume, for simplicity, that p_c^s is given and fixed. However, when this single node model is used in a network model, p_c^s will be a fixed-point parameter that depends on the state of the traffic loading in the network.

Because we have M input links, then at any point in time, we could potentially have a total of M input links, each of which is either idle, or *active*, that is, transmitting/dumping a burst into the switch or allocated for a circuit (long or short). Since we have L output links and we assume $M > L$ (because otherwise we have the trivial case where the blocking probability is equal to zero), then there exist time periods during which the number of active input links is greater than the number of output links. Such time periods are henceforth called *congestion* periods. During congestion periods, the system is said to be *congested*.

In our model, when the system is congested, bursts and circuit requests are treated differently. In particular, during congestion periods bursts are dumped, but circuit requests may be queued if the number of other circuits that are already queued is lower than the number of bursts currently transmitted. Otherwise, the circuit is rejected. The above is applied to both types of circuits—long and short. However, there is a difference between the long and short circuits from the modeling point of view. A short circuit may be terminated while it is waiting in the queue, or holding a link, while a long circuit can only be terminated when it is holding a link.

Having two types of circuits (short or long) and two types of bursts (in progress or dumped) gives rise to a 4-D Markov chain. A reduction to a 3-D state-space is achieved, in our model, by simply ignoring the short circuits (i.e., setting p_c^s to zero).

This can be justified by the following:

1) normally there are several orders of magnitude more bursts than circuits;
2) to provide acceptable QoS, it is reasonable to assume that the number of circuits blocked is smaller by two or more order of magnitudes than the number of circuits arriving;
3) a short circuit is only created when a circuit is blocked;
4) the holding time of a short circuit is several orders of magnitude less than that of a long circuit and is likely to be even smaller than a burst.

Clearly, the effect of the short circuits on burst or circuit blocking probability is negligible.

IV. Exact Solution

Let $\pi_{i,j,k}$ ($i, j, k \geq 0, i \leq L, j \leq L, k \leq M - L$) be the steady state probability that i input links are used for bursts transmission, j are used for circuits (including for waiting circuits), and k are used for dumping blocked bursts. The number of idle input links is, therefore, given by $M - i - j - k$ (≥ 0). The $\pi_{i,j,k}$ values satisfy the following steady state equations. For $i + j < L$

$$\begin{aligned}\pi_{i,j,k}&\left[(i+k)\mu_b + j\mu_c + (M-i-j-k)\lambda\right]\\ &= \pi_{i+1,j,k}(i+1)\mu_b + \pi_{i,j+1,k}(j+1)\mu_c\\ &\quad + \pi_{i,j,k+1}(k+1)\mu_b + \pi_{i-1,j,k}\left[M-(i-1+j+k)\right]\lambda_b\\ &\quad + \pi_{i,j-1,k}\left[M-(i+j-1+k)\right]\lambda_c \end{aligned} \tag{1}$$

and for $i + j \geq L$

$$\begin{aligned}\pi_{i,j,k}&\left[(i+k)\mu_b+(L-i)\mu_c+(M-i-j-k)(\lambda_b+\delta\lambda_c)\right]\\ &= \pi_{i+1,j,k}(i+1)\mu_b+\pi_{i,j+1,k}(L-i)\mu_c\\ &\quad + \pi_{i,j,k+1}(k+1)\mu_b+\pi_{i-1,j,k}\left[M-(i-1+j+k)\right]\gamma\lambda_b\\ &\quad + \pi_{i,j-1,k}\left[M-(i+j-1+k)\right]\lambda_c\\ &\quad + \pi_{i,j,k-1}\left[M-(i+j+k-1)\right]\lambda_b \end{aligned} \tag{2}$$

where $\delta = 1$ if $j < L$ and $\delta = 0$, otherwise, and $\gamma = 1$ if $i - 1 + j < L$ and $\gamma = 0$, otherwise.

Introducing also the normalization equation

$$\sum_{i,j,k} \pi_{i,j,k} = 1$$

gives rise to a complete set of equations from which the $\pi_{i,j,k}$ probabilities can be obtained. In the following, we provide formulae for traffic measures that include *offered traffic load* and *carried traffic load* for the burst traffic and for the circuit traffic. These are measured in units of Erlang and represent the mean number of wavelength channels required to serve the traffic. The carried traffic load represents the mean number of wavelength channels required to serve the traffic excluding the blocked traffic while the offered traffic load is the mean number of wavelength channels required to serve the entire offered traffic including also the blocked traffic. The blocking probability is, therefore, one minus the ratio of the carried traffic load to the offered traffic load. To obtain the blocking probability of the circuits and of the bursts, we will first derive their offered and carried traffic load from which their blocking probability will be readily available.

As in the classical derivation of the Engset formula [39], the circuit offered traffic load, denoted O_c, is the weighted sum of the offered traffic load in each state. The weights in that weighted sum are the steady-state probabilities of being in each state. We, therefore, obtain

$$O_c = \sum_{i,j,k} (M-i-j-k)(\lambda_c/\mu_c)\pi_{i,j,k}.$$

There are two ways to derive the circuit carried traffic load, denoted C_c. One way is to consider it at the output. Recalling that j is the total number of circuits both in service and in queue and the i is the number of bursts in service. If $L \geq i+j$, or $j \leq L-i$, then there are no queued circuits so j is also the number of circuits in service, but if $L < i + j$, or $j > L - i$, then there are $i+j-L$ circuits in the queue and $j-(i+j-L) = L-i$ circuits in service. In conclusion, in any state (i, j, k) the circuit carried traffic is $\min(j, L-i)$. Therefore, to obtain C_c, we consider the weighted sum on the circuit carried traffic in each state and we obtain

$$C_c = \sum_{i,j,k} \min(j, L-i)\pi_{i,j,k}. \tag{3}$$

An alternative way to derive the circuit carried traffic load is to consider the input by considering the states at which input traffic is not blocked. Namely, C_c is equal to the weighted sum of all

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

nonblocking states. This gives

$$C_c = \sum_{i,j<L,k} (M-i-j-k)(\lambda_c/\mu_c)\pi_{i,j,k}. \quad (4)$$

To confirm the consistency of (3) and (4), we consider our 3-D Markov chain and divide its states into a comprehensive and mutually exclusive sets of $L+1$ "superstates". The jth superstate is associated with a given number of circuits in the system j and it comprises all states (i,j,k) for all i and k. Observe that there are only transitions from superstate j to superstate $j+1$ (for all $j, L-1 \geq j \geq 0$), or to superstate $j-1$ (for all $j, L \geq j \geq 1$). Therefore, considering the conservation of flows [42], and summing up all the probability flux associated with transitions from j to $j-1$ for $j = 1,2,\cdots L$, and equate them with the sum of all the probability flux associated with transitions from j to $j+1$ for $j = 0,2,\cdots L-1$, and divide both sides of the resulting equation by μ_c, we obtain

$$\sum_{i,j<L,k} (M-i-j-k)(\lambda_c/\mu_c)\pi_{i,j,k} = \sum_{i,j,k} \min(j, L-i)\pi_{i,j,k}.$$

Note that in the derivation of the circuit carried traffic load, we did not consider the queued circuits. By only considering traffic that is actually carried we avoid double counting. Notice that queued circuits will eventually become carried circuits so they will be counted.

Having obtained Q_c and C_c, the circuit blocking probability B_c is given by

$$B_c = 1 - \frac{C_c}{O_c}.$$

Following similar arguments, the burst offered traffic load O_b is given by

$$O_b = \sum_{i,j,k} (M-i-j-k)(\lambda_b/\mu_b)\pi_{i,j,k}$$

and the burst carried traffic load C_b is

$$C_b = \sum_{i,j,k} i\pi_{i,j,k}$$

or

$$C_b = \sum_{i+j<L,k} (M-i-j-k)(\lambda_b/\mu_b)\pi_{i,j,k}.$$

Thus, the burst blocking probability B_b is evaluated by

$$B_b = 1 - \frac{C_b}{O_b}.$$

The utilization of the output trunk is given by

$$U = \frac{1}{L} \sum_{\forall i,j,k} \min(i+j, L)\pi_{i,j,k}. \quad (5)$$

The circuit carried traffic load comprises two parts. One part is traffic that is never queued and as it is generated, it immediately finds a free wavelength channel and the second part, denoted C_q, is traffic that is queued before a free wavelength channel becomes available to it. Derivation of C_q is useful in deriving the proportion of circuits that are delayed and the mean circuit queueing delay and will be derived in the following. Notice that the conditions for a circuit to be queued require that all L wavelength channels are busy, but not all used by circuits (in which case the circuit is blocked). Specifically, these conditions are $j+i \geq L$ and $j \neq L$. Therefore, we obtain

$$C_q = \sum_{\forall j<L, i+j \geq L} (M-i-j-k)(\lambda_c/\mu_c)\pi_{i,j,k}.$$

The proportion of circuits that are delayed, denoted as D_p, is

$$D_p = \frac{C_q}{C_c}.$$

The average number of circuits in the queue, denoted as N_{qc}, is given by

$$N_{qc} = \sum_{\forall i+j>L,k} (i+j-L)\pi_{i,j,k}.$$

Let us define a system made of the circuit queue. This system may have $0,1,2,\ldots B$ circuits. To obtain the mean circuit queueing delay, we will use Little's formula. To this end, we will need to know the net arrival rate (excluding the blocked circuits) into this system which is equal to its output rate. The output rate, denoted μ_{qc}, is given by the following sum:

$$\mu_{qc} = \sum_{\forall i+j>L,k} [i\mu_b + (L-i)\mu_c]\,\pi_{i,j,k}. \quad (6)$$

Notice that μ_{qc} is the output rate of the circuit queue and not the service rate as explained in the following. In (6), we weight all possible service rates by their relative probabilities. Each of these service rates is associated with an event where an item (a circuit or a burst) terminates its service at the switch, and immediately a queued circuit leaves our system of queued circuits. During the time in which the circuit queue is empty, the output rate is equal to zero, so the events associated with empty circuit queue are not included in the sum of (6). If one is interested in the service rate of the circuit queued then μ_{qc} will need to be normalized by the sum $\sum_{\forall i+j>L,k} \pi_{i,j,k}$.

Therefore, by Little's formula, the mean delay in the circuit queue, which is also the mean queueing delay of a circuit, is given by

$$Q_c = \frac{N_{qc}}{\mu_{qc}}.$$

Because our queued circuit system includes also the state zero when there are no circuits queued, Q_c is the mean queueing delay of all circuits including those that do not suffer any queueing delay. Therefore, the mean queueing delay for a delayed circuit, denoted as Q_{qc}, is given by

$$Q_{qc} = \frac{Q_c}{D_p}.$$

Solving (1) and (2) is not scalable for large L and M, so scalable approximations are derived next.

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

V. Approximation

The approximation consists of two separate and independent stages. The first stage yields the blocking probability and state distribution for circuits. By state distribution, it is meant the set of probabilities $\{p_j : j = 0 \ldots, L\}$, where p_j is the probability that j circuits are in progress at steady-state. The second stage approximates the burst blocking probability by conditioning on the state distribution for circuits computed in the first stage.

The reason that we can decouple the circuit and the burst processes is that by the nature of our protocol, if bursts are assumed to be much shorter than circuits, the circuits are not significantly affected by the bursts. As mentioned above, circuits are not blocked as a direct result of existence of bursts in the system. The only effect bursts have on circuits is a slight delay assuming circuits are much longer than bursts. This could potentially indirectly increase circuit blocking because it increases circuit delay which effectively increases circuit holding time which, in turn, increases circuit blocking probability. However, as we demonstrate later (see Fig. 6), it turns out that our approximation based on decoupling the burst and circuit processes is still reasonably accurate even if the mean length of a burst is equal to that of a circuit. Therefore, from the point of view of circuit blocking probability, circuits can be assumed to have preemption priority over bursts, so their blocking probability can be evaluated as if there are no bursts in the system.

However, before we consider a system completely empty of bursts, certain adjustments are required to be made for the circuit arrival process to consider times at which circuit cannot arrive on an input wavelength channel because that channel is either busy by an arriving burst or by a burst that is being dumped. From the point of view of the circuit traffic, there is no difference between the two. In both cases the input link appears *active* from a point of view of a circuit request.

The first stage of the approximation makes use of the fact that from a circuit viewpoint, an input link is either active or inactive. We do not distinguish between an active input link that transmits a successful burst or a dumped one. We do distinguish, however, between input links used for circuits and those used for bursts.

As in the well-known Engset model, during the period that an input link is active, neither another burst nor another circuit can use this input link. On an inactive input link, a burst arrives with probability λ_b/λ, while a circuit arrives with probability λ_c/λ. As the first part of the approximation focus on the circuits, it is important to capture accurately the effect of burst traffic. In particular, the effect of burst traffic is viewed here as being a part of the off-periods on the circuits in a system as if it is only made of circuits. To this end, the effect of burst arrivals is taken into account by increasing the mean off-period between two successive circuits. Let the modified mean off-period between two circuits be $1/\lambda'$, which as in [3] is given by

$$1/\lambda' = (\lambda_c/\lambda)(1/\lambda) + (\lambda_b/\lambda)(1/\lambda + 1/\mu_b + 1/\lambda') \tag{7}$$

or

$$1/\lambda' = 1/\lambda + (\lambda_b/\lambda_c)(1/\lambda + 1/\mu_b).$$

The first term $1/\lambda$ in (7) is the mean off-period given that the next arrival is a circuit, which occurs with probability λ_c/λ, while the second term $1/\lambda+1/\mu_b+1/\lambda'$ is the mean off-period given that the next arrival is a burst, which occurs with probability λ_b/λ. Thus, the circuit blocking probability is given by

$$\mathrm{Eng}(\lambda', \mu_c, M, L) \triangleq \frac{\binom{M-1}{L}(\lambda'/\mu_c)^L}{\sum_{i=0}^{L}\binom{M-1}{i}(\lambda'/\mu_c)^i}$$

which is the standard Engset formula, and the state distribution is given by

$$p_j = \frac{\binom{M}{j}(\lambda'/\mu_c)^j}{\sum_{i=0}^{L}\binom{M}{i}(\lambda'/\mu_c)^i}, \quad j = 0, \ldots, L.$$

The second stage involves approximating the burst blocking probability. To evaluate the burst blocking probability, we note the difference between the present case of the nonpreemptive protocol and the case of the preemptive protocol. In the case of preemption priority to circuits, bursts can be lost in two ways. Either they are blocked on arrival if the system is full (no free wavelength channel is available), or they are preempted by an arriving circuit. In the present case, we do not have bursts being preempted, but they are still blocked if the system is full when they arrive. This means that while for circuit blocking probability evaluation, it is appropriate to use the preemptive priority model, it is not appropriate for the bursts.

Since bursts are not preempted in the present case, and assuming that circuits are significantly longer than bursts, accurate burst blocking probability can be obtained using the so-called "quasi-stationary" approach. Under this approach, burst blocking probability is evaluated by conditioning on the state distribution $\{p_j : j = 0 \ldots, L\}$. Namely, with probability p_j, an arriving burst finds j output links busy with circuits and if it finds less than $L-j$ bursts in the system, it will not be blocked. To consider the fact that the number of input links is limited, the burst blocking probability is computed given that j $(= 0, \ldots, L)$ circuits are in progress using the approximation based on the Engset formula with modified off-period as given by [3] (10). That is, the approximation given in [3] (10) is applied $L+1$ times to compute the burst blocking probability given j $(= 0, \ldots, L)$ circuits are in progress.

In particular, let $P_{\mathrm{blocked}}(j)$, $j = 0, \ldots, L$, be the probability that an incoming burst on an input link is blocked given j circuits are in progress. Furthermore, let $1/\lambda^*(j)$, $j = 0, \ldots, L$, be the modified mean off-period between two successive bursts given j circuits are in progress. From [3], we have

$$1/\lambda^*(j) = (1-P_{\mathrm{blocked}}(j))\,(1/\lambda_b) + P_{\mathrm{blocked}}(j)(1/\mu_b + 1/\lambda_b). \tag{8}$$

An alternative formula to evaluate $\lambda^*(j)$ appears in [8] in the context of a Generalized Engset model that applies to a pure OBS system and gives rise to another approximation of $\lambda^*(j)$ for the present problem. Using (6) in [8], we have

$$1/\lambda^*(j) = (1-P_{\mathrm{blocked}}(j))\,(1/\lambda_b) + P_{\mathrm{blocked}}(j)\,(1/\mu_b + 1/\lambda_b + 1/\lambda^*(j)) \tag{9}$$

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

or

$$1/\lambda^*(j) = \frac{(\mu_b + \lambda_b P_{\text{blocked}}(j))}{\lambda_b \mu_b (1 - P_{\text{blocked}}(j))}. \tag{10}$$

Equation (9) is similar to (8) except that $1/\lambda^*(j)$ has been added to the second term. This modification allows for the possibility that further "dumped burst" periods may follow the initial dumped burst period, whereas (8) relies on the restrictive assumption that a blocked burst is always followed by a successful burst.

Given that j circuits are in the system, an arriving burst is only blocked if there is a total of $L - j$ bursts in progress. Therefore

$$P_{\text{blocked}}(j) = \text{Eng}\left(\lambda^*(j), \mu_b, M - j, L - j\right). \tag{11}$$

In addition, conditioning on having j circuits in the system, the distribution of having i successful bursts in the system being transmitted is given by

$$p_{i,j} = \frac{\binom{M-j}{i} \left(\lambda^*(j)/\mu_b\right)^i}{\sum_{k=0}^{L-j} \binom{M-j}{k} \left(\lambda^*(j)/\mu_b\right)^k}, \quad i = 0, \ldots, L - j.$$

The functional relation between $P_{\text{blocked}}(j)$ of (11) and $1/\lambda^*(j)$ of (8) or (9) gives rise to a fixed-point equation. The fixed-point, i.e., consistent values for $P_{\text{blocked}}(j)$ and $1/\lambda^*(j)$, can be computed by the same successive substitution algorithm defined in [3]. The successive substitution algorithm is applied $L + 1$ times to compute $P_{\text{blocked}}(j)$, $j = 0, \ldots, L$. The burst blocking probability is then approximated by un-conditioning on j to give $\sum_{j=0}^{L} p_j P_{\text{blocked}}(j)$. However, as noted in [8], the successive substitution algorithm in [3] cannot guarantee convergence to a unique solution.

In the following, we will provide a proof of existence and uniqueness of a fixed-point solution of (9). Note that the other case for the solution of (8) is very similar and we do not present it here for brevity. A specific algorithm that solves that fixed-point equation(s) together with a proof of its convergence to the unique solution is provided in the Appendix. The proof of existence and uniqueness of a fixed-point solution of (9) can be achieved by replacing Π, m, λ, μ, N and K of (1) and (6) in [8] by $P_{\text{blocked}}(j)$, $1/\lambda^*(j)$, λ_b, μ_b, $M - j$ and $L - j$, respectively. In addition, according to Section D in [8], the following theorem follows.

Theorem 1: A unique solution exists for the set of coupled equations defined by (11) together with (9).

Having the steady state probabilities $p_{i,j}$'s, we can now derive the utilization of the output trunk, the proportion of circuits that are delayed and the mean queueing delay for a delayed circuit.

The utilization of the output trunk is given by

$$U = \frac{1}{L} \sum_{\forall i,j} \min(i + j, L) p_j p_{i,j}. \tag{12}$$

The carried load of circuits in the queue is

$$C_q = \sum_{\forall j<L, i+j \geq L} (M - j)(\lambda'/\mu_c) p_j p_{i,j}.$$

Therefore, the proportion of circuits that are delayed, denoted as D_p, is

$$D_p = \frac{C_q}{C_c}$$

where $C_c = \sum_{j>0} j p_j$. Assume that the distributions of circuits in the system and bursts in progress are independent. Then, the probability that there are i bursts in progress, denoted by p_i^*, is given by

$$p_i^* = \sum_{\forall j} p_{i,j} p_j.$$

Therefore, the average number of circuits in the queue, denoted as N_{qc}, is

$$N_{qc} = \sum_{\forall i+j>L} (i + j - L) p_j p_i^*.$$

The average output rate of the circuit queue, denoted as μ_{qc}, is given by

$$\mu_{qc} = \sum_{\forall i+j>L} \left[i\mu_b + (L - i)\mu_c\right] p_j p_i^*.$$

By Little's formula, the mean circuit queueing delay is given by

$$Q_c = \frac{N_{qc}}{\mu_{qc}}$$

and the mean queueing delay for a delayed circuit, denoted Q_{qc}, is

$$Q_{qc} = \frac{Q_c}{D_p}.$$

VI. Erlang B With Nonpreemptive Priorities

In this section, we consider the above approximation for the case where the number of input links M approaches infinity. In this case, the mean circuit/burst off-period is very long compared to the burst mean dumping/transmission period, and, therefore, the effect of the latter on the modified off period can be ignored. In the first stage, (7) becomes

$$\lambda' = \lambda_c. \tag{13}$$

The circuit blocking probability is given by the Erlang B formula as follows:

$$\text{Erlang}(\lambda', \mu_c, L) \triangleq \frac{\frac{1}{L!}(\lambda_c/\mu_c)^L}{\sum_{i=0}^{L} \frac{1}{i!}(\lambda_c/\mu_c)^i}$$

and the state distribution is given by

$$p_j = \frac{\frac{1}{j!}(\lambda_c/\mu_c)^j}{\sum_{i=0}^{L} \frac{1}{i!}(\lambda_c/\mu_c)^i}.$$

In the second stage, (8) and (9) become

$$\lambda^*(j) = \lambda_b.$$

Equation (11) becomes

$$\begin{aligned} P_{\text{blocked}}(j) &= \text{Erlang}\left(\lambda^*(j), \mu_b, L - j\right) \\ &= \frac{\frac{1}{(L-j)!}(\lambda_b/\mu_b)^{L-j}}{\sum_{i=0}^{L-j} \frac{1}{i!}(\lambda_b/\mu_b)^i}. \end{aligned} \tag{14}$$

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

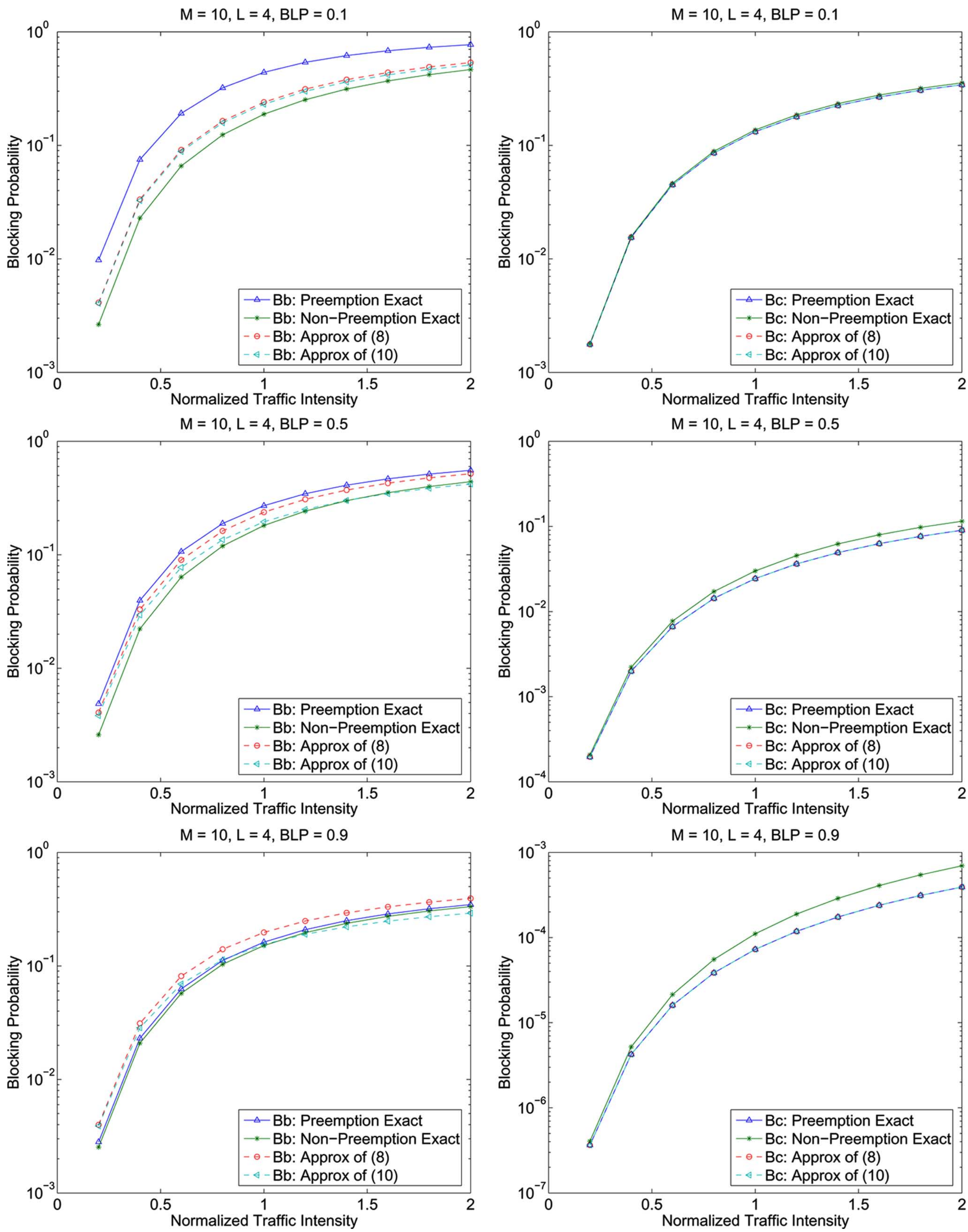


Fig. 6. Burst blocking probability (left) and circuit blocking probability (right) versus normalized combined traffic intensity varying burst loading proportion (BLP $= 0.1$, 0.5, and 0.9) for $\mu_b/\mu_c = 1$.

In addition, the corresponding steady state distribution is given by

$$p_{i,j} = \frac{\frac{1}{(i)!}(\lambda_b/\mu_b)^i}{\sum_{k=0}^{L-j}\frac{1}{k!}(\lambda_b/\mu_b)^k}, \quad i = 0, \ldots, L-j.$$

Note that, in this case, a fixed-point procedure is not required to obtain $P_{\mathrm{blocked}}(j)$ as it can be obtained directly by the Erlang B formula given by (14).

VII. Numerical Results and Discussions

This section has two aims. First, we use the exact solution for the steady state probabilities for the case of nonpreemp-

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

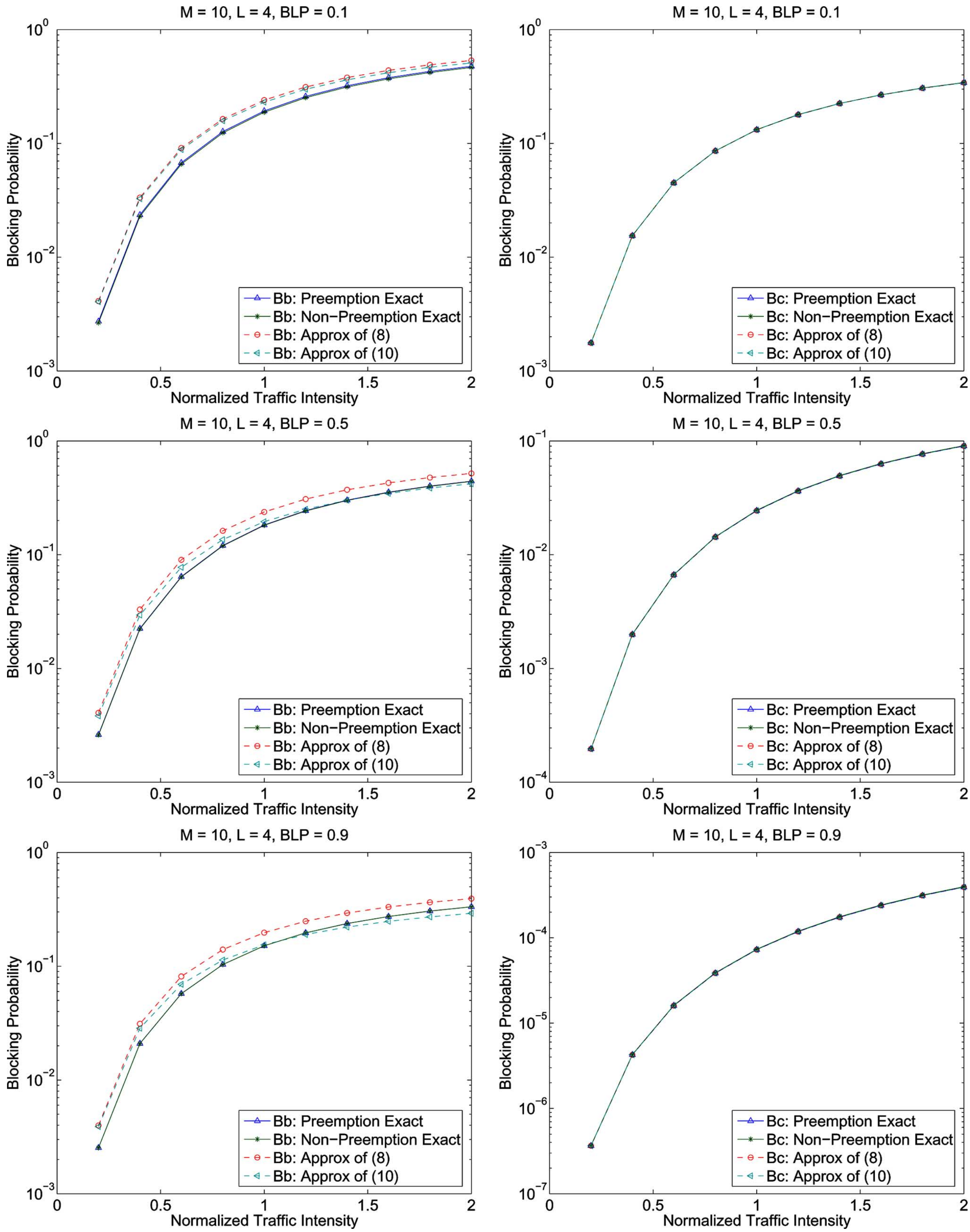


Fig. 7. Burst blocking probability (left) and circuit blocking probability (right) versus normalized combined traffic intensity varying burst loading proportion (BLP $= 0.1$, 0.5, and 0.9) for $\mu_b/\mu_c = 100$

tive circuit priority obtained by (1) and (2) to verify the accuracy of our approximations. Second, the numerical results enable us to draw conclusions on blocking probability and delay which provide prediction of link utilization. It is important, for example, to compare the expected link utilization for a given required performance for the case of having full wavelength conversion versus having no wavelength conversion. For comparison, we include the exact solution for preemptive circuit priority obtained from [3]. Results are presented here for the blocking probability and delay versus what we call the

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

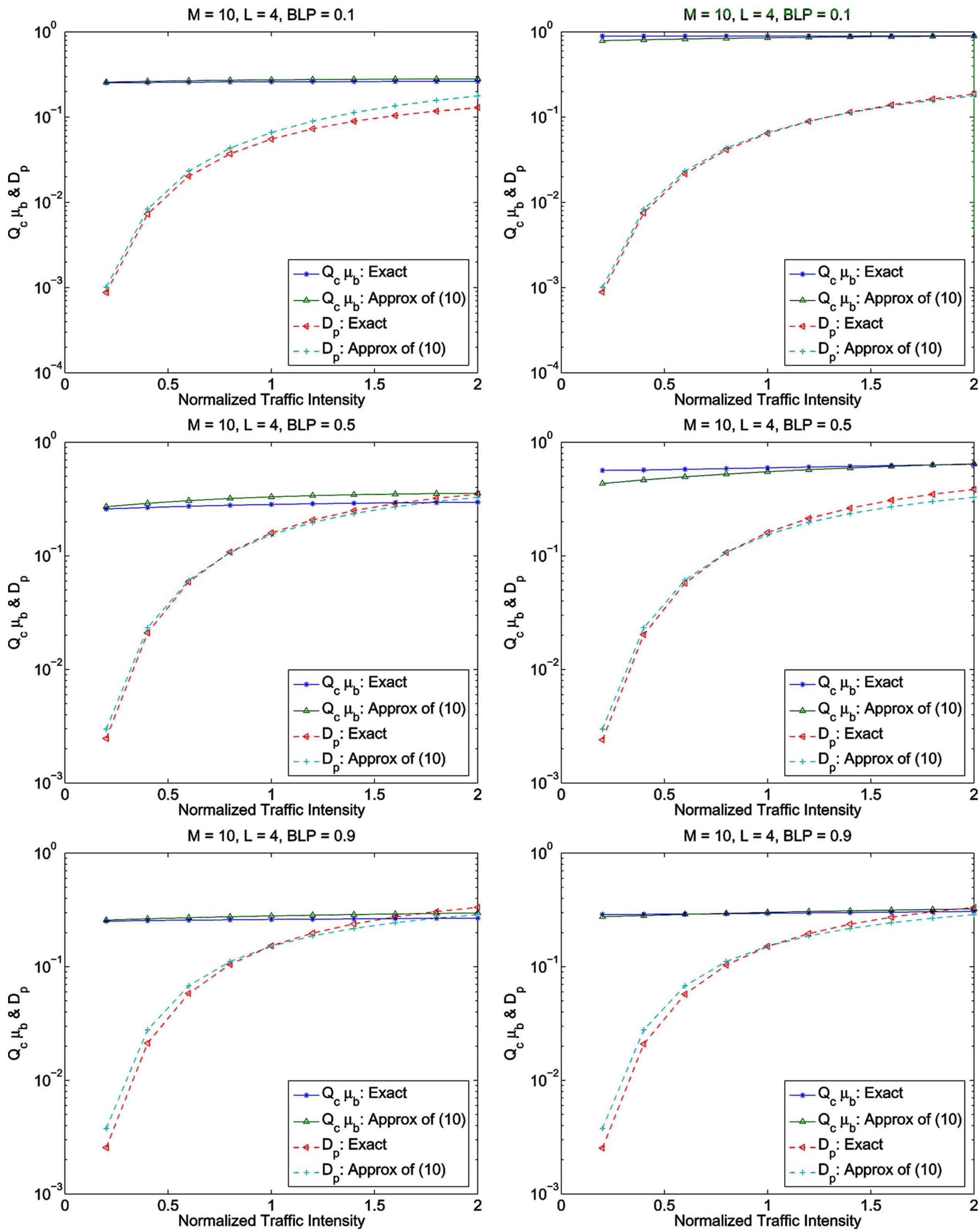


Fig. 8. Circuit queueing delay as a fraction of the mean burst transmission time $(Q_c\mu_b)$ and the proportion of circuits that are delayed (D_p) versus normalized combined traffic intensity varying burst loading proportion (BLP $= 0.1$, 0.5 and 0.9) for $\mu_b/\mu_c = 1$ (left) and $\mu_b/\mu_c = 100$ (right), respectively.

normalized combined traffic intensity, defined by

$$\frac{M}{L}\left(\frac{\lambda_b}{\mu_b}+\frac{\lambda_c}{\mu_c}\right).$$

The burst loading proportion (BLP) is defined as

$$\mathrm{BLP} = \frac{\frac{\lambda_b}{\mu_b}}{\frac{\lambda_b}{\mu_b}+\frac{\lambda_c}{\mu_c}}.$$

In all our examples, we consider $1/\mu_b = 1$ s.

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

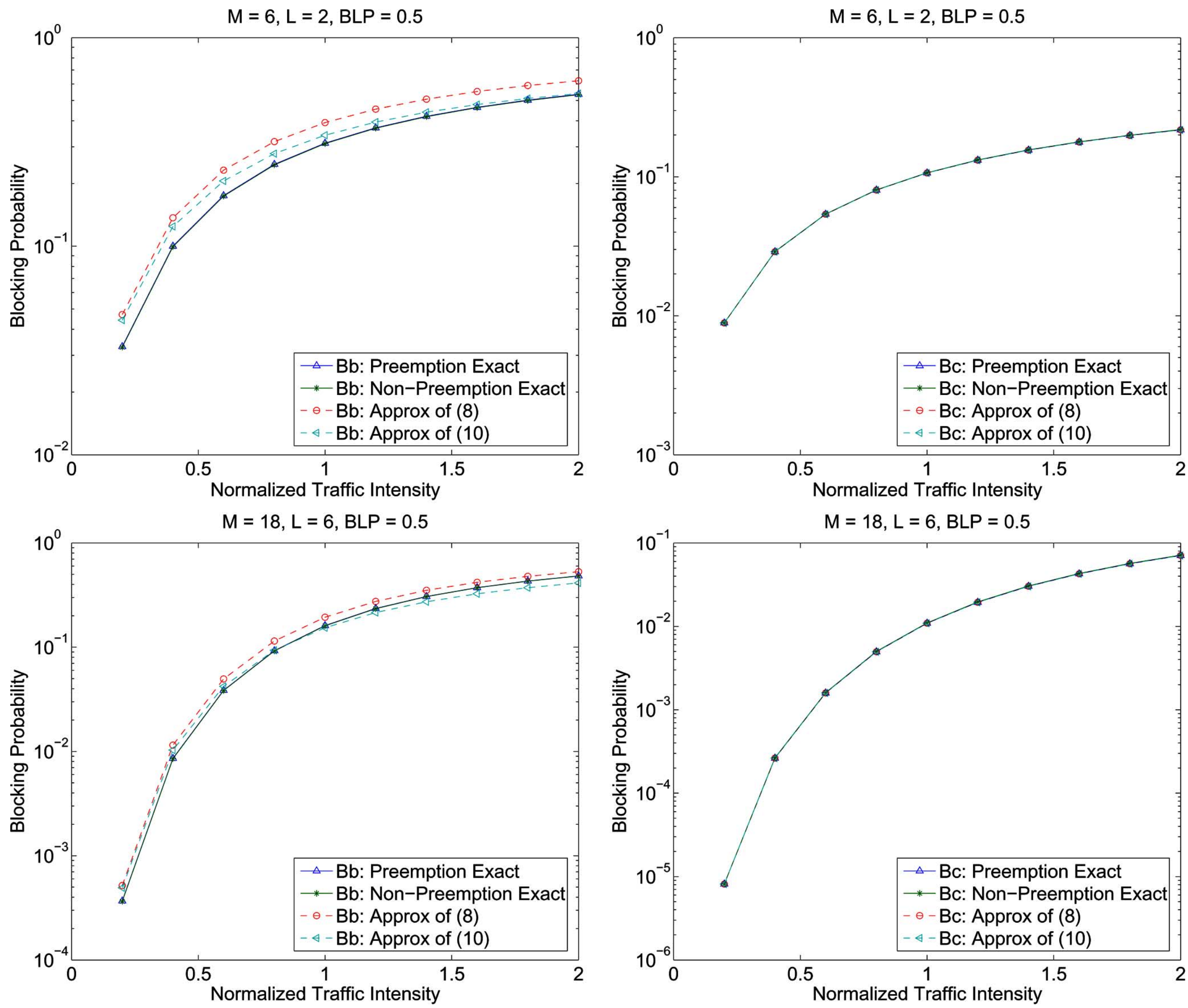


Fig. 9. Burst blocking probability (left) and circuit blocking probability (right) versus normalized combined traffic intensity varying M and L for $M/L = 3$, BLP $= 0.5$, and $\mu_b/\mu_c = 100$.

In all scenarios studied, regardless of the values of M, L, BLP, the traffic intensity, and the ratio of μ_b to μ_c, our numerical results show that the approximations, in general, agree reasonably well with the exact solutions as demonstrated, for example, in Figs. 6–10. The agreements demonstrated in Figs. 6–10 have also been observed in many other cases we considered; however, for brevity, we do not present them here.

Fig. 6 shows the burst blocking probability (left) and the circuit blocking probability (right) versus normalized combined traffic intensity varying burst loading proportion (BLP $= 0.1$, 0.5, and 0.9) for $\mu_b/\mu_c = 1$. As we can see, for small BLP the burst blocking probabilities are quite different between the exact solutions of circuit preemption and circuit nonpreemption protocols. However, this difference reduces as BLP increases. As expected, the circuits do not affect burst blocking probability significantly if the burst traffic dominates. In addition, the approximation of (10) performs better than the approximation of (8). It is interesting to observe that the burst blocking probability of the approximation of (10) is closer to that of circuit nonpreemption than to that of circuit preemption. It is because the "quasi-stationary" approach does not consider burst loss due to preemption as discussed above. Concerning the circuit blocking probabilities, they are close to each other for all four curves except for the case BLP $= 0.9$. In this case, the burst traffic dominate. As mentioned before, the effect bursts have on circuits is the increase of circuit blocking probability because the bursts increase circuit delay which increases circuit effective holding time which, in turn, increase circuit blocking probability. However, this effect is not captured by the approximations of both (10) and (8) as well as the exact solution of circuit preemption protocol. It is also interesting to notice that, in many cases, the approach of approximation of [3] intended for the preemptive protocol is more suitable for the present nonpreemptive protocol than for the preemptive protocol for which it was originally designed.

Fig. 7 shows the burst blocking probability (left) and the circuit blocking probability (right) versus normalized combined traffic intensity varying burst loading proportion (BLP $= 0.1$, 0.5, and 0.9) for $\mu_b/\mu_c = 100$. Here we focus our attention on effects different to those in the previous figure. First, the

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

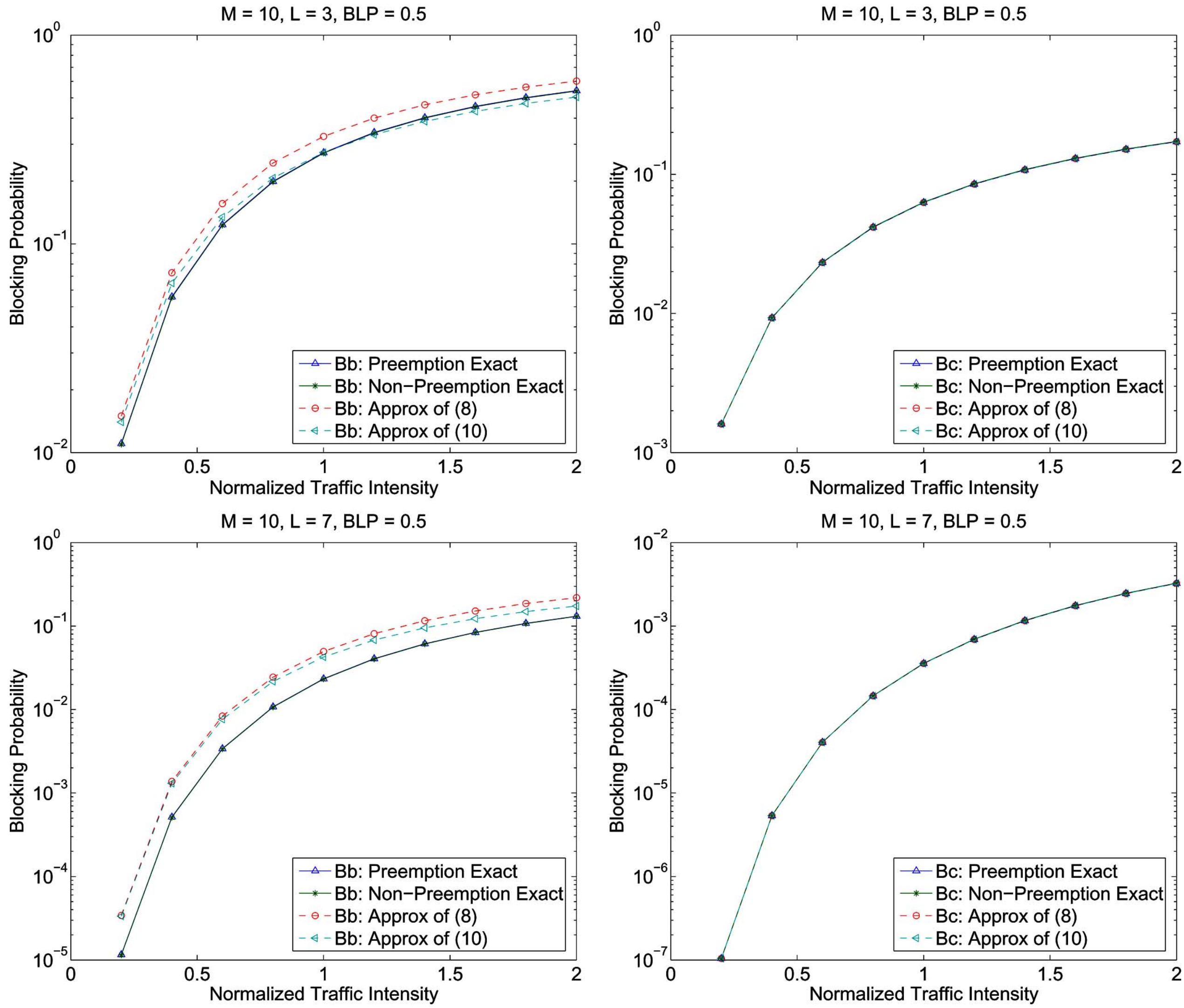


Fig. 10. Burst blocking probability (left) and circuit blocking probability (right) versus normalized combined traffic intensity varying L for $M = 10$, BLP $=$ 0.5, and $\mu_b/\mu_c = 100$.

burst blocking probabilities are very much the same between the exact solutions of circuit preemption and circuit nonpreemption protocols. This is because a circuit is now much longer (100 times longer) than a burst. This means that from the bursts viewpoint, circuits rarely arrive and hence choosing between circuit preemption versus circuit nonpreemption protocols makes no much difference. Again, we can see that the approximation of (10) performs better than the approximation of (8). Concerning the circuit blocking probability, the circuit blocking probabilities for all four curves are close to each other for all three BLP cases. As explained before, it is because, for the case that a circuit is much longer than a burst, the only effect bursts have on circuits in the nonpreemptive protocol is a slight delay which effectively increases circuit holding time, thus, slightly increases circuit blocking probability. Therefore, there is no much difference in circuit blocking probability between the nonpreemptive and the preemptive protocols.

Fig. 8 shows circuit queueing delay as a fraction of the mean burst transmission time ($Q_c\mu_b$) and the proportion of circuits that are delayed (D_p) versus normalized combined traffic intensity varying burst loading proportion (BLP $=$ 0.1, 0.5, and 0.9) for $\mu_b/\mu_c = 1$ (left) and $\mu_b/\mu_c = 100$ (right), respectively. Regarding the delay, the approximation of (10) performs quite well in terms of the proportion of circuits that are delayed and the mean queueing delay of the circuits. Note that the mean queueing delay of the circuits is always less than a single mean burst transmission time as expected given the exponential burst distribution (recall that the mean residual burst transmission time is equal to its mean) and the fact that the circuit will only wait for its "own burst(s)" to clear.

Fig. 9 shows the burst blocking probability (left) and the circuit blocking probability (right) versus normalized combined traffic intensity varying M and L for $M/L = 3$, BLP $= 0.5$, and $\mu_b/\mu_c = 100$. Again, the approximation of (10) performs well and better than the approximation of (8). Note that in terms of burst blocking probability, the approximation of (10) performs better for large M and L. This is a desirable feature because realistic systems usually have large M and L.

Fig. 10 shows the burst blocking probability (left) and the circuit blocking probability (right) versus normalized combined traffic intensity varying L for $M = 10$, BLP $=$ 0.5, and $\mu_b/\mu_c = 100$. Again, the approximation of (10) performs

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

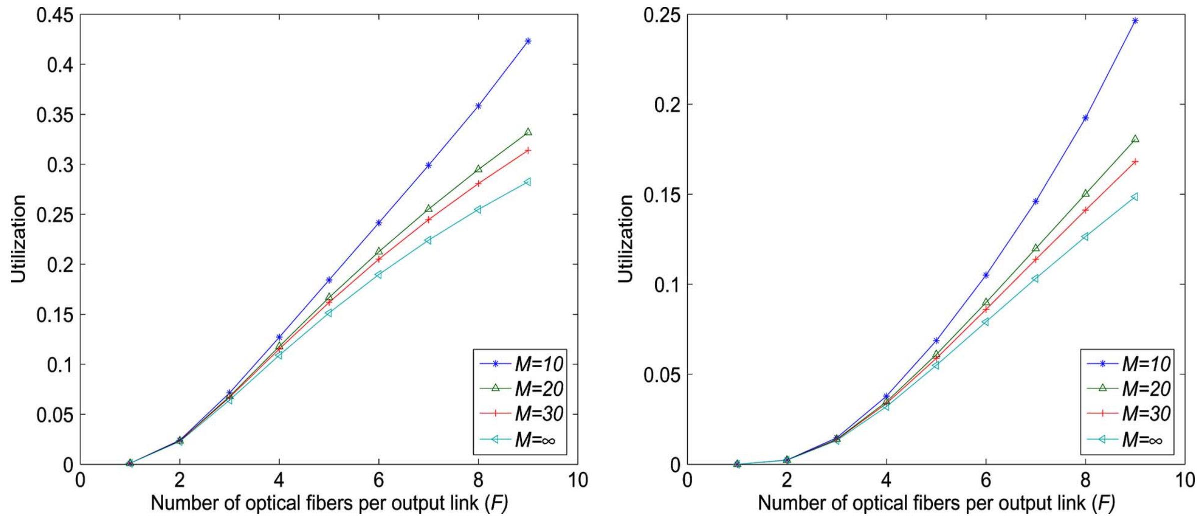


Fig. 11. Utilization versus the number of optical fibers per output trunk without wavelength conversion for BLP $= 0.5$, as well as $B_b = 10^{-3}$ (left) and $B_b = 10^{-5}$ (right).

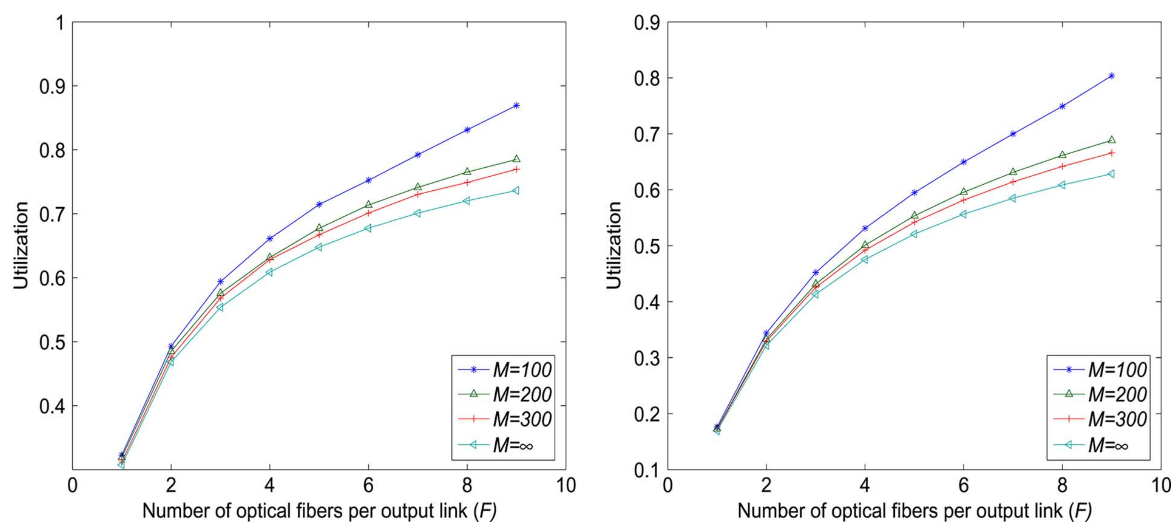


Fig. 12. Utilization versus the number of optical fibers per output trunk with wavelength conversion for BLP $= 0.5$, as well as $B_b = 10^{-3}$ (left) and $B_b = 10^{-5}$ (right).

well and better than the approximation of (8). In particular, we observe that in term of burst blocking probability the approximation of (10) performs better if M is much larger than L.

It is known that wavelength conversion can help considerably to increase utilization. Let us now demonstrate this fact by comparing between the cases of full wavelength conversion and no wavelength conversion. Consider, for example, the case of $F = 5$ and $W = 10$, that is, each trunk carries five optical fibers with 10 wavelengths per fiber, altogether each trunk carries $5 \times 10 = 50$ wavelengths. Therefore, for the case of full wavelength conversion we will have $L = 50$ and an incoming burst on any input wavelength has the flexibility to choose any of the 50 wavelengths. For the case of no wavelength conversion we will have $L = 5$ because an incoming burst on a given input wavelength can choose only the same wavelength, and there are five such wavelengths, one on each of the five fibers of the output trunk. The number of the relevant input wavelengths is bounded above by the number of input ports times the relevant L.

Then, we consider the case of no wavelength conversion (i.e., $L = F$) and $W = 10$ and a range of M values and F. For each case, using our results for burst blocking probability, we find the maximal λ value such that the burst blocking probability does not exceed 10^{-3} and 10^{-5}, respectively. Then we compute the utilization (of the output trunk) based on (12). For each M we plot the utilization as a function of F in Fig. 11. When the burst blocking probability requirement becomes more stringent and changes from 10^{-3} to 10^{-5}, the utilization decreases correspondingly, as expected. In addition, we can see that the results presented approach the $M = \infty$ curve as M increases. This gives us additional confidence in our approximation as well as a conservative bound for the utilization.

In Fig. 12, we present equivalent results presented in Fig. 11, but here we consider the case of full wavelength conversion (i.e., $L = 10F$), $W = 10$ and BLP $= 0.5$ with $B_b = 10^{-3}$ and $B_b = 10^{-5}$, respectively. As expected, the utilization is much higher. Again, we can see that the results presented approach the $M = \infty$ curve as M increases.

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

VIII. Conclusion

We have considered an optical hybrid switch and proposed a new implementation whereby circuits have nonpreemptive priority over bursts. We presented an analysis based on a 3-D Markov chain that provides exact results and computationally scalable approximations for the blocking probabilities of bursts and circuits, the proportion of circuits that are delayed and the mean delay of the circuits that are delayed. Extensive numerical results have demonstrated the accuracy of the approximations. Accordingly, they can be used as part of a dimensioning tool for practical scenarios involving hundreds or even thousands of wavelength channels per trunk where neither exact solution is possible, nor simulation results can give accurate results. Applying the results of this paper to each bottleneck gives a conservative methodology for buffer and trunk dimensioning in an OHS network. Our approximations have been also used to provide qualitative results to evaluate utilization and burst blocking probability tradeoff for cases of with and without wavelength conversion.

Appendix
Binary Search Algorithm to Find the Unique Solution

By setting $x = 1/\lambda^*(j)$ in (9) and moving the left-hand side of (9) to the right, we define the function

$$f(x) = \frac{1}{\lambda_b} + P_{\text{blocked}}(j,x)\left(\frac{1}{\mu_b} + x\right) - x, \quad x \geq 0 \qquad (15)$$

where we have written $P_{\text{blocked}}(j,x)$ instead of $P_{\text{blocked}}(j)$ to emphasize that $P_{\text{blocked}}(j)$ is functionally dependent through (11) on the mean OFF time, x.

Let x^* be the unique solution of $f(x) = 0$. We define the monotonically decreasing transformation $\Gamma(x) \longrightarrow x$ such that

$$\Gamma(x) = \frac{\mu_b + \lambda_b P_{\text{blocked}}(j,x)}{\left(1 - P_{\text{blocked}}(j,x)\right)\mu_b \lambda_b}.$$

We show that the binary search algorithm specified in Algorithm 1 finds the unique solution of $f(x) = 0$ for an absolute error criterion of ϵ.

Algorithm 1 Calculate Solution of $f(x) = 0$ for an Absolute Error Criterion of ϵ

1: $x^- = (1/\lambda_b)$, $x^+ = \Gamma(1/\lambda_b)$ Initial lower/upper bounds
2: **while** $x^+ - x^- > \epsilon$ **do**
3: $\quad x = (x^+ + x^-)/2$ Halve the search interval
4: $\quad$ **if** $\Gamma(x) > x$ **then**
5: $\quad\quad x^- = x$ Tighten lower bound
6: $\quad$ **else**
7: $\quad\quad x^+ = x$ Tighten upper bound
8: $\quad$ **end if**
9: **end while**
10: **return** $x = (x^+ + x^-)/2$ ϵ satisfied, thus return x

Due to the monotonicity of $\Gamma(x)$, at each iteration of Algorithm 1, if $x < x^*$, then $\Gamma(x) > x^*$ and, thus, $\Gamma(x) > x^* > x$. Conversely, if $x > x^*$, then $\Gamma(x) < x^*$ and, thus, $\Gamma(x) < x^* < x$. Consequently, x^* lies in the interval $[x^-, x^+]$ at each iteration of Algorithm 1. Furthermore, this interval halves at each iteration, thereby ensuring x^* is sandwiched within an interval whose eventual length does not exceed ϵ. Note that lines 4, 5, and 7 in Algorithm 1 are different from the corresponding lines in the algorithm of [8] and the differences here guarantee convergence.

Acknowledgment

The authors would like to thank the anonymous reviewers for thorough reviews and valuable suggestions.

References

[1] C. Xin, C. Qiao, Y. Ye, and S. Dixit, "A hybrid optical switching approach," in *Proc. IEEE GLOBECOM*, San Francisco, CA, Dec. 2003, vol. 7, pp. 3808–3812.
[2] G. M. Lee, B. Wydrowski, M. Zukerman, J. K. Choi, and C. H. Foh, "Performance evaluation of an optical hybrid switching system," in *Proc. IEEE GLOBECOM*, San Francisco, CA, Dec. 2003, vol. 5, pp. 2508–2512.
[3] H. L. Vu, A. Zalesky, E. W. M. Wong, Z. Rosberg, S. M. H. Bilgrami, M. Zukerman, and R. S. Tucker, "Scalable performance evaluation of a hybrid optical switch," *J. Lightw. Technol.*, vol. 23, no. 10, pp. 2961–2973, Oct. 2005.
[4] E. W. M. Wong and M. Zukerman, "Performance evaluation for an optical hybrid switch with circuit queued reservations," *Opt. Exp.*, vol. 13, no. 23, pp. 9446–9459, Nov. 2005.
[5] E. W. M. Wong and M. Zukerman, "Analysis of an optical hybrid switch," *IEEE Commun. Lett.*, vol. 10, no. 2, pp. 108–110, Feb. 2006.
[6] E. W. M. Wong and M. Zukerman, "Performance evaluation of an optical hybrid switch with circuit queued reservations and circuit priority preemption," *Opt. Exp.*, vol. 14, no. 23, pp. 11043–11070, Nov. 2006.
[7] E. W. M. Wong and M. Zukerman, "Bandwidth and buffer tradeoffs in optical packet switching," *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4790–4798, Dec. 2006.
[8] E. W. M. Wong, A. Zalesky, and M. Zukerman, "On generalizations of the Engset model," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 360–362, Apr. 2007.
[9] K. J. Barker *et al.*, "On the feasibility of optical circuit switching for high performance computing systems," in *Proc. ACM/IEEE Supercomputing (SC)*, Seattle, WA, Nov. 2005, p. 16.
[10] Z. Rosberg, A. Zalesky, and M. Zukerman, "Packet delay in optical circuit-switched networks," *IEEE/ACM Trans. Netw.*, vol. 14, no. 2, pp. 341–354, Apr. 2006.
[11] A. Ghafoor, M. Guizani, and S. Sheikh, "Architecture of an all-optical circuit-switched multistage interconnection network," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 8, pp. 1595–1607, Oct. 1990.
[12] W. S.-Y. Yuen, A. Chen, A. K.-S. Wong, and C.-T. Lea, "Buffer-efficient routing strategies for optical circuit switching networks," in *Proc. BroadNets*, San Jose, CA, Oct. 2004, pp. 328–330.
[13] C. Salisbury and R. Melhem, "Multicast control in optical circuit-switched Banyan networks," in *Proc. 5th Int. Conf. Massively Parallel Processing*, Las Vegas, NV, Jun. 1998, pp. 85–93.
[14] C. Qiao and M. Yoo, "Optical burst switching (OBS): A new paradigm for an optical internet," *J. High Speed Netw.*, vol. 8, no. 1, pp. 69–84, Jan. 1999.
[15] M. Yoo, C. Qiao, and S. Dixit, "Optical burst switching for service differentiation in the next-generation optical internet," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 98–104, Feb. 2001.
[16] T. Battestilli and H. Perros, "An introduction to optical burst switching," *IEEE Commun. Mag.*, vol. 41, no. 8, pp. S10–S15, Aug. 2003.
[17] A. Detti, V. Eramo, and M. Listanti, "Performance evaluation of a new technique for IP support in a WDM optical network: Optical composite burst switching (OCBS)," *J. Lightw. Technol.*, vol. 20, no. 2, pp. 154–165, Feb. 2002.
[18] Y. Chen, C. Qiao, and X. Yu, "Optical burst switching (OBS): A new area in optical networking research," *IEEE Netw. Mag.*, vol. 18, no. 3, pp. 16–23, May/Jun. 2004.

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.

[19] J. Turner, "Terabit burst switching," *J. High Speed Netw.*, vol. 8, no. 1, pp. 3–16, Mar. 1999.
[20] S. Verma, H. Chaskar, and R. Ravikanth, "Optical burst switching: A viable solution for terabit IP backbone," *IEEE Netw. Mag.*, vol. 14, no. 6, pp. 48–53, Nov./Dec. 2000.
[21] M. Yoo and C. Qiao, "Just-enough-time (JET): A high speed protocol for bursty traffic in optical networks," in *Proc. IEEE/LEOS Conf. Technologies For a Global Information Infrastructure*, Montreal, QC, Canada, Aug. 1997, pp. 26–27.
[22] M. Yoo and C. Qiao, "Supporting multiple classes of services in IP over WDM networks," in *Proc. IEEE GLOBECOM*, San Francisco, CA, Dec. 1999, pp. 1023–1027.
[23] K. Dolzer, C. Gauger, J. Späth, and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching," *AEÜ Int. J. Electron. Commun.*, vol. 55, no. 1, pp. 1–8, Jan. 2001.
[24] H. L. Vu and M. Zukerman, "Blocking Probability for Priority Classes in Optical Burst Switching Networks," *IEEE Commun. Lett.*, vol. 6, no. 5, pp. 214–216, May 2002.
[25] I. Baldine, G. N. Rouskas, H. G. Perros, and D. Stevenson, "Jump-Start: A just-in-time signaling architecture for WDM burst-switched networks," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 82–89, Feb. 2002.
[26] N. Barakat and E. H. Sargent, "Analytical modeling of offset-induced priority in multiclass OBS networks," *IEEE Trans. Commun.*, vol. 53, no. 8, pp. 1343–1352, Aug. 2005.
[27] H. Overby, "Performance modelling of optical packet switched networks with the Engset traffic model," *Opt. Exp.*, vol. 13, no. 5, pp. 1685–1695, Mar. 2005.
[28] A. Zalesky, E. W. M. Wong, M. Zukerman, H. L. Vu, and R. S. Tucker, "Performance analysis of an OBS edge router," *IEEE Photon. Technol. Lett.*, vol. 16, no. 2, pp. 695–697, Feb. 2004.
[29] M. Duser and P. Bayvel, "Analysis of a dynamically wavelength routed optical burst switched network architecture," *J. Lightw. Technol.*, vol. 20, no. 4, pp. 574–585, Apr. 2002.
[30] J. White, M. Zukerman, and H. L. Vu, "A framework for optical burst switching network design," *IEEE Commun. Lett.*, vol. 6, no. 6, pp. 268–270, Jun. 2002.
[31] A. Zalesky, H. L. Vu, Z. Rosberg, E. W. M. Wong, and M. Zukerman, "OBS contention resolution performance," *Perf. Eval.*, vol. 64, no. 4, pp. 357–373, May 2007.
[32] A. Zalesky, H. L. Vu, Z. Rosberg, E. W. M. Wong, and M. Zukerman, "Stabilizing deflection routing in optical burst switched networks," *IEEE J. Sel. Areas Commun., Suppl. Opt. Commun. Netw.*, vol. 25, no. 6, pp. 3–19, Aug. 2007.
[33] C. Qiao, W. Wei, and X. Liu, "Extending generalized multiprotocol label switching (GMPLS) for polymorphous, agile, and transparent optical networks (PATON)," *IEEE Commun. Mag.*, vol. 44, no. 12, pp. 104–114, Dec. 2006.
[34] G. L. Choudhury, D. M. Lucantoni, and W. Whitt, "Squeezing the most out of ATM," *IEEE Trans. Commun.*, vol. 44, no. 2, pp. 203–217, Feb. 1996.
[35] Z. Rosberg, H. L. Vu, M. Zukerman, and J. White, "Performance analyses of optical burst-switching networks," *J. Sel. Areas Commun.*, vol. 21, no. 7, pp. 1187–1197, Sep. 2003.
[36] Z. Rosberg, A. Zalesky, H. L. Vu, and M. Zukerman, "Analysis of OBS networks with limited wavelength conversion," *IEEE/ACM Trans. Netw.*, vol. 14, no. 5, pp. 1118–1127, Oct. 2006.
[37] N. Akar, Y. Gunalay, N. Akar, and Y. Gunalay, "Stochastic analysis of finite population bufferless multiplexing in optical packet/burst switching systems," *IEICE Trans. Commun.*, vol. E90-B, no. 2, pp. 342–345, 2007.
[38] N. Barakat and E. H. Sargent, "Dual-header optical burst switching: A new architecture for WDM burst-switched networks," presented at the IEEE INFOCOM, Miami, FL, Mar. 2005.
[39] T. Engset, "Die wahrscheinlichkeitsrechnung zur bestimmung der wahleranzahl in automatischen fernsprechamtern," *Elektrotechnische Zeitschrift*, vol. 39, no. 31, pp. 304–306, Aug. 1918.
[40] J. Hui, *Switching, and Traffic Theory for Integrated Broadband Networks*. Norwell, MA: Kluwer, 1990.
[41] M. Zukerman, E. W. M. Wong, Z. Rosberg, G. M. Lee, and H. L. Vu, "On teletraffic application to OBS," *IEEE Commun. Lett.*, vol. 8, no. 2, pp. 116–118, Feb. 2004.
[42] L. Kleinrock, *Queueing Systems, Volume 1—Theory*. New York: Wiley, 1975.
[43] A. Leon-Garcia, R. H. Kwong, and G. F. Williams, "Performance evaluation methods for an integrated voice/data link," *IEEE Trans. Commun.*, vol. 30, no. 8, pp. 1848–1858, Aug. 1982.
[44] M. Zukerman, "Circuit allocation and overload control in a hybrid switching system," *Comput. Netw. ISDN Syst.*, vol. 16, no. 4, pp. 281–298, March 1989.

Eric W. M. Wong (S'87–M'90–SM'00) received the B.Sc. and M.Phil. degrees in electronic engineering from the Chinese University of Hong Kong, Hong Kong, in 1988 and 1990, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Massachusetts, Amherst, in 1994.

In 1994, he joined the City University of Hong Kong, where he is now an Associate Professor with the Department of Electronic Engineering. His research interests include the analysis and design of telecommunications networks and video-on-demand systems. His most notable research work involved the first workable model for state dependent dynamic routing. Since 1991, the model has been used by AT&T to design and dimension its telephone network that uses real-time network routing.

Moshe Zukerman (M'87–SM'91–F'07) received the B.Sc. degree in industrial engineering and management and the M.Sc. degree in operations research from The Technion—Israel Institute of Technology, Haifa, Israel, and the Ph.D. degree in electrical engineering from the University of California, Los Angeles, in 1985.

He was an independent Consultant with the IRI Corporation and a Postdoctoral Fellow with the University of California, Los Angeles, in 1985–1986. In 1986–1997, he was with the Telstra Research Laboratories (TRL), first as a Research Engineer and, in 1988–1997, as a Project Leader. He also taught and supervised graduate students at Monash University in 1990–2001. During 1997–2008, he was with the University of Melbourne, Victoria, Australia. In 2008, he joined the City University of Hong Kong where he is a Professor (Chair) and a group leader. He has over 200 publications in scientific journals and conference proceedings.

Dr. Zukerman has served on various editorial boards such as *Computer Networks*, the *IEEE Communications Magazine*, the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS, the IEEE TRANSACTIONS ON NETWORKING, and the *International Journal of Communication Systems*.

Authorized licensed use limited to: CityU. Downloaded on February 17, 2009 at 10:36 from IEEE Xplore. Restrictions apply.