Stabilizing Deflection Routing in Optical Burst Switched Networks

Andrew Zalesky, Hai Le Vu, Zvi Rosberg, Moshe Zukerman and Eric W. M. Wong

***Abstract*— This paper studies the blocking performance of optical burst switching (OBS) networks using a sequential office control (SOC) state-independent deflection routing policy. We show that unprotected deflection routing may destabilize OBS, resulting in higher blocking probabilities than if bursts were not deflected but simply blocked. This study was motivated by the well-known destabilizing effect that alternative routing has on circuit switching in classical telephony networks. We propose two forms of protection to guard against destabilization: 1) wavelength reservation, which is analogous to trunk reservation in circuit switching; and, 2) preemptive priority, which is a new form of protection where bursts that have not been deflected are given preemptive priority over bursts that have been deflected. Our main contribution is a one-moment reduced-load approximation to evaluate the blocking performance of OBS networks using deflection routing protected by either wavelength reservation or preemptive priority. Our reduced-load approximation relies on the usual assumptions of link independence and Poisson distributed link arrivals. We quantify the error admitted in making these two assumptions via simulation. Using our reduced-load approximation, we evaluate the blocking performance of protected and unprotected deflection routing in several randomly generated networks. The chief conclusion of our study is that deflection routing in OBS should be given some form of protection to avoid destabilization resulting from upward load variations, and in terms of blocking performance, preemptive priority is the best form of protection for OBS. Our reduced-load approximation may be used as a fast approach to provision capacity or evaluate the blocking performance of large OBS networks using deflection routing.**

***Index Terms*— Optical burst switching, deflection routing, stability, reduced-load approximation, wavelength reservation, preemptive priority.**

I. INTRODUCTION

DEFLECTION routing has featured prominently in the literature covering optical burst switching (OBS) over the last four to five years. However, in all of this literature, it has been tacitly assumed that deflection routing does not destabilize OBS in the same way as it is well-known to destabilize circuit switching in conventional telephony networks. This destabilizing effect may result in higher blocking probabilities than if bursts were not deflected but simply blocked.

Most of the seminal literature [27], [28], [29], [35], [39] describes the workings of OBS in detail as well as recent work [2], [4], [9], [14], we therefore only give a brief description. OBS has many traits in common with tell-and-go switching [33], [36], [42] in ATM networks as well as modern-day optical packet switching [5].

The basic switching entity in OBS is a burst. A burst is train of packets that is transmitted from a source to a destination via an all-optical route that may traverse several intermediate nodes[1]. Associated with each burst is a header. The key feature distinguishing OBS from optical packet switching is that a burst is separated from its header by an offset time. An offset time eliminates the need to optically buffer a burst during the time required to process its header at each intermediate node.

The term *node* may refer to any of an intermediate node, a source node or a destination node. Any pair of nodes may be interconnected via a link, which consists of several fibers aligned in the same direction, each of which in turn contain many wavelength channels.

At its source node, a burst that intends traversing N links, or equivalently, $N+1$ nodes (source node and destination node inclusive), must be separated from its header by an offset time of at least $N\delta$, where δ is the time required for a node to process a header. Since a header encounters a delay δ at each intermediate node as well as its destination node, its offset time is incrementally reduced by δ. More precisely, at node $n = 1, \dots, N+1$, a burst is separated from its header by an offset time of at least $(N-n+1)\delta$, where the '+1' appears because n is an index beginning at 1. Therefore, at its destination, a burst catches-up to its header and they are no longer separated. A timing diagram of a burst and its header is shown in Fig. 1, which was originally presented in [28].

As soon as a header arrives at node $n = 1, \dots, N+1$, it seeks to reserve an appropriate outgoing wavelength for a time interval that begins $(N+2-n)\delta$ into the future, which is the time at which its associated burst is expected to arrive and is referred to as the residual offset time. At a given node, residual offset time may vary from header-to-header, since several different routes, corresponding to different source and destination pairs, may traverse that node. This gives rise to the need for burst scheduling algorithms [24], [37], [43] to efficiently allocate bursts to the so-called voids that lie within the fragmented bandwidth of a wavelength.

A. Zalesky, H. L. Vu and M. Zukerman are with the ARC Special Research Centre for Ultra-Broadband Information Networks (CUBIN), Department of Electrical and Electronic Engineering, The University of Melbourne, Vic. 3010, Australia; email: {a.zalesky,h.vu,m.zukerman}@ee.unimelb.edu.au.

CUBIN is an affiliated program of National ICT Australia.

Z. Rosberg is with the Department of Communication Systems Engineering, Ben Gurion University, Beer-Sheva, 84105, Israel; email: rosberg@bgumail.bgu.ac.il.

E. W. M. Wong is with the Optoelectronics Research Centre, Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China; email: e.wong@cityu.edu.hk.

An earlier version of this paper was presented in part at INFOCOM'04. This paper differs from its conference version chiefly by considering preemption as an alternative method of stabilizing deflection routing.

[1]We should really be referring to intermediate nodes as optical cross-connects.

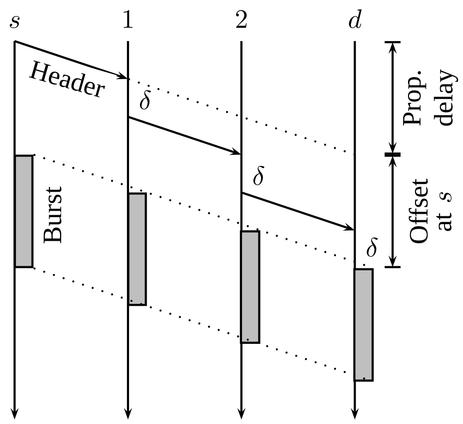


Fig. 1. Timing diagram for a burst traversing two intermediate nodes from it source s to its destination d

Even with the most efficient burst scheduling algorithms, a burst may be blocked at an intermediate node in the case that two or more headers seek to reserve overlapping time intervals on the same wavelength. With native OBS, in this case, one of the contending bursts must be blocked and subsequently retransmitted.

High blocking probabilities are probably one of the biggest technical stumbling blocks that OBS must overcome before it considered a commercially viable technology. To reduce blocking probabilities, numerous approaches of resolving wavelength contention have been proposed. These include: burst segmentation [38]; deflection routing; fiber delay lines to delay a burst that would otherwise be blocked [28]; wavelength conversion to allow for relaxation of the wavelength continuity constraint [31]; and, state-of-the-art scheduling algorithms [24]. Some of these approaches are often considered impractical as they mandate the use of costly optical technology such as fiber delay lines and wavelength converters.

In this paper, we consider deflection routing. Deflection routing in the context of OBS has received a lot of attention recently. In [10], [17], the presence of deflection routing in a single node was modeled by a multidimensional Markov process. Blocking probabilities were computed by numerically solving the associated local balance equations. In [22], [40], simulations were used to evaluate the performance of deflection routing in OBS networks. Some of these studies claim that using particular deflection routing policies may reduce blocking probabilities by more than one order of magnitude. Efforts have also been devoted to dynamically optimizing deflection routes based on network state information [23]. Several approaches of resolving wavelength contention, including deflection routing, have been compared in terms of blocking probabilities via simulation studies [15], [44].

It is well-known that deflection routing may destabilize circuit switching in conventional telephony networks [1], [16] as well as optical packet switched networks [6]. Instabilities associated with deflection routing may manifest simply as a sudden downturn in utilization that is instigated by a minimal load increase or as a complex set of equilibria between which a network fluctuates. They can be intuitively explained in terms of unstable positive feedback. In particular, since first-choice routes and deflection routes may use common links, a deflection from one first-choice route may trigger a spate of subsequent deflections from other first choice routes, each of which in turn may trigger further deflections.

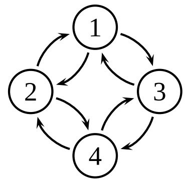


Fig. 2. Four-node ring network

We are interested in determining if deflection routing may also destabilize OBS. This issue has been glossed over in most of the recent literature treating deflection routing in OBS [10], [15], [22], [23], [40]. Although OBS is in many ways different to circuit switching as well as optical packet switching, it does not seem unreasonable to suspect that deflection routing may destabilize OBS. As a matter of fact, intuition does suggest that this is indeed the case, since there is no reason indicating that unstable positive feedback instigated by a deflection is somehow mitigated in OBS.

To give credence to this intuition, we simulated a form of OBS in the four-node ring network shown in Fig. 2. (The form of OBS as well as the deflection routing policy we consider in this paper will be described in the next section.) It was assumed bursts arrive according to independent Poisson processes with the same rate at each source and destination pair for which there is a one-hop first-choice route. A three-hop deflection route for each of these source and destination pairs is thus uniquely determined. To preserve symmetry, it was further assumed that all other source and destination pairs are not used.

Using this simulation, we plot blocking probability and carried load as a function of offered load in Fig. 3. The abrupt downturn in carried load evident in Fig. 3 is highly undesirable and definitely suggests that instabilities may be present. Furthermore, the downturn occurs over a range of blocking probabilities (10^{-3} to 10^{-2}) that can be considered quite realistic in the context of OBS. This result prompts further study and will lead us to develop a new tractable methodology to evaluate the performance of OBS networks using deflection routing.

Two different approaches have been used to protect circuit switching and optical packet switching against destabilization. To protect circuit switching, calls that have been deflected are barred from engaging an idle trunk on any trunk group for which the total number of busy trunks on that trunk group exceeds a predefined threshold. This approach is referred to as trunk reservation [1], [16] and is a form of admission control that intentionally limits the amount of deflection. One drawback of trunk reservation is the lack of rigorous criteria to determine the reservation threshold. See [19] for details.

To protect optical packet switching, several approaches have been suggested, all of which are based on the idea of using fiber delay lines in a recirculating delay loop setup to delay a packet that would otherwise be deflected. These approaches have been found especially useful in stabilizing asynchronous

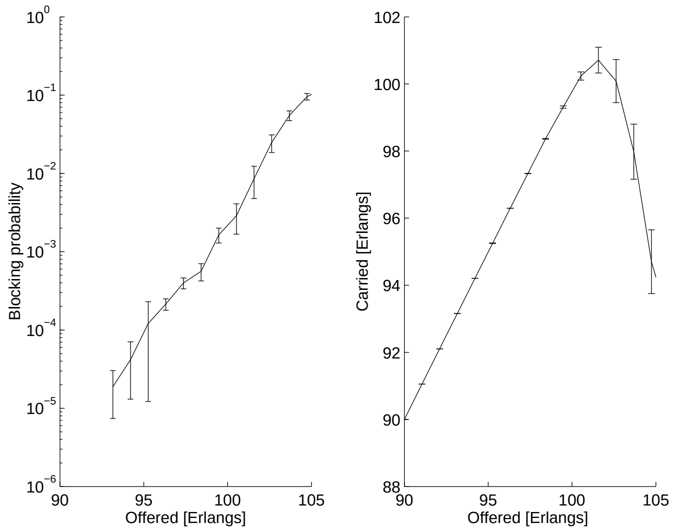


Fig. 3. Blocking probability and carried load as a function of offered load

(un-slotted) optical packet switching and have been compared in [6]. Further protection can be added by purging packets exceeding a certain hop-count.

In principle, it seems both these approaches may also be used to protect OBS, though approaches relying on fiber delay lines would probably be ruled out at the outset by many due to technological barriers. In this paper, we propose and evaluate the performance of a new approach to protect OBS networks against destabilization. This approach is based on enforcing preemptive priority between first-choice bursts and deflected bursts, where a first-choice burst is defined as burst that has not been deflected and a deflected burst is defined complementarily. With this approach, a header associated with a first-choice burst is given the right to preempt a reservation (overlapping time interval) that has been scheduled for a deflected burst. Preemption is always a last resort in the sense that a header associated with a first-choice burst always seeks to reserve a time interval without resorting to preemption.

Preemptive priority is unsuitable for circuit switching in telephony networks since it is unacceptable from a quality of service point of view to preempt a call that is in progress. This would obviously be perceived by users as an unexpected call termination. However with OBS, a burst that is preempted suffers the same fate as a burst that is blocked at an intermediate node. We discuss this point in greater detail in Section III.

We first considered preemptive priority in [7] in the context of a hot-potato routing policy. In this paper, we develop a new reduced-load approximation to evaluate the performance of OBS networks that have been stabilized with either wavelength reservation or preemptive priority. Wavelength reservation is analogous to trunk reservation in circuit switching. Using our approximation, we empirically show that preemptive priority consistently yields lower blocking probabilities than wavelength reservation. We also argue that preemptive priority is guaranteed to stabilize deflection routing, whereas the stabilizing properties of trunk reservation are highly dependent on the choice of reservation threshold.

In Section II, we discuss the form of OBS considered in this paper and define a simple deflection routing policy. In Section III, we confirm the downturn in carried load evident in Fig. 3 is indeed a result of destabilization. We then show that either wavelength reservation or preemptive priority correct this downturn. In Section IV, we present our reduced-load approximation. In Section V, our reduced-load approximation is used to evaluate the blocking performance of unprotected and protected deflection routing in several randomly generated networks.

II. A Deflection Routing Policy for OBS

In this paper, we consider a form of OBS called dual-header OBS. The greatest advantage of dual-header OBS is that the residual offset time at each intermediate node does not vary from header-to-header. This greatly simplifies the complexity of scheduling algorithms. Further details regarding dual-header OBS can be found in [3].

The reason we consider dual-header OBS is chiefly because it is difficult to accurately model native forms of OBS, since with native OBS, residual offset time may vary from header-to-header at each intermediate node. Therefore, this leads to the unsolved problem of calculating blocking probabilities in a finite server queue where the time at which a customer arrives is separated from the time at which it requests service by a random time. See [34] for further insight. Some rough approximations for this problem have been presented in [26] and later used in the context of OBS in [18].

Although we consider dual-header OBS, our results can be treated as an optimistic approximation for native forms of OBS. This type of optimistic approximation has been shown to be quite accurate for just-enough-time OBS [28] with void filling in [3] and [32].

We further assume full-wavelength conversion is available at all nodes. Apart from this assumption, we adopt a conservative stance by assuming burst segmentation, fiber delay lines and all other enhancements discussed in the literature are unavailable. We are not concerned with burst scheduling algorithms as they are not required for dual-header OBS.

We continue by describing the deflection routing policy considered in this paper.

Deflection routing policies in general can be categorized as either originating office control (OOC) or sequential office control (SOC). See [16] for a detailed description of this categorization. SOC is fast reacting and permits immediate deflection at any node at which contention is encountered by allowing a header to seek to reserve a time interval on an outgoing link that is alternative to the first-choice link. OBS is restricted to SOC policies. Using OOC policies in OBS would require excessively long offset times to allow for crank-back of a header to its source.

Let $\mathcal{L}$ be the set of all links. Consider an arbitrary source and destination pair. Suppose its first-choice route traverses N links, or equivalently, $N+1$ nodes and let its first-choice route be denoted as the ordered set $\mathbf{r} = (r_1, \dots, r_N)$, where $r_1, \dots, r_N \in \mathcal{L}$. For link $l \in \mathcal{L}$, let l^- denote the node that link l is incident from and let l^+ denote the node that link l is

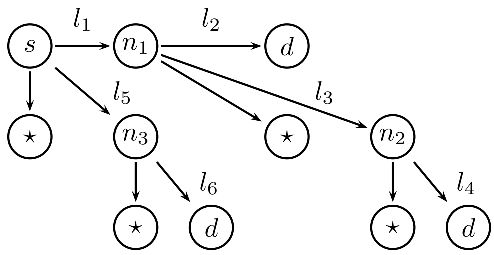


Fig. 4. Augmented route tree where loss nodes have been depicted with an asterisk

incident to. To ensure contiguity of $\mathbf{r}$, for all $n = 1, \dots, N-1$: $r_n^+ = r_{n+1}^-$; $r_1^- = s$; and, $r_N^+ = d$.

As soon as a header arrives at node r_n^-, say at time t, it seeks to reserve a wavelength in link r_n for an interval beginning at time $t+\Delta_n$ into the future and ending at time $t+\Delta_n+L/\mu$, where Δ_n is the residual offset time at node r_n^-, L is the size of its associated burst and μ is the transmission rate of a wavelength. Reservations that overlap time interval $[t+\Delta_n, t+\Delta_n+L/\mu]$ may have already been scheduled to all wavelengths in link r_n. In this case, link r_n is said to be in contention with respect to this time interval.

For each node r_n^-, $n = 1, \dots, N$, define a *deflection route* to be the ordered set $\mathbf{d}(n) = \big(d_1(n), \dots, d_{M_n}(n)\big)$, where $d_1(n), \dots, d_{M_n}(n) \in \mathcal{L}$ and $d_1(n) \neq r_n$. To ensure the contiguity of $\mathbf{d}$, for all $m = 1, \dots, M_n - 1$: $d_m(n)^+ = d_{m+1}(n)^-$; $d_1(n)^- = r_n^-$; and, $d_{M_n}(n)^+ = d$.

With deflection routing, a header arriving at node r_n^- that finds link r_n in contention may seek to reserve a wavelength in link $d_1(n)$, which is by definition a link incident from node r_n^- but is alternative to link r_n. Therefore, a header is blocked at node r_n^- if and only if all wavelengths in link r_n and $d_n(1)$ are in contention with respect to time interval $[t+\Delta_n, t+\Delta_n+L/\mu]$. However, without deflection routing, a header is blocked at node r_n^- if and only if all wavelengths in link r_n are in contention with respect to this time interval.

To avoid excessive hop-counts and to guard against the so-called ring-around-the-rosie problem [16], we only permit one deflection per header. That is, a deflection from a deflection route is forbidden.

The augmented route tree shown in Fig. 4 is used to clarify our notation. See [8], [16], [25] for discussions on augmented route trees. For this augmented route tree, we have $\mathcal{L} = \{l_1, \dots, l_6\}$, $\mathbf{r} = (r_1, r_2) = (l_1, l_2)$, $l_1^- = s$, $l_2^- = n_1$, $N = 2$, $M_1 = M_2 = 2$ and

$$\mathbf{d}(n) = \begin{cases} (l_5, l_6), & n = 1, \\ (l_3, l_4), & n = 2. \end{cases}$$

The main drawback of deflection routing in OBS is the so-called insufficient offset time problem that has been discussed in [17]. This problem refers to the situation in which a header is deflected and traverses more nodes than it would have on its first-choice route. Additional processing delays of δ encountered at each extra node may decrease a header's residual offset time to zero before it has reached its destination.

A few different approaches have been suggested to combat this problem. We adopt the most conservative approach of adding extra offset time. In particular, at its source, a burst is separated from its header by an offset time of at least $N_{\max}\delta$, where $N_{\max}$ is the maximum possible number of links a burst can expect to traverse and is given by

$$N_{\max} = \max\left(\max_{n=1,\dots,N} (M_n + n - 1), N \right).$$

For the augmented route tree shown in Fig. 4, we have $N_{\max} = 3$.

We must emphasize that we have described a rather simple deflection routing policy for OBS. Other more dynamic policies based on state-dependent routing [11], [16] may turn out to offer superior performance. They have not been studied in the context of OBS. We have simulated policies where multiple deflections are permitted per header, however no noteworthy benefit was observed relative to the case we consider in this paper where only one deflection is permitted per header.

III. Stabilizing OBS

In this section, we confirm the downturn in carried load evident in Fig. 3 is indeed a result of destabilization. We then show that either wavelength reservation or preemptive priority correct this downturn.

To this end, we propose to analyze the four-node ring network shown in Fig. 2 based on the following assumptions:

- A.1) Bursts arrive at each source and destination pair according to independent Poisson processes.
- A.2) A header itself does not offer any load.
- A.3) Burst size follows an independent exponential distribution.
- A.4) A blocked burst is cleared and never returns.
- A.5) The distribution of the number of busy wavelengths in a link is mutually independent of any other link.
- A.6) The total traffic offered to a link is the superposition of several independent Poisson processes and is therefore itself a Poisson process.

The last two assumptions are probably the most noteworthy. They are synonymous with the usual reduced-load approximation and have been discussed in this context and to some degree justified in [11], [12], [20], [21], [41]. All of these assumptions will also be used in our reduced-load approximation.

We will briefly outline some consequences of the last two assumptions. The last assumption allows for a one-moment analysis where the total traffic offered to a link is characterized solely in terms of its mean; more precisely, the mean of the distribution of the number of busy wavelengths on a link if it were to contain a hypothetical infinite number of wavelengths. However, the variance of this distribution as well as other higher moments may be vastly different from the variance and corresponding higher moments of a Poisson process. With a one-moment analysis, variance and other higher moments are not considered and are simply assumed to follow the variance and corresponding higher moments of a Poisson process. For further details, see discussions in [16] regarding the equivalent random method as well as Hayward's method.

The second last assumption is commonly referred to as the independence assumption. It allows for decoupling of a network into into its constituent links by ignoring any dependence between blocking events from link-to-link. This kind of independence assumption has been widely used in many types of network analyses.

Since a burst always follows the routing of its header and since it has been assumed a header itself offers no load, we are henceforth able to abstract by ignoring the presence of headers and working only in terms of bursts.

At a time instant in steady-state, assuming steady-state eventually prevails, let the random variable $X_l \in \{0,\dots,C_l\}$ denote the number of busy wavelengths in link $l \in \mathcal{L}$, where C_l is the total number of wavelengths in that link. Also, let $\mathbf{X} = (X_l)_{l\in\mathcal{L}}$. Then according to the independence assumption (see A.5), we can write

$$\mathbb{P}(\mathbf{X} = \mathbf{x}) = \prod_{l\in\mathcal{L}} \mathbb{P}(X_l = x),$$

for all $\mathbf{x} \in \{0,\dots,C\} \times \cdots \times \{0,\dots,C\}$.

For the remaining part of this section, we will concentrate specifically on the four-node ring network that we have already discussed. Since the four-node ring network is completely symmetric, it is sufficient to work in terms of an arbitrary link, and thus it is possible to write $X = X_l$ and $C = C_l$ for all $l \in \mathcal{L}$.

Recall that bursts only arrive at each source and destination pair for which there is a one-hop first-choice route. A three-hop deflection route for each of these source and destination pairs is thus uniquely determined. Also recall that all other source and destination pairs are not used.

Let λ be the burst arrival rate at each source and destination pair. Accordingly, the load offered to each source and destination pair is $\mathbb{E}(L)\lambda/\mu$ Erlangs, where L is a random variable representing burst size and μ is the wavelength transmission rate. Let $\overline{a} = \mathbb{E}(L)\lambda/\mu$ and let a denote the *total* load offered to a link, which is assumed to be the sum of several independent Poisson processes (see A.6). The probability that a burst is blocked at a link is then given by the Erlang B formula[2],

$$\begin{aligned} b &= \mathbb{P}(X = C) \\ &= \mathbf{E}(a, C) \triangleq \frac{a^C}{C!}\left(\sum_{i=0}^{C}\frac{a^i}{i!}\right)^{-1}. \end{aligned} \tag{1}$$

We are interested in calculating the blocking probability perceived by a burst, which will be denoted as p. Summing the total load carried by a link gives

$$(1-b)a = \big((1-b) + (1-b)b + (1-b)^2 b + (1-b)^3 b\big)\overline{a}. \tag{2}$$

Note that with circuit switching, we would write $(1-b)a = \big((1-b) + 3(1-b)^3 b\big)\overline{a}$ instead of (2), since the load carried by each of the three links comprising a deflection route must be equal for circuit switching.

Rearranging (2) gives

$$\overline{a} = \frac{a}{1 + 2b - 6b^2 + 4b^3 - b^4}. \tag{3}$$

[2]The Erlang B formula can be efficiently computed via the recursion $\mathbf{E}(a,c) = \frac{a\mathbf{E}(a,c-1)}{c + a\mathbf{E}(a,c-1)}$ for $c = 1,\dots,C$, where $\mathbf{E}(a,0) = 1$.

It can then be straightforwardly verified that

$$p = 3b^2 - 3b^3 + b^4. \tag{4}$$

To confirm the simulation results presented in Fig. 3, we plot p and $(1-p)\overline{a}$ as a function of $\overline{a}$ in Fig. 5 as solid lines labeled ‘unprotected’. These two plots can be generated as follows: for each of several values of a, compute b via (1) and then compute $\overline{a}$ and p based on this value of b via (3) and (4), respectively.

It turns out that neither p nor $(1-p)\overline{a}$ are proper functions of $\overline{a}$ because the mapping from $\overline{a}$ to p is not one-to-one. This definitely confirms that deflection routing may destabilize OBS. For some values of $\overline{a}$, there are up to three equilibria that may exist in steady-state. It is not clear if one equilibria is dominant or if there are oscillations between all three equilibria. The plots shown in Fig. 3 generally do not match up well with their counterparts in Fig. 5. This is most likely because simulation relies on long-run averaging, which yields averages lying somewhere in between these three equilibria. That is, we are trying to simulate behavior that is inherently non-stationary. It is however satisfying to note that the downturn in carried load occurs at approximately the same value of $\overline{a}$ in Fig. 3 and Fig. 5.

In the next two subsections, we present a parallel analysis of wavelength reservation and preemptive priority. Any notation that we reuse continues to bear the same definition as above.

A. Wavelength Reservation

Recall that with wavelength reservation, deflected bursts are barred from engaging an idle wavelength on any link for which the total number of busy wavelengths on that link exceeds a predefined threshold. Let that threshold be denoted as K. Therefore, a deflected burst cannot be scheduled to a link for which K or more of its wavelengths are busy.

Let $\hat{a}$ denote the deflected load offered to a link. The total load offered to a link is the sum of loads it is offered by deflected bursts and first-choice bursts. Since a first-choice route is associated with one unique link, it is not hard to see that

$$\hat{a} = a - \overline{a}. \tag{5}$$

Treating a link as a simple one-dimensional birth-and-death process, we have a recursion of the form

$$\begin{aligned} \pi_i &= \mathbb{P}(X = i) \\ &= \begin{cases} a^i \pi_0 / i!, & i = 1,\dots,K, \\ (a-\hat{a})^{i-K} a^K \pi_0 / i!, & i = K+1,\dots,C, \end{cases} \end{aligned} \tag{6}$$

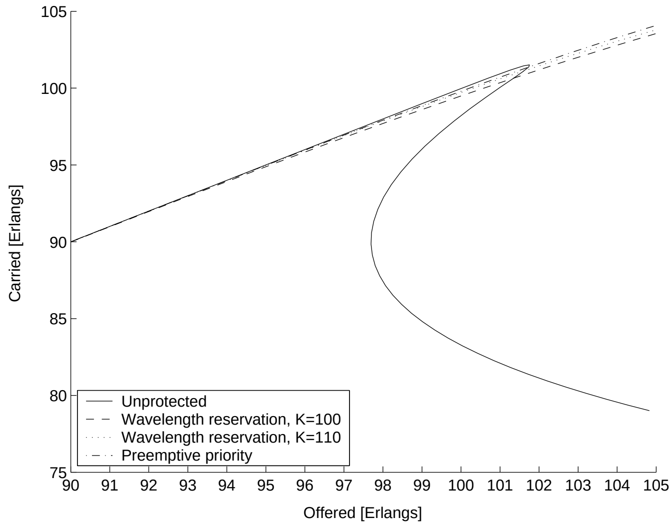
where the normalization constant π_0 is determined as usual via $\sum_{i=0}^{C} \pi_i = 1$. The probability that a first-choice burst is blocked at a link is given by $b = \pi_C$, while the probability that a deflected burst is blocked at a link is given by $q = \sum_{i=K}^{C} \pi_i$.

Analogous to (2), summing the total load carried by a link gives

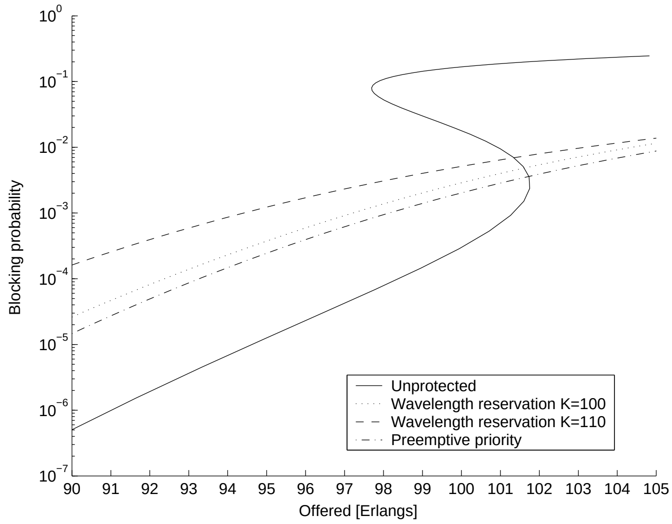
$$(1-b)a = \big((1-b) + (1-q)b + (1-q)^2 b + (1-q)^3 b\big)\overline{a}, \tag{7}$$

which after rearrangement can be rewritten as

$$\overline{a} = \frac{(1-b)a}{1 + 2b - 6bq + 4bq^2 - bq^3} \tag{8}$$



(a) Carried load



(b) Blocking probability

Fig. 5. Stabilizing a four-node symmetrical OBS network

It can then be verified that

$$p = 3bq - 3bq^2 + bq^3. \quad (9)$$

As a check, by setting $q = b$, it can be seen that (9) reduces to its counterpart (4).

There are no rigorous criteria governing the choice of reservation threshold. See [19] for details. Choosing K too large results in performance that is not much better than if deflection was not permitted at all, while choosing K too small does not correct the downturn in carried load. We choose K by iteratively incrementing its value until the mapping from $\overline{a}$ to p appears to be one-to-one. Through trial and error, for $C = 120$, we found that $K \in [100, 110]$ was a good choice.

To show that wavelength reservation with a sufficiently large reservation threshold can correct the downturn in carried load shown in Fig. 3, we again plot p and $(1-p)\overline{a}$ as a function of $\overline{a}$ in Fig. 5 as dotted and dashed lines labeled 'wavelength reservation'. The dashed line is for $K = 100$ and the dotted line is for $K = 110$. These two plots can be generated using a fixed-point iterative procedure as follows. For each of several values of a, arbitrarily choose $\hat{a}$ and compute $\overline{a}$ via (5) based on this arbitrary choice of $\hat{a}$. After this initialization phase, compute b as well as q by solving the recursion given by (6). Then recompute $\overline{a}$ via (8) and check if the absolute difference between the old value of $\overline{a}$ and its recomputed value is sufficiently close to satisfy a prescribed error criterion. These set of steps comprise one iteration. Subsequent iterations are continued until the error criterion is satisfied by updating the value of $\hat{a}$ according to (5). Using the values of b and q computed during the final iteration, p is determined via (9).

Based on Fig. 5, it may be tempting to consider reducing the value of K to improve performance, however, if K is reduced below 110, a kink begins to appear in Fig. 5(b) signaling the onset of destabilization.

B. *Preemptive Priority*

Preemptive priority is a new approach that we propose to protect OBS against destabilization that is based on enforcing preemptive priority between first-choice bursts and deflected bursts. With this approach, a first-choice burst is given the right to preempt a reservation that has been scheduled for a deflected burst. Peremption is a last resort in the sense that a first-choice burst foremost seeks an appropriate idle wavelength.

Almost all the equations presented in the preceding subsection treating wavelength reservation also hold for preemptive priority. The exception is that the probability that a first-choice burst is blocked at a link and the probability that a deflected burst is blocked at a link, which we have denoted as b and q, respectively, can no longer be computed via the recursion given by (6). Instead, we compute b and q as follows.

A first-choice burst is oblivious to the presence of deflected bursts and only perceives other first-choice bursts. It follows that $b = \mathbf{E}(a - \hat{a}, C)$ and

$$q = \frac{a\mathbf{E}(a, C) - (a - \hat{a})\mathbf{E}(a - \hat{a}, C)}{\hat{a}}. \quad (10)$$

The numerator of (10) is equal to the deflected burst load *carried* by a link, while the denominator is by definition the deflected burst load offered to a link. Taking their ratio gives the probability that a deflected burst is blocked at a link.

For the case of preemptive priority, we plot p and $(1-p)\overline{a}$ as a function of $\overline{a}$ in Fig. 5 as an interchanging dotted/dashed line labeled 'preemption'. The same fixed-point iterative procedure described in the preceding subsection can be used to generate these plots but b and q are now computed via (10).

We can conclude that preemptive priority may yield marginally lower blocking probabilities than wavelength reservation. Although the benefit of preemptive priority is unremarkable for $K = 110$, a marked disparity is evident for $K = 100$, especially at low to moderate loads.

A key advantage of preemptive priority is that it is guaranteed to stabilize deflection routing in OBS as well as circuit switching and optical packet switching, though we have already discussed that some attributes of preemptive priority render it an inappropriate form of protection for circuit switching. Preemptive priority guarantees stability because it ensures performance that is no worse than if bursts were not deflected but simply blocked. This property is a consequence of the impossibility of a deflected burst to alter the fate of a first-choice burst. Moreover, we know that OBS is stable without deflection routing. Consequently, protecting OBS with preemptive priority guarantees stability. On the contrary, the stabilizing properties of trunk reservation are highly dependent on the choice of reservation threshold.

With preemptive priority, a preempted burst is not necessarily blocked in its entirety. For example, a burst may suffer preemption at a link well after its head has been transmitted on that link. In this case, packets residing in its tail are blocked but those residing in its head are unaffected by preemption and continue as normal. The reverse case where packets residing in its head are blocked but those residing in its tail are unaffected is also possible. This results in the presence of truncated bursts and is reminiscent of burst segmentation [38].

A problem may arise when a truncated burst arrives at its destination. Although in principle it is possible to recover packets from a truncated burst, this is complicated since knowledge of a truncation is localized to the intermediate node at which it occurred. Therefore, each destination anticipates a complete burst with well-defined packet boundaries. In this paper, we have adopted a conservative stance by assuming that it is not possible to recover packets from a truncated burst.

An alternative would be to assume a more sophisticated node architecture that is capable of salvaging packets from a truncated burst. Although this leads to a remarkable increase in node throughput [38], signaling complexity also increases because a packet delineation protocol that includes functionality to check the integrity of each packet, such as the simple data link (SDL) protocol discussed in [13], is essential.

IV. Reduced-Load Approximation for OBS

In this section, we develop a new reduced-load approximation to evaluate the performance of OBS networks that have been stabilized with either wavelength reservation or preemptive priority.

Assumptions A.5 and A.6 will play a key role. They were defined and discussed in the preceding section. We will use simulation to quantify the error admitted in making these two assumptions. Assumptions A.1 to A.4 will also be reinvoked.

The reduced-load approximation was conceived in 1964 [12] for the analysis of circuit-switched networks and has remained a cornerstone of network performance evaluation. See [11], [12], [20], [21], [41] and references therein for details on the reduced-load approximation and its many applications. In [31], [32], we presented a reduced-load approximation for OBS networks where each source and destination pair is assigned a single fixed route. That is, OBS networks without deflection routing.

At this point, it may be worthwhile recalling notation presented in Section II as it will be used extensively in this section.

A. Step One: Link Offered Loads

The first step is to decompose the network into its constituent links. In particular, assumptions A.5 and A.6 permit each link to be treated as an independent birth-and-death process that is Markovian. To compute the steady-state distribution $\pi_i = \mathbb{P}(X = i)$, $i = 0, \dots, C$, for this kind of birth-and-death process, it suffices to know the load that it is offered, which is the ratio of the birth rate to the death rate. Therefore, we must determine the load offered to each link $l \in \mathcal{L}$. The difficulty is that the load offered to a given link is a function of the steady-state distributions at all other links, which are unknown.

We first compute the load offered to each link $l \in \mathcal{L}$ that is owing to an arbitrary source and destination pair by assuming $\mathbf{r} \cap \mathbf{d}(1) \cap \cdots \cap \mathbf{d}(N) = \emptyset$. We then continue by relaxing this temporary assumption and presenting an algorithm to compute the load offered to each link $l \in \mathcal{L}$ that is owing to all source and destination pairs.

Consider an arbitrary source and destination pair and let $\overline{a}$ be the load it is offered. Furthermore, for the sake of clarity, assume $\mathbf{r} \cap \mathbf{d}(1) \cap \cdots \cap \mathbf{d}(N) = \emptyset$, which we call the disjointedness assumption. To begin with, suppose b_l and q_l are known for all $l \in \mathcal{L}$. It then follows that the load offered to $r_n \in \mathbf{r}$ owing to this source and destination pair is given by

$$a_{r_n} = \overline{a}(1 - b_{r_1}) \cdots (1 - b_{r_{n-1}}), \quad n = 1, \dots, N, \qquad (11)$$

and for $d_m(n) \in \mathbf{d}(n)$, $n = 1, \dots, N$, we have

$$\begin{aligned} a_{d_m(n)} &= \hat{a}_{d_m(n)} \\ &= \overline{a}(1 - b_{r_1}) \cdots (1 - b_{r_{n-1}}) b_{r_n} \beta_m(n), \quad (12) \end{aligned}$$

for all $m = 1, \dots, M_n$, where

$$\beta_m(n) = (1 - q_{d_1(n)}) \cdots (1 - q_{d_{m-1}(n)}). \qquad (13)$$

The equality $a_{d_m(n)} = \hat{a}_{d_m(n)}$ is an immediate consequence of the disjointedness assumption. The probability that a burst is not blocked at the links preceding link $d_m(n) \in \mathbf{d}(n)$ is given by $\beta_m(n)$. Equation (12) concerns the intersection of three events: 1) a burst is not blocked at the links preceding link r_n, which occurs with probability $(1-b_{r_1}) \cdots (1-b_{r_{n-1}})$; 2) a burst is blocked at link r_n, which occurs with probability b_{r_n}; and, 3) a burst is not blocked at the links preceding link $d_m(n)$, which occurs with probability $\beta_m(n)$. It is the probability of the intersection of these three events that is of interest. By the independence assumption (see A.5) any two of these events are mutually independent and thus (12) follows.

To relax the disjointedness assumption, we need to take care of the possibility that

$$\Omega_m(n) = \{d_1(n), \dots, d_{m-1}(n)\} \cap \{r_1, \dots, r_n\} \neq \emptyset$$

by conditioning the probability $\beta_m(n)$ as specified in (14) (see inset next page). The expression given in (14) can be simplified based on the independence assumption and the following fact.

Fact 1: The conditional probability that a first-choice burst is not blocked at link $l \in \mathbf{d}$ given that a deflected burst is not blocked at that same link $l \in \mathbf{r}$ for some $l \in \Omega_m(n)$ is given by

$$\begin{aligned} &\mathbb{P}(\text{not blocked at } l \in \mathbf{d} | \text{ not blocked at } l \in \mathbf{r}) \\ &= \frac{\mathbb{P}(\text{not blocked at } l \in \mathbf{d})}{\mathbb{P}(\text{not blocked at } l \in \mathbf{r})} = \frac{1 - q_l}{1 - b_l}. \end{aligned}$$

$$\beta_m(n) \quad = \quad \mathbb{P}\big(\text{not blocked at } d_1(n), \dots, d_{m-1}(n) | \text{ blocked at } r_n \cap \text{not blocked at } r_1, \dots, r_{n-1}\big) \tag{14}$$

Proof: This fact holds for wavelength reservation as well as preemptive priority. Its proof is elementary after establishing that $\{\text{not blocked at } l \in \mathbf{d}\} \subseteq \{\text{not blocked at } l \in \mathbf{r}\}$. To establish this inclusion consider the following. With wavelength reservation, a deflected burst is not blocked at link $l \in \mathbf{d}$ if and only if $X_l < K$, but a first-choice burst is not blocked at that same link $l \in \mathbf{r}$ if and only if $X_l < C$. Since $X_l < C$ implies $X_l < K$, this inclusion follows immediately. Similarly, with preemptive priority, if a deflected burst is not blocked at link $l \in \mathbf{d}$, then $X_l < C$, which is sufficient to ensure a first-choice burst is not blocked at that same link $l \in \mathbf{r}$. ∎

Based on Fact 1 and the independence assumption, (14) can be rewritten as

$$\begin{aligned} \beta_m(n) \quad &= \quad \frac{\mathbb{P}\big(\text{not blocked at } d_1(n), \dots, d_{m-1}(n)\big)}{\mathbb{P}\big(\text{not blocked at } r_1, \dots, r_{n-1} \in \Omega_m(n)\big)} \\ &= \quad \frac{(1 - q_{d_1(n)}) \cdots (1 - q_{d_{m-1}(n)})}{\prod_{l \in \Omega_m(n)} (1 - b_l)} \end{aligned} \tag{15}$$

See the appendix for details. Henceforth we relax the disjointedness assumption by computing $\beta_m(n)$ according to (15) instead of (13).

Let $\mathcal{J}$ be the set of all source and destination pairs. When we are required to distinguish between source and destination pairs, we will superscript existing notation with a j to denote it pertains to source and destination pair $j \in \mathcal{J}$. For example, $\overline{a}^j$ is the load offered to source and destination pair $j \in \mathcal{J}$. Using (11), (12) and (15), we are able to formulate Algorithm 1, which computes the load offered to each link $l \in \mathcal{L}$ that is owing to all source and destination pairs. The complexity of Algorithm 1 is bounded by $O(JL^2)$, where $J = |\mathcal{J}|$ and $L = |\mathcal{L}|$.

In Algorithm 1, at iteration n of the $n = 1, \dots, N^j$ for-loop, the auxiliary variable x is scaled by $(1 - b_i)$, where $i = r_n^j$. Thus, according to (11), x equals the reduced-load offered to link r_{n+1}^j that pertains to first-choice bursts of source and destination pair $j \in \mathcal{J}$.

Similarly, at iteration m of the $m = 1, \dots, M^j(n)$ for-loop, the auxiliary variable y is scaled by $(1 - q_i)/(1 - b_i)$, where $i = d_m^j(n)$, if $r_i \in r_1^j, \dots, r_n^j$. Otherwise, y is scaled by $1 - q_i$. Thus, according to (12) and (15), y equals the reduced-load offered to link $d_{m+1}^j(n)$ pertaining to deflected bursts of source and destination pair $j \in \mathcal{J}$.

Algorithm 1 Calculate $a_l, \hat{a}_l \ \forall l \in \mathcal{L}$

Require: $b_l, q_l \ \forall l \in \mathcal{L}$; $\mathbf{r}^j, \mathbf{d}^j(n), \Omega_m^j(n) \ \forall j \in \mathcal{J}, n = 1, \dots, N^j, m = 1, \dots, M_n^j$

$a_l, \hat{a}_l = 0 \ \forall l \in \mathcal{L}$ // Initialization
for $j \in \mathcal{J}$ **do**
  $x = \overline{a}^j$
  for $n = 1, \dots, N^j$ **do**
    $i = r_n^j$; $a_i = a_i + x$
    $y = xb_i$; $x = x(1 - b_i)$;
    for $m = 1, \dots, M^j(n)$ **do**
      $i = d_m^j(n)$; $a_i = a_i + y$; $\hat{a}_i = \hat{a}_i + y$
      if $i \in \Omega_m^j(n)$ **then**
        $y = y(1 - q_i)/(1 - b_i)$
      else
        $y = y(1 - q_i)$
      end if
    end for
  end for
end for
return $a_l, \hat{a}_l \ \forall l \in \mathcal{L}$

TABLE I

FORMULAE TO COMPUTE b_l AND q_l

	b_l	q_l
No protection	$\mathbf{E}(a_l, C_l)$	$\mathbf{E}(a_l, C_l)$
Wavelength reservation	$\pi_{C,l}$	$\sum_{i=K}^{C} \pi_{i,l}$
Preemptive priority	$\mathbf{E}(\omega_l, C)$	$\big(a_l \mathbf{E}(a_l, C) - \omega_l \mathbf{E}(\omega_l, C)\big)/\hat{a}_l$

B. Step Two: Link Blocking Probabilities

Computation of the blocking probabilities b_l and q_l at each link $l \in \mathcal{L}$ differs according to the type of protection used to guard against destabilization and was considered for each of the three cases of no protection, wavelength reservation and preemptive priority in Section III. In particular, refer to (1), (6) and (10), respectively. For convenience, we provide a brief summary of the formulae used to compute b_l and q_l for each type of protection in Table I, where for brevity, we have defined $\omega_l = a_l - \hat{a}_l$. It may be worth recalling that for the case of wavelength reservation, the steady-state distribution $\pi_{i,l} = \mathbb{P}(X_l = i)$ is computed according to the recursion

$$\pi_{i,l} = \begin{cases} a_l^i \pi_0 / i!, & i = 1, \dots, K, \\ (a_l - \hat{a}_l)^{i-K} a_l^K \pi_0 / i!, & i = K+1, \dots, C. \end{cases}$$

Let $\mathbf{b} = \{b_l\}_{l \in \mathcal{L}}$, $\mathbf{q} = \{q_l\}_{l \in \mathcal{L}}$, $\mathbf{a} = \{a_l\}_{l \in \mathcal{L}}$ and $\hat{\mathbf{a}} = \{\hat{a}_l\}_{l \in \mathcal{L}}$. Also, let the mapping $g : (\mathbf{b}, \mathbf{q}) \to (\mathbf{a}, \hat{\mathbf{a}})$ represent the operation of Algorithm 1 and let the mapping $f : (\mathbf{a}, \hat{\mathbf{a}}) \to (\mathbf{b}, \mathbf{q})$ represent the operation of an algorithm that computes link blocking probabilities according to the formulae shown in Table I. This is admittedly a rather non-rigorous definition of g and f, but it will be sufficient for our purposes. We are interested in finding a solution $(\mathbf{b}, \mathbf{q}, \mathbf{a}, \hat{\mathbf{a}})$ to

$$\begin{cases} (\mathbf{b}, \mathbf{q}) = f(\mathbf{a}, \hat{\mathbf{a}}), \\ (\mathbf{a}, \hat{\mathbf{a}}) = g(\mathbf{b}, \mathbf{q}). \end{cases} \tag{16}$$

Since f and g are non-linear, it is difficult to determine if (16) has a unique solution $(\mathbf{b}, \mathbf{q}, \mathbf{a}, \hat{\mathbf{a}})$. It has been proved that the analogous form of (16) for circuit-switched networks using fixed routing does have a unique solution [21], though it is well-known that multiple solutions are possible for circuit-switched networks using deflection routing. We discuss solution uniqueness in the next section.

Presupposing that a solution $(\mathbf{b}, \mathbf{q}, \mathbf{a}, \hat{\mathbf{a}})$ for (16) does indeed exist, it may be determined via Algorithm 2. Algorithm 2 is a fixed-point iterative algorithm which terminates once $\mathbf{b}$ and $\mathbf{q}$ satisfy a prescribed error criterion and are thus said to have converged to a fixed-point. Fixed-point iterative algorithms have been used prevalently in the context of the reduced-load approximation. See [11], [20], [41], [31], [32], [42] for various examples. Although convergence of this kind of algorithm is not a certainty, divergence is rare in practice and can often be overcome by periodically re-initializing with a convex combination of the most recent iterations.

Algorithm 2 Calculate $b_l, q_l\ \forall l \in \mathcal{L}$

Require: ϵ; $c_1, c_2 \geq 0$ such that $c_1 + c_2 = 1$; $\mathbf{r}^j, \mathbf{d}^j(n), \Omega^j_m(n)\ \forall j \in \mathcal{J}, n = 1, \ldots, N^j$

```
b_l = 1, q_l = 1 ∀l ∈ L // Initialization
b'_l = 0, q'_l = 0 ∀l ∈ L // Initialization
while ∃l ∈ L such that |b_l − b'_l| > ϵ or |q_l − q'_l| > ϵ do
    for l ∈ L do
        b'_l = c_1 b_l + c_2 b'_l // Convex combination
        q'_l = c_1 b_l + c_2 q'_l
    end for
    b' = {b'_l}_{l∈L}; q' = {q'_l}_{l∈L}
    (a, â) = g(b', q') // Algorithm 1
    (b, q) = f(a, â) // Update link blocking probabilities
end while
```

In Algorithm 2, the error criterion is denoted as $\epsilon > 0$ and the outdated values of $\mathbf{b}$ and $\mathbf{q}$ are denoted as $\mathbf{b}'$ and $\mathbf{q}'$, respectively. Furthermore, the coefficients used to form a convex combination of the two most recent values of $\mathbf{b}$ and $\mathbf{q}$ are denoted by $c_1, c_2 \geq 0$, where $c_1 + c_2 = 1$.

C. Step Three: End-to-End Blocking Probabilities

Given that b_l and q_l are known for all $l \in \mathcal{L}$, it is possible to compute the end-to-end blocking probability for each source and destination pair. Let p^j denote the end-to-end blocking probability for source and destination pair $j \in \mathcal{J}$.

For the moment, we suppress the superscript j and thereby consider an arbitrary source and destination pair. Let γ_n be the probability of the intersection of the following three events: 1) a burst is not blocked at the links preceding link r_n, which occurs with probability $(1 - b_{r_1}) \cdots (1 - b_{r_{n-1}})$; 2) a burst is blocked at link r_n, which occurs with probability b_{r_n}; and, 3) a burst is not blocked at links $d_1(n), \ldots, d_{M_n}(n)$, which occurs with probability $\beta_{M_n+1}(n)$. Note that the '$+1$' appears in $\beta_{M_n+1}(n)$ to annihilate the '-1' appearing in its definition, which is given by (13), otherwise, without the '-1', $d_{M_n}(n)$ would be missed. It can be verified that a burst is not blocked if and only if: 1) all three of these events occur for some $n = 1, \ldots, N$; or, 2) a burst is not blocked at links $r_1, \ldots, r_N$. Therefore, we can write

$$p = 1 - (1 - b_{r_1}) \cdots (1 - b_{r_N}) - \sum_{n=1}^{N} \gamma_n, \tag{17}$$

where

$$\begin{aligned} \gamma_n &= \mathbb{P}(\text{not blocked at } r_1 \ldots r_{n-1}) \mathbb{P}(\text{blocked at } r_n) \\ &\quad \times \beta_{M_n+1}(n) \\ &= (1 - b_{r_1}) \cdots (1 - b_{r_{n-1}}) b_{r_n} \beta_{M_n+1}(n). \end{aligned} \tag{18}$$

As a check, comparing (18) with (12) reveals that $\hat{a}_{d_{M_n+1}(n)}/\overline{a} = \gamma_n$, as expected. Using this relation, we can compute p within Algorithm (1) simply by initializing $p^j = 1$ for all $j \in \mathcal{J}$ and executing the following operation immediately after line 14

$$p^j \leftarrow p^j - \frac{y}{\overline{a}^j},$$

as well as the following operation immediately after line 15

$$p^j \leftarrow p^j - \frac{x}{\overline{a}^j}.$$

Recall that x and y are auxiliary variables defined in Algorithm 1.

Finally, we let P denote the average blocking probability across all source and destination pairs, which is computed as

$$P = \left(\sum_{j \in \mathcal{J}} \overline{a}^j \right)^{-1} \sum_{j \in \mathcal{J}} \overline{a}^j p^j. \tag{19}$$

In concluding this section, we remark that our reduced-load approximation can be straightforwardly extended to any SOC routing policy that can be represented with an augmented route tree. To realize this extension, we would use the recursive approach outlined in [8], [25], [30] to compute the probability that a blocking or completion route of an augmented route tree is used given that the load offered to each link is known. This approach relies on a recursion that is commonly used in the field of system's reliability analysis. Although the computational complexity of this recursion may be high, it can simplified for SOC routing, as remarked in [30]. (In writing (12), we have in fact implicitly used the simplification alluded to in [30].)

This extension would allow us to study policies where more than one deflection is permitted per header or deflections from deflection routes are permitted. We have chosen not to pursue this extension because we have simulated policies in which multiple deflections are permitted per header and observed an unremarkable improvement. See the conference version of this paper for empirical results substantiating this claim.

V. Empirical Performance Evaluation

In this section, we will use simulation to quantify the error admitted in making assumptions A.5 and A.6. We will then use our reduced-load approximation to evaluate the performance of deflection routing in randomly generated networks. In particular, with respect to average blocking probability, which is given by (19), we will compare the performance of unprotected deflection routing and deflection routing protected with either wavelength reservation or preemptive priority.

Unless otherwise specified, all the results presented in this section pertain to networks that have been randomly generated according to the specifications shown in Table II, where

TABLE II
SPECIFICATIONS OF RANDOMLY GENERATED NETWORK

Parameter	Value
Number of source and destination pairs	$J = 50$
Number of links	$L = 30$
First choice route hop-count	$N \sim \mathbf{U}[1, 4]$
Additional hop-count	$\kappa \sim \mathbf{U}[1, 8]$
Reservation threshold	$K_l = \lfloor 0.8C_l \rfloor \ \forall l \in \mathcal{L}$

$\mathbf{U}[a, b]$ denotes the discrete uniform distribution taking values on the integers $a, a + 1, \dots, b$. The parameter referred to as additional hop-count and denoted as κ in Table II needs further clarification. It governs the total hop-count of each deflection route $\mathbf{d}(n)$, $n = 1 \dots, N$, which we have already denoted as M_n, so that

$$M_n = N - n + \kappa, \quad n = 1 \dots, N. \tag{20}$$

Computing the total hop-count of a deflection route according to (20) ensures that the hop-count of a deflection route is at least no less than the hop-count of its corresponding first-choice route. This is usually the case (but not always) in practice, since if $M_n < N$ for some $n = 1, \dots, N$, it is probably preferable to use $\mathbf{d}(n)$ as a first-choice route instead of $\mathbf{r}$, unless $\mathbf{d}(n)$ traverses links that are heavily congested.

An algorithm to generate a random network takes the parameters shown in Table II and returns the ordered sets $\mathbf{r}^j$ and $\mathbf{d}^j(n)$ for $j = 1, \dots, J$ and $n = 1, \dots, N^j$. We do not specify details of such an algorithm as it would take us too far afield. However, we remark that no bias was given to any particular link or source and destination pair in our implementation of this algorithm.

To reduce the number of free parameters, we assume $\overline{a}^j = \overline{a}$ for all $j \in \mathcal{J}$. Once the ordered sets $\mathbf{r}^j$ and $\mathbf{d}^j(n)$ have been generated, we provision capacity based on an iterative heuristic that aims at achieving a target link blocking probability of 10^{-2} for a nominally chosen value of $\overline{a}$. At each iteration of this heuristic, our reduced-load approximation is used to compute the link blocking probabilities for the current wavelength vector $(C_l)_{l \in \mathcal{L}}$. Then for each link $l \in \mathcal{L}$, if

$$\frac{(a_l - \hat{a}_l)b_l + q_l \hat{a}_l}{a} > 10^{-2},$$

the current value of C_l is incremented by unity, otherwise it is decremented by unity. This completes one iteration. We stop iterating as soon as all link blocking probabilities are sufficiently close to 10^{-2}. Although this provisioning heuristic does not ensure link blocking probabilities will converge to a prescribed target, it turned out to perform well for most of the networks we studied. Unless otherwise stated, we aimed at selecting a nominal value of $\overline{a}$ that resulted in $\sum_{l \in \mathcal{L}} C_l / L \approx 30$.

To quantify the error admitted in making assumptions A.5 and A.6, we generated several random networks and used our reduced-load approximation as well as simulation, which does not rely on these two assumptions, to compute the average blocking probability for several values of $\overline{a}$. The values of $\overline{a}$ were chosen to lie uniformly in an interval centered about the

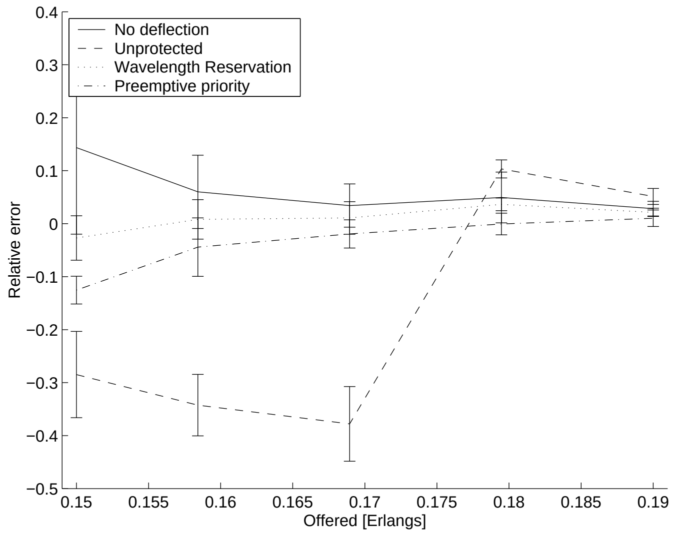


Fig. 7. Relative error in estimating blocking probability as a function of load offered to each source and destination pair for a randomly generated network; confidence intervals are commensurate to one standard deviation

nominal value of $\overline{a}$ for which dimensioning was performed. The results for one particular random network are shown in Fig. 6 and Fig. 7, where RLM and SM denote our reduced-load approximation and simulation, respectively. In particular, we plot P as a function of $\overline{a}$ in Fig. 6 for unprotected deflection routing, wavelength reservation and preemptive priority. To serve as a benchmark to gauge the performance gains of deflection routing, we also plot P as a function of $\overline{a}$ for no deflection routing. In Fig. 7, we plot relative error as a function of $\overline{a}$ for each of these cases, where relative error is defined in the usual way as

$$\frac{P \text{ as computed by RLA } - P \text{ as computed by SM}}{P \text{ as computed by SM}}.$$

The conclusions to be drawn are:

- Unprotected deflection routing may destabilize OBS. Destabilization may result in higher blocking probabilities than if bursts were not deflected but simply blocked.
- Destabilization manifests at loads that are considered moderate to high in the context of OBS. In particular, loads that are commensurate to an average blocking probability that is greater than or in the order of 10^{-2}.
- At low loads, unprotected deflection routing may yield better performance than protected deflection routing. However, the converse is true at high loads. It follows that protection may be counterproductive for an over-provisioned network. According to this observation, it seems reasonable to dynamically activate/deactivate protection, or adjust the reservation threshold in the case of wavelength reservation, on an hourly or daily basis in accordance with anticipated load conditions. In particular, during busy periods, protection would be activated to guard against destabilization, while during quiet periods, it would be deactivated to improve blocking performance.
- Preemptive priority consistently yields better blocking performance than wavelength reservation.

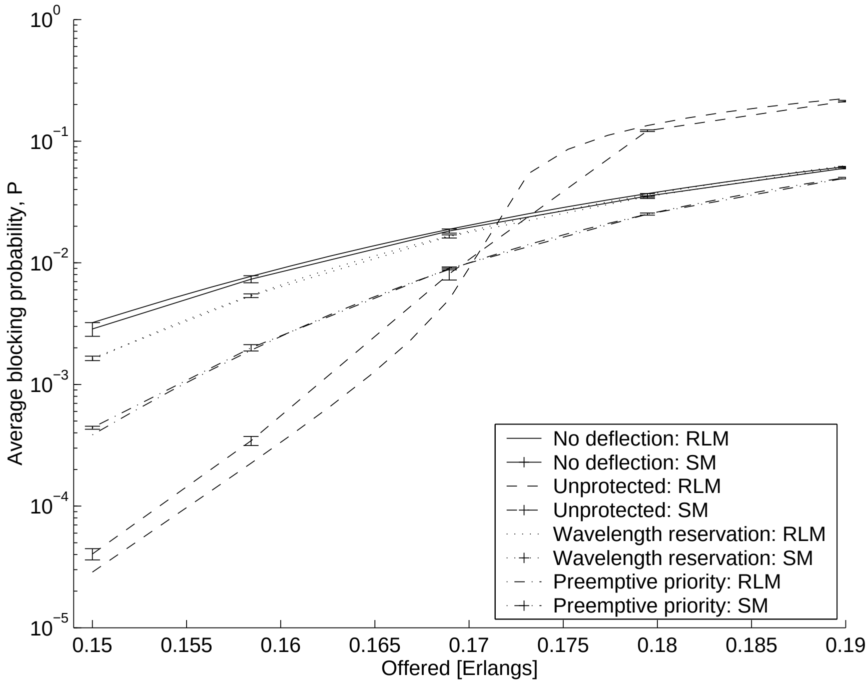


Fig. 6. Blocking probability as a function of load offered to each source and destination pair for a randomly generated network; confidence intervals are commensurate to one standard deviation

- In terms of blocking performance, deflection routing is a viable approach of resolving wavelength contention in OBS. At low loads, it may yield reductions in blocking probability of more than one order in magnitude compared to no deflection.
- The accuracy of our reduced-load approximation deteriorates for the case of unprotected deflection routing. This inaccuracy may in fact be a consequence of the difficulty in accurately simulating unprotected deflection routing. As we alluded to earlier, using simulation to predict non-stationary behavior associated with unprotected deflection routing may yield unpredictable results. Furthermore, since the amount of deflection is greatest for the case of unprotected deflection routing, it is this case that violates the Poisson assumption (see A.6) the most. In particular, the variance of the load offered to a deflection route is always larger than its mean, which is not the case for a Poisson process. Apart from the case of unprotected deflection routing, our reduced-load approximation is remarkably accurate. Therefore, assumptions A.5 and A.6 do not admit significant error.

To plot P as a function of $\overline{a}$, we repeatedly used our reduced-load approximation to explicitly compute a unique value of P for each given value of $\overline{a}$. However, this presupposes that the mapping from $\overline{a}$ to P is one-to-one, which we know may not be the case for unprotected deflection routing. Therefore, results pertaining to this case must be viewed with some caution as they may reflect the 'average' blocking probability over multiple stable equilibria that exist in steady-state. Recall that there were three stable equilibria evident in the four-node ring network studied in Section III. The approach we used to identify these three stable equilibria relied on indirectly computing blocking probability, as well as the corresponding value of $\overline{a}$, as a function of the load offered to a *link*, rather than explicitly computing blocking probability as a function of $\overline{a}$. However, this indirect approach does not generalize to asymmetric networks.

For unprotected deflection routing, we occasionally found that Algorithm 2 failed to converge or periodically cycled between multiple fixed-points. Cycling was quite rare and disappeared as soon as sufficient protection was added. We speculate that cycling and divergence of Algorithm 2 is probably closely tied to the fact that (16) may have multiple solutions. This issue is specifically discussed in the context of wavelength reservation in the conference version of this paper.

To conclude this section, we study the sensitivity of blocking performance to two effects: variation in the hop-count of deflection routes; and, variation in the wavelength reservation threshold. We study each of these two effects independently by considering two experiments where we vary the additional hop-count parameter κ and the wavelength reservation threshold K, respectively.

To this end, we generated 20 random networks and dimensioned each of them independently based on the heuristic described earlier in this section. Using our reduced-load approximation, we then computed P as a function of κ for a fixed value of $\overline{a}$ and P as a function of $\overline{a}$ for different values of K. To separate spurious randomness from underlying trends, we averaged P over all 20 random networks. We plot P as a function of κ in Fig. 8 and P as a function of $\overline{a}$ for different values of K in Fig 9.

Based on Fig. 8, we conclude that unprotected deflection routing is highly sensitive to hop-count variation. This high sensitivity may have ramifications if rerouting is performed (to

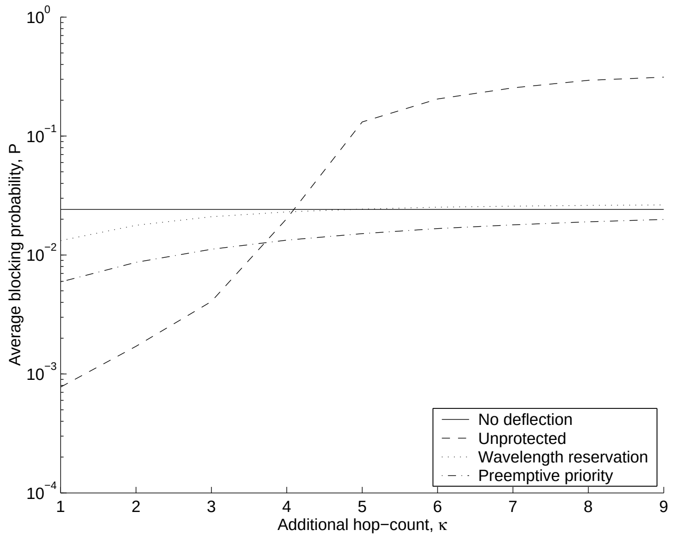


Fig. 8. Average blocking probability as a function of additional hop-count of deflection routes

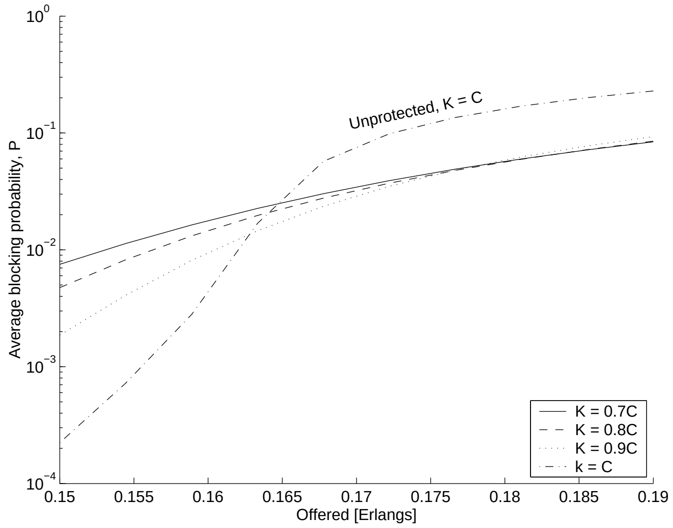


Fig. 9. Average blocking probability as a function of load offered to each source and destination pair for varying reservation thresholds

bypass severed fibers for example) and results in an increased hop-count. Wavelength reservation and preemptive priority are more robust to hop-count variation, however, at low loads, they yield poorer blocking performance than unprotected deflection.

In conclusion, we remark that this paper provides strong evidence recommending that OBS using deflection routing should be given some form of protection to guard against destabilization resulting from upward load variations. Our empirical results reveal that in terms of blocking performance and insensitivity to variation in hop-count, preemptive priority is the best form of protection for OBS. The chief contribution of this paper is our reduced-load approximation, which provides a fast and versatile approach to provision capacity or evaluate the blocking performance of large OBS networks using deflection routing.

References

[1] J. M. Akinpelu, "The overload performance of engineered networks with hierarchial and nonhierarchial routing," in *Proc. ITC-10*, June 1983.

[2] 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 Communications Mag.*, vol. 2, no. 2, pp. 82-89, Feb. 2002.

[3] N. Barakat and E. H. Sargent, "Dual-header optical burst switching: a new architecture for WDM burst-switched networks," in *Proc. IEEE INFOCOM'05*, March 2005.

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

[5] D. Blumenthal, P. Prucnal and J. Sauer, "Photonic packet switches - architectures and experimental implementations," *IEEE Proceedings*, vol 82, Nov. 1994, pp. 1650-1667.

[6] F. Borgonovo, L. Fratta and J. A. Bannister, "On the design of optical deflection-routing networks," in *Proc. INFOCOM'94*, vol. 1, June 1994, pp. 120-129.

[7] C. Cameron, A. Zalesky and M. Zukerman, "Prioritized deflection routing in optical burst switching networks," to appear in *IEICE Communication Trans.*

[8] W. S. Chan, "Recursive algorithms for computing end-to-end blocking in a network with arbitrary routing plan," *IEEE Trans. on Communications*, vol. COM-28, no. 2, Feb. 1980, pp. 153-164.

[9] Y. Chen, C. Qiao and X. Yu, "Optical burst switching (OBS): A new area in optical networking research," *IEEE Network*, vol. 18, no. 3, May-Jun., 2004, pp. 16-23.

[10] Y. Chen, H. Wu, D. Hu, C. Qiao, "Performance analysis of optical burst switched node with deflection rotuing," in *Proc. ICC'03*, vol. 2, May 2003, pp. 1355-1359.

[11] S. P. Chung, A. Kashper and K. W. Ross, "Computing approximate blocking probabilities for large loss networks with state-dependent routing," *IEEE/ACM Trans. on Networking*, vol. 1, no. 1, Feb. 1993, pp. 105-115.

[12] R. B. Cooper and S. Katz, "Analysis of alternate routing networks with account taken of nonrandomness of overflow traffic," Technical Report, Bell Telephone Lab. Memo., 1964.

[13] A. Detti, V. Earmo and M. Listanti, "Performance evaluation of a new technique for IP support in a WDM optical network: optical composite burst switching (OCBS)," *IEEE J. Lightwave Tech.*, vol. 20, no. 2, Feb. 2002, pp. 154–165.

[14] M. Dueser and P. Bayvel, "Analyisis of a dynamically wavelength-routed optical burst switched network architecture," *IEEE J. Lightwave Tech.*, vol. 20, April 2002, pp. 574-585.

[15] C. M. Gauger, M. Kohn and M. Scharf, "Comparison of contention resolution strategies in OBS network scenarios," in *Proc. of ICTON'04*, vol. 1, July 2004, pp. 18-21.

[16] A. Girard, *Routing and Dimensioning in Circiuit-Switched Networks*, Addison-Wesley, 1990.

[17] C. F. Hsu, T. L. Liu, N. F. Huang, "Performance analysis of deflection routing in optical burst switched networks," in *Proc. INFOCOM'02*, vol. 1, June 2002, pp. 55-73.

[18] A. Kaheel, H. Alnuweiri and F. Gebali, "Analytical evaluation of blocking probability in optical burst switching networks," in *Proc. IEEE ICC'04*, vol. 3, June 2004, pp. 1548-1553.

[19] J. Kaniyil, N. Hagiya, S. Shimamoto, Y. Onozato and S. Noguchi, "Structural stability aspects of alternate routing in non-hierarchical networks under reservation," *IEEE GLOBECOM'92*, vol. 2, Dec. 1992, pp. 797-801.

[20] F. P. Kelly, "Blocking probabilities in large circuit-switched networks," *Advances in Applied Probability*, vol. 18, pp. 473-505, 1986.

[21] ———, "Loss networks," *The Annals of Applied Probability*, vol. 1, no. 3, Aug. 1991, pp. 319-378,

[22] S. Kim, N. Kim and M. Kang, "Contention resolution for optical burst switching networks using alternate routing," in *Proc. IEEE ICC'02*, vol. 5, April 2002, pp. 2678-2681.

[23] S. K. Lee, K. Siriam, H. S. Kim and J. S. Song, "Performance improvement of deflection routing in optical burst switching networks," *Lecture Notes in Computer Science*, vol. 2659, Aug. 2003, pp. 445-452.

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

[25] P. M. Lin, B. J. Leon and C. R. Stewart, "Analysis of circuit-switched networks employing originating-office control with spill-forward," *IEEE Trans. on Communications*, vol. COM-26, no. 6, June 1978, pp. 754-765.

[26] Y. Liang, K. Liao, J. W. Roberts, and A. Simonian, "Queueing models for reserved set up telecommunications services," in *Proc ITC-12*, June 1988.

[27] J. Y. Wei and R. I. McFarland, "Just-in-time signaling for WDM optical burst switching networks," *IEEE J. Lightwave Tech.*, vol. 18, no. 12, Dec. 2000, pp. 2019-2037.

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

[29] C. Qiao, "Labeled optical burst switching for IP-over-WDM integration," *IEEE Communications Mag.* vol. 38, Sep. 2000, pp. 104-114.

[30] S. Qu, "A new recursive algorithm for calculating end-to-end blocking probability with an arbitrary fixed nonhierarchical routing," *IEEE Trans. on Communications*, vol. 45, no. 12, Dec. 1997, pp. 1517-1519.

[31] Z. Rosberg, H. L. Vu and M. Zukerman, "Performance evaluation of optical burst switching networks with limited wavelength conversion," in *Proc. ONDM'03*, vol. 2, Feb. 2003, pp. 1155-1169.

[32] Z. Rosberg, H. L. Vu, M. Zukerman and J. White, "Performance analyses of optical burst-switching networks," *IEEE J. Selected Areas Communications*, vol. 21, no. 7, Sept. 2003, pp. 1187 - 1197.

[33] H. Shimonishi, T. Takine, M. Murata and H. Miyahara, "Performance analysis of fast reservation protocol with generalized bandwidth reservation method," in *Proc. IEEE INFOCOM'96*, vol. 2, March 1996, pp. 758-767.

[34] P. Taylor and R. Maillardet, "Queues with reservations," presented at *ANZIAM'05*.

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

[36] E. Varvarigos and V. Sharma, "The ready-to-go virtual circuit protocol: a loss-free protocol for multigigabit networks using FIFO buffers," *IEEE/ACM Trans. on Networking*, vol. 5, no. 5, Oct. 1997, pp. 705-718.

[37] V. M. Vokkarane, G. P. V. Thodime, V. U. B. Challagulla, and J. P. Jue, "Channel scheduling algorithms using burst segmentation and FDLs for optical burst-switched networks," in Proc. IEEE ICC'03 vol. 2, May 2003, pp. 1443-1447.

[38] V. M. Vokkarane and J. P. Jue, "Prioritized burst segmentation and composite burst-assembly techniques for QoS support in optical burst-switched networks," *IEEE J. Selected Areas Communications*, vol. 21, no. 7, Sept. 2003, pp. 1198 - 1209.

[39] S. Verma, H. Chaskar and R. Ravikanth, "Optical burst switching: a viable solution for the terabit IP backbone," *IEEE Network*, vol. 14, no. 6, Nov. 2000, pp. 48-53.

[40] X. Wang, H. Morikawa and T. Aoyama, "Deflection routing protocol for burst switching WDM mesh networks," in *Proc. SPIE/IEEE Terabit Optical Networking*, Nov. 2000, pp. 242-252.

[41] W. Whitt, "Blocking when service is required from several facilities simulataneously," *AT&T Technical Journal*, vol. 64, 1985, pp. 1807-1856.

[42] I. Widjaja, "Performance analysis of burst admission-control protocols," *IEE Proc.-Communications*, vol. 142 no. 1, Feb. 1995, pp. 7-14.

[43] J. Xu, C. Qiao, J. Li and G. Xu, "Efficient burst scheduling algorithms in optical burst-switched networks using geometric techniques," *IEEE J. Selected Areas Communications*, vol. 22, no. 9, Nov. 2004, pp. 1796-1811.

[44] S. Yao, B. Muhkerjee, S. J. B. Yoo and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *IEEE J. of Lightwave Tech.*, vol .21, no .3, March 2003, pp. 672-683.

APPENDIX

In this appendix, the details of simplifying the expression for $\beta_m(n)$ from (14) to (15) are shown (see sideways inset). The second equality is because $r_n \notin d_1(n), \ldots, d_{m-1}(n)$ by definition, while the third equality is an immediate consequence of Fact 1.

$$
\begin{aligned}
\beta_m(n) &= \mathbb{P}\big(\text{not blocked at } d_1(n), \ldots, d_{m-1}(n) \,|\, \text{blocked at } r_n \cap \text{not blocked at } r_1, \ldots, r_{n-1}\big) \\
&= \frac{\mathbb{P}\big(\text{not blocked at } d_1(n), \ldots, d_{m-1}(n) \cap \text{not blocked at } r_1, \ldots, r_{n-1}\big)\mathbb{P}\big(\text{blocked at } r_n\big)}{\mathbb{P}\big(\text{not blocked at } r_1, \ldots, r_{n-1}\big)\mathbb{P}\big(\text{blocked at } r_n\big)} \\
&= \frac{\mathbb{P}\big(\text{not blocked at } d_1(n), \ldots, d_{m-1}(n)\big)\mathbb{P}\big(\text{not blocked at } r_1, \ldots, r_{n-1} \notin \Omega_m(n)\big)}{\mathbb{P}\big(\text{not blocked at } r_1, \ldots, r_{n-1}\big)} \\
&= \frac{\mathbb{P}\big(\text{not blocked at } d_1(n), \ldots, d_{m-1}(n)\big)}{\mathbb{P}\big(\text{not blocked at } r_1, \ldots, r_{n-1} \in \Omega_m(n)\big)} \\
&= \frac{\big(1 - q_{d_1(n)}\big) \cdots \big(1 - q_{d_{m-1}(n)}\big)}{\prod_{l \in \Omega_m(n)} \big(1 - b_l\big)}
\end{aligned}
$$