Delay Analysis of Optical Burst Switching Networks

Hai Le Vu, Andrew Zalesky, Moshe Zukerman, Zvi Rosberg, Jun Guo and Tai W. Um

Abstract— **This paper proposes a new optical burst switching (OBS) paradigm known as dynamic two-way reservation OBS (DTWR/OBS), in which the burst length is dynamically determined at an ingress router, according to the minimum length reservation period available on each wavelength channel within a lightpath, from the ingress router to the egress router. The dynamic nature of DTWR/OBS ensures the packet transmission rate is controlled at the ingress routers, based on the level of network congestion. This increases network utilization and eliminates burst blocking due to wavelength contention. An analytical model for a single isolated optical cross-connect (OXC) is derived to determine the waiting time distribution of an arbitrary IP packet at the ingress router using the proposed DTWR/OBS. The accuracy of the analytical model is confirmed by simulations.**

I. INTRODUCTION

Optical Burst Switching (OBS) [5], [6], [9] encompasses a broad range of experimental switching paradigms available to support Internet Protocol (IP) over wavelength division multiplexing (IP-over-WDM) in future all-optical telecommunications networks. An all-optical network consists of optical cross-connects (OXCs) interconnected via fibers that carry hundreds of wavelength channels. IP edge routers with buffering capabilities are located at the perimeter of the network. IP packets arriving at an ingress edge router are transported across the network to an egress edge router in accordance with the protocols governing the deployed switching paradigm.

Many types of OBS paradigms are available for deployment in future all-optical networks. The concept underpinning most types of OBS is as follows. IP packets destined for a common egress router are assembled into data *bursts* at an ingress router. A *control packet* precedes each burst by a time *offset*; the control packet is electronically processed at a sequence of OXCs to reserve a contiguous sequence of wavelength channels, known as a *lightpath* or *route*, for the pending burst. Then, depending on the type of OBS deployed, the pending burst is entirely switched in the optical domain either *with* or *without* acknowledgement; that is, two-way or one-way reservation, respectively. OBS types using one-way reservation, such as just-enough-time (JET) [5], just-in-time (JIT) [11] and Horizon [8], send a burst before it is confirmed that a lightpath can be reserved, from the ingress router to the egress router. Conversely, OBS types using two-way reservation, such as wavelength-routed OBS (WR/OBS) [2], delay a burst at the ingress router until an acknowledgement propagates from the egress router to the ingress router, thus confirming a successful lightpath reservation. These OBS types overcome burst blocking arising from wavelength contention since if the control packet fails to reserve a lightpath, the burst is electronically buffered at the ingress router and rescheduled for later transmission. However, two-way reservation prolongs the packet delay since the burst must remain idle until an acknowledgement is received.

This work was supported by the Australian Research Council.

H. L. Vu, A. Zalesky, J. Gou and M. Zukerman are with the ARC Special Research Center for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, The University of Melbourne, Melbourne, VIC 3010, Australia. CUBIN is an affiliated program of National ICT Australia.

T. W. Um is with the Information and Communications University, P. O. Box 77, Yusong, Taejon 305-348 Korea.

Z. Rosberg is with the Department of Communication Systems Engineering, Ben Gurion University, Beer-Sheva, 84105, Israel.

Extensive analysis and performance evaluation of OBS types using one-way reservation is presented in [1], [5], [7]. The mean burst blocking probability is considered the most important measure of quality of service (QoS) in OBS types using one-way reservation. For OBS types using two-way reservation, the most important measures of QoS is queuing delay at ingress routers. WR/OBS is the only OBS type using two-way reservation for which performance evaluation is considered, and presented in [2].

For all current OBS types, using both one-way and two-way reservation, the burst length is determined unilaterally at the ingress router. Once the burst length is determined, the control packet must reserve a reservation period of length at least equal to the burst length on each wavelength channel within a lightpath, from the ingress router to the egress router, otherwise, the pending burst is blocked. Note that the precise length of the reservation period depends on the reservation mechanism adopted, however, for both the immediate and delayed mechanisms, the length of the reservation period must be *at least* equal to the burst length. For OBS types using one-way reservation, such a strict requirement on the length of the reservation period gives rise to a high burst blocking probability, and for OBS types using two-way reservation, this requirement prolongs the delay of packets, as a consequence of burst rescheduling.

To reduce the level of burst rescheduling, it appears reasonable to reduce the effective length of a burst to a length which is equal to the *minimum* available resource within a lightpath. Consequently, only a minimal portion of the burst is rescheduled for transmission at a later time, thus reducing the mean waiting time of packets and increasing network utilization. Indeed a somewhat similar approach is adopted in an OBS type known as burst segmentation OBS (BS/OBS) [10]. However, since BS/OBS is an OBS type using one-way reservation, the length of a burst must be reduced at

0-7803-8938-7/05/$20.00 (C) 2005 IEEE

each OXC and the portion of the burst that cannot be sent is simply blocked.

This paper presents a new OBS type known as *dynamic two-way reservation* OBS (DTWR/OBS), in which the burst length is dynamically determined at an ingress router, according to the minimum length reservation period available on each wavelength channel within a lightpath. In particular, the burst length is reduced as the level of congestion is increased, and thus the packet transmission rate is appropriately regulated via the minimum length reservation period.

In a DTWR/OBS network there is no loss due to reservation failure since a burst is sent at a confirmed time with an appropriate length that fits into the reserved period. The cost of this no-blocking network is a packet delay at the ingress router. The focus of this paper is to derive an analytical model for a single isolated OXC, which is used to determine the waiting time distribution of an arbitrary packet at the ingress router using DTWR/OBS.

II. Dynamic Two-Way Reservation OBS

The operation of DTWR/OBS is fundamentally different from other OBS types using two-way reservation. To clarify the operation of DTWR/OBS, consider the signaling that is required between the ingress router and the egress router before a burst can be sent. It is assumed that all OXCs use a common timing, and the propagation delay between OXCs is fixed and known. The signaling cycle is described as follows. IP packets arriving at an ingress router are stored in an electronic buffer. Let time be divided into consecutive fixed length intervals each of size T time units. At the end of each time interval, the ingress router sends a *reservation packet* to the egress router. The time interval T is a design parameter controlling the waiting time of IP packets.

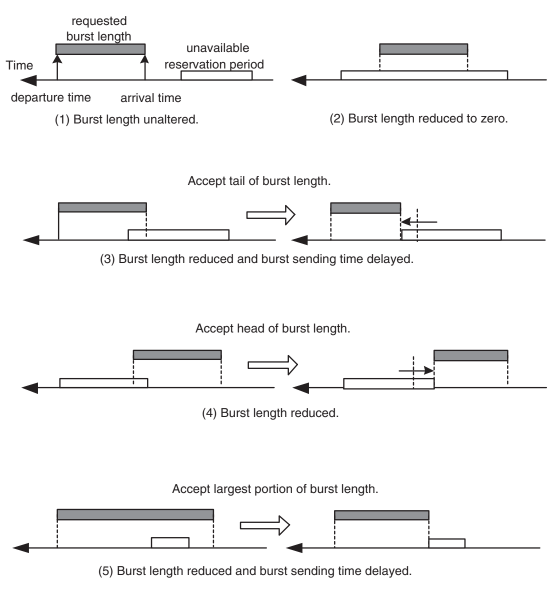

Fig. 1. Reduction of the effective burst length based on the length of the most appropriate reservation period available on outgoing wavelength channels.

The reservation packet contains the information of an ingress/egress address, the scheduled burst sending time, and most importantly, the requested burst length. As shown in Fig. 1, at each OXC along the route, from the ingress router to the egress router, the effective burst length, also called the *acknowledged reservation period*, is reduced based on the length of the most appropriate reservation period available on outgoing wavelength channels. The five possible scenarios in which the length of a requested burst can be reduced, or unaltered, are as follows.

1) If a reservation period is available from the burst arrival time to the burst departure time, it is reserved and the reservation packet is forwarded to the subsequent OXC. The requested burst length is unaltered.

2) If a reservation period is unavailable, the OXC returns a *withdraw packet* to the ingress router and the pending burst is rescheduled for transmission with the next burst. The signaling cycle for the next burst will start at a time T after the commencement of the signaling cycle for the current burst.
3) If the most appropriate reservation period can only accommodate the tail of the burst, the burst sending time is delayed and the requested burst length is reduced. Then, the reservation period is reserved and the reservation packet is forwarded to the subsequent OXC with the adjusted burst sending time and the reduced burst length.

4) If the most appropriate reservation period can only accommodate the head of the burst, the requested burst length is reduced. Then, the reservation period is reserved and the reservation packet is forwarded to the subsequent OXC with the reduced burst length.

5) If reservation periods are available to accommodate both the head and the tail of the burst but not the middle portion, the longest available reservation period is reserved, corresponding to either the head or the tail. Similarly, the reservation packet is forwarded to the subsequent OXC with the adjusted burst sending time and the reduced burst length.

As the reservation packet traverses a route to the egress router, the burst sending time may be increased and the requested burst length may be reduced, based on the length of the most appropriate reservation period available on outgoing wavelength channels. Consequently, OXCs located immediately after the ingress router typically reserve reservation periods of longer length in comparison with those OXCs located immediately before the egress router. This over-provisioning of reservation lengths, can be amended by sending a *confirmation packet*, from the egress router to the ingress router.

Upon receiving the reservation packet, the egress router sends a confirmation packet to the ingress router via the same route traversed by the reservation packet with a final *acknowledged reservation period* – ARP, and *acknowledged sending time* – AST of the burst.

When an OXC receives the confirmation packet, it compares the AST and ARP values with the values initially allocated while processing the reservation packet. The over-provisioned portion of reservation periods can thus be released. When the ingress router receives the confirmation packet, it sends a portion of the burst with length equals to the ARP value at the time determined by AST value received from the confirmation packet. The remaining portion of the burst is rescheduled for transmission with the next signaling cycle.

III. Modelling and Analysis

To facilitate an amenable analytical model, consider M ingress routers connected to a single isolated OXC serviced by K, $K < M$, wavelength channels. Let time be divided into consecutive fixed length intervals each of size T time units. To ensure the fairness between different users all M ingress routers are synchronized in time; that is, all buffers simultaneously send a reservation packet at the end of a time interval.

After a round-trip propagation delay, T_{RTP}, a confirmation packet is received by each of the M ingress routers and the transmission period commences. To simplify the analytical model, assume $T_{RTP} = 0$. (Note that this assumption can be relaxed if we assume that the ingress router sends the request T_{RTP} time units before an end of the time interval.) This way, each time interval can be considered as a signaling cycle; namely, at the end of each interval, all M ingress routers request a wavelength resource to transmit all their enqueued packets, and immediately receive a notification of how much they can transmit and when.

In particular, at time kT, an ingress router which has j packets in its queue will ask to transmit these j packets as a single burst within $(kT, kT + (\Delta T)_j)$ where $(\Delta T)_j$ is the time required to transmit the j packets. A notification (acknowledgement) will be immediately received by the ingress router of the ARP within $\big(kT, kT + (\Delta T)'_j\big)$ where $0 \leq (\Delta T)'_j \leq (\Delta T)_j$. If $(\Delta T)'_j = 0$, then the reservation fails for this cycle.

Note that as a consequence of synchronization, reservation packets cannot be processed in the order at which they arrive at the OXC since all arrive at precisely the same time. Therefore, the order that reservation packets are processed is randomly selected. (In practice, DTWR/OBS does not mandate a synchronized signaling cycle for all ingress routers.) The order that the reservations are processed will henceforth be called *random allocation order*.

Assume that packets arrive at each ingress router according to an independent Poisson process with rate λ packets per time unit. By normalizing the packet transmission rate, for a wavelength channel, to one packet per time unit, it follows that a maximum of T packets can be sent during each time interval T. Let B, $B \leq T$, be the maximum number of packets an ingress router can store in its buffer. If there are fewer than B packets in the buffer, an arriving packet is admitted, otherwise it is lost. In the numerical evaluation presented in Section IV, we set $B = T$. According to this normalization, we have $(\Delta T)_j = j \leq T$, and $(\Delta T)'_j = i$, $0 \leq i \leq j$ is the number of packets transmitted out of the j requested packets.

Consider an arbitrary edge router. At the end of a time interval T, the buffer will contain j, $0 \leq j \leq B$, packets. The confirmation packet will confirm that i, $0 \leq i \leq j$, of the buffered packets can be sent. The remaining $j - i$ packets will remain in the buffer and be considered for transmission in the subsequent signaling cycle. Suppose that with probability $d(i, j)$, i of the j packets can be sent on a first come first served (FCFS) basis. For $0 \leq j \leq B$, the probabilities $\{d(i, j)\}$ satisfy $\sum_{i=0}^{j} d(i, j) = 1$.

A packet admitted during the time interval $[nT, (n+1)T]$, $n = 1, 2, 3, \ldots$, which represents signaling cycle n, may be scheduled to send at one of the subsequent frames $[kT, (k+1)T]$, $k \geq n + 1$. Therefore, if a packet is admitted at time $(n+1)T - u$, $0 \leq u \leq T$, its waiting time until being scheduled for transmission at the ingress router is defined as the sum of u and $(k - n - 1)T$.

Let J_n be the number of packets in the buffer at time nT^- just before the end of the time interval when a request is sent. $\{J_n\}$ is then a Markov chain with state space $\{0, 1, \ldots, B\}$. Its probability transition matrix $\{P_{ij}\}$, $0 \leq i, j \leq B$, is given by

$$P_{i,j} = P(J_{n+1} = j | J_n = i) = \begin{cases} \displaystyle\sum_{\nu=\max(0,i-j)}^{i} d(\nu, i)\, e^{-\lambda T} \frac{(\lambda T)^{j-i+\nu}}{(j-i+\nu)!}, & 0 \leq j \leq B-1; \\ \displaystyle\sum_{\nu=0}^{i} d(\nu, i) \left[1 - \sum_{q=0}^{B-i+\nu-1} e^{-\lambda T} \frac{(\lambda T)^q}{q!}\right], & j = B. \end{cases} \quad (1)$$

The stationary probabilities, $[\pi_0, \pi_1, \ldots, \pi_B]$, for $\{J_n\}$ can be easily computed.

Given that each ingress router evolves independently, the probability that an arbitrary ingress router contains at least one packet at time kT^- is given by $\overline{P} = 1 - \pi_0$. The probability of packet loss due to buffer overflow at the edge router is quantified by $1 - S/(\lambda T)$, where $S = \sum_{j=1}^{B} \pi_j \sum_{i=1}^{j} i d(i, j)$ is the expected number of packets served (and also is equivalent to the mean amount of packets admitted) during one signaling cycle.

We will now derive an expression for the probabilities $\{d(i, j)\}$, which may be solved iteratively using a successive substitution algorithm. Given the probabilities $\{d(i, j)\}$, the analysis of [3] can be applied to determine the queueing waiting time distribution of an arbitrary packet from the time of arriving until the time it has been scheduled for transmission.

A. Queueing Delay Distribution

Some preliminary probabilities require consideration before an expression can be provided for $\{d(i, j)\}$. They are as follows.

At time kT^-, given that an arbitrary ingress router contains $j > 0$ packets, this tagged router can either

- send all j packets with probability one if there are strictly less than K other *active* ingress routers containing one or more packets. The probability that there are strictly less than K active ingress routers at time kT^{-}, is given by

$$\Gamma = \sum_{k=0}^{K-1} \binom{M-1}{k} \overline{P}^{k} (1-\overline{P})^{M-1-k}. \quad (2)$$

- send all j packets with probability $\frac{K}{K+s+1}$, given that $K+s,\ 0 \le s \le M-K-1$, other active ingress routers at time kT^{-}. The probability that there are $K+s$ other active ingress routers present, is given by

$$\Gamma_s = \binom{M-1}{K+s} \overline{P}^{K+s} (1-\overline{P})^{M-1-K-s}. \quad (3)$$

  Note that $\frac{K}{K+s+1}$ is the probability that a tagged router is among the K "lucky" routers out of the $K+s+1$ active routers who have less than K active routers ahead of them in the random allocation order.
- send $i,\ \ 0 \le i < j$, packets with probability denoted $\mathcal{P}(s,i,j)$, given that there are $K+s,\ 0 \le s \le M-K-1$, other ingress routers contain one or more packets at time kT^{-}. In the following, the probability $\{\mathcal{P}(s,i,j)\}$ are specified.

For any $m,\ 0 \le m \le s$, the probability that a tagged router sees exactly $K+m$ out of the total $K+s$ active routers ahead of it in the random allocation order is $\frac{1}{K+s+1}$ (independent of m).

Let $Y_1, Y_2, \ldots, Y_{K+v},\ 0 \le v \le m$ be the random variables representing the number of packets that have been allocated for transmission among those $K+m$ active routers. Note that for any s value if $m = 0$ then $v = 0$; i.e. the tagged router sees K out of the total $K+s+1$ active routers ahead of it in the random allocation order with $Y_1, Y_2, \ldots, Y_K$ packets have been already allocated for transmission.

Let η_j be the stationary probabilities $\pi_j,\ j = 1, 2, \ldots, B$ of the Markov chain $\{J_n\}$ conditional upon $j > 0$. Thus, $Y_1, Y_2, ..., Y_K$ are independent discrete random variables and identically distributed according to η_j; i.e., $P(Y_i = j) = \eta_j, i = 1, 2, ..., K$.

The distribution of $Y_{K+v}, 1 \le v \le m$, is not distributed according to η_j, in fact its tail is heavier than η_j. However, to simplify the model by recasting it in the form of an order statistic problem, it is assumed that $Y_{K+v}, 1 \le v \le m$, is in fact distributed according to η_j. Hence, $Y_1, Y_2, ..., Y_{K+v}$ is a sequence of i.i.d. random variables, which as will be shown is the exact form required to consider the use of order statistics. The error introduced by this assumption is quantified in Section IV.

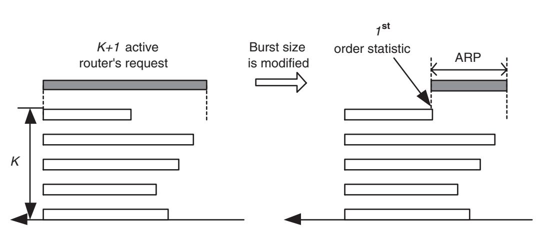

Fig. 2. Allocation of *acknowledged reservation period* (ARP) to the $K+1$ active router

As shown in Fig. 2, upon its request arrival, if a tagged router contains j packets and sees $K, (m = 0)$ other active routers ahead of it in the random allocation order, it will be allocated

$$i = (j - X_K(1))^{+} \quad (4)$$

packets, where $X_K(1) = \min\{Y_1, Y_2, \ldots, Y_K\}$ and $(X)^{+} \equiv max(X, 0)$. The random variable $X_K(1)$ is also known as the 1^{st} order statistic of $K,\ \{Y_1, Y_2, \ldots, Y_K\}$, discrete random variables.

In general, the probability density function (pdf) $f_{X_K(m)}(x),\ 1 \le x \le B$, where $X_K(m)$ is the m^{th} order statistic of $K,\ \{Y_1, Y_2, \ldots, Y_K\}$, discrete random variables, is given by (5) below.

Observe that in (4), if $X_K(1) \ge j$ then no packets of the tagged router will be allocated for transmission. The probability of this even is given by

$$q(1) = \sum_{x=j}^{B} f_{X_K(1)}(x). \quad (6)$$

Similarly, if a tagged router sees $K+1, (m = 1)$ other active routers ahead of it in the random allocation order, it will be then allocated

$$i = \begin{cases} (j - X_K(1))^{+} & \text{with prob. (w.p.)} \quad q(1); \\ (j - X_{K+1}(2))^{+} & \text{w.p.} \quad p(1) = 1 - q(1), \end{cases}$$

where $X_{K+1}(2) = \min\{\{Y_1, Y_2, \ldots, Y_{K+1}\} - \{X_{K+1}(1)\}\}$.

The random variable $X_{K+1}(2)$ is the 2^{nd} order statistic of $K+1,\ \{Y_1, Y_2, \ldots, Y_K, Y_{K+1}\}$, discrete random variables.

Let us introduce the following notation extending (6)

$$q(m+1) = 1 - p(m+1) = \sum_{x=j}^{B} f_{X_{K+m}(m+1)}(x), 0 \le m \le s. \quad (7)$$

In general, if a tagged router sees $K+m,\ 0 \le m \le s$, other active routers ahead of it in the random allocation order, it will be then allocated i packets equals to

$$\begin{cases} (j - X_K(1))^{+} \quad \text{w.p.} \quad q(1)^m; \\ (j - X_{K+1}(2))^{+} \quad \text{w.p.} \quad p(1) \sum_{k=0}^{m-1} q(1)^k q(2)^{(m-1)-k}; \\ \ldots \\ (j - X_{K+h-1}(h))^{+} \quad \text{w.p.} \quad \Psi(m,h); \\ \ldots \\ (j - X_{K+m}(m+1))^{+} \quad \text{w.p.} \quad \prod_{k=1}^{m} p(k), \end{cases} \quad (8)$$

where

$$X_{K+m}(m+1) = \min\{\{Y_1, Y_2, \ldots, Y_{K+m}\} - \{X_{K+m}(1), X_{K+m}(2), \ldots, X_{K+m}(m)\}\}, \quad (9)$$

and

$$1 \le \forall h \le m+1, \quad \Psi(m,h) = \prod_{i=1}^{h-1} p(i) \sum_{l_1+l_2+\ldots+l_h = m+1-h} q(1)^{l_1} q(2)^{l_2} \ldots q(h)^{l_h}, \quad (10)$$

$$f_{X_K(m)}(x) = \begin{cases} \sum_{u=m}^{K} \binom{K}{u} \eta_x^u \left(\sum_{v=2}^{B} \eta_v \right)^{K-u}, & x = 1; \\ \sum_{u=1}^{K} \binom{K}{u} \eta_x^u \left(\sum_{r=(m-u)^+}^{\min\{m-1,K-u\}} \binom{K-u}{r} \left(\sum_{v=1}^{x-1} \eta_v \right)^r \left(\sum_{w=x+1}^{B} \eta_w \right)^{(K-u-r)^+} \right), & x = 2, \ldots, B-1; \\ \sum_{u=1}^{m} \binom{K}{u-1} \eta_x^{K-u+1} \left(\sum_{v=1}^{B-1} \eta_v \right)^{u-1}, & x = B. \end{cases} \quad (5)$$

the notation $\sum_{l_1+l_2+...+l_h=m+1-h}$ is to be understood as the sum over all combinations of $l_1, l_2, ...l_h$, such that $l_i, i = 1, 2, ..., h$ is a nonnegative integer and $l_1 + l_2 + ... + l_h = m + 1 - h$. The random variable $X_{K+m}(m+1)$ is the $(m+1)^{th}$ order statistic of $K + m$, $\{Y_1, Y_2, \ldots, Y_{K+m}\}$, discrete random variables. Knowing the probabilities η_j, the probability distribution function of $X_{K+m}(m+1)$ is determined.

From (8) for any m and h, the tagged router can transmit $i = j - X_{K+h-1}(h) > 0$ with probability $\Psi(m,h) f_{X_{K+h}(h+1)}(j-i)$ and $i = 0$ with probability $\Psi(m,h) q(h)$.

Using (5), (7), (8) and (10) the probability $\mathcal{P}(s,i,j)$ is given by

$$\mathcal{P}(s,i,j) = \begin{cases} 1, & i = j = 0; \\ \Theta \sum_{m=0}^{s} \sum_{h=1}^{m} \Psi(m,h) q(h), & i = 0 < j; \\ K\Theta, & i = j \neq 0; \\ \Theta \sum_{m=0}^{s} \sum_{h=1}^{m} \Psi(m,h) f_{X_{K+h}(h+1)}(j-i), & 0 < i < j, \end{cases} \quad (11)$$

where $\Theta = \Gamma_s/(K+s+1)$.

Based on (11), it follows that the probabilities $\{d(i,j)\}$ are given by

$$d(i,j) = \begin{cases} 1, & i = j = 0; \\ \Gamma + \sum_{s=0}^{M-1-K} \mathcal{P}(s,i,j), & i = j \neq 0; \\ \sum_{s=0}^{M-1-K} \mathcal{P}(s,i,j), & 0 \leq i < j. \end{cases} \quad (12)$$

Combining (1), (11) and (12) gives rise to a coupled system of nonlinear algebraic equations. A unique solution to this coupled system is termed the fixed-point (FP) and represents the true values of the probabilities $\{d(i,j)\}$. The successive substitution algorithm detailed below is an efficient method that may find the FP. The superscript $t = 1, 2, \ldots$ affixed to a variable is used to denote the value of that variable at iteration t.

ALGORITHM: SUCCESSIVE SUBSTITUTION

- **0) Initialize**: For every $j \leq B$ and $i \leq j$ set $\{d(i,j)\}$ to some arbitrary distribution on $[0,1]$.
- **1) Compute**: Set $t = t + 1$. Compute the new steady-state probabilities $[\pi_0^t, \pi_1^t, \ldots, \pi_B^t]$ based on (1) with $\{d(i,j)\}^{t-1}$. Compute the new probabilities $\{\mathcal{P}(s,i,j)\}^t$ based on (11) with $[\pi_0^t, \pi_1^t, \ldots, \pi_B^t]$.
- **2) Update**: Update the probabilities $\{d(i,j)\}^t$ based on (12) with $\{\mathcal{P}(s,i,j)\}^t$. If $\{d(i,j)\}^t$ is sufficiently close to $\{d(i,j)\}^{t-1}$, then stop. Otherwise go to step 1.

Once the probabilities $d(i,j)$ are determined, the queueing waiting time distribution $F_\theta(\tau), \tau \geq 0$ of an arbitrary packet from the time it arrives until the time it is scheduled for transmission can be obtained by a complex accounting of the queueing process on successive time intervals. That fairly involved derivation was presented in [3] and will not be repeated here for brevity.

B. Total Delay Distribution

To determine the remaining delay distribution of an arbitrary packet from the time it is scheduled for transmission until its successful transmission, let the random variable Z represents the delay between packet's scheduling and its transmission.

Recall that at time kT^- when the scheduling takes place we know the stationary probabilities $\pi_j, j = 0, .., B$ that there are j packets in the buffer. Condition upon $j > 0$ we have the stationary probabilities η_j whose cumulative distribution function is denoted by $F_J(j), j > 0$. Using π_j and the $\{d(i,j)\}$ values calculated from the above successive algorithm we can compute the probability $P(Y = i > 0)$ that an arbitrary router will be allocated $i > 0$ of its own packets for transmission. Let $F_Y(y)$ be the cumulative distribution function (cdf) of Y. Depending on the number of active routers and how many of its packets are allocated for transmission the remaining delay in question is different. Accordingly in the following we consider two cases where the calculation is slightly different.

i) Case 1: if the tagged router sees $i, i \leq K$, other active routers ahead of it in a random allocation order, then all its

j packets will be scheduled to transmit in one batch during a total j unit of time. The remaining time in question of an arbitrary packet until its transmission is the same as its position in the batch. Let $P(U = n), n = 1, .., j \le T$, be the probability that a randomly selected packet is the n^{th} in its own batch. This probability is given by a forward recurrence time in renewal theory concept [4] as follows.

$$P(U = n) = (1 - F_J(k-1))/E[J], \tag{13}$$

where $E[J]$ is the expected value of J random variable. In this case the pdf $f_Z^i(n), 1 \le i \le K$ of the random variable Z is just $P(U = n)$.

ii) Case 2: if the router sees $K + m, m = 0, 1, .., M - K - 1$ other active routers ahead of it in a random allocation order, then the remaining delay is a summation of the above defined forward recurrence time U and the h^{th} order statistic $X_{K+h-1}(h), 1 \le h \le m+1$, random variable defined in (8). Furthermore, in this case the forward recurrence time distribution is determined based on (13) using $F_Y(y)$ rather than $F_J(j)$, since the tagged router can at most be allocated $(j - X_{K+h-1}(h))^+ < j$ packets for transmission. The pdf $f_Z^{K+m}(n)$ of the random variable Z for a given $(K + m)$ number of active routers is a convolution between the pdf of the two random variables U and $X_{K+h-1}(h)$ above.

Knowing $f_Z^i(n) \quad (i = 1, .., K, .., K + m)$ for all possible $0 \le m \le M - K - 1$, and $1 \le h \le m + 1$ values the general pdf $f_Z(n) = P(Z = n)$ is determined.

Finally the continuous cdf function of the total delay D of an arbitrary packet is given by.

$$\begin{aligned} F_D(t) &= 1 - P(D \ge t) \\ &= 1 - \sum_{n=0}^{T} P(Z = n)[1 - F_\theta(t - n)], \end{aligned} \tag{14}$$

where $F_\theta(t)$ is a continuous cdf of packet's delay until its scheduling for transmission, and $P(Z = n)$ is a discrete pdf from the moment of scheduling until its successful transmission.

IV. Numerical Evaluation

The objective of this section is twofold: to quantify the error introduced by the assumptions in the analytical model using simulation; and, to evaluate the waiting time distribution of an arbitrary packet at the ingress router using DTWR/OBS.

Numerical evaluation will be considered for five ingress routers connected to a single isolated OXC, which is serviced by three wavelength channels; i.e., $M = 5$ and $K = 3$. A maximum of 10 packets can be sent during each transmission period; i.e., $T = 10$, and each ingress router can store a maximum of 10 packets in its buffer; i.e., $B = 10$. In this example, instead of the exact value as presented in case 2 of the previous section, we compute the upper bound of the $f_Z^{K+m}(n), m = 0, 1, .., M - K - 1$ pdf function by setting $h = m + 1$ only, i.e. all the $(K + m)$ routers ahead of the tagged router in the random allocation order are allocated $i > 0$ of their own packets for transmission.

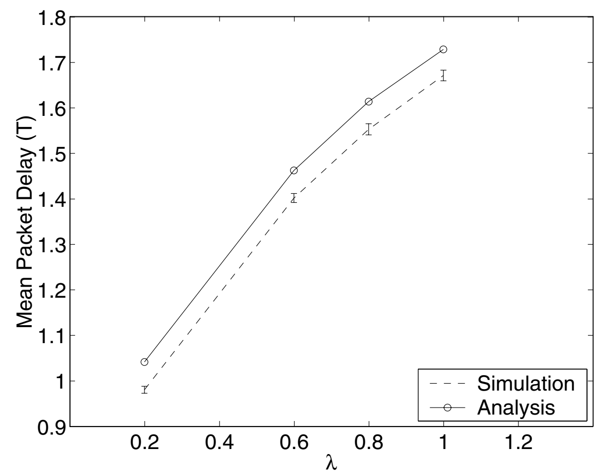

Fig. 3. Mean IP packet delay generated by analytical model and simulation with 95% confidence intervals.

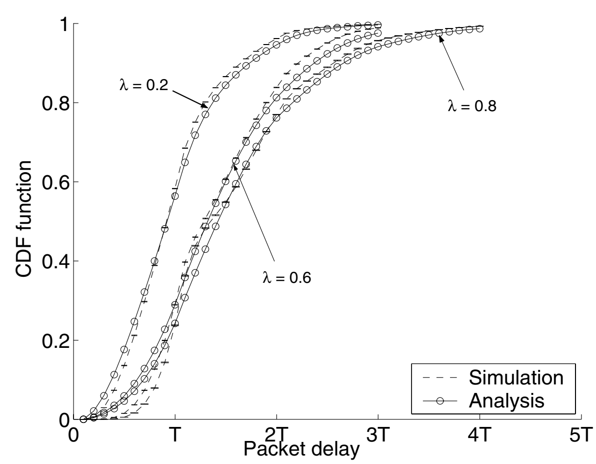

Fig. 4. Cumulative distribution function (cdf) for IP packet delay generated by analytical model and simulation with 95% confidence intervals.

Figure 3 shows the average packet delay calculated by the analytical model and simulation with 95% confidence intervals at different traffic loads ($\lambda = 0.2, 0.4, 0.6, 0.8, 1$).

Observe that the analytical values of the average delay are in good agreement with the simulation results and indeed provide a tight upper bound for the average packet's delay.

In Fig. 4 we plot the graphs of the corresponding delay cumulative distributions at several arrival rates obtained by computation and simulation. Our results demonstrate that the probability of an arbitrary packet waiting for more than two transmission periods is quite small, particularly for a low arrival rate.

Note that for IP traffic, packets that are lost due to burst contention in OBS types using one-way reservation will have to be retransmitted, which introduces perhaps much more delay at the application level due to the behavior of TCP protocol and the retransmission process involved. Therefore using DTWR/OBS may significantly reduce the overall delay of IP flows in an end-to-end application.

V. Conclusion

We proposed a new OBS paradigm, known as dynamic two-way reservation OBS (DTWR/OBS). Unlike current OBS types for which the burst length is determined unilaterally at the ingress router, in DTWR/OBS the effective burst length is dynamically determined at the ingress router, according to the minimum length reservation period available on each wavelength channel within a lightpath. The dynamic nature of DTWR/OBS increases network utilization and ensures the packet transmission rate is controlled at the ingress routers, based on the level of network congestion.

We have developed an analytical model to evaluate the waiting time distribution of an arbitrary packet at the ingress router in OBS networks using DTWR/OBS. The accuracy of the analytical model was confirmed by simulations. The numerical evaluation we considered, demonstrated that the probability of the waiting time for an arbitrary packet exceeding more than two transmission periods is small.

References

[1] K. Dolzer, C. Gauger, J. Spath and S. Bodamer, "Evaluation of reservation mechanisms for optical burst switching", *AEU Int. J. of Electron. and Commun.*, vol. 55, no. 1, 2001.

[2] M. Duser and P. Bayvel, "Analysis of a Dynamically Wavelength Routed Optical Burst Switched Network Architecture," *IEEE J. Lightwave Tech.*, vol. 20, no. 4, April 2002, pp. 574–585.

[3] M. F. Neuts, J. Guo, M. Zukerman, and H. L. Vu, "The Waiting Time Distribution for a TDMA Model with a Finite Buffer," *Proc. IEEE INFOCOM 2003*, April 2003, pp. 177–185.

[4] P. G. Potter and M. Zukerman, "Analysis of a Discrete Multipriority queuing System Involving a Central Shared Processor Serving Many Local Queues," *IEEE Jsac*, vol. 9, no. 2, Feb. 1991, pp.194-202.

[5] C. Qiao, 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.

[6] C. Qiao, "Labeled Optical Burst Switching for IP-over-WDM Integration," *IEEE Mag.*, Sept. 2000, pp. 104–114.

[7] Z. Rosberg, H. L. Vu, M. Zukerman, J. White, "Performance Analyses of Optical Burst Switching Networks", *IEEE Journal of Selected Areas in Communications*, vol. 21, no. 7, Sept. 2003, pp. 1187-1197.

[8] J. S. Turner, "Terabit Burst Switching," *Journal of High Speed Networks*, vol. 8, no. 1, Jan. 1999, pp. 3–16.

[9] S. Verma, H. Chaskar, R. Ravikanth, "Optical Burst Switching: A Viable for Terabit IP Backbone," *IEEE Network*, Nov./Dec. 2002, pp. 48–53.

[10] V. M. Vokkarane, J. P. Jue, S. Sitaraman, "Burst Segmentation: An Approach for Reducing Packet Loss in Optical Burst Switched Network," *Proc. IEEE ICC 2002*, New York, April 2002.

[11] J. Y. Wei and R. I. McFarland, "Just-In-Time Signaling for WDM Optical Burst Switching Networks," *IEEE J. of Lightwave Tech.*, vol. 18, no. 12, Dec. 2000, pp. 2109–2031.