

An Analytical Model for Shared Fiber-Delay Line Buffers in Asynchronous Optical Packet and Burst Switches

Tao Zhang, Kejie Lu, and Jason P. Jue

Department of Computer Science

The University of Texas at Dallas, Richardson, TX 75083

Abstract—Packet contention is a major issue in asynchronous optical packet and burst switching networks. Optical buffering, which is implemented by *fiber delay lines (FDLs)*, is fundamental to many optical switch implementations for resolving contention. Most existing optical buffering implementations are output-based and require a huge amount of FDLs as well as larger switch sizes, which impose extra cost on the overall system. In this paper, we consider shared optical buffering which can reduce the buffer size at a switch. Since no previous study is available to analyze the performance of asynchronous architectures with shared buffers, we propose an analytical model to evaluate the packet loss probability and the average delay for shared buffers at a single switch. We then compare the performance of output buffers to shared buffers under different granularities of FDLs. We observe that, by choosing an appropriate granularity, the shared buffering scheme can significantly reduce packet loss with much smaller switch sizes and fewer FDLs than the output buffering architecture. The accuracy of the analytical model is also confirmed by extensive simulation.

I. INTRODUCTION

Driven by increasing Internet traffic, next-generation communication networks are expected to provide huge bandwidth as well as support for diverse service demands. These requirements can be fulfilled by implementing Dense Wavelength Division Multiplexing (DWDM) systems, which can now provide more than 2 Tbit/s in a single fiber. To fully utilize the fiber capacity, packet-based optical packet switching (OPS) and burst switching (OBS) have been proposed as switching paradigms in the optical domain to eliminate the processing bottleneck in the electronic domain [1].

In general, packet-based optical networks can be divided into two categories: time-slotted/synchronous networks with fixed-length packets and unslotted/asynchronous networks with fixed-size or variable-size packets. In the literature, most early studies on packet switching technologies focused on bit-synchronous, fixed-size packet switching. However, in optical networks with extremely high data rates, it is very difficult to synchronize all incoming packets in the optical domain. On the other hand, it is more natural for variable-size packets to match varying IP packet sizes. For these reasons, we focus on asynchronous switching in this work.

A major issue in asynchronous packet switching networks is packet contention, which occurs when two or more incoming packets contend for the same output at the same time. In the

optical domain, approaches for resolving contention include wavelength conversion, deflection routing, and buffering.

In wavelength conversion, a contending packet can be converted from one wavelength to another in order to avoid conflict. In deflection routing, a packet is forwarded to an alternative output port if its primary output port is occupied by another packet. Thus, the links in the network are acting as dynamic buffers without increasing the cost of the switch. In buffering, contending packets are temporarily stored and are forwarded at a later time. In the optical domain, *fiber delay lines (FDLs)* [2], [3] are generally utilized to delay packets for a fixed amount of time. In this article, we study the performance of FDL buffers in asynchronous optical packet and burst switches.

In the literature, most existing optical buffering implementations are output-based, which means that a set of FDLs are dedicated to a specific output port. Since each port requires the same number of FDLs and each FDL requires one switch port, the output-based buffering scheme will require a huge amount of FDLs as well as a very large switching fabric, if the total number of switch ports is large. To reduce the overall system cost, shared buffering architectures are widely used in traditional switching architectures [4], [5]. In this work, we consider a well-known shared optical buffer architecture, SLOB [2], which has been proposed for providing various delays in the optical domain, and which consists of multiple smaller size switching elements instead of a single large switching fabric. The main contribution of our study is that we develop an accurate analytical model to evaluate the packet loss and delay performance of shared FDL buffers.

A fundamental difficulty in analyzing the performance of shared buffering is that FDL buffers can only provide a discrete set of delays. Therefore, the behavior of FDL buffers is essentially different from traditional electronic buffers. The performance of such an architecture cannot be captured accurately using conventional queuing models. To the best of our knowledge, no previous work has analyzed shared buffers with variable-size packets in asynchronous systems. The works in [4] and [5] analyze shared buffering, but both models analyze synchronous systems with fixed-size packets. In the literature, there are a few works on output FDL buffering in asynchronous optical switches [6]–[8]. In [6], the authors provided an approximate analytical model based on the $M/M/1$ queue with balking, which means that packets will not enter

This work was supported in part by the National Science Foundation (NSF) under grant ANI-01-33899.

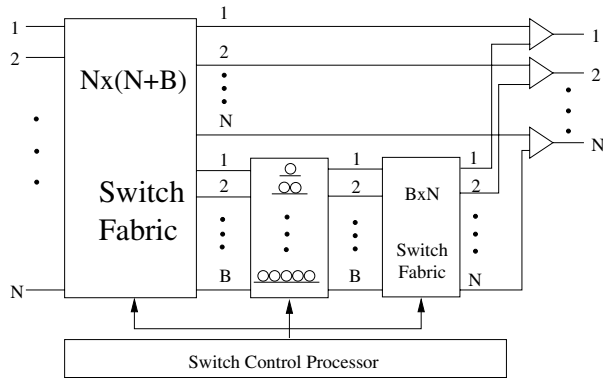


Fig. 1. Shared buffer architecture.

the buffer if the total length of packets in the buffer exceeds a certain threshold. Recently, accurate analytical models have been proposed in [7] and [8]. In [7], the time is divided into equally-sized small slots, and it is assumed that packets arrive at slot boundaries and that the lengths of packets are multiples of the slot length. Therefore, the buffering behavior can be formulated as a difference equation and can be solved in the Z domain. [8] derived an exact finite-state queuing model, in which the states indicate the FDL that an incoming packet will enter and whether the buffer will become full immediately after the packet enters the FDL.

In this paper, we develop an analytical model to evaluate the packet loss and delay performance of a shared buffer architecture. By choosing an appropriate granularity of FDLs, the shared buffering scheme can achieve significantly lower packet loss with much smaller switching fabrics and fewer FDLs. The simulation and analysis results also reveal that our analytical model is highly accurate for different granularities of FDLs under different traffic loads.

The rest of this paper is organized as follows. In Section II, we elaborate an asynchronous shared buffer architecture. An analytical model is developed in Section III to evaluate the blocking probability and the average delay. Section IV provides numerical results and discussions. Finally, Section V concludes the paper.

II. SWITCH ARCHITECTURE

In this section, we briefly describe a shared buffering switch architecture in asynchronous optical networks. The architecture, shown in Fig. 1, is a special case of SLOB [2], which was originally proposed for providing variable length delays. In this architecture, packet contention can be resolved by a one-stage feed-forward buffering scheme, in which a set of FDLs connect the outputs of the first switch to the inputs of the second switch. As the first step of our study, we consider a switch without wavelength converters. Thus, packets transmitted on different wavelengths will be switched separately.

The advantage of the shared buffering scheme is that packets headed for different outputs can share an FDL at different

times. Thus, it is more efficient than a dedicated output buffering scheme with the same or even fewer buffers. On the other hand, we also notice that this architecture requires two switching fabrics. Therefore, it is important to study the performance of the shared buffering scheme to better understand the trade-off between the switch and buffer sizes and the performance.

In addition to being affected by the number of FDLs, the packet loss performance can be affected by two other parameters. The first parameter is the distribution of the lengths of FDLs. In this study, we assume that the lengths of FDLs are *degenerate* [6], which means that the lengths of FDLs are consecutive multiples of a certain *granularity* D . For example, the delay of the first FDL is D time units, the delay of the second FDL is $2D$ time units, etc. Another important factor is the packet forwarding policy. In this paper, we consider a *non-void-filling* scheduling scheme, in which there are no conflicts when packets leave the second stage switching fabric, and in which the shortest feasible FDL that does not violate packet ordering will be selected.

III. ANALYTICAL MODEL

In this section, we present an analytical model to evaluate the packet loss and delay performance of the shared FDL buffers in asynchronous optical switches. We extend the model proposed in [8], which is an accurate finite state queuing model for output-based FDL buffers. However, we notice that the output-based model can not be directly extended for shared buffers because it may result in a multi-dimensional Markov chain, which may not be feasible to solve in practice. To deal with this problem, we propose an approximate model for the shared FDL buffers. Similar approaches have been widely used to analyze the performance of networks with arbitrary topologies.

The main idea of the analysis is to solve the problem in an iterative manner. Specifically, we assume that the outputs are independent of one another so that we can first analyze the loss and delay behaviors for a single tagged output. However, a certain FDL may not be available to packets headed for the tagged output due to packets headed for other outputs occupying the delay line. Therefore, we calculate the probability that packets headed for other outputs will block packets headed for the tagged output on each FDL. With these probabilities, we recalculate the performance of the tagged output. This process is repeated until the results converge.

We make the following assumptions:

- The total number of outputs is N and index n ($1 \leq n \leq N$) indicates the n -th output.
- The total number of FDLs is B , and L_i ($i = 1, 2, \dots, B$) denotes the length of the i -th FDL in units of time.
- There are no wavelength converters in the switch, thus we consider only one wavelength plane.
- When a packet arrives, the scheduling policy discussed in Section II will be applied.
- For any output n , the packets arrive according to a Poisson process with rate λ_n .

- Incoming packets have variable sizes, and the average length of packets is one time unit.
- For any output n , the packet length in units of time has an arbitrary distribution $p_n(\tau)$.

We now consider the queuing model for a tagged output n . Let τ_n be the length of the new packet to n , which is a random variable with *p.d.f.* $p_n(\tau)$. Following [8], we define the state of the system when a new packet successfully enters the buffer as

$$S = \{A_0, A_1, \dots, A_{B-1}, F_0, F_1, \dots, F_B\} \quad (1)$$

where state A_i means that the packet is forwarded to FDL i and $L_i + \tau_n \leq L_B$, state F_i denotes that the packet is forwarded to FDL i and $L_i + \tau_n > L_B$. We can then formulate an embedded Markov chain for the buffering system as in [8].

To conduct the analysis, we further define the following parameters:

- τ_{ni} denotes the random variable of packet length in units of time given that the state is A_i when the packet to n enters the buffer. Let $p_{ni}(\tau)$ be the *p.d.f.* of τ_{ni} .
- τ'_{ni} denotes the random variable of packet length in units of time given that the state is F_i when the packet to n enters the system. Let $p'_{ni}(\tau)$ be the *p.d.f.* of τ'_{ni} .
- t_n denotes the random variable of the inter-arrival time of packets to output n , which follows the exponential distribution with parameter λ_n .
- q_i denotes the steady state probability that the i -th FDL ($0 < i \leq B$) is busy. Here we define that a FDL is busy if a certain packet has been forwarded into the delay line while the tail of the packet has not entered. In other words, a new packet cannot be forwarded into a delay line if it is busy; otherwise a conflict will occur.
- q_{ni} denotes the steady state probability that the i -th FDL ($0 < i \leq B$) is occupied by a packet to output port n .
- ξ_n denotes the interval from the present time to the time at which the tail of the last packet for output port n will leave the buffer.
- $g_{j|s}^n$ denotes the conditional probability that, when a new packet to output n arrives, ξ_n is greater than L_{j-1} and smaller than or equal to L_j , given that the previous state of the system is s ($s \in S$). In other words, j is the shortest delay line into which the packet can be forwarded.
- h_{jk}^n denotes the probability that a packet to output n is scheduled to delay line j given that $L_{k-1} < \xi_n \leq L_k$ when the packet arrives.
- y_k^n denotes the probability that a packet to output n must be dropped given that $L_{k-1} < \xi_n \leq L_k$ when the packet arrives.
- P_s^n ($s \in S$) denotes the steady state probability.
- P_{s_1, s_2}^n ($s_1, s_2 \in S$) denotes the state transition probability from state s_1 to s_2 .

In the rest of this section, we first discuss how to calculate the state transition probabilities. We then provide the calculation of average packet loss and delay, followed by the calculation of other parameters as q_i . Finally, we will present the framework of the algorithm.

A. State Transition Probabilities

1) P_{A_i, A_j}^n and P_{A_i, F_j}^n : Obviously, if $\xi_n < 0$, then the packet will be forwarded to the output directly. Therefore,

$$P_{A_i, A_0}^n = \int_0^{L_B} p_n(t) dt \cdot g_{0|A_i}^n \quad (2)$$

$$P_{A_i, F_0}^n = \int_{L_B}^{\infty} p_n(t) dt \cdot g_{0|A_i}^n. \quad (3)$$

We now consider the probability that the new packet will be forwarded into FDL j ($j > 0$). In the output buffer case, a packet entering delay line j is equivalent to $L_{j-1} < \xi_n \leq L_j$. However, in the shared buffer case, delay line j may not be available since it may be occupied by packets to other output ports. Notice that a new packet forwarded to delay line j means $0 < \xi_n \leq L_j$. The probability that a new packet will enter delay line j given that the current state is A_i can be expressed as

$$\Pr[\text{packet enters delay line } j | A_i] = \sum_{k=1}^j \left(g_{k|A_i}^n \times h_{jk}^n \right) \quad (4)$$

where $g_{j|A_i}^n$ can be further derived as

$$\begin{aligned} g_{j|A_i}^n &= \Pr[L_{j-1} < \xi_n \leq L_j | A_i] \\ &= \Pr[L_{j-1} < L_i + \tau_{ni} - t_n \leq L_j]. \end{aligned} \quad (5)$$

We can observe that the right hand side of Eq. (5) is a convolution of random variable τ_{ni} and t_n . Detailed discussion on how to calculate $g_{j|A_i}^n$ can be found in [8].

Finally, for all $j > 0$, we have

$$P_{A_i, A_j}^n = \int_0^{L_B - L_j} p_n(t) dt \cdot \sum_{k=1}^j \left(g_{k|A_i}^n \times h_{jk}^n \right) \quad (6)$$

$$P_{A_i, F_j}^n = \int_{L_B - L_j}^{\infty} p_n(t) dt \cdot \sum_{k=1}^j \left(g_{k|A_i}^n \times h_{jk}^n \right). \quad (7)$$

2) P_{F_i, A_j}^n and P_{F_i, F_j}^n : These two types of transition probabilities can be calculated by using a similar scheme as discussed above. The only difference is that we must use $g_{k|F_i}^n$ to replace $g_{k|A_i}^n$ in Eq. (2), Eq. (3), Eq. (6) and Eq. (7). We observe that if $t_n > L_B$ when a packet arrives, the packet will be dropped according to the scheduling policy. Since the arrival process is memoryless, we have

$$g_{j|F_i}^n = \int_{L_B - L_j}^{L_B - L_{j-1}} \lambda_n e^{-\lambda_n t} dt. \quad (8)$$

B. Packet Loss Probability and Average Delay

For a given output port, the average delay of packets can be easily derived through

$$\bar{D} = \sum_i L_i P_{A_i} + \sum_i L_i P_{F_i} \quad (9)$$

which has the same form as the output based buffer.

Although the analysis for delay is easy, the loss probability is much more complex. There are two kinds of packet loss in the shared buffering system:

- Loss due to packet schedule constraint;
in this case, suppose a packet to n arrives and the state shifts to F_i . Then an incoming packet to n will be dropped as long as $\xi_n > L_B$.
- Loss due to full buffer;
in this case, the packet will be dropped even though it can be successfully scheduled on the certain output.

The first type of loss has been studied in [8], which is

$$\sum_{i=0}^B P_{F_i} \lambda_n r_{ni} \left[1 + \sum_{i=0}^B P_{F_i} \lambda_n r_{ni} \right]^{-1}$$

where

$$r_{ni} = E[\tau'_{ni}] + L_i - L_B$$

and $E[\tau'_{ni}]$ is the mean value of random variable τ'_{ni} .

For the second type of loss, we can use

$$\sum_{\forall s \in S} P_s \left[\sum_{k=1}^B (g_{k|s}^n \times y_k^n) \right].$$

In summary, we have

$$P_{nb} = \sum_{\forall s \in S} P_s \left[\sum_{k>0} (g_{k|s}^n \times y_k^n) \right] + \left[\frac{\sum_i P_{F_i} \lambda_n r_{ni}}{1 + \sum_i P_{F_i} \lambda_n r_{ni}} \right]. \quad (10)$$

C. Other parameters

1) q_i and q_{ni} : To calculate q_i and q_{ni} , we first consider a large amount of time T . Suppose during T , T_i is the total amount of time that FDL i is occupied and T_{ni} is the total amount of time that delay line i is occupied by packets to output port n . We can then derive $q_i = \frac{T_i}{T}$ and $q_{ni} = \frac{T_{ni}}{T}$. Since $T_i = \sum_n T_{ni}$, we now focus on the calculation of T_{ni} . Suppose that during T , the total number of incoming packets to output n is I_n . According to the definition of the states, we have

$$T_{ni} = I_n \times (1 - P_{nb}) \times (P_{A_i}^n E[\tau_{ni}] + P_{F_i}^n E[\tau'_{ni}]). \quad (11)$$

Since $\lambda_n = \frac{I_n}{T}$, we have

$$q_{ni} = \frac{T_{ni}}{T} = \lambda_n \times (1 - P_{nb}) \times (P_{A_i}^n E[\tau_{ni}] + P_{F_i}^n E[\tau'_{ni}]) \quad (12)$$

and

$$q_i = \sum_n \frac{T_{ni}}{T} = \sum_n q_{ni}. \quad (13)$$

2) h_{jk}^n and y_k^n : To calculate h_{jk}^n and y_k^n , we first consider the situation that a packet headed for output n can not be scheduled to delay line k given that $\xi_n \leq L_k$. From the discussion above we can see that such situation can happen if delay line k is occupied by a packet to outputs other than n . In other words, we need to calculate the probability that delay line k is busy given that it is not occupied by packet to output port n . Similar to the discussion of q_{ni} , this probability can be expressed as

$$\frac{T_k - T_{nk}}{T - T_{nk}} = \frac{q_k - q_{nk}}{1 - q_{nk}}.$$

Therefore, if $L_{k-1} < \xi_n \leq L_k$ and the packet can not be forwarded into any delay line, then

$$y_k^n = \prod_{l=k}^B \frac{q_l - q_{nl}}{1 - q_{nl}}. \quad (14)$$

Here, to simplify the analysis, we make an important assumption that the occupancy on delay line i does not depend on the occupancy on delay line i' ($i' < i$).

If the packet is finally forwarded to delay line j ($j > k$), then we know that delay line j is free given that it is not occupied by packets to n and all delay lines from k to $j-1$ are occupied by packets to output other than n given that it is not occupied by packets to n . Therefore, we have

$$h_{jk}^n = \begin{cases} \frac{1 - q_j}{(1 - y_k^n)(1 - q_{nj})} & j = k \\ \frac{(1 - q_j)}{(1 - y_k^n)(1 - q_{nj})} \times \prod_{l=k}^{j-1} \frac{q_l - q_{nl}}{1 - q_{nl}} & j > k. \end{cases} \quad (15)$$

D. Framework

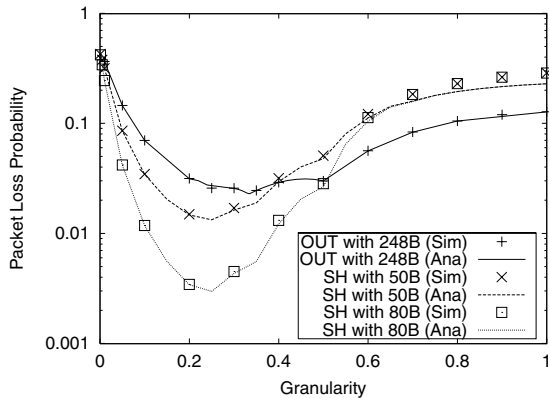
We now provide the framework to calculate the packet loss and delay performance, which is an iterative algorithm as follows:

- 1) initiate all q_i and q_{ni} to be 0;
- 2) calculate all h_{jk}^n ;
- 3) calculate all state transition probabilities;
- 4) calculate all steady state probabilities;
- 5) calculate the loss probabilities for all output ports and stop if all results are converged;
- 6) calculate all q_i and q_{ni} ; then go to Step 2.

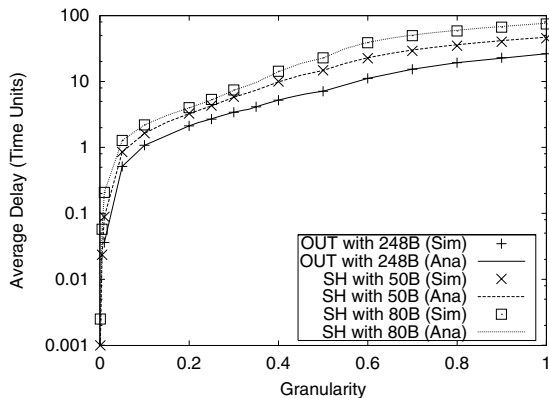
IV. NUMERICAL RESULTS

In this section, we evaluate the performance of shared buffers through simulation and analysis. An 8×8 switch with a single wavelength plane is considered. There is a single class of traffic, and packet arrivals are Poisson. Traffic is uniformly distributed over all switch outputs. Due to limited space, we consider only fixed-size incoming packets in this paper and the packet length is normalized to one time unit. Similar results were found for uniform and exponential distribution packet lengths.

In Fig. 2, we compare the packet loss and delay performance of shared buffering to output buffering scenarios under different FDL granularities with fixed-size packets. For the output buffering architecture, each output has 31 buffers; thus the total number of FDLs is 248, and a single stage 8×256 switching fabric is used. In shared buffering, we consider two cases. In the first case, we have 50 FDLs, thus the switch in the first stage requires an 8×58 switching fabric, and the switch in the second stage requires an 50×8 switching fabric. In the second case, we have 80 FDLs, thus the switch in the first stage requires an 8×88 switching fabric and the switch in the second stage requires an 80×8 switching fabric. We can see from Fig. 2(a) that the minimum packet loss probability of shared buffering with 50 or 80 FDLs is lower than that of dedicated output buffering with 248 FDLs when the traffic



(a) Packet Loss



(b) Average Delay

Fig. 2. Performance of shared buffering with different granularities (fixed-size packets, $\rho = 0.8$).

load is fixed to 0.8 Erlang. We also observe the optimum granularity D is about 0.35 and 0.25 for output buffers and shared buffers, respectively. From Fig. 2(b) we note that with the optimum granularity, the shared buffering scheme can provide almost the same delay performance as that of the output buffering scheme. Fig. 2(a) also illustrates that the performance of shared buffering is worse than output buffering if the granularity is large, particularly when the granularity is equal to 1. This result suggests that, for fixed-size packets, simply setting the granularity to 1 should be avoided in the shared buffering implementation.

In Fig. 3, we compare shared buffering with output buffering under different traffic loads with fixed-size packet. In output buffering, 31 buffers are dedicated for each output port, and the granularity is fixed at 0.35 and 1, respectively. Here we choose two different granularities for the output buffering scheme because the optimum granularity for output buffers is dependent of the traffic load. Particularly, the optimum granularity becomes 1 when the traffic load is lower than 0.6 Erlang. In shared buffering, we use 80 buffers, and the granularity is fixed at 0.2. From Fig. 3 we observe that, within the interested range of packet loss ($10^{-8} \rightarrow 1$), the shared buffering scheme with 80 FDLs always outperforms that of

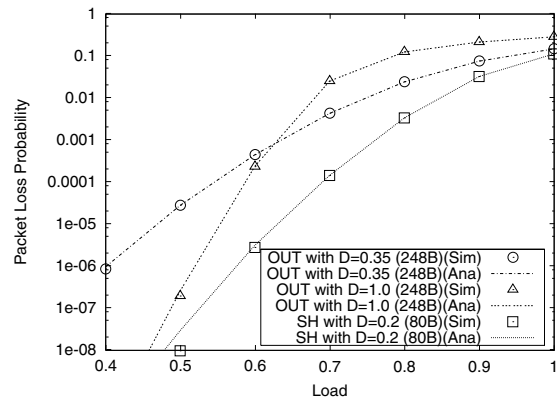


Fig. 3. Packet loss vs. arrival rate for output and shared buffering.

the dedicated output buffering with 248 FDLs under all traffic loads and with different granularities. This result is important because, for output buffering, the granularity of FDLs is fixed while the arrival rate may vary. Thus, an optimum granularity for a certain traffic load may degrade the performance under other traffic loads. However, even though the granularity is fixed, shared buffering still outperforms output buffering under a large range of traffic loads.

V. CONCLUSION

In this paper, we discuss a shared FDL buffering architecture to resolve contention in asynchronous optical switches. An analytical model is provided to evaluate the packet loss and delay performance of shared buffers. By exploiting the granularity of FDLs, we observe that shared buffers can significantly reduce the packet loss with much smaller size of switching fabrics and much fewer FDLs. The simulation and analysis results also show that our analytical model is highly accurate with different granularities of FDLs under various traffic loads.

REFERENCES

- [1] G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, "Optical switching: Switch fabrics, techniques, and architectures," *IEEE/OSA J. Lightwave Technol.*, vol. 21, no. 2, pp. 384–405, Feb. 2003.
- [2] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, and I. Andonovic, "SLOB: A switch with large optical buffers for packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 16, no. 10, pp. 1725–1736, Oct. 1998.
- [3] T. Zhang, K. Lu, and J. P. Jue, "Differentiated contention resolution for QoS in photonic packet-switched networks," *IEEE/OSA J. Lightwave Technol.*, vol. 22, no. 11, pp. 2523–2535, Nov. 2004.
- [4] P. D. Bergstrom, Jr., M. A. Ingram, A. J. Vernon, J. L. A. Hughes, and P. Tetali, "A markov chain model for an optical shared-memory packet switch," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1593–1603, Oct. 1999.
- [5] J. Diao and P. L. Chu, "Analysis of partially shared buffering for WDM optical packet switching," *IEEE/OSA J. Lightwave Technol.*, vol. 17, no. 4, pp. 525–534, Apr. 1999.
- [6] F. Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, no. 9, pp. 292–294, Sept. 2000.
- [7] K. Laevens and H. Bruneel, "Analysis of a single-wavelength optical buffer," in *Proc. IEEE INFOCOM 2003*, vol. 3, San Francisco, CA USA, Mar. 2003, pp. 2262–2267.
- [8] R. C. Almeida, Jr., J. U. Pelegri, and H. Waldman, "Delay-line buffer modeling for asynchronous optical networks," in *Proc. SPIE OptiComm '03*, Dallas, TX USA, Oct. 2003, pp. 381–391.