A Construction of 1-to-2 Shared Optical Buffer Queue with Switched Delay Lines

Xiaoliang Wang, Xiaohong Jiang, Achille Pattavina, and Susumu Horiguchi

***Abstract*—Optical buffering is fundamental to contention resolution in optical networks. The current works on this line mainly focus on the emulation of dedicated input/output buffer queue by using switched fiber delay lines (SDL). It is notable that the shared buffer queue, where a common buffer pool is shared by all the input/output ports of a switch, has the potential to significantly reduce the overall buffer capacity requirement. As far as we know, however, no related work is available yet on the exact emulation of a shared optical buffer queue with SDLs.**

In this paper, we focus on the design of first in first out (FIFO) shared optical buffer queue based on the optical feedback SDL construction. The construction considered consists of an $(M+2)\times(M+2)$ switch fabric and M fiber delay lines $\mathrm{FDL}_1, \ldots, \mathrm{FDL}_M$, where FDL_i connects the i^{th} output of the switch fabric with its i^{th} input. We show that by setting the length of FDL_i as $\min(M+1-i, i)$, $i=1,\ldots,M$, such a construction can actually work as an 1-to-2 shared buffer queue. We then extend this emulation to the more general N-to-2 case.

***Index Terms*—Shared buffer queue, optical buffer, switched delay line (SDL), fiber delay line (FDL), optical packet switching (OPS).**

I. Introduction

ALL-OPTICAL[1] switching is attractive since it can eliminate the quite expensive optical-electronic-optical conversions and help us to make good use of the enormous bandwidth of optical networks. Time sliced (synchronous) optical switching is a simple yet cost-effective technology for implementing all-optical packet switching [1]–[5], where optical buffers are required to resolve packets contention. Since optical-RAM is not available yet, the optical fiber delay line (FDL) is usually adopted to implement the buffering function. Unlike the traditional electronic memories with random access, a packet entering a FDL must propagate for a fixed amount of time and can not be retrieved anytime earlier. As such, designing FDL-based optical buffers with the same throughput and delay performance as that of the electronic ones is still a challenging issue now.

Basically, we have three possibilities for packet buffering in a switch, namely input buffer queuing (buffering packets at the input side), output buffer queuing (buffering packets at the output side) and shared buffer queuing (buffering packets internally) [6], [7]. Early works on the construction of optical buffer with switches and fiber delay lines (SDL) mainly focus on the feasibility study of such constructions, see, for example, the shared-memory optical ATM switch by Karol [8], CORD (contention resolution by delay lines) by Chlamtac et al. [3], COD (cascaded optical delay line) by Cruz et al. [2] and SLOB (switch with large buffer) by Hunter et al. [4]. Recently, C. S. Chang et al. demonstrated that it is possible for us to exactly emulate various optical buffer queues with SDL [9]–[15]. These works have been successful in implementing the optical counterparts of input buffer queue and output buffer queue, and some typical implementations among them are as follows.

- *1-to-1 FIFO queue*. Via a concatenation of the 2×2 feedback switch elements, an interesting construction of 1-to-1 FIFO queue was suggested in [10]. Such an optical FIFO queue can achieve a buffer size of $B=2^n-1$ by using $2n$ switch elements and $3\cdot 2^{n-1}-2$ fiber length.
- *2-to-1 FIFO Multiplexer*. A multiplexer works like a 'hopper', i.e., it always has a departure packet whenever it is nonempty. It has been proved in [14] that an $(M+2)\times(M+2)$ feedback switch has the capability to emulate a 2-to-1 FIFO multiplexer of size $O(2^M)$.
- *1-to-1 Priority queue*. In a priority queue, the packet with the highest priority is always sent to output link when a departure request comes, while the packet with the lowest priority will be dropped whenever buffer overflow happens. The recent research indicated that the $(M+1)\times(M+1)$ feedback switch can be used to emulate an 1-to-1 priority queue of size $O(M^3)$ [13].

The shared buffer queuing, which consists of a common memory shared by all the inputs and outputs, is another attractive structure for implementing electronic ATM switches and IP routers [7], [16]. In comparison with the input/output buffer queuing built on SDL, the corresponding shared buffer queuing structure has the potential to significantly reduce both the buffer capacity requirement and switch size[2]. However, no work is available yet on how to use SDL to exactly emulate shared buffer queue, which is the focus of this paper. Our main finding is that, by applying a slightly modified switching strategy to the feedback switch system proposed in [12], such a system can actually work as an 1-to-2 optical shared buffer queue. This result is further extended to the N-to-2 case with N inputs. The work of this paper lays the foundation for the general design of share buffer queue, and the 1-to-2 (and also

Paper approved by T. T. Lee, the Editor for Switching Architecture Performance of the IEEE Communications Society. Manuscript received June 25, 2008; revised December 1, 2008 and February 12, 2009.

X. Wang, X. Jiang, and S. Horiguchi are with the Graduate School of Information Sciences, Tohoku University, Sendai, 980-8579, Japan (e-mail: {waxili, jiang, susumu}@ecei.tohoku.ac.jp). X. Jiang is also a visiting researcher of the State Key Laboratory for Novel Software Technology, Nanjing University, China.

A. Pattavina is with the Department of Electronics and Information, Politecnico di Milano, 20133 Milano, Italy (e-mail: pattavina@elet.polimi.it).

This work is partially supported by JSPS grants (B) 21300018.

Digital Object Identifier 10.1109/TCOMM.2009.12.080309

[1]"all-optical" implies that the data transmission is in the optical domain, but the switching control can be electrical.

[2]In SDL based buffer queue, one FDL requires one dedicated switch port.

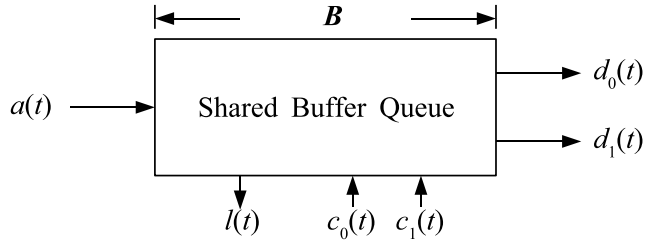

Fig. 1. An 1-to-2 shared buffer queue with buffer size B.

N-to-2) modules to be examined have the potential to serve as the basic building blocks for the future construction of large-scale shared buffer structures.

The remainder of this paper is organized as follows: Section II presents the definition of 1-to-2 shared buffer queue, and Section III provides a construction for it via SDL. In Section IV, a lower bound on the buffer size of our construction is derived. Finally, this construction is further extended to the more general N-to-2 case in Section V.

II. Preliminaries

In this section, we first formally define an 1-to-2 shared buffer queue and then introduce a trivial construction of it.

A. 1-to-2 shared buffer queue

To simplify the design and operation of optical buffer queue, we assume that the time is sliced and synchronized, i.e., the boundaries of arrival packets are aligned with their corresponding time slots. To implement synchronization, we need the function of adjusting packets arrival time, which can be accomplished by using packet recognizer and a set of delay lines [17]–[19]. Without loss of generality, we further assume that the packet size is fixed, a packet can be transmitted within one time slot, and the length of a delay line is equal to an integer number of time slots.

Based on these assumptions, an 1-to-2 shared buffer queue can be formally defined as follows.

Definition 1 (1-to-2 shared buffer queue): An 1-to-2 shared buffer queue is a network element that has one input link, two control inputs, two output links for departure packets, and one output link for lost packets due to buffer overflow (as illustrated in Fig. 1). For time t and output link i, $i \in \{0, 1\}$, we introduce the following notations:

- $a_i(t)$: the set of arrival packet destined for output link at the time.
- $d_i(t)$: the set of departure packets from output link at the time.
- $c_i(t)$: the control input state of output link at the time (the output port is enabled if $c_i(t) = 1$; disabled, otherwise).
- $q_i(t)$: the set of packets stored in the buffer and destined for the output link by the end of the time.
- $l(t)$: the set of lost packets at the time.

As the input port is unique, $|a_0(t)|$ and $|a_1(t)|$ can not be 1 simultaneously. Then the 1-to-2 shared buffer queue with buffer size B satisfies the following properties:

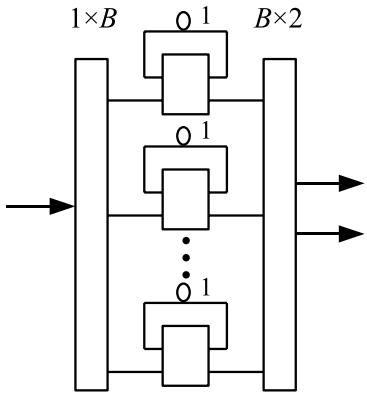

Fig. 2. A simple construction of 1-to-2 shared buffer queue.

(P1) Flow conservation: an arriving packet from the input link is either stored in the buffer or transmitted through one output link, i.e.,

$$q_0(t) \cup q_1(t) = q_0(t-1) \cup q_1(t-1) \cup a_0(t) \cup a_1(t) \backslash (d_0(t) \cup d_1(t) \cup l(t)) \tag{1}$$

(P2) Non-idling: if one output is enabled, there is always a departing packet from this output if there is at least one packet destined for the output in the buffer or the input, i.e.,

$$d_0(t) = \begin{cases} 1 & \text{if } c_0(t) = 1 \text{ and } |q_0(t-1) \cup a_0(t)| > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$d_1(t) = \begin{cases} 1 & \text{if } c_1(t) = 1 \text{ and } |q_1(t-1) \cup a_1(t)| > 0 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

(P3) Maximum buffer usage: an arriving packet is lost only when the buffer is full and the corresponding output of this packet is disabled, i.e.,

$$l(t) = \begin{cases} 1 & \text{if } |q_0(t-1) \cup q_1(t-1) \cup a_0(t)| > B \\ & \quad \text{and } c_0(t) = 0 \\ & \text{or } |q_0(t-1) \cup q_1(t-1) \cup a_1(t)| > B \\ & \quad \text{and } c_1(t) = 0 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

(P4) FIFO: packets destined for the same output link depart in the First-In-First-Out (FIFO) order.

B. A simple construction of 1-to-2 shared buffer queue

As we know, the queue with buffer size $B = 1$ can be constructed by using a 2×2 feedback switch element with one fiber delay line of length 1 [10]. Therefore, an 1-to-2 optical shared buffer queue with buffer size B can be exactly emulated by using B parallel feedback switch elements sandwiched between an $1 \times B$ crossbar and a $B \times 2$ crossbar (see Fig. 2). The newly arrived packets can be sent to any idle feedback switch elements through the $1 \times B$ crossbar, and the buffered packets can be read out from feedback switch elements and sent to the corresponding output links through the $B \times 2$ crossbar. If all the B feedback switches are occupied and there are no departure requests, any newly arrived packets have to be dropped. Since the switch size is almost the same as the buffer size and only the delay lines of length 1 are adopted

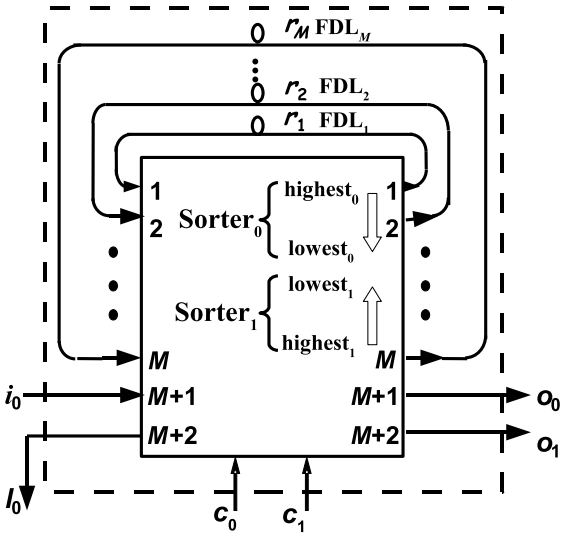

Fig. 3. A feedback construction of 1-to-2 shared buffer queue.

here, this structure has a high hardware complexity and also involves a large number of packet circulations[3].

III. An Effective Construction of Shared Buffer Queue

In this section, we introduce a more efficient construction of shared buffer queue by exploring the feedback switch architecture and a new packet switching strategy for it.

A. Architecture and switching algorithm

The feedback switch structure has been widely explored for the efficient construction of optical priority queues [11], [12]. Here, we extended the symmetric feedback switch considered in [12] for the design of an 1-to-2 shared buffer queue. One such structure is illustrated in Fig. 3, which consists of an $(M+2)\times(M+2)$ switch fabric, one input port (i_0), one lost port (l_0), two output ports (o_0 and o_1), two control inputs (c_0 and c_1) and M delay lines connecting M outputs back to M inputs of the switch fabric.

To schedule packets properly inside the switch, all packets are classified into two flows (f_0 and f_1) based on their destined output ports. The packets of one flow are assigned with priorities according to their arrival time, such that they depart in the FIFO order. To ensure that the oldest packet of each flow is always reachable in each time slot, the packet scheduling should satisfy the following rule:

(R1) The packet with priority k can never be switched to a delay line with length longer than k.

By applying the rule (R1) to the switch system, in each time slot the set of packets available at the inputs of the switch will contain the highest priority packets of both flows f_0 and f_1. These packets are further sorted by two dedicated sorters (each for one flow) according to the order of their priorities. After sorting, the highest priority packet of a flow will be directly sent out if the corresponding output port is enabled, and the rest packets of this flow are sent to consecutive fiber delay lines in sequence, following the descending order of their priorities; On the other hand, if the output port is disabled, all the sorted packets of this flow need to be rebuffered and will be sent to fiber delay lines in the same way above. To share the common buffer properly among two flows, we assign their packets to the fiber delay lines in the back-to-back manner, as illustrated in Fig. 3. More formally, the switching algorithm adopted here can be summarized as following:

Switching Algorithm (S1)

Classify all the packets appearing at the $M+1$ inputs of switch into two sets according to their output ports.
Denote by α the set of packets belonging to f_0.
Denote by β the set of packets belonging to f_1.
Sort packets in α based on their priorities.
Sort packets in β based on their priorities.
if $c_0(t) == 1$ **then**
  Remove the highest priority packet from α and send it to Output O_0.
  Decrease the priority of all packets belonging to f_0 in the system by 1.
if $c_1(t) == 1$ **then**
  Remove the highest priority packet from β and send it to Output O_1.
  Decrease the priority of all packets belonging to f_1 in the system by 1.
Send packets in α to FDL$_1$, FDL$_2$, ... following the descending order of their priorities.
Send packets in β to FDL$_M$, FDL$_{M-1}$, ... following the descending order of their priorities.

According to the above switching algorithm, if a packet of f_0 is sent to the $(i+1)^{th}$ delay line, the delay lines indexed from 1 to i must have been occupied by i packets of the same flow with higher priorities. Similarly, for a packet of f_1, if it is assigned to the $(M-i)^{th}$ delay line, there must be i packets of this flow with higher priorities in the delay lines indexed from M to $M+1-i$. Therefore, based on (R1), the length of i^{th} delay line can be set as

$$r_i = \min(i, M+1-i),\ i = 1, 2, \ldots, M \tag{5}$$

It is notable that at most M packets can be simultaneously inserted into the fiber delay lines, so *conflict* may happen if $M+1$ packets come to the inputs of the switch at the same time slot (i.e., one newly arrival packet to the input port and M feedback packets from fiber delay lines).

***Definition** 2:* **(Conflict)** For the switch system in Fig. 3, we say it is in conflict if at one time slot there are $M+1$ packets at the inputs of the switch but no departure is requested for any of them.

For an exactly emulation of 1-to-2 share buffer queue based on the switch system in Fig. 3, we need to determine the condition under which the conflict can never happen.

B. Buffer size

To avoid conflict, the number of accommodated packets in the system must be constrained such that the acceptance of

[3]Packet circulation through switch fabric causes significant attenuation of optical signal[5], [20].

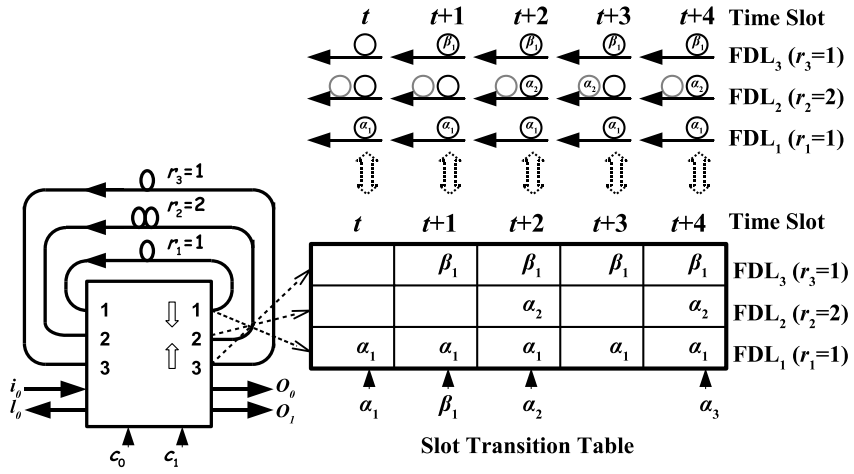

Fig. 4. Illustration for the buffer of size 3 only ($M = 3$).

any newly arrival packet will not lead the system into the state of conflict defined above. For this purpose, we introduce the following definition.

***Definition** 3:* **(Buffer size B)** For the 1-to-2 shared buffer queue construction in Fig. 3 with scheduling scheme (S1) and delay line setting in Equation (5), we define its buffer size as the maximum buffer capacity B the construction can provide so that conflicts will never happen under any scenario.

Let X and Y denote the set of packets buffered for flows f_0 and f_1, so $|X|$ and $|Y|$ refer to the total number of packets buffered for two flows, respectively. We further use $\mathbf{x}^D$ and $\mathbf{y}^D$ to denote the number of packets in X and Y that appear at the inputs of switch and require for buffering. Then we can construct a packet conflict set R as

$$R = \{(X, Y) | X, Y \text{ conflict one another, i.e., } \mathbf{x}^D + \mathbf{y}^D > M\} \tag{6}$$

Now, the buffer size B is actually determined as

$$B = \min_{(X,Y)\in R} (|X| + |Y|) - 1 \tag{7}$$

Obviously, the construction in Fig. 3 can exactly emulate any 1-to-2 shared buffer queue with buffer size not larger than B.

Since a delay line with delay length r can accommodate r packets, one may wonder if $B = \sum_{i=1}^{M} r_i$ for the construction in Fig. 3. We use a counterexample here to show that when $M = 3$, the buffer size B of the construction in Fig. 3 is 3 although the total length of delay lines is 4 there.

***Example** 1:* For the construction in Fig. 3 with M delay lines ($M = 3$ here), the *Slot Transition Table* in Fig. 4 is adopted to illustrate the slot occupation state (i.e., the first slot states) of these M delay lines after each switching operation [1]. Let α_i (resp. β_i) be the i^{th} arrival packet of f_0 (resp. f_1). Assume that the construction started from an empty system and there was no departure request from t to $t + 4$. As illustrated in Fig. 4, the packet α_1 arrived at time slot t and was sent to FDL_1, while the packet β_1 arrived at time slot $t+1$ and was sent to FDL_3. At time slot $t+2$, the packet α_2 arrived and had to be sent to FDL_2 of length 2 according to the switching algorithm (S1). At time slot $t + 4$, the packets α_1, α_2 and β_1 emerged from the outputs of delay lines and a new packet α_3 arrived simultaneously. Since the output ports were sill disabled at $t + 4$, to avoid conflict, α_1, α_2 and β_1 were rebuffered in FDL_1, FDL_2 and FDL_3, but the packet α_3 had to be dropped. Thus, the buffer size B of this construction is no more than 3 (please refer to Equation (7)). On the other hand, given that there are 3 distinct delay lines in this construction, it can always accommodate at least 3 packets without conflict. Therefore, the buffer size B of this shared buffer queue construction is just 3.

In the following, we will establish a lower bound B^* on the buffer size of the 1-to-2 shared buffer queue construction in Fig. 3, such that this construction can be used to exactly emulate an 1-to-2 shared buffer queue with size B^*. The following lemma is going to be used throughout the paper (we omit the proof due to its simplicity).

***Lemma** 1:* For any positive integer a, we have

(i) $\left\lceil \frac{a-1}{2} \right\rceil = \left\lfloor \frac{a}{2} \right\rfloor$

(ii) $\left\lceil \frac{a}{2} \right\rceil = \left\lfloor \frac{a-1}{2} \right\rfloor + 1$

IV. A Lower Bound of Buffer Size

From Equation (7) we can see that if we know how many packets have already been buffered in the network when conflict happens, then the buffer size B can be determined by finding the minimum number of buffered packets among all possible conflict scenarios. It is notable, however, that although we can easily find all the conflict scenarios at the M inputs of delay lines, it is very hard to determine all the possible sets of packets (and thus the number of packets) stored in the network for one given conflict scenario. Hereafter, we derive a lower bound of buffer size, which can serve as a sufficient condition for the construction of non-conflict shared buffer queues.

A. Buffered packets of one flow

For convenience, let "*rebuffered packets*" denote those packets that emerge from the delay lines and are rebuffered in the network after sorting. In the following, we first establish a lemma (Lemma 2) regarding one important property of the construction in Fig. 3, then we determine in Lemma 3 that if

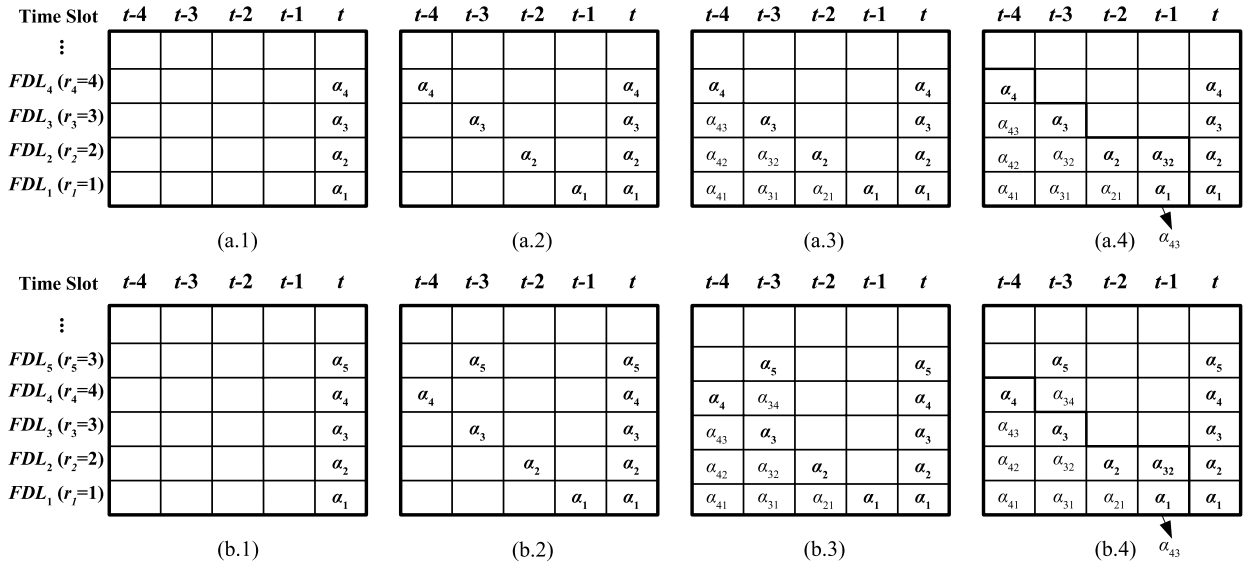

Fig. 5. Examples of packets scheduling ($M = 7$).

there exist i rebuffered packets of one flow, at least how many packets from the same flow must have been buffered in the delay lines.

***Lemma** 2:* For the construction in Fig. 3, suppose that at a given time we see i ($i \leq M$) rebuffered packets that belong to the same flow and come from i consecutive FDLs, starting with the shortest delay line. Then we know that at least $Q^*(i)$ packets (including these i rebuffered packets) of this flow are currently being buffered in the network, where $Q^*(i)$ is given by:

$$Q^*(i) = \begin{cases} \left\lceil \frac{i}{2} \right\rceil \left\lfloor \frac{i}{2} \right\rfloor + 1, & \text{if } 0 < i \leq \left\lceil \frac{M}{2} \right\rceil \\ \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil \left\lfloor \frac{\lceil M/2 \rceil}{2} \right\rfloor + i^2 - Mi + \left\lceil \frac{M}{2} \right\rceil \left\lfloor \frac{M}{2} \right\rfloor + 1, & \\ \qquad \text{if } \left\lceil \frac{M}{2} \right\rceil < i \leq \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil & \\ \left(\left\lfloor \frac{M}{2} \right\rfloor + 2 \right) \cdot i - \left(\left\lfloor \frac{M}{2} \right\rfloor \right)^2 - \left\lfloor \frac{M}{2} \right\rfloor - M, & \\ \qquad \text{if } \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil < i \leq M & \end{cases} \tag{8}$$

Proof: This lemma will be proved with the help of the slot transition table. For ease of comprehension, we first consider one example of the switch system with $M = 7$. We assume that at time t, there are four packets of one flow that emerged from FDL_1 to FDL_4 and are now rebuffered in these delay lines. Then the delay line occupation state at this time can be illustrated by the slot transition table in Fig. 5 (a.1). Since these four packets came from FDL_1 to FDL_4 respectively, they must have been inserted into their corresponding delay lines at time slot $t-1$, $t-2$, $t-3$ and $t-4$, as illustrated in Fig. 5 (a.2). Let us first consider the time slot $t-4$. As α_4 occupied FDL_4 at this time, from the switching algorithm (S1) we know that the delay lines FDL_3, FDL_2 and FDL_1 must have already been occupied by three other packets with higher priorities. We denote these three packets by α_{43}, α_{42} and α_{41}, as shown in Fig. 5 (a.3). Similarly, we can deduce that two packets (marked as α_{32} and α_{31}) with priorities higher than that of α_3 must have occupied FDL_2 and FDL_1 at time slot $t-3$, and one packet (marked as α_{21}) with priority higher than that of α_2 must have occupied FDL_1 at time slot $t-2$. Notice that the FDL_i can delay a packet for r_i time slots only, so the packets α_{43} and α_{32} would emerge from FDL_3 and FDL_2 at time slot $t-1$. Since the departure request can come at any time slot, it may happen that at time slot $t-1$ one packet (assume α_{43} here) will depart from the output link, as shown in Fig. 5 (a.4). Thus, we can conclude that by time slot t at least five packets of the same flow are buffered in the system, namely α_1, α_2, α_3, α_4 and the packet α_{32} in FDL_2. Extending the idea of this example, we can get the following general results.

(a) $0 < i \leq \left\lceil \frac{M}{2} \right\rceil$. By referring to the j^{th} ($j \leq i$) row of the delay line occupation state shown in Fig. 5 (a.3), we know that from time slot $t-i$ to time slot $t-1$ at least $i+1-j$ packets have been inserted into the j^{th} delay line. Notice that a delay line with length j can only accommodate j packets simultaneously, so $\sum_{j=1}^{i} \min(i+1-j, j)$ packets have ever been buffered before time slot t. Since some of these buffered packets may come out from their delay lines after $t - \left\lfloor \frac{i-1}{2} \right\rfloor - 1$, and one flow have at most one departure packet in each time slot, then $\left\lfloor \frac{i-1}{2} \right\rfloor$ packets may depart from output link from $t - \left\lfloor \frac{i-1}{2} \right\rfloor$ to $t-1$. Therefore, we conclude that by time slot t at least $Q^*(i)$ packets have been buffered in FDLs, where

$$\begin{aligned} Q^*(i) &= \sum_{j=1}^{i} \min(i+1-j, j) - \left\lfloor \frac{i-1}{2} \right\rfloor \\ &= \left\lceil \frac{i}{2} \right\rceil \times \left(1 + \left\lfloor \frac{i}{2} \right\rfloor \right) - \left(\left\lceil \frac{i}{2} \right\rceil - 1 \right) \\ &= \left\lceil \frac{i}{2} \right\rceil \times \left\lfloor \frac{i}{2} \right\rfloor + 1 \end{aligned} \tag{9}$$

(b) $\left\lceil \frac{M}{2} \right\rceil < i \leq M$. To explain the condition in this scenario, we adopt another example shown in Fig. 5 (b) , where $i = 5$ ($> \lceil 7/2 \rceil$) packets that emerged from FDL_1 to FDL_5 are rebuffered at time slot t. Similar to the above scenario, by checking the $t - k^{th}$ ($k \in [1, i]$) time slot in Fig. 5 (b.3), we can see that at least k packets were inserted into delay lines at this slot. To determine at least how many packets have been

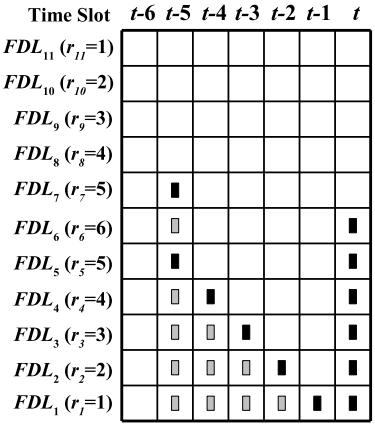

Fig. 6. An overlap arrival case (M=11).

buffered in the system by time slot t, we further divide this scenario into four cases:

Case 1. M is odd and $\lceil \frac{M}{2} \rceil < i \leq \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil = \frac{M-1}{2} + \lceil \frac{M+1}{4} \rceil$

By comparing Tables (b.4) and (a.4) in Fig. 5, we can see that the packets in Fig. 5 (b.4) contain all the packets in Fig. 5 (a.4) and two new packets α_5 and α_{34}. Therefore, all the buffered packets in this case can be divided into two sets: set 1 that contains the packets deduced from the occupation states of delay lines 1 to $\lceil \frac{M}{2} \rceil$ at time slot t, and set 2 that contains the packets deduced from the occupation states of delay lines $\lceil \frac{M}{2} \rceil + 1$ to i at time slot t. Let us consider the time slot $t - 2\lceil \frac{M}{2} \rceil + k$, $\lceil \frac{M}{2} \rceil < k \leq \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil$. If the k^{th} delay line is occupied at this time, there should be $\left(\lceil \frac{M}{2} \rceil - (k - \lceil \frac{M}{2} \rceil)\right)$ packets belonging to set 1 and $k - \left(\lceil \frac{M}{2} \rceil - (k - \lceil \frac{M}{2} \rceil)\right) = 2(k - \lceil \frac{M}{2} \rceil)$ packets belonging to set 2, as shown in Fig. 5 (b.4). Based on the packets in set 1 and set 2, we can determine the number of packets buffered in the system by time slot t as:

$$\begin{aligned} Q^*(i) &= Q^*\left(\left\lceil \frac{M}{2} \right\rceil\right) + \sum_{k=\lceil \frac{M}{2} \rceil + 1}^{i} 2\left(k - \left\lceil \frac{M}{2} \right\rceil\right) \\ &= \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil \times \left\lfloor \frac{\lceil M/2 \rceil}{2} \right\rfloor + 1 + i^2 - M \times i + \frac{M^2 - 1}{4} \\ &= \left\lceil \frac{M+1}{4} \right\rceil \times \left\lfloor \frac{M+1}{4} \right\rfloor + i^2 - M \times i + \frac{M^2 + 3}{4} \end{aligned} \quad (10)$$

In particular, when $i = \lceil \frac{M}{2} \rceil$, the Equation (10) reduces to Equation (9), so

$$Q^*\left(\left\lceil \frac{M}{2} \right\rceil\right) = \left\lceil \frac{M+1}{4} \right\rceil \left\lfloor \frac{M+1}{4} \right\rfloor + 1 \quad (11)$$

Case 2. M is odd and $\frac{M-1}{2} + \lceil \frac{M+1}{4} \rceil < i \leq M$

Following the similar idea as that of scenario (a), we can determine that from time slot $t - \lceil \frac{M}{2} \rceil$ to $t-1$ at least $\min(i + 1 - j, i - \lceil \frac{M}{2} \rceil + 1)$ packets have been inserted into the j^{th} $(0 < j \leq i)$ delay line. Since some of these packets buffered before time slot t will come out from their delay lines after $t - (2\lceil \frac{M}{2} \rceil - i)$, then at most $2\lceil \frac{M}{2} \rceil - i - 1$ packets may depart from output link. Notice that the j^{th} delay line here can only accommodate $\min(j, M + 1 - j)$ packets simultaneously, we conclude that in this case at least $Q^*(i)$ packets are buffered in FDLs, where

$$\begin{aligned} Q^*(i) &= \sum_{j=1}^{i} \min\left(i + 1 - j, i - \left\lceil \frac{M}{2} \right\rceil + 1, j, M + 1 - j\right) \\ &\quad - \left(2\left\lceil \frac{M}{2} \right\rceil - i - 1\right) \\ &= \left\lceil \frac{M}{2} \right\rceil \times (i - \left\lceil \frac{M}{2} \right\rceil + 1) - (M + 1 - i - 1) \\ &= \frac{M+1}{2} \times i - \frac{M^2 - 1}{4} - M + i \\ &= \frac{M+3}{2} \times i - \frac{M^2 + 4M - 1}{4} \end{aligned} \quad (12)$$

For the case when M is even, the proof is similar and we only give the finial results as follows:

Case 3. M is even and $i \leq \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil = \frac{M}{2} + \lceil \frac{M}{4} \rceil$

$$\begin{aligned} Q^*(i) &= Q^*\left(\left\lceil \frac{M}{2} \right\rceil\right) + \sum_{j=1}^{i - \lceil \frac{M}{2} \rceil} (2j - 1) \\ &= \left\lceil \frac{M}{4} \right\rceil \times \left\lfloor \frac{M}{4} \right\rfloor + i^2 - M \times i + \frac{M^2 + 4}{4} \end{aligned} \quad (13)$$

In particular, when $i = \frac{M}{2}$, the Equation (13) reduces to Equation (9), so

$$Q^*\left(\frac{M}{2}\right) = \left\lceil \frac{M}{4} \right\rceil \left\lfloor \frac{M}{4} \right\rfloor + 1 \quad (14)$$

Case 4. M is even and $\frac{M}{2} + \lceil \frac{M}{4} \rceil < i \leq M$

$$\begin{aligned} Q^*(i) &= \sum_{j=1}^{i} \min\left(i + 1 - j, i - \left\lceil \frac{M}{2} \right\rceil, j, M + 1 - j\right) \\ &\quad - \left[\left\lceil \frac{M}{2} \right\rceil - (i - \left\lceil \frac{M}{2} \right\rceil)\right] \\ &= \frac{M+4}{2} \times i - \frac{M^2 + 6M}{4} \end{aligned} \quad (15)$$

Combining the above results together, the Equation (8) follows. ∎

The above analysis focuses on the case that all the i rebuffered packets come from *consecutive delay lines* (from delay line 1 to delay line i) only. In practice, however, the rebuffered packets in the feedback construction may come from non-consecutive delay lines as well. Take the slot transition table in Fig. 6 as an example ($M = 11$), where six packets of one flow are rebuffered in the delay lines from FDL_1 to FDL_6 at time slot t. From Fig. 6 we can see that if five of these packets came from the 1^{st} to 5^{th} consecutive delay lines but another packet came from the 7^{th} delay line, there should be at least $Q^*(5) + 2 = 9$ packets buffered by slot t, including the two packets inserted into FDL_6 and FDL_7 at time slot $t - 5$. On the other hand, if all these six rebuffered packets came from the 1^{st} to 6^{th} consecutive delay lines, we know from Lemma 2 that at least $Q^*(6) = 10$ packets have been buffered at slot t. This example indicates that the number of buffered packets in the non-consecutive case may be less than that of the consecutive case. This is due to the *"packets overlap"* phenomena in the non-consecutive case,

TABLE I
COMPARISON OF OPTICAL QUEUE CONSTRUCTIONS

	1×2 shared buffer queue	1×1 FIFO queue	2×1 FIFO Multiplexer	1×1 Priority queue
Feedback Construction	$(M+2) \times (M+2)$	$2n$ 2×2	$(M+2) \times (M+2)$	$(M+1) \times (M+1)$
Buffer Size	$O(M^2)$	$O(2^n)$	$O(2^M)$	$O(M^3)$
FDLs length	$O(M^2)$	$O(2^n)$	$O(2^M)$	$O(M^3)$

i.e., two packets may feedback through two delay lines of same length. The observation in above example leads to the following lemma.

***Lemma** 3:* For the construction in Fig. 3, if there are i ($i \leq M$) rebuffered packets of the same flow at one time, then the number of buffered packets (including these i rebuffered packets) of this flow is at least $P^*(i)$ by this time, where When M is odd

$$P^*(i) = \begin{cases} Q^*\left(i - \left\lceil \frac{i-3}{5} \right\rfloor - 1\right) + \sum_{j=1}^{\left\lceil \frac{i-3}{5} \right\rfloor} 2j, \\ \qquad \text{if } 0 < i \leq \left\lceil \frac{M}{2} \right\rceil \\ Q^*\left(\left\lceil \frac{M}{2} \right\rceil - \left\lceil \frac{5M-8i-1}{10} \right\rfloor - 1\right) + \sum_{j=1}^{i-\left\lceil \frac{M}{2} \right\rceil + \left\lceil \frac{5M-8i-1}{10} \right\rfloor} 2j, \\ \qquad \text{if } \left\lceil \frac{M}{2} \right\rceil < i \leq \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil \\ Q^*(i), \qquad \text{if } \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil < i \leq M \end{cases}$$

When M is even

$$P^*(i) = \begin{cases} Q^*\left(i - \left\lceil \frac{i-1}{5} \right\rfloor - 1\right) + \sum_{j=1}^{\left\lceil \frac{i-1}{5} \right\rfloor} (2j-1), \\ \qquad \text{if } 0 < i \leq \left\lceil \frac{M}{2} \right\rceil \\ Q^*\left(\left\lceil \frac{M}{2} \right\rceil - \left\lceil \frac{5M-8i-2}{10} \right\rfloor - 1\right) + \sum_{j=1}^{i-\left\lceil \frac{M}{2} \right\rceil + \left\lceil \frac{5M-8i-2}{10} \right\rfloor} (2j-1), \\ \qquad \text{if } \left\lceil \frac{M}{2} \right\rceil < i \leq \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil \\ Q^*(i), \qquad \text{if } \left\lfloor \frac{M}{2} \right\rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil < i \leq M \end{cases}$$

here $\lceil a \rfloor$ denotes the integer that is most close to a.

Proof: See Appendix A. ■

B. A lower bound of buffer size B

Since the number of buffered packets $P^*(i)$ derived in Lemma 3 serves as a lower bound of actually buffered packets of one flow, in this section we apply $P^*(i)$ (instead of $|X|$ and $|Y|$ in (7)) to obtain a lower bound on the buffer size. Notice that if there are i packets of flow f_0 occupying delay lines indexed from 1 to i ($0 \leq i \leq \lceil \frac{M}{2} \rceil$), there must be $M-i$ packets of flow f_1 occupying other delay lines after sorting. Therefore, a lower bound B^* on buffer size can be achieved by solving the following integer linear programming problem:

$$B^* = \min_i \left(P^*(i) + P^*(M-i)\right) \tag{16}$$

$$\textit{Subject to}: 0 \leq i \leq \left\lceil \frac{M}{2} \right\rceil \tag{17}$$

$$i \text{ is an integer} \tag{18}$$

By finding the close form solution of above problem, the following Theorem follows (please refer to Appendix B for the proof).

***Theorem** 1:* The construction in Fig. 3 is an 1×2 shared buffer queue with buffer size $B \geq B^*$, where

$$B^* = \begin{cases} \max\left(\left\lfloor \frac{M^2-2M+1}{10} \right\rfloor - 1, M\right), & \text{if } M \text{ is odd} \\ \max\left(\left\lfloor \frac{M^2-4M+9}{10} \right\rfloor - 1, M\right), & \text{if } M \text{ is even} \end{cases} \tag{19}$$

Since the construction can accommodate at most $\sum_{i=1}^{M} d_i$ packets, the setting in Equation (5) implies that a shared buffer queue with at most $O(M^2)$ buffer size can be constructed via the $(M+2) \times (M+2)$ switch. Then we have the following corollary from the conclusion in Theorem 1:

***Corollary** 1:* The construction in Fig. 3 can exactly emulate a shared optical queue with $O(M^2)$ buffer size.

In comparison with the simple construction in Fig. 2, where the achievable buffer size is the same as the number of delay lines M, our feedback construction in Fig. 3 can provide a much larger buffer size ($O(M^2)$). A further comparison among the implementations of different optical queues is shown in Table I.

C. Practical considerations

Now we investigate some practical issues for building an 1-to-2 shared buffer queue. The core of our construction is an $(M+2) \times (M+2)$ optical switch fabric, which in principle can be implemented by a non-blocking optical space switch. Two promising switching technologies for building high speed optical switch are arrayed-waveguide-granting (AWG) and directional-coupler (DC), because they can switch at the speed of nanoseconds and thus are suitable for supporting packet switching inside the shared buffer queue [21], [22]. To build a large size switch fabric, the multistage switch architectures like Clos and Cantor are usually adopted to achieve a good scalability (The recent advances on the high speed optical switch design can be found in [23], [24] and the references there in).

In an 1-to-2 shared buffer queue, it may happen that a packet needs to be recirculated many times in the system, since the departure time of the packet can not be determined in advance. Therefore, the optical signal may be significantly attenuated after many times of circulation in the system, and thus the optical amplifiers are necessary at the output ports of system for compensating the signal loss. It is notable, however, that the optical amplifier will introduce additional accumulated spontaneous emission noise and signal-level fluctuation [20]. Also, the crosstalk introduced by switch devices (such as AWG and DC) will further degrade the signal and increase the bit error rate [15]. How these constrains limit the maximum buffering time in and the practical design of shared buffer queue still deserve deliberate studies.

TABLE II
BUFFER CAPACITY OF N-TO-2 SHARED BUFFER QUEUE ACHIEVED BY THEOREM 2

M	1	2	3	8	15	32	63	128	255	512	1023
N=1	1	2	3	8	18	89	383	1587	6450	26609	104447
N=2	1	2	3	8	16	84	372	1563	6400	25908	104244
N=5	1	2	3	8	15	71	339	1491	6253	25607	103635

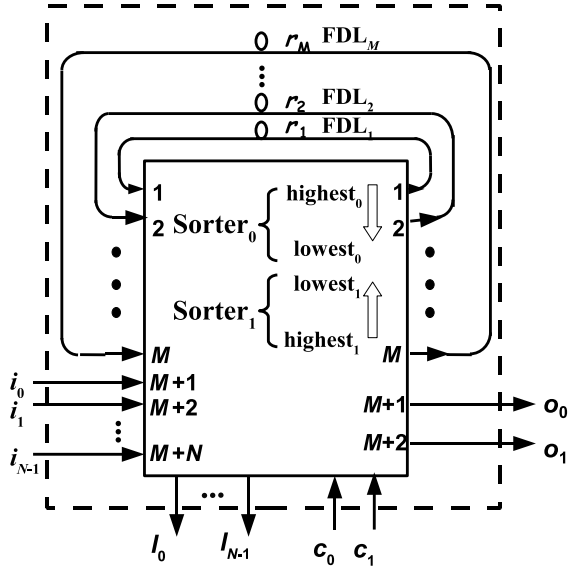

Fig. 7. An N-to-2 construction of shared buffer queue.

V. THE CONSTRUCTION OF N-TO-2 SHARED BUFFER QUEUE

In this section, we extend the result in Section IV to the more general N-to-2 shared buffer queue with N input links, 2 control inputs, 2 departure links and N lost links (see Fig. 7). At time t, let $c(t)$ be the states of the control inputs, $a(t)$ be the set of arrival packets and $q(t)$ be the set of departure packets at this time. Similar to the construction of the 1-to-2 shared buffer queue, the N-to-2 construction also needs to satisfy the same properties of flow conservation (P1), non-idling (P2) and FIFO (P4). About the maximum buffer usage property (P3), it should be replaced by the following property (P3$'$) due to the multiple inputs here:

(P3$'$): if $|q(t-1) \cup a(t) \backslash c(t)| > B$, then $|q(t-1) \cup a(t) \backslash c(t)| - B$ packets will be dropped.

To guarantee the above properties, the N-to-2 construction still adopts the same switching algorithm (S1) and delay setting (5) as that of its 1-to-2 counterpart. Here, we also explore a lower bound on the buffer size of the N-to-2 shared buffer queue to guarantee no conflict for it under any scenario.

Since in one time slot up to N packets (denoted as p_1, p_2,..., p_N here) may arrive in the queue, without loss of generality we only need to consider the following two conflict cases:

Case 1. All packets $p_1, \ldots, p_N$ are destined for output link 1

Assume that at time slot t, after scheduling the packet p_N is dropped to avoid conflict and packets p_1,..., p_{N-1} are assigned to delay lines $(i+1)$, ..., $(i+N-1)$ $(0 \le i \le M-N+1)$. Based on the Lemma 3 we know that at least $P^*(i) + P^*(M-(i+N-1)) + N - 1$ packets have been buffered in the system by this time.

Case 2. k (resp. $N-k$) packets are destined for output link 1 (resp. 2), $1 \le k < N$

If none of those N arrival packets is dropped at the inputs, after sorting two packets of different flows may compete for one common delay line. Suppose this competition happens at delay line $(i+1)$ $(0 \le i \le M-N+1)$. In this case, only i rebuffered packets of f_0, $M-(i+N-1$ rebuffered packets of f_1) and $N-1$ newly arrived packets can be inserted into delay lines. Again, we know that at least $P^*(i)+P^*(M-(i+N-1)) + N - 1$ packets have been buffered in the system.

As the results in above two cases are the same, a lower bound B^* on the buffer size of the N-to-2 shared buffer queue can be obtained by solving a linear programming problem similar to (16). Thus, we have the following theorem (Please refer to Appendix C for the proof).

***Theorem** 2:* If $r_i = \min[i, M+1-i]$ for all $i = 1, 2, \ldots, M$ and $M \ge 2$, then the construction in Fig. 7 is an $N \times 2$ shared buffer queue with buffer size $B \ge B^*$, here

$$B^* = \begin{cases} M, & M \le N-1 \\ \max\left(B', M\right), & \text{otherwise} \end{cases} \tag{20}$$

and

$$B' = \begin{cases} \left\lfloor \frac{(M-N)^2}{10} \right\rfloor + N - 2, & M \text{ is odd} \\ \left\lfloor \frac{(M-N-1)^2+5}{10} \right\rfloor + N - 2, & M \text{ is even} \end{cases} \tag{21}$$

To illustrate the condition developed in Theorem 2, Table II shows the buffer capacity B^* for combinations of different M and N. We can see that for a given M, the buffer capacity actually decreases as N increases. This is because as N increases, more packets may require buffering within one time slot. On the other hand, for a given N, the buffer capacity grows monotonously as M increases, and the growth of buffer capacity becomes significant when the value of M is large enough. We can also see from Table II that when M is small (say, less than 20), the lower bound is almost the same as the value of M.

VI. CONCLUSIONS

In this paper, we studied the exact emulation of an 1-to-2 FIFO shared buffer queue based on the optical feedback SDL construction. The construction consists of an $(M+2)\times(M+2)$ space switch and M fiber delay lines connecting M outputs of the switch fabric back to its M inputs. We showed that by setting the length of the i^{th} delay line as $\min(i, M+1-i)$, $i = 1, \ldots, M$, such a construction can exactly emulate an 1-to-2 shared buffer queue with $O(M^2)$ buffer size. We then extended this construction to the more general N-to-2 case.

Note that this paper only studied the design of the simple 1-to-2 (and also N-to-2) shared buffer queues. How to extend the single stage construction in this paper directly to the general N-to-N case, and how to use the 1-to-2 (and N-to-2) modules studied here as building blocks and apply multistage structures for constructing the N-to-N shared buffer queues can be some interesting future works.

APPENDIX A
PROOF OF LEMMA 3

We consider here the non-consecutive case with packets overlap. First, we separate all the i rebuffered packets into two groups: group 1 that contains the rebuffered packets from the first half of delay lines (from 1 to $\lceil \frac{M}{2} \rceil$), and group 2 contains the rebuffered packets from the second half of delay lines (from $\lceil \frac{M}{2} \rceil + 1$ to M). Analogous to the proof of Lemma 2 (Case 1), all the buffered packets in the system now can be divided into two sets: set 1 that contains the buffered packets deduced from the rebuffered packets in group 1, and set 2 that contains the buffered packets deduced from the rebuffered packets in group 2. Thus, a lower bound of buffer size can be achieved by finding the minimum numbers of buffered packets in set 1 and set 2. Without loss of generality, we assume that there are $\min(\lceil \frac{M}{2} \rceil, i) - k$ rebuffered packets in group 1 (or equally there are k delay lines that do not have rebuffered packets), $k < \min(i, \lceil \frac{M}{2} \rceil)$. Correspondingly, there should be $i - (\min(\lceil \frac{M}{2} \rceil, i) - k)$ rebuffered packets in group 2. From the proof of Lemma 2 we know that there are at least $Q^*(\min(\lceil \frac{M}{2} \rceil, i) - k - 1)$ buffered packets in set 1 now. For set 2, we can easily see from the slot transition table that it contains all the packets above the diagonal line from (FDL_1, $t-1$) to ($\text{FDL}_{\lceil M/2 \rceil}$, $t - \lceil \frac{M}{2} \rceil$). Thus, the number of buffered packets in set 2 will be minimum if all the $i - (\min(\lceil \frac{M}{2} \rceil, i) - k)$ rebuffered packets come from consecutive delay lines, starting from $\text{FDL}_{\lceil M/2 \rceil + 1}$. Then we can deduce the minimum number of all the buffered packets in terms of k as follows:

(a) M is odd.

Case 1. $0 < i \le \lceil \frac{M}{2} \rceil$

In this case, there are $i - k$ rebuffered packets from group 1 and k rebuffered packets from group 2. Then we have that the number of buffered packets $P(k)$ in the system satisfies the following inequality,

$$\begin{aligned} P(k) &\ge Q^*(i-k-1) + \sum_{j=1}^{k} 2j \\ &= \left\lceil \frac{i-k-1}{2} \right\rceil \left\lfloor \frac{i-k-1}{2} \right\rfloor + 1 + k^2 + k \\ &= \frac{5}{4}k^2 - \frac{i-3}{2}k + \begin{cases} \frac{i^2-2i+4}{4}, & i-k-1 \text{ is odd} \\ \frac{i^2-2i+5}{4}, & i-k-1 \text{ is even} \end{cases} \end{aligned} \quad (22)$$

The minimum value of (22) is attained when $k = \lceil \frac{i-3}{5} \rceil$.

Case 2. $\lceil \frac{M}{2} \rceil < i \le \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil$

$$\begin{aligned} P(k) &\ge Q^*\left(\left\lceil \frac{M}{2} \right\rceil - k - 1\right) + \sum_{j=1}^{i-\lceil \frac{M}{2} \rceil + k} 2j \\ &= \left\lceil \frac{\lceil M/2 \rceil - k - 1}{2} \right\rceil \left\lfloor \frac{\lceil M/2 \rceil - k - 1}{2} \right\rfloor + 1 \\ &\quad + (i - \left\lceil \frac{M}{2} \right\rceil + k - 1)(i - \left\lceil \frac{M}{2} \right\rceil + k) \\ &= \frac{5}{4}k^2 + (2i - \frac{5}{4}M + \frac{1}{4})k + i^2 - Mi \\ &\quad + \begin{cases} \frac{5M^2-2M+9}{16}, & \lceil \frac{M}{2} \rceil - k - 1 \text{ is odd} \\ \frac{5M^2-2M+13}{16}, & \lceil \frac{M}{2} \rceil - k - 1 \text{ is even} \end{cases} \end{aligned} \quad (23)$$

The minimum value of (23) is attained when $k = \lceil \frac{5M-8i-1}{10} \rfloor$.

It is notable that the Equation (23) reduces to Equation (22) when $i = \lceil \frac{M}{2} \rceil$, which is

$$\sum_{j=1}^{\lceil \frac{M-5}{10} \rfloor} 2j + Q^*\left(\left\lceil \frac{M}{2} \right\rceil - \left\lceil \frac{M-5}{10} \right\rceil - 1\right) \quad (24)$$

Case 3. $\lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil < i \le M$

If a rebuffered packet comes from delay line j ($j > \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil$) at time t, then we can deduce from the slot transition table that at time slot $t - (2\lceil \frac{M}{2} \rceil - j)$, the number of inserted packets between delay line j and delay line $2\lceil \frac{M}{2} \rceil - j$ satisfies

$$\begin{aligned} j - \left[2\left\lceil \frac{M}{2} \right\rceil - j\right] &= 2\left(j - \left\lceil \frac{M}{2} \right\rceil\right) \\ &\ge 2\left(\left\lceil \frac{M}{2} \right\rceil + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil - \left\lceil \frac{M}{2} \right\rceil\right) \ge \left\lceil \frac{M}{2} \right\rceil \end{aligned} \quad (25)$$

Therefore, the minimum number of buffered packets in this case is achieved when all the i rebuffered packets come from consecutive delay lines 1 to i, which is just $Q^*(i)$ as shown in Lemma 2.

(b) M is even.

For the sake of brevity, we omit the proof which is very similar to the cases when M is odd.

Summarizing the above results together, we have Lemma 3.

APPENDIX B
PROOF OF THEOREM 1

We solve the linear programming problem in (16) based on the results of Lemma 3.

(a) M is odd

Case 1. $\left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil \le i \le \lceil \frac{M}{2} \rceil$ and $\lceil \frac{M}{2} \rceil < M - i \le \lfloor \frac{M}{2} \rfloor + \left\lceil \frac{\lceil M/2 \rceil}{2} \right\rceil$

Since $\frac{5M-8(M-i)-1}{10} = \frac{8i-3M-1}{10}$, we have

$$
\begin{aligned}
&P^*(i)+P^*(M-i)\\
=&Q^*\left(i-\left\lceil\frac{i-3}{5}\right\rceil-1\right)+\sum_{j=1}^{\lceil\frac{i-3}{5}\rceil}2j\\
&+Q^*\left(\left\lceil\frac{M}{2}\right\rceil-\left\lceil\frac{8i-3M-1}{10}\right\rceil-1\right)+\sum_{j=1}^{M-i-\lceil\frac{M}{2}\rceil+\lceil\frac{8i-3M-1}{10}\rceil}2j\\
>&\left\lfloor Q^*\left(i-\frac{i-3}{5}-1\right)+\sum_{j=1}^{\frac{i-3}{5}}2j\right.\\
&\left.+Q^*\left(\left\lceil\frac{M}{2}\right\rceil-\frac{8i-3M-1}{10}-1\right)+\sum_{j=1}^{M-i-\lceil\frac{M}{2}\rceil+\frac{8i-3M-1}{10}}2j\right\rfloor-2\\
\geq&\left\lfloor\frac{5}{4}\left(\frac{i-3}{5}\right)^2-\frac{i-3}{2}\frac{i-3}{5}+\frac{i^2-2i+4}{4}\right.\\
&+\frac{5}{4}\left(\frac{8i-3M-1}{10}\right)^2+\left[2(M-i)-\frac{5}{4}M+\frac{1}{4}\right]\\
&\left.\cdot\left(\frac{8i-3M-1}{10}\right)+i^2-Mi+\frac{5M^2-2M+9}{16}\right\rfloor-2\\
=&\left\lfloor\frac{2}{5}i^2-\frac{2}{5}Mi+\frac{4M^2-4M+22}{20}\right\rfloor-2\\
=&\left\lfloor\frac{2}{5}i^2-\frac{2}{5}Mi+\frac{2M^2-2M+1}{10}\right\rfloor-1 \qquad (26)
\end{aligned}
$$

As the stationary point of the term enclosed within above floor function is $i=\frac{M}{2}$, the minimum value of Case 1 is

$$\left\lfloor\frac{M^2-2M+1}{10}\right\rfloor-1. \qquad (27)$$

Case 2. $0<i<\left\lceil\frac{\lceil M/2\rceil}{2}\right\rceil$ and $\lfloor\frac{M}{2}\rfloor+\left\lceil\frac{\lceil M/2\rceil}{2}\right\rceil\leq M-i\leq M$

$$
\begin{aligned}
&P^*(i)+P^*(M-i)\\
=&Q^*\left(i-\left\lceil\frac{i-3}{5}\right\rceil-1\right)+\sum_{j=1}^{\lceil\frac{i-3}{5}\rceil}2j+Q^*(M-i)\\
\geq&\left\lfloor Q^*\left(i-\frac{i-3}{5}-1\right)+\sum_{j=1}^{\frac{i-3}{5}}2j+Q^*(M-i)\right\rfloor\\
\geq&\left\lfloor\frac{5}{4}\left(\frac{i-3}{5}\right)^2-\frac{i-3}{2}\frac{i-3}{5}+\frac{i^2-2i+4}{4}\right.\\
&\left.+\frac{M+3}{2}(M-i)-\frac{M^2+4M-1}{4}\right\rfloor\\
=&\left\lfloor\frac{4i^2-(10M+34)i+5M^2+10M+16}{20}\right\rfloor \qquad (28)
\end{aligned}
$$

The stationary point of the term enclosed within above floor function is $i=\frac{5M+17}{4}$. Since $\frac{5M+17}{4}>\lceil\frac{M+1}{4}\rceil=\left\lceil\frac{\lceil M/2\rceil}{2}\right\rceil$, the minimum value of Case 2 is attained when $i=\lceil\frac{M+1}{4}\rceil-1$.

Case 3. $i=0$ and $M-i=M$

$$
\begin{aligned}
&P^*(i)+P^*(M-i)=P^*(M)=Q^*(M)\\
=&\frac{M+3}{2}M-\frac{M^2+4M-1}{4}=\frac{M^2+2M+1}{4} \qquad (29)
\end{aligned}
$$

Now, we will find out the minimum value of Function (16) when M is odd. First, by setting $i=1$ in Equation (28), we have

$$\left\lfloor\frac{4-(10M+34)+5M^2+10M+16}{20}\right\rfloor=\left\lfloor\frac{5M^2-14}{20}\right\rfloor$$

which is smaller than the value of Equation (29). Thus, the minimum value of Case 2 is smaller than that of Case 3. Second, to show the relationship between Case 1 and Case 2, we set $i=\lceil\frac{M+1}{4}\rceil$ in Equation (26) and set $i=\lceil\frac{M+1}{4}\rceil-1$ in Equation (28). Then we have

$$
\begin{aligned}
&\frac{4i^2-(10M+34)i+5M^2+10M+16}{20}\Big|_{i=\lceil\frac{M+1}{4}\rceil-1}\\
&-\frac{8i^2-8Mi+2M^2-2M+1}{10}\Big|_{i=\lceil\frac{M+1}{4}\rceil}\\
=&\frac{1}{20}\left[-4\left\lceil\frac{M+1}{4}\right\rceil^2-(2M+42)\left\lceil\frac{M+1}{4}\right\rceil+M^2+24M+52\right]\\
\geq&\frac{1}{20}\left[-4(\frac{M+1}{4}+1)^2-(2M+42)\left(\frac{M+1}{4}+1\right)\right.\\
&\left.+M^2+24M+52\right]\\
=&\frac{M^2+34M-27}{80} \qquad (30)
\end{aligned}
$$

Since $\frac{M^2+34M-27}{80}>0$, the above expression indicates that the minimum value of Case 1 is smaller than that of Case 2. Summarizing the above results, we know that when M is odd the minimum value of Function (16) is given by (27).

(b) M is even

The proof is similar to the above cases when M is odd, here we omit it.

APPENDIX C
PROOF OF THEOREM 2

Generally, the computation of the lower bound B^* on buffer size can be expressed as the following integer linear programming problem:

$$B^*=\min_i\left(P^*(i)+P^*(M-i-N+1)+N-1\right) \qquad (31)$$

$$\textit{Subject to}:\ \left\lceil\frac{M}{2}\right\rceil\leq i\leq M \qquad (32)$$

$$i \text{ is an integer} \qquad (33)$$

When $N-1\geq M$, the maximum acceptable buffer size is M, so we only need to consider the condition $N-1<M$ in the following proof, i.e.,

$$M+1-N>0 \qquad (34)$$

(a) M is odd

Case 1. $0<i\leq\lceil\frac{M}{2}\rceil$ and $0<M-i-N+1\leq\lceil\frac{M}{2}\rceil$, i.e.,

$$\begin{cases}0<i\leq\frac{M+1}{2}\\ \frac{M+1}{2}-N\leq i<M-N+1\end{cases} \qquad (35)$$

$$
\begin{aligned}
&P^*(i)+P^*(M-i-N+1)+N-1\\
&\geq\left\lfloor\frac{5}{4}\left(\frac{i-3}{5}\right)^2-\frac{i-3}{2}\cdot\frac{i-3}{5}+\frac{i^2-2i+4}{4}\right.\\
&\quad+\frac{5}{4}\left(\frac{M-i-N+1-3}{5}\right)^2-\frac{M-i-N+1-3}{2}\\
&\quad\left.\cdot\frac{M-i-N+1-3}{5}+\frac{1}{4}\left[(M-i-N+1)^2\right.\right.\\
&\quad\left.\left.-2(M-i-N+1)+4\right]\right\rfloor+N-3\\
&\geq\left\lfloor\frac{2}{5}i^2-\frac{2(M-N+1)}{5}i\right.\\
&\quad\left.+\frac{2(M-N+1)^2-2(M-N+1)+1}{10}\right\rfloor+N-2 \quad (36)
\end{aligned}
$$

The minimum value of Equation (36) is attained when $i=\frac{M-N+1}{2}$, which meets the constraint (35).

Case 2. $\left\lceil\frac{M+1}{4}\right\rceil\leq i\leq\left\lceil\frac{M}{2}\right\rceil$, $\left\lceil\frac{M}{2}\right\rceil<M-i-N+1\leq\left\lfloor\frac{M}{2}\right\rfloor+\left\lceil\frac{M+1}{4}\right\rceil$ and $N-1<\frac{M+1}{2}$

$$
\begin{aligned}
&P^*(i)+P^*(M-i-N+1)+N-1\\
&\geq\left\lfloor\frac{5}{4}\left(\frac{i-3}{5}\right)^2-\frac{i-3}{2}\frac{i-3}{5}+\frac{i^2-2i+4}{4}\right.\\
&\quad+\frac{5}{4}\left(\frac{8i+8N-3M-9}{10}\right)^2\\
&\quad+\left[2(M-N+1-i)-\frac{5}{4}M+\frac{1}{4}\right]\left(\frac{8i+8N-3M-9}{10}\right)\\
&\quad+(M-N+1-i)^2-M(M-N+1-i)+\left(\frac{M+1}{2}\right)^2\\
&\quad\left.-\frac{M+1}{2}+1+\frac{M^2-2M-3}{16}\right\rfloor+N-3\\
&=\left\lfloor\frac{2}{5}i^2-\frac{2(M-N+1)}{5}i\right.\\
&\quad\left.+\frac{2(M-N+1)^2-2(M-N+1)+1}{10}\right\rfloor+N-2 \qquad (37)
\end{aligned}
$$

Since the stationary point of the term enclosed within above floor function is $i=\frac{M-N+1}{2}$, the minimum value of Case 2 is

$$\left\lfloor\frac{(M-N)^2}{10}\right\rfloor+N-2. \qquad (38)$$

Case 3. $0<i<\left\lceil\frac{M+1}{4}\right\rceil$, $\frac{M-1}{2}+\left\lceil\frac{M+1}{4}\right\rceil<M-i-N+1\leq M$ and $N-1<\left\lceil\frac{M+1}{4}\right\rceil$

$$
\begin{aligned}
&P^*(i)+P^*(M-i-N+1)\\
&\geq\left\lfloor\frac{5}{4}\left(\frac{i-3}{5}\right)^2-\frac{i-3}{2}\frac{i-3}{5}+\frac{i^2-2i+4}{4}\right.\\
&\quad\left.+\frac{M+3}{2}(M-i-N+1)-\frac{M^2+4M-1}{4}\right\rfloor\\
&=\left\lfloor\frac{4i^2-(10M+34)i+5M^2+10M+16}{20}\right.\\
&\quad\left.-\frac{(M+3)(N-1)}{2}\right\rfloor \qquad (39)
\end{aligned}
$$

The stationary point of the term enclosed within above floor function is $i=\frac{5M+17}{4}>\left\lceil\frac{M+1}{4}\right\rceil$. Since $\frac{5M+17}{4}$ is out of the range of i in this case, the minimum of Case 3 is achieved when $i=\left\lceil\frac{M+1}{4}\right\rceil-1$ (or $M-i=\left\lfloor\frac{M}{2}\right\rfloor+\left\lceil\frac{M+1}{4}\right\rceil+1$).

From (35) and (37), we can see that Case 1 and Case 2 have the same results. To show the relationship between Case 2 and Case 3, we set $i=\left\lceil\frac{M+1}{4}\right\rceil$ in (37) and set $i=\left\lceil\frac{M+1}{4}\right\rceil-1$ in (39). Then we have

$$
\begin{aligned}
&\left[\frac{4i^2-(10M+34)i+5M^2+10M+16}{20}\right.\\
&\qquad\left.-\frac{(M+3)(N-1)}{2}\right]\Big|_{i=\left\lceil\frac{M+1}{4}\right\rceil-1}\\
&-\left[\frac{2}{5}i^2+\frac{2(N-M-1)}{5}i\right.\\
&\qquad\left.+\frac{4(M-N+1)^2-4(M-N+1)+2}{20}\right]\Big|_{i=\left\lceil\frac{M+1}{4}\right\rceil}\\
&+M^2+24M+52+(2M+34)(1-N)-4(1-N)^2\Big]\\
&\geq\frac{1}{20}\left[-4\left\lceil\frac{M+1}{4}\right\rceil^2-\left(2M+42+8\left\lceil\frac{M+1}{4}\right\rceil\right)\left\lceil\frac{M+1}{4}\right\rceil\right.\\
&\quad\left.+M^2+24M+52+(2M+34)\left\lceil\frac{M+1}{4}\right\rceil-4\left\lceil\frac{M+1}{4}\right\rceil^2\right]\\
&\geq\frac{1}{20}\left[-16\left(\frac{M+5}{4}\right)^2-8\left(\frac{M+5}{4}\right)+M^2+24M+52\right]\\
&=\frac{1}{20}(12M+22)>0 \qquad (40)
\end{aligned}
$$

Summarizing the above expressions, we know that when M is odd the minimum value of Function (31) is given by (38).

(b) M is even

We omit the proof since it is quite similar to the proof when M is odd.

Summarizing the above results together, the Theorem 2 follows.

REFERENCES

[1] J. Ramamirtham and J. Turner, "Time sliced optical burst switching," in *Proc. IEEE INFOCOM 2003*, vol. 3, pp. 2030-2038, 2003.

[2] R. L. Cruz and J. T. Tsai, "Cod: alternative architectures for high speed packet switching," *IEEE/ACM Trans. Networking*, vol. 4, pp. 11-20, 1996.

[3] I. Chlamtac, A. Fumagalli, L. G. Kazovsky, P. Melman, W. H. Nelson, P. Poggiolini, M. Cerisola, A. N. M. M. Choudhury, T. K. Fong, R. T. Hofmeister, C. L. Lu, A. Mekkittikul, D. J. M. S. IX, C. J. Suh, and E. W. M. Wong, "Cord: contention resolution by delay lines," *IEEE J. Sel. Areas Commun.*, vol. 14, pp. 1014-1029, 1996.

[4] D. K. Hunter, W. D. Cornwell, T. H. Gilfedder, A. Franzen, and I. Andonovic, "Slob: a switch with large optical buffers for packet switching," *J. Lightwave Technol.*, vol. 16, no. 10, pp. 1725-1736, Oct. 1998.

[5] S. Yao, B. Mukherjee, and S. Dixit, "Advances in photonic packet switching: an overview," *IEEE Commun. Mag.*, vol. 38, no. 2, pp. 84-94, Feb. 2000.

[6] D. K. Hunter, M. C. Chia, and I. Andonovic, "Buffering in optical packet switches," *J. Lightwave Technol.*, vol. 16, no. 12, pp. 2081-2094, Dec. 1998.

[7] M. Arpaci and J. A. Copeland, "Buffer management for shared-memory ATM switches," *IEEE Commun. Surveys Tutorials*, vol. 3, no. 1, pp. 2-10, 2000.

[8] M. J. Karol, "Shared-memory optical packet (ATM) switch," *SPIE: Multigigabit Fiber Commun. Syst.*, vol. 2024, pp. 212-222, July 1993.

[9] C. S. Chang, D. S. Lee, and C. K. Tu, "Recursive construction of optical multiplexers with switched delay lines," *IEEE Trans. Inf. Theory*, vol. 50, pp. 3221-3233, Dec. 2004.

[10] C. S. Chang, Y. T. Chen, and D.-S. Lee, "Constructions of optical FIFO queues," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2838-2843, 2006.

[11] A. D. Sarwate and V. Anantharam, "Exact emulation of a priority queue with a switch and delay lines," *Queueing Systems: Theory Applications*, vol. 53, pp. 115-125, July 2006.

[12] H. C. Chiu, C. S. Chang, J. Cheng, and D. S. Lee, "A simple proof for the constructions of optical priority queues," *Queueing Systems: Theory Applications*, vol. 56, pp. 73-77, 2007.
[13] ——, "Using a single switch with $o(m)$ inputs/outputs for the construction of an optical priority queue with $o(m^3)$ buffer," in *Proc. IEEE INFOCOM Mini Symposium*, 2007.
[14] C. C. Chou, C. S. Chang, D. S. Lee, and J. Cheng, "A necessary and sufficient condition for the construction of 2-to-1 optical FIFO multiplexers by a single crossbar switch and fiber delay lines," *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4519-4531, 2006.
[15] H. Lan, C. Chang, J. Cheng, and D. Lee, "Constructions and analysis of crosstalk-free optical queues," in *Proc. IEEE International Conf. High Performance Switching Routing (HPSR'08)*, Shanghai, China, May 2008, pp. 27-32.
[16] S. Iyer, R. Zhang, and N. Mckeown, "Routers with a single stage of buffering," in *Proc. SIGCOMM 2002*, Pittsburgh, PA, USA, Aug. 2002.
[17] A. Franzen, D. K. Hunter, and I. Andonvic, "Synchronization in optical packet switched networks," in *Proc. WAON98, Workshop All-Optical Networks*, Zagreb, Croatia, May 1998.
[18] T. Sakamoto, A. Okada, M. Hirayama, Y. Sakai, O. Moriwaki, I. Ogawa, R. Sato, K. Noguchi, and M. Matsuoka, "Optical packet synchronizer using wavelength and space switching," *IEEE Photonics Technol. Lett.*, vol. 14, no. 9, pp. 1360-1362, Sep. 2002.
[19] V. T. Nguyen, R. L. Cigno, and Y. Ofek, "Tunable laser-based design and analysis for fractional lambda switches," *IEEE Trans. Commun.*, vol. 56, no. 6, pp. 957-967, 2008.
[20] R. S. Tucker and W. D. Zhong, "Photonic packet switching: an overview," *IEICE Trans. Commun.*, vol. E82-B, no. 2, pp. 254-264, 1999.
[21] G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, "Optical switching: switch fabrics, techniques, and architectures," *J. Lightwave Technol.*, vol. 21, no. 2, pp. 384-405, Feb. 2003.
[22] S. J. B. Yoo, "Optical packet and burst switching technologies for the future photonic Internet," *J. Lightwave Technol.*, vol. 24, no. 12, pp. 4468-4492, Dec. 2006.
[23] H. Ngo, D. Pan, and C. Qiao, "Constructions and analyses of nonblocking WDM switches based on arrayed waveguide grating and limited wavelength conversion," *IEEE/ACM Trans. Networking*, vol. 14, no. 1, pp. 205-217, Feb. 2006.
[24] X. Jiang, A. Pattavina, and S. Horiguchi, "Strictly nonblocking f-cast photonic networks," *IEEE/ACM Trans. Networking*, vol. 16, no. 3, pp. 732-745, June 2008.

Xiaoliang Wang received his B.S., M.S. degrees in 2003 and 2006 respectively, all from Xidian University, Xi'an, China. He is currently pursing the Ph.D degree in the Graduate School of Information Sciences, Tohoku University, Japan. His research interests include optical switching networks, scheduling algorithm and router design.

Xiaohong Jiang received his B.S., M.S. and Ph.D degrees in 1989, 1992, and 1999 respectively, all from Xidian University, Xi'an, China. He is currently an Associate Professor in the Department of Computer Science, Graduate School of Information Science, TOHOKU University, Japan. Before joining TOHOKU University, Dr.Jiang was an assistant professor in the Graduate School of Information Science, Japan Advanced Institute of Science and Technology (JAIST), from Oct.2001 to Jan.2005. Dr. Jiang was a JSPS (Japan Society for the Promotion of Science) postdoctoral research fellow at JAIST from Oct.1999-Oct. 2001. He was a research associate in the Department of Electronics and Electrical Engineering, the University of Edinburgh from Mar.1999-Oct.1999. Dr. Jiang's research interests include optical switching networks, routers, network coding, WDM networks, VoIP, interconnection networks, IC yield modeling, timing analysis of digital circuits, clock distribution and fault-tolerant technologies for VLSI/WSI. He has published over 120 referred technical papers in these areas. He is a member of IEEE.

Achille Pattavina received the Dr. Eng. degree in Electronic Engineering from University "La Sapienza" of Rome (Italy) in 1977. He was with the same University until 1991 when he moved to "Politecnico di Milano", Milan (Italy), where he is now Full Professor. He has been author of more than 100 papers in the area of Communications Networks published in leading international journals and conference proceedings. He has been guest or co-guest editor of special issues on switching architectures in IEEE and non-IEEE journals. He has been engaged in many research activities, including European Union funded projects. Dr. Pattavina has authored two books, *Switching Theory, Architectures and Performance in Broadband ATM Networks* (New York: Wiley, 1998) and *Communication Networks* (McGraw-Hill, 2002, in Italian). He has been Editor for Switching Architecture Performance of the IEEE TRANSACTIONS ON COMMUNICATIONS since 1994 and Editor-in-Chief of the EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS since 2001. He is a Senior Member of the IEEE Communications Society. His current research interests are in the area of optical networks and traffic modeling.

Susumu Horiguchi (M'81-SM'95) received the B.Eng the M.Eng and PhD degrees from Tohoku University in 1976, 1978 and 1981 respectively. He is currently a Full Professor in the Graduate School of Information Sciences, Tohoku University. He was a visiting scientist at the IBM Thomas J. Watson Research Center from 1986 to 1987. He was also a professor in the Graduate School of Information Science, JAIST (Japan Advanced Institute of Science and Technology). He has been involved in organizing international workshops, symposia and conferences sponsored by the IEEE, IEICE, IASTED and IPS. He has published over 150 papers technical papers on optical networks, interconnection networks, parallel algorithms, high performance computer architectures and VLSI/WSI architectures. Prof. Horiguchi is members of IEICE, IPS and IASTED.