Low-Loss Switching Fabric Design for Recirculating Buffer in WDM Optical Packet Switching Networks Using Arrayed Waveguide Grating Routers

Zhenghao Zhang and Yuanyuan Yang, *Senior Member, IEEE*

***Abstract*—In this paper, we give a new switching fabric design for the recirculating buffer in optical packet switching networks. We note that since a packet to be buffered can be routed to *any* delay lines, the switching fabric connecting packets to the delay lines can be simplified. We give a design based on the arrayed waveguide grating router, and give a simple linear time-control algorithm for assigning buffer locations to the packets. To the best of our knowledge, this is the first switching fabric specifically designed for recirculating buffers which takes advantage of the fact that packets can be routed to any delay lines.**

***Index Terms*—Arrayed waveguide grating router (AWGR), concentrator, optical packet switching (OPS), recirculating buffer.**

I. INTRODUCTION AND BACKGROUND

OPTICAL networks with wavelength division multiplexing (WDM) are now widely regarded as the candidate for future backbone networks because of their nearly unlimited bandwidth. In this paper, we study the optical packet switching (OPS) network, as it has better flexibility and is more bandwidth-efficient [7].

In OPS networks, output contention may arise when packets on the same wavelength are destined to the same output fiber. Output contention can be resolved by temporarily storing packets in a buffer. Currently, optical buffers can be emulated with fiber delay lines (FDLs) which delay the incoming signal for a certain amount of time proportional to the lengths of the FDLs. FDLs are large in size, and as suggested by [1], [2], [8], [10], and [12], it is more cost-effective to let the FDLs be shared by all outputs to reduce the cost and size of the switch. The basic idea of a WDM switch with shared buffer is shown in Fig. 1(a). As in [5] and [7], we assume that the OPS network is time-slotted, the duration of a packet is one time slot, and the packets arrive at the switch at the beginning of time slots. The switch has N input fibers and N output fibers. Inside the switch, there are B FDLs, each capable of delaying a packet for one time slot. To ensure an acceptable throughput, $B \geq N$, as can be seen in the results in [12], where N is usually around 4 but B often needs to be no less than 16. On each fiber, there are k wavelengths. The input composite signal is first demultiplexed, then the separated signals on distinct wavelengths are fed into wavelength converters to be converted to proper wavelengths. The converted signals are then sent to a switching fabric. If there is no contention, an arriving packet is routed to its destination output fiber by the switching fabric; otherwise, it is routed to one of the delay lines, also by the switching fabric. After being delayed for one time slot, the buffered packet will come out of the delay line. If the packet still cannot be sent to the output fiber, it will be routed to the delay lines again; otherwise, it will be sent to the output fiber. The switch is named "recirculating" because a packet will keep circulating in the switch until being sent out to the output fiber.

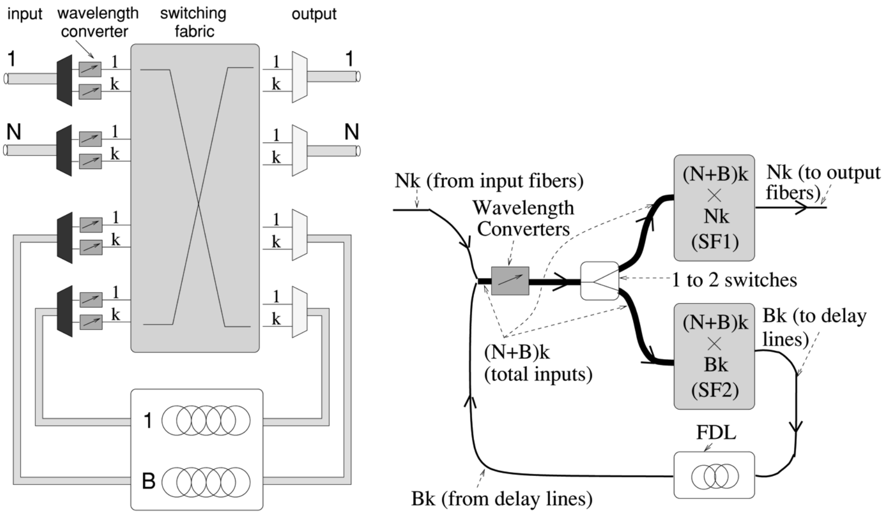
Note that the switching fabric has two types of inputs: the "input fiber inputs," which are packets arriving at the input fibers of the switch, and the "delay line inputs," which are packets coming out of the delay lines. It also has two types of outputs: the "output fiber outputs," which are the wavelength channels on the output fibers of the switch, and the "delay line outputs," which are the wavelength channels on the delay lines. It is convenient to consider the switching fabric as a single piece of switching fabric, that is, to assume that it connects both types of inputs to both types of outputs. However, we note that the requirements for connecting the inputs to the two types of outputs are quite different, because while a packet that is sent to the "output fiber outputs" must be sent to the correct output fiber, a packet that is sent to the "delay line outputs" can be sent to *any* delay line, since any delay line will give it a one-time-slot delay, as required. As a result, the switching fabric connecting the inputs to the "delay line outputs" can be much simpler than the the switching fabric connecting the inputs to the "output fiber outputs." Based on this observation, we propose to "break" the switching fabric into two, as has been conceptually shown in Fig. 1(b), in which one fabric, Switching Fabric 1 (SF1), is used for connecting the inputs to the "output fiber outputs," and the other fabric, Switching Fabric 2 (SF2), is used for connecting the inputs to the "delay line outputs." As in the figure, after wavelength conversion, each input will be sent to a 1×2 switch which will forward the packet to one of the switching fabrics. Packets that can be sent to the output fibers are sent to SF1, which will route them out of the switch. Packets that can be buffered are sent to SF2, which will route them to wavelength channels on the delay lines. The exits of the delay lines are connected back to the input side of the switch, such that after a packet has been delayed for one time slot, it can be switched again, either to the output fibers to exit the switch or to the delay lines one more

---

Paper approved by A. Pattavina, the Editor for Switching Architecture Performance of the IEEE Communications Society. Manuscript received September 29, 2004; revised October 13, 2005 and January 25, 2006. This work was supported in part by the U.S. National Science Foundation under Grants CCR-0073085 and CCR-0207999.

The authors are with the Department of Electrical and Computer Engineering, State University of New York, Stony Brook, NY 11794 USA (e-mail: yang@ece.sunysb.edu).

Digital Object Identifier 10.1109/TCOMM.2006.878838



(a) (b)

Fig. 1. (a) Optical packet switch with recirculating buffering and wavelength conversion. (b) Using two switching fabrics.

time. It should be noted that although a total of $(N+B)k$ additional 1×2 switches are needed, this new design will reduce the overall cost of the switch, since B is usually much larger than N, and SF2 can be much simpler than a nonblocking switching fabric.

SF1 can be a crossbar. In this paper, we focus on the design of SF2. SF2 has $(B+N)k$ inputs and Bk outputs. Note that at any time, the number of packets routed by SF2 should not exceed Bk, since the buffer can hold only Bk packets. Since the total number of packets can be as high as $(B+N)k$, we assume that there is an algorithm that selects no more than Bk packets to be sent to SF2 based on criteria such as quality of service and fairness, etc. Packets that are not sent to SF2 are either sent to SF1 to be routed to the output fibers or simply dropped. In this paper, we only focus on the design of SF2 and realize the function of the following. *Given* no more than Bk packets, either coming from the input fibers or from the delay lines, route each packet to a wavelength channel on the delay lines. Note that it does not guarantee that a packet can always be routed to a specific wavelength channel; instead, it only guarantees that a packet can always be routed to *some* wavelength channel. In other words, it is a *concentrator* [9], and in the rest of the paper, we will also refer to SF2 as the WDM concentrator.

We will use the arrayed waveguide grating router (AWGR) to implement SF2. AWGR is a low-loss all-optical switching device, and is now commercially available. It is desirable to use a low-loss device to build SF2, since a buffered packet has to be switched by SF2, and thus, experiences switching loss every time it is recirculated through the switch. One major limitation of an AWGR is that the number of wavelengths that can be routed from an input fiber to an output fiber is often much less than k (the total number of wavelengths), while in practice, the number of packets going from one input fiber to an output fiber can be as many as k. As a result, it is relatively complicated to use AWGR to build nonblocking switching fabrics [4], [11]. However, in a concentrator, this is no longer a constraint, since the packets do not have specific destinations. Therefore, AWGR can be considered as an ideal device for implementing SF2.

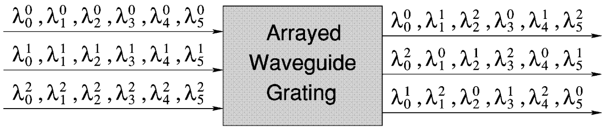


Fig. 2. AWGR with three input/output fibers and six wavelengths per fiber.

The rest of the paper is organized as follows. Section II describes the routing functions of AWGR. Section III describes the hardware and control algorithm of the concentrator using AWGR. Finally, Section IV concludes this paper.

II. ARRAYED WAVEGUIDE GRATING ROUTER

AWGR is a device that provides fixed routing patterns between inputs and outputs. The lengths of the paths from an input to the outputs are carefully calibrated, such that at an output, only one wavelength will "add in phase" while other wavelengths will "cancel out." As a result, different wavelengths will be routed to different output fibers, depending on the wavelengths. Note that signals are not split in an AWGR, and thus will not experience splitting loss as in broadcast-and-select switches.

It is usually assumed that in an AWGR with B input/output fibers, wavelength λ_j on the ith input fiber is routed to the $[(j-i)\mathrm{mod}B]$th output fiber [11]. As in [11], we also assume that B is small, as compared with k, and k is a multiple of B. Given the large number of wavelengths a fiber can carry, practically, this is not a restriction. Fig. 2 shows an AWGR with three input fibers and three output fibers, with each fiber carrying six wavelengths, where λ_j^i represents λ_j on the ith input fiber. In an AWGR, to send a packet from to an input fiber to an output fiber, the wavelength of the packet should be converted to one of the wavelengths that will be routed to this output fiber.

III. THE WDM CONCENTRATOR

A. Hardware Setup

Our design of SF2 is shown in Fig. 3. The inputs to SF2 are the Nk wavelength channels from the input fibers (input fiber

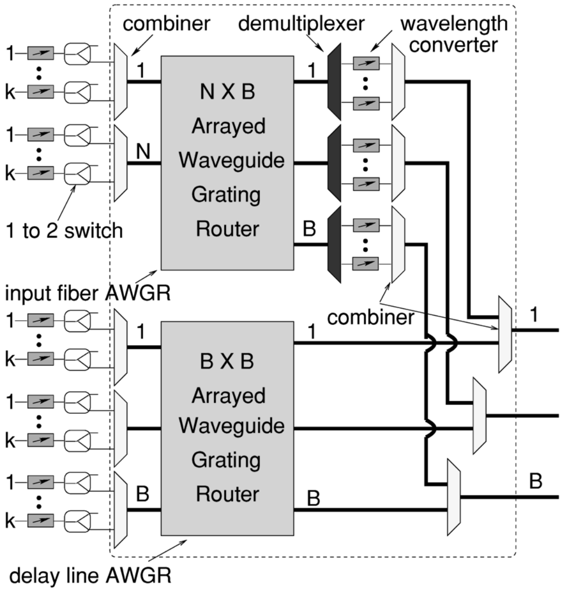


Fig. 3. $(N+B)k \times Bk$ WDM concentrator.

TABLE I
CONTROL ALGORITHM FOR THE CONCENTRATOR

```
$P \leftarrow 0$;
for $i := 0$ to $B-1$ do (checking the DL-AWGR)
  for $j := 0$ to $k-1$ do
    if there is a packet on $\lambda_j$ on the $i_{th}$ input fiber of the DL-AWGR
      Convert $\lambda_j$ to a wavelength that will be routed to the $P_{th}$
      output fiber of the DL-AWGR.
      $P \leftarrow (P+1) \mod B$;
    end if
  end for
end for
for $i := 0$ to $N-1$ do (checking the IF-AWGR)
  for $j := 0$ to $k-1$ do
    if there is a packet on $\lambda_j$ on the $i_{th}$ input fiber of the IF-AWGR
      Convert $\lambda_j$ to a wavelength that will be routed to the $P_{th}$
      output fiber of the IF-AWGR.
      if this wavelength has been used in the $P_{th}$ delay line
        Convert it to a free wavelength in the $P_{th}$ delay line
      end if
      $P \leftarrow (P+1) \mod B$;
    end if
  end for
end for
```

inputs), and Bk wavelength channels from the delay lines (delay line inputs), which are shown in the upper and lower part of the input side, respectively. Before entering the concentrator, the wavelength of a packet should have been converted to a proper wavelength, which is determined by the control algorithm to be described later. Packets coming from the same input fiber or the same delay line are recombined into one composite signal to be routed by AWGRs. There are two AWGRs. One has N input fibers and B output fibers, called the "input fiber AWGR" (IF-AWGR), to which packets from the input fibers are sent. The other has B input fibers and B output fibers, called the "delay line AWGR" (DL-AWGR), to which packets from the delay lines are sent. The routing pattern of the DL-AWGR is the same as the AWGR in Section II. The routing pattern of the N input fibers of the IF-AWGR is the same as the first N input fibers of the DL-AWGR. At the output side of the IF-AWGR, the signal on an output fiber is first demultiplexed, then the signal on each wavelength is sent to a wavelength converter to be converted to a proper wavelength. After that, the signals are recombined into one composite signal, which will be further combined with the signal at the corresponding output fiber of the DL-AWGR and then be sent to the delay line.

The wavelength conversion following the IF-AWGR is for resolving possible contentions, because, as can be seen in the figure, packets on the ith output fiber of both AWGRs are sent to the ith delay line. If there is no wavelength conversion, contention will arise; for example, when there is a packet on λ_0 on the first input fiber of both the IF-AWGR and the DL-AWGR, both will be routed to the first delay line on λ_0. Clearly, this problem can be solved by converting the wavelength of one packet to another wavelength. Note that there are only Nk/B nonidle wavelengths on an output fiber of the IF-AWGR, and therefore, only Nk wavelength converters are needed. This is because, for example, suppose $N=2, B=3, k=6$. Since the routing patterns of the input fibers of the IF-AWGR are the same as the first two input fibers of a 3×3 AWGR, the first output fiber of the IF-AWGR will have four nonidle wavelengths, which are λ_0^0 and λ_3^0 (from input fiber 0) and λ_1^1 and λ_4^1 (from input fiber 1), as can be derived from Fig. 2, and λ_2 and λ_5 on this output fiber will never receive any packets. This is one of the reasons for us to choose to convert the wavelengths of packets from the IF-AWGR rather than from the DL-AWGR, since otherwise, Bk wavelength converters are needed where $B > N$. Another reason is that more devices, e.g., more demultiplexers and combiners, means more insertion loss. Packets from the input fibers will experience this additional loss in the additional wavelength conversion stage; however, they experience this loss only once, since after that, they will enter the delay lines, and packets from the delay lines need not go through the additional wavelength conversion stage, and need not suffer this additional loss every time they are circulated in the switch. Also note that the switching time of the switch is mainly determined by the tuning time of the wavelength converters, which are fixed-input/tunable-output converters. The tuning time of the inside tunable laser of the converter is on the order of several tens of nanoseconds [13], therefore, the switching time of the switch is on the order of several tens of nanoseconds.

B. Control Algorithm

With this hardware setup, a control algorithm is needed to assign the wavelength channels on the delay lines to the packets. Recall that SF2 should be able to connect any input to some output if the total number of active inputs is no more than Bk. Our algorithm is shown in Table I, where it can be easily seen that it runs in $O(N+B)k$ time which is linear to the input size. In the following, we explain how it works.

First, we will consider inputs to the DL-AWGR. We will check input fibers to the DL-AWGR one by one, and when checking an input fiber, we check the wavelengths one by one. When a packet is found, its wavelength will be converted to a wavelength that will be routed to the Pth output fiber of the DL-AWGR, after which $P \leftarrow (P+1)\text{mod}B$, where P is an integer initially set to zero.

We now explain why this routing mechanism is correct. First, note that packets from different input fibers of the DL-AWGR are routed to different wavelength channels. Thus, the routing mechanism is correct if it does not attempt to route more than k/B packets on one input fiber of the DL-AWGR to the same output fiber of the DL-AWGR, since, as can be seen in Fig. 2, at most k/B packets can be sent from one input fiber to one output fiber. However, clearly this will not happen, since when checking an input fiber, we will distribute the packets over all output fibers in a round-robin manner, and no output fiber will receive more than k/B packets from one input fiber.

After routing the inputs to the DL-AWGR, inputs to the IF-AWGR are routed in a similar manner, that is, we always convert the wavelength of a packet to a wavelength, say, λ_a, that will be routed to the Pth output fiber of the IF-AWGR. However, if there has already been another packet that will be routed to λ_a on the Pth delay line, that is, if this wavelength channel has already been occupied, we will convert the wavelength of the packet currently under consideration from λ_a to a free wavelength on the the Pth delay line by the wavelength converters following the IF-AWGR. Note that the routing mechanism for the IF-AWGR is correct if it does not route a packet to a delay line already full, i.e., with no free wavelength channels. However, this will not happen, since due to the round-robin manner of the algorithm, the Pth output fiber of the IF-AWGR will be connected to the Pth delay line which always has the minimum number of occupied wavelength channels. That is, if the Pth delay line is full, the buffer is full. However, if the buffer is full, there cannot be more packets that should be routed by SF2, since the total number of packets is no more than Bk.

Combining the above discussions, we have the following theorem.

Theorem 1: The switching fabric shown in Fig. 3 is capable of assigning a wavelength channel on the delay lines to each packet if the total number of packets is no more than Bk.

IV. CONCLUSION

In this paper, we gave a new switching fabric design for the recirculating buffer in OPS networks. We used the fact that a packet to be buffered can be routed to *any* delay lines to simplify the switching fabric. We gave a design based on AWGR, and gave a simple linear time control algorithm for assigning buffer locations to the packets.

REFERENCES

[1] G. Bendeli, "Performance assessment of a photonic ATM switch based on a wavelength controlled fiber loop buffer," in *OFC Tech. Dig.*, 1996, pp. 106–107.

[2] D. K. Hunter, "WASPNET: A wavelength switched packet network," *IEEE Commun. Mag.*, vol. 37, no. 3, pp. 120–129, Mar. 1999.

[3] J. Cheyns, C. Develder, E. V. Breusegem, A. Ackaert, M. Pickavet, and P. Demeester, "Routing in an AWG based optical packet switch," *Photon. Netw. Commun.*, vol. 5, no. 1, pp. 69–80, Jan. 2003.

[4] J. Ramamirtham and J. Turner, "Design of wavelength converting switches for optical burst switching," in *Proc. IEEE INFOCOMM*, New York, NY, 2002, vol. 2, pp. 362–370.

[5] S. L. Danielsen, C. Joergensen, B. Mikkelsen, and K. E. Stubkjaer, "Analysis of a WDM packet switch with improved performance under bursty traffic conditions due to tunable wavelength converters," *J. Lightw. Technol.*, vol. 16, no. 5, pp. 729–735, May 1998.

[6] R. Ramaswami and K. N. Sivarajan, *Optical Networks: A Practical Perspective*. New York: Academic, 2001.

[7] L. Xu, H. G. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 136–142, Jan. 2001.

[8] D. K. Hunter and I. Andronovic, "Approaches to optical Internet packet switching," *IEEE Commun. Mag.*, vol. 38, no. 9, pp. 116–122, Sep. 2000.

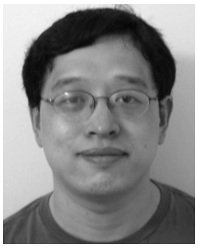
[9] A. Y. Oruc and H. M. Huang, "Crosspoint complexity of sparse crossbar concentrators," *IEEE Trans. Inf. Theory*, vol. 42, no. 9, pp. 1466–1471, Sep. 1996.

[10] F. Xue and S. J. Ben Yoo, "High capacity multiservice optical label switching for the next-generation Internet," *IEEE Commun. Mag.*, vol. 42, no. 5, pp. 16–22, May 2004.

[11] 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. Netw.*, vol. 14, no. 1, pp. 205–217, Feb. 2006.

[12] S. Yao, B. Mukherjee, S.J.B. Yoo, and S. Dixit, "A unified study of contention-resolution schemes in optical packet-switched networks," *IEEE J. Lightw. Technol.*, vol. 21, no. 3, pp. 672–683, Mar. 2003.

[13] J. Gripp, M. Duelk, J. E. Simsarian, A. Bhardwaj, P. Bernasconi, O. Laznicka, and M. Zirngibl, "Optical switch fabrics for ultra-high-capacity IP routers," *IEEE J. Lightw. Technol.*, vol. 21, no. 11, pp. 2839–2850, Nov. 2003.



Zhenghao Zhang received the B.Eng. and M.S. degrees in electrical engineering from Zhejiang University, Hangzhou, China, in 1996 and 1999, respectively. Since 2001, he has been working toward the Ph.D. degree in the Department of Electrical and Computer Engineering, State University of New York at Stony Brook.

From 1999 to 2001, he worked in industry as a software engineer for embedded systems design. His research interest includes scheduling and performance analysis of optical networks.

Yuanyuan Yang (S'91–M'91–SM'98) received the B.Eng. and M.S. degrees in computer science and engineering from Tsinghua University, Beijing, China, and the M.S.E. and Ph.D. degrees in computer science from The Johns Hopkins University, Baltimore, MD.

She is currently a Professor of Computer Engineering and Computer Science with the State University of New York at Stony Brook. Her research interests include parallel and distributed computing and systems, high speed networks, optical and wireless networks and high performance computer architecture. She has published extensively in major journals and refereed conference proceedings, and holds six U.S. patents. She is an Editor for the *Journal of Parallel and Distributed Computing*. She has served on National Science Foundation review panels and program/organizing committees of numerous international conferences in her areas of research.

Dr. Yang is an Editor for the IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS.