High-Capacity Packet-Switched Optical Ring Network

A. Smiljanić, M. Boroditsky, and N.J. Frigo
AT&T Laboratories, Middletown, NJ 07748, USA
aleks,boroditsky,njf@research.att.com

Abstract— **We propose novel architecture, MAC and admission control protocols for the high-capacity packet-switched optical ring network. In this network link capacity significantly exceeds node bit rate. Nodes transmit and receive packets on multiple wavelengths in parallel by using novel optical techniques. Network control is simple since the load is balanced over wavelengths at the physical layer. MAC protocol is based on credits, and derived admission control protocol has similar complexity as in a single channel network. Consequently, network can follow fast traffic changes typical in data networks.**

I. Introduction

Regional access networks with ring topology are attractive because they easily recover from a single failure. In single-channel ring networks where nodes operate at the link bit-rate, network control is simple. In FDDI [4], MAC protocol is based on token, and admission control protocol ensures that the sum of all requested bit-rates is smaller than the link bit-rate. In ring network with spatial reuse [2], [3], MAC protocol is based on credits, and admission control protocol ensures that the sum of requested bit-rates passing through any link is smaller than its capacity. With a development of WDMA technology, the total throughput of a packet-switched ring network can be significantly increased. New network architecture requires appropriate MAC and admission control protocols.

An architecture for WDMA packet-switched ring network has been proposed in [5], [6], [7]. Each node is equipped with a fast tunable laser and slow tunable optical filter, because fast tunable receivers are unavailable. For this architecture, network control becomes formidable task. Whenever traffic pattern changes, allocation of wavelengths to the receivers should be recalculated and changed accordingly. The bandwidth request might be rejected even though involved tranceivers and links have enough of the spare capacity, but because the particular wavelength is overloaded.

We propose new network architecture that features simple MAC and admission control protocols. A "composite" packet is transmitted on all wavelengths in parallel by using fast tunable laser, wavelength dependent delay line, and 2 × 2 optical switch. This composite packet is received by using 2 × 2 optical switch, and wavelength dependent delay line. Since each packet is transmitted and received on all wavelengths, the wavelengths are evenly utilized. Because there is no partitioning of the capacity , the bandwidth can be allocated as long as transceivers and links in question have enough of spare capacity. We will show that the complexity of admission control protocol is almost as simple as in a single channel network.

II. Network Architecture and MAC Protocol

The node architecture is shown in Figure 1. Time is divided into slots of duration equal to T_p. Fast tunable laser at some node starts transmission W time slots before its scheduled time slot, where W is the number of wavelengths. In each following time slot, it transmits data on different wavelength. Signal passes through the array of fiber gratings separated by delay lines so that the data transmitted at different wavelengths are aligned in time. We demonstrated this novel technique, termed as wavelength stacking, in [1]. Packet is then transmitted to the network on all wavelengths in parallel by setting switch S to the cross state. On the receiver side, the reverse procedure is performed. Packet is received when switch S is in the cross state, and then unstacked when passing through the same array of fiber-gratings and delay lines.

Because wavelength stacking takes W time slots, a node should decide in advance when to access the medium. Separate wavelength is used as a control channel for the advanced reservations. Time slots are grouped into cycles of length W. Each node may transmit and receive at most one packet within each cycle which matches its transmission and reception speeds. Switches T and R in Fig. 1 synchronize wavelength stacking and unstacking. Time diagram in Figure 2 shows transmissions and receptions of node i. Whenever node i reserves a time slot, its laser starts transmission in the next cycle. Wavelength stacking is completed in the last time slot of this cycle, and the packet is stored into the buffer (delay line) by setting switch T in the cross state. Packet is stored as long as switch T is in the bar state. Packet is transmitted to the network by setting switches T and S in the cross states exactly $2W$ time slots after the reservation. Whenever node i recognizes its address on the control channel, it stores

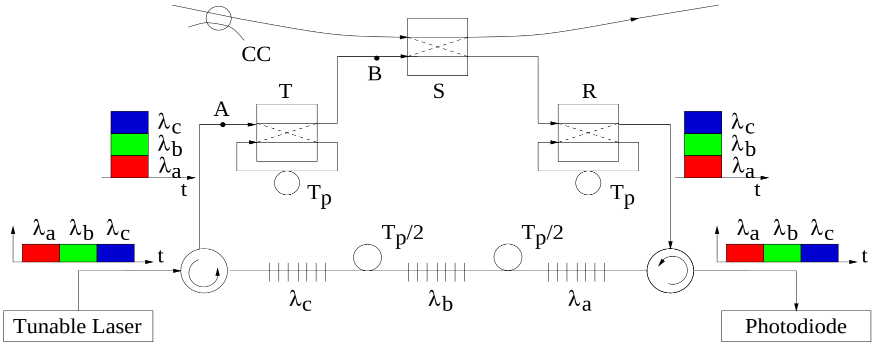

Fig. 1. Node architecture, $W = 3$.

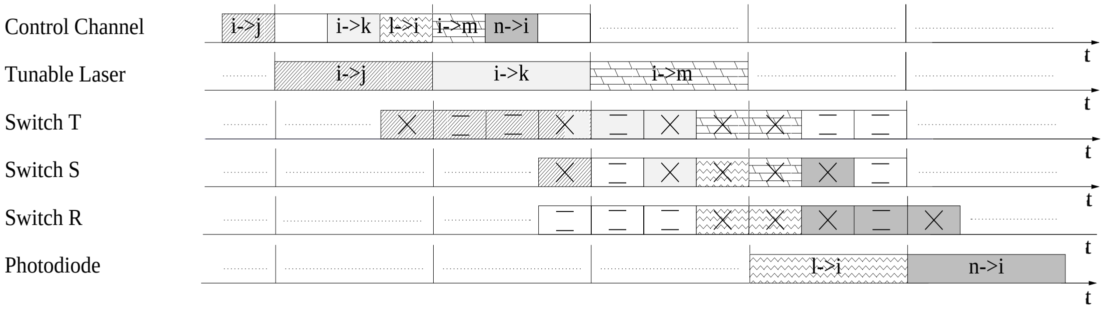

Fig. 2. Time diagram, $W = 3$.

the packet into the buffer by setting switches S and R in the cross states $2W$ time slots after the adress notification. Node starts unstacking of the packet at the beginning of the next cycle by setting switch R in the cross state. Node deletes a packet that it receives as well as its reservation.

III. Admission Control

Explicit bandwidth reservations are provided through credits. Each source-destination pair uses a specified number of credits per frame period. Frame period lasts until all busy queues have used their credits. One priority bit is associated to each packet, which is set to one if the credit was used for this packet transmission and to zero otherwise. A packet that uses the credit, can replace a packet that does not and whose priority bit equals zero. Source transmits a packet to the destination for which it has credits with a priority. Admission controller decides if new credits can be allocated to some source-destination pair.

Initially, nodes negotiate the maximum frame length, F_{max}, in terms of time slots. The frame duration should be much longer than the ring latency in order to provide high efficiency, but short enough to respond to the fast traffic changes. Then, a credit of one time slot per frame guarantees to the particular queue a bandwidth of $W \cdot B/F_{max}$, where B is the bit-rate at one wavelength. An admission controller is placed at some node, and it analyzes if the newly requested bandwidth could be allocated to the particular source-destination pair. More precisely, an admission controller calculates if MAC protocol ensures new Δa_{ij} credits to the node pair (i, j) in each frame for the existing credit allocation a_{kl}, $1 \le k, l \le N$, where N is the number of nodes.

Theorem: Credits $a_{ij} > 0$ assigned to source-destination pairs (i, j), $1 \le i, j \le N$, will be used within a frame of length F_{max} if it holds that:

$$W \cdot (s_k + D_k) + l_k \le F_{max},\ \ 1 \le k \le N, \qquad (1)$$

where $s_k = \sum_m a_{km}$ is the number of credits assigned to source k, $D_k = \max_{l, a_{kl}>0} d_l$ where $d_l = \sum_n a_{nl}$ is the number of credits assigned to destination l, and $l_k = \sum_{\substack{m,n \\ m \to k \to n}} a_{mn}$ is the number of credits for packets passing by link k.

Proof: We use similar methodology as in [8], [9]. Assume that t_{max} is the last time slot assigned to source-destination pair (i, j) which is in cycle c_{max}. In any cycle $c \le c_{max}$ either node i transmits a packet, or node j receives a packet, or all time slots passing node i have been taken. There are at most $s_i + d_j - a_{ij} - 1 \le s_i + D_i - a_{ij} - 1$ cycles before c_{max} in which either source i or destination j are busy. The rest of the cycles are fully occupied with less than l_i packets. So, $t_{max} < W \cdot (s_i + D_i - a_{ij}) + l_i \le F_{max}$ and the claim of the theorem directly follows. □

When the bandwidth is requested, tentative parameters s'_k, D'_k, l'_k are calculated, and inequalities (1) are checked in parallel. In a single channel ring network with the spatial reuse, it should be checked if links from source to destination have enough of bandwidth, which also involves up to N inequalities. Calculation of D_k, $1 \leq k \leq N$, which takes $\log_2 N$ steps, is the only additional complexity of the proposed admission control protocol when compared to the complexity of the admission control protocol in a single channel network. If the request is accepted, parameters in question are updated.

Allowed bandwidth allocation in the network is easily derived from admission condition (1). Let $t_k = s_k \cdot WB/F_{max}$ denotes the bandwidth reserved for transmission by source k, $R_k = \max_{a_{kl}>0} r_l$ where $r_l = d_l \cdot WB/F_{max}$ denotes the bandwidth reserved for reception by destination k, and $p_k = d_k \cdot WB/F_{max}$ denotes the bandwidth reserved on link k. Then, admission condition (1) becomes:

$$W \cdot (t_k + R_k) + p_k \leq WB, \; 1 \leq k \leq N, \qquad (2)$$

IV. The Network Performance

Our proposed network is equivalent to an ideal WDMA packet-switched ring network in terms of the accessible capacity. In this ideal network users would be equipped with fast tunable lasers and filters. In each time slot, source learns from the control channel about available destinations and wavelengths in the next time slot. So, source can transmit a packet to some destination as long as there is an idle wavelength and no other source already transmits to this destination in the given time slot. Let us assume without loss of generality that a time slot in the exemplary network overlaps with a cycle in our network. In both networks, in each time slot, or cycle, a transmitter can transmit at most one packet, and a receiver can receive at most one packet due to the speed limitation. Also, in both networks, in each time slot, or cycle, at most W packets can be transmitted. Since constraints for the media access are identical in both networks, capacities that can be utilized are the same. Also, it follows that the admission condition for bandwidth reservations in the network with fast tunable lasers and filters would be the same as in our proposed network.

The admission control protocol can be further simplified. Namely, the admission condition (2) can be split, so that the node and link capacities allocated for reservations are determined in advance:

$$t_k \leq T, \; r_k \leq R, \; p_k \leq P, \; 1 \leq k \leq N,$$

$$W \cdot (T + R) + P = WB. \qquad (3)$$

Possible choice of parameters T, R and P would equilize utilizations of transmitters, receivers and links:

$$\frac{T}{B} = \frac{R}{B} = \frac{P}{WB} \qquad (4)$$

From (3,4) it follows that $T = R = P/W = B/3$, and the admission condition is:

$$t_k \leq \frac{B}{3}, \; r_k \leq \frac{B}{3}, \; p_k \leq \frac{WB}{3}. \qquad (5)$$

If the number of nodes in the network is large, the node capacity allocated for reservations can be further decreased, and the link capacity to be reserved would increase according to equation (3). In the given example, a third of the network capacity can be reserved. In the network with $W = 32$ wavelengths, the bandwidth that can be reserved is 10 times larger than in a single channel network. The remaining bandwidth can be used for best effort traffic as we explained earlier. Best effort packets have zero priority bits and can be replaced by packets with reservations.

V. Conclusions

In this paper, we combine architecture, MAC and admission control protocols to flexibly utilize the high-capacity packet-switched ring network. The wavelength stacking and unstacking allow each node to access full ring capacity. Network control requires minimal complexity because the traffic load is balanced over the wavelengths at the physical layer.

References

[1] M. Boroditsky *et al*, "Experimental Demonstration of Composite Packet Switching on a WDM Photonic Slot Routing Network," *Optical Fiber Communication Conference*, March 2001, ThG6.

[2] I. Cidon, and Y. Ofek, "MetaRing-A Full-Duplex Ring with Fairness and Spatial Reuse," *IEEE Transactions on Communications*, vol. 41, no. 1, January 1993, pp. 110-120.

[3] L. Georgiadis, R. Guerin, and I. Cidon, "Throughput Properties of Fair Policies in Ring Networks," *IEEE/ACM Transactions on Networking,* vol. 1, no. 6, December 1993, pp. 718-728.

[4] Raj Jain, *FDDI Handbook*, Addison-Wesley Publishing Company, 1994.

[5] M. A. Marsan, A. Bianco, E. Leonardi, F. Neri, and S. Toniolo, "An Almost Optimal MAC Protocol for All-Optical WDM Multi-Rings with Tunable Transmitters and Fixed Receivers," *ICC'97*, pp. 437-442.

[6] M. A. Marsan, A. Bianco, E. Leonardi, M. Meo, and F. Neri, "MAC Protocols and Fariness Control in WDM Multirings with Tunable Transmitters and Fixed Receivers," *Journal on Lightwave Technology*, vol. 14, no. 6, June 1996, pp. 1230-1244.

[7] M. A. Marsan, A. Bianco, E. Leonardi, A. Morabito, and F. Neri, "All-Optical WDM Multi-Rings with Differentiated QoS," *IEEE Communication Magazine*, February 1999, pp. 58-66.

[8] A. Smiljanić, "Flexible Bandwidth Allocation in Terabit Packet Switches," *IEEE Conference on High-Performance Switching and Routing,* (Best Paper Award), June 2000, pp. 233-241.

[9] A. Smiljanić, "Flexible Bandwidth Allocation in High-Capacity Packet Switches," *to appear in IEEE/ACM Transactions on Networking.*

[10] A. Smiljanić, M. Boroditsky, and N. Frigo, "Optical Packet-Switched Network with Flexible Bandwidth Allocation," *IEEE Workshop on High-Performance Switching and Routing,* May 2001, pp. 83-101.