Novel MAC protocol with idle detection for all-optical WDM ring networks

Wen-Ping Chen,[1,*] Jih-Hsin Ho,[2] Wen-Shyang Hwang,[1] and Ce-Kuen Shieh[3]

[1]*Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung 807, Taiwan*
[2]*Department of Information Management, Diwan University, Tainan 721, Taiwan*
[3]*Department of Electrical Engineering, National Cheng Kung University, Tainan 701, Taiwan*
Corresponding author: pen@mail.ee.kuas.edu.tw

Received August 14, 2008; revised October 21, 2008; accepted November 17, 2008; published January 12, 2009 (Doc. ID 100181)

A novel media access control (MAC) protocol named carrier sense multiple access with idle detection (CSMA/ID) is proposed to handle variable-length packets over an all-optical ring network. To evaluate optimal utilization of channel bandwidth, we study packet scheduling based on three transmitting queue discipline (TQD) architectures and four idle space allocation (ISA) algorithms with regard to their impact on performance. For numerical evaluation of performance, an analytical model is developed by a preclassification queue with weighted round-robin (PCQ_WRR) architecture and a random algorithm. Moreover, three related MAC protocols are examined and compared, namely, multitoken, carrier sense multiple access/collision avoidance (CSMA/CA) and carrier sense multiple access/collision preemption (CSMA/CP). Simulation results indicate that, of the TQDs, better performance is obtained by PCQ_WRR compared with first-in-first-out and preclassification queues. The first fit space (FFS) algorithm has the best performance of the ISAs. The 12 combinations of TQDs–ISAs are then considered. It is found that the combination of PCQ_WRR with FFS provides the greatest efficiency and has the lowest packet latency, providing better throughput than three different MAC protocols under either symmetric or asymmetric traffic load on all-optical ring networks. © 2009 Optical Society of America

OCIS codes: 060.0600, 060.4262, 060.4259.

1. Introduction

The exploding levels of information traffic from applications such as electronic commerce, multimedia, and voice-over-Internet protocol (VoIP) have generated an enormous increase in Internet bandwidth requirements. The need for a transmission medium with the bandwidth ability to handle such vast amounts of information is paramount. In an attempt to provide the bandwidth necessary to fulfill this ever-increasing demand, copper networks are being upgraded or replaced with optical fiber networks. Recent advances in solid-state and optical amplifiers, for example, erbium-doped fiber amplifiers (EDFAs) that allow dense-wavelength division multiplexing (DWDM) on a single fiber, have been emerging as the technology of choice for increasing the transmission capacity of carrier networks. The bit rate on a wavelength has exceeded 40 Gbits/s, and a light signal can travel for several hundred kilometers through a single-mode fiber (SMF) without amplification. Further, the number of wavelengths in a fiber has grown to more than 1000 and this clearly is not a limit [1]. DWDM technology now offers a solution for bandwidth insatiability, promising a backbone area network that can support more than 10.2 Tbits/s [2].

IP packet traffic dominates Internet utilization. However, IP packets now are transferred, switched, and manipulated through complex protocol stacks such as IP/ATM/SONET/WDM and IP/HDLC/SONET/WDM [3]. These extended network stacks generate high overhead, complicate system infrastructure, increase cost, and have made redundancy an important research issue. In addition, wavelength division multiplexing (WDM) systems are being deployed in wide area networks (WANs), thereby pushing the bottleneck of communications from backbone networks to local access networks (LANs). In these circumstances, the metropolitan area networks

1536-5379/09/020112-18/$15.00

(MANs) play a critical role in the overall expansion of network services [4]. MANs provide services within individual metropolitan areas and serve as gateways for wide-area national- and international-scale networks. As a result, applying WDM to LANs and MANs has attracted much research interest.

Optical packet switching (OPS) [5] appears to be the switching platform of choice for the arbitrarily fine transmission and switching granularity needed for the all-optical Internet. Although most research has focused on star and mesh OPS architecture, the recent decade has seen increased attention to the OPS ring networks [4,6–16]. Today, packet transmission in the OPS metro ring network has two major research fields: synchronous and asynchronous. Synchronous transmission is suited for fixed-length packets such as MetaRing [6], the Hybrid Optoelectronic Ring Network (HORNET) [7,8], and other slot-based ring networks [4,10–12]. Asynchronous transmission is for variable-length packets such as token-based rings [14,15], HORNET [carrier sense multiple access/collision avoidance (CSMA/CA) with backoff] [7,9], carrier sense multiple access/collision preemption (CSMA/CP) [13], and IEEE 802.17 RPR [16,17]. The latter is well-suited to IP traffic and therefore is considered in this study. The multitoken interarrival time (MTIT) [14] uses the FT^W–FR^W (where FT stands for fixed transmission and FR for fixed receiver) system architecture that requires an additional FT–FR pair for the add–drop control channel (W_C) but is used only to transmit token packets, an approach that is costly and has little scalability. Moreover, the multitoken protocol is very difficult to implement in metro ring networks [18].

In 1999 Stanford University proposed HORNET using hybrid optical-electronic technology to transmit packets, therefore constraining the transmission rate. HORNET uses tunable transmitter (TT) FR architecture and the CSMA/CA protocol [7,9] with a backoff process to handle variable-size packets. The architecture has poor performance, however, because the access point (AP) immediately ends packet transmission and sends a jamming signal to notify the next AP to discard the incomplete packet when the AP is transmitting a packet and another packet is arriving on the same wavelength at its input. In 2000 the Institute of Electrical and Electronic Engineers (IEEE) attempted to achieve spatial wavelength reuse and improve fault tolerance by extending Ethernet-based packets from LAN to MAN. To do so they defined a dual-ring-based network protocol standard IEEE 802.17 named the resilient packet ring (RPR) [16,17]. To solve the problem of packet collision, after optical-to-electrical (O/E) conversion and when the node is transmitting a packet, RPR stores the packet that has arrived from a channel in a particular transit queue, thereby constraining the transmission rate of the network. In 2007, the author proposed CSMA/CP [13], which uses packet fragmentation to solve the problem of packet collision in WDM metro ring networks. CSMA/CP offers improved efficiency and lower packet latency but presents two drawbacks. First, a node must monitor the channel status all the time, regardless of whether searching idle space or transmitting packets, including the time required to transmit all its fragmented packets. Second, the CSMA/CP fragmentation scheme requires packet header overhead and guard-band distance, with further processing overhead from fragmentation and reassembly. In general, fragmentation schemes are not allowed in optical access networks such as the Ethernet optical passive network (EPON) [19,20].

To overcome the above issues, this study proposes a novel media access control (MAC) protocol designated carrier sense multiple access with idle detection (CSMA/ID). In CSMA/ID, nodes are equipped with a fixed-length fiber delay line (FDL) to handle variable-size packets and to avoid packet collision and fragmentation in the optical domain. The nodes monitor channel statuses when searching for an available idle space to transmit a packet. After the packet is transmitted, the monitor action immediately stops. In the quest to improve channel bandwidth, we will evaluate and compare three different transmitting queue discipline (TQD) architectures and four different idle space allocation (ISA) algorithms with regard to their performance in the proposed protocol in all-optical WDM ring networks. To verify and demonstrate that the proposed CSMA/ID protocol offers high performance, this paper also considers three different MAC protocols and compares them with CSMA/ID.

The all-optical WDM ring network architecture for the proposed protocol is presented in Section 2. A novel MAC protocol is proposed for all-optical WDM ring networks with TQD structures and ISA algorithms as described in Section 3. Section 4 presents a performance analysis. Section 5 gives simulation results and further discusses the proposed protocol. Finally, a few concluding remarks are given in Section 6.

2. System Architecture

The all-optical WDM ring network is a candidate for MANs. We give an overview of such a network and node architecture, and then use it to discuss the proposed network system.

2.A. Network Architecture

The network architecture presented here is a single and unidirectional fiber ring network that connects N nodes. Each optical fiber between the nodes is composed of m channels $(W_1, W_2, W_3, \ldots, W_m)$ as shown in Fig. 1. The ring is assumed to cover a metropolitan area, with a ring circumference of about 100 km. Thus the system is referred to as a WDM metro ring. In the network, a node called the AP is used for connecting the LAN to the MAN, while a node called a point of presence (PoP) is used for connecting the MAN to the WAN. In the fibers, each channel makes use of one specific wavelength to convey optical data. Therefore, based on WDM technology, channels can work independently without mutual interference. Logically, the network can be treated as a multiring network.

In the network, each AP is equipped with a tunable transmitter (TT) and m fixed receivers ($\mathrm{TT}-\mathrm{FR}^m$) where every receiver is dedicated to a particular channel to eliminate unfair access to node position and to increase the chance of packet transmission. The packets in an AP may come from/pass to the attached LAN, gigabit Ethernet (GbE), passive optical networks (PONs), or wireless LAN. The packets are transmitted on/received from any one of the channels by the AP node and travel as optical signals along the ring without any electro-optic conversion at intermediate nodes.

2.B. Node Architecture

The node architecture of a WDM metro ring network is shown in Fig. 2, where the packets from the LAN are first inserted into one of the transmission queues according to the destination address, class of service (CoS), or other criteria. For increased channel utilization, this paper uses three different TQD architectures in a single node:

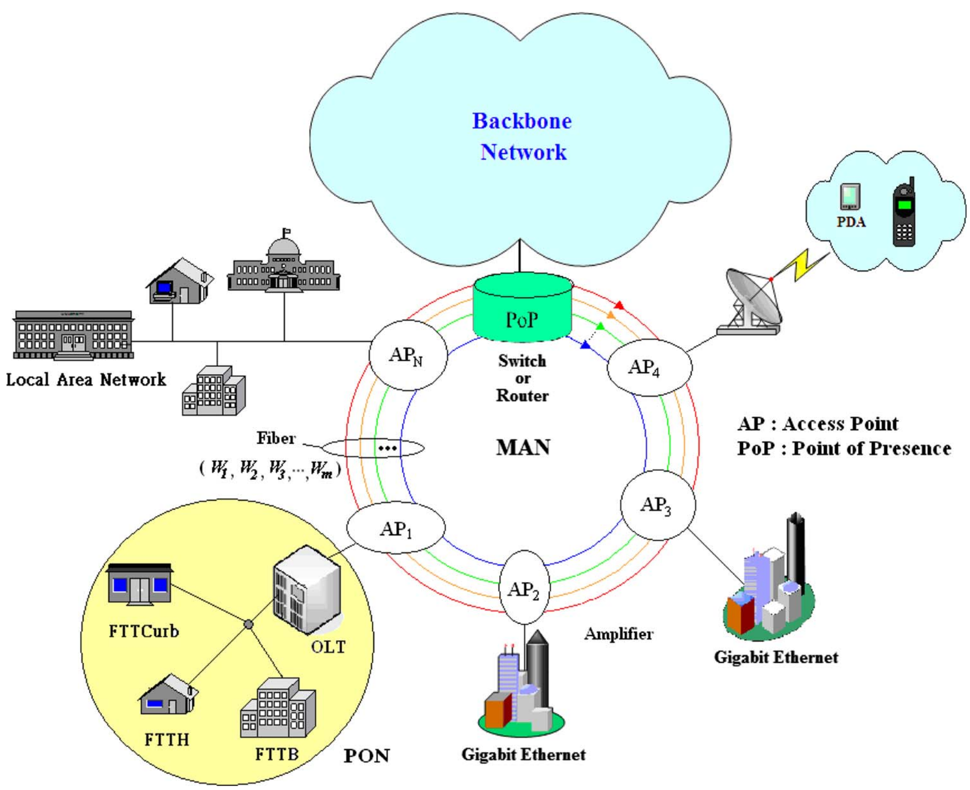


Fig. 1. Architecture of WDM metro ring network.

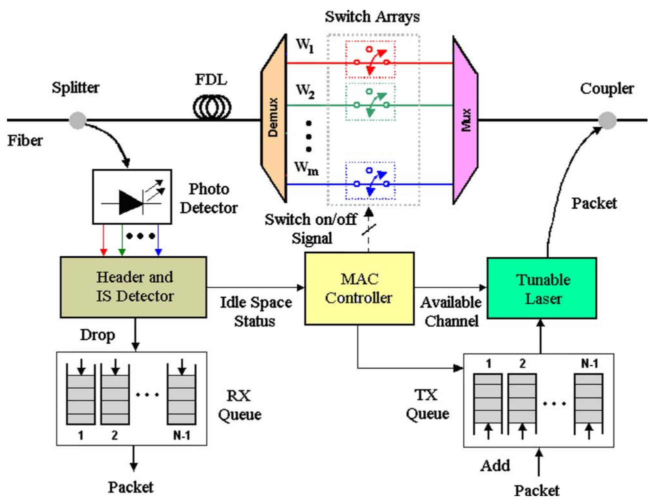


Fig. 2. Node architecture of CSMA/ID.

first-in-first-out (FIFO), preclassification queue (PCQ), and preclassification queue with weighted round-robin (PCQ_WRR) scheduler. For simplifying the description, the AP node presented below will use only the FIFO discipline to represent the add–drop packet operations from LAN or MAN.

Each AP node listens to all wavelengths by continuously monitoring the subcarrier [21] or baseband [22] optical signals. It is similar to HORNET [7] in using the subcarrier multiplexing label technology, in which a packet header is carried on the subcarrier for easy pickup. When an optical signal comes from the upstream AP nodes, the splitter taps off a small portion of its optical power into two parts. The first part is used to check the destination address of the optical packet. If the address matches the node address, the packet is inserted into the receiver (RX) queue. Meanwhile, the MAC controller is signaled to activate the optical switch arrays for removing the major optical signal portion of the received packet in the corresponding channel. This mechanism is called destination removal and has more network capacity than the source removal mechanism because of spatial wavelength reuse. In addition, the energy of the optical signal was dropped where the splitter was installed; therefore the optical amplifiers were installed in every ten nodes to ensure the quality of the optical signal. The number used for the EDFAs and amplifier noise accumulation is reported in minimum [23].

The protocol processing time of the MAC controller is long enough because the FDL that is employed at each node installed on the transmission path delays the packets from the upstream AP nodes for a fixed time T_{FDL}. The length of FDL is equal to the maximum transfer unit (MTU) plus guard-band distance; thus the packet collision is avoidable even if the MTU packet is transmitting. If the address is different than the AP node address, the received packet is ignored and the optical signal in the FDL is bypassed to the next AP node. In addition, the signal in the FDL is used for measuring the idle space duration. When an AP node has packets to transmit, it measures and searches all the idle space lengths of the channels for idle space suitable for its waiting packets. The large FDL increases the ring length and the propagation delay naturally. The added ring length (R_{FDL}) and the propagation delay (T_{pro}) on balance traffic load are listed as follows:

$$R_{\mathrm{FDL}} = \left(N \times \frac{8L}{C} + T_g \right) \times V, \tag{1}$$

$$T_{\mathrm{pro}} = \frac{L_R}{2V}, \tag{2}$$

where N is the total node number, T_g is the time of guard-band distance, L is the MTU length in bytes, C is the channel speed, and V is the light velocity in fiber. However, it only increases the delay slightly. For example, assume that $N=16$, MTU=1500 bytes, T_g=4 ns, C=10 Gbits/s, and $V=2\times10^5$ km/s in the fiber, then the R_{FDL} and T_{pro} are only about 3.84 km and 9.6 μs.

Both the optical switch array and the tunable laser are major components used to realize packet drop–add and destination removal. The switching time of the tunable laser is less than a previously demonstrated 8 ns [7,24]. The switching time of the optical switch [25] obviously affects system performance. The currently popular optical switching technology is of the microelectromechanical system (MEMS) [26,27] type, which uses tiny mirrors to deflect light from an input to a particular output port. MEMS switches have both 2D and 3D variants, offer low rates of signal distortion, and have switching times in the range of milliseconds. However, MEMS switching is currently unsuitable for use in OPS because the switching time needs to be in the range of nanoseconds. Two widespread approaches used for the OPS switch are the semiconductor optical amplifier (SOA) [28] and the arrayed waveguide grating (AWG) with tunable wavelength converters (TWCs) [29]. AWG requires precise thermal control since it suffers from temperature-dependence issues, whereas SOA offers high-speed operation over 40 Gbits/s and supports switching for optical signals on time scales as short as 1 ns [30]. Therefore, SOAs are expected to be universal building blocks of future all-optical networks. Recently, a proof-of-principle 10×10 radio-over-fiber (RoF) SOA-based cross-point switch array has been demonstrated [31]. Hence we assume an array of SOA suitable for our node architecture to drop packets by removing the major portion of the optical power in the fiber.

3. CSMA/ID and Packet Scheduling

The frame format of the proposed MAC protocol CSMA/ID is described in detail here. To better understand the various influences on performance, the packet-scheduling and ISA algorithms and TQD architectures are considered.

3.A. CSMA/ID Protocol

CSMA/ID is a collision avoidance MAC protocol based on carrier sensing, by which an AP senses all wavelengths (channels) and detects idle space status during the detection window T_{FDL} as seen in Fig. 3. An AP is permitted to transmit its packet only when it has detected an idle space that is long enough for the packet. Generally, the idle space status message will pass to the MAC controller for deciding to select a packet from the transmitting queues according to ISA and TQD. After the decision had been made, the transmission of an AP in the CSMA/ID network can be divided into add–drop packet models described as follows:

(1) In the drop packet model, each AP listens to all wavelengths by monitoring either subcarriers or optical signals. If the address of a received packet matches the AP address, the packet is inserted into one of the RX queues according to its attribute. Then the MAC controller is signaled to activate an optical switch to remove the optical signal part of the received packet in the fiber. If the address does not match the AP address, the optical signal in the FDL is bypassed to the next AP node.

(2) In the add packet model, the MAC controller measures all idle spaces in every channel during the window T_{FDL} for any transmitter (TX) queue that has a packet to transmit. The idle space must be equal to or larger than the packet size. Afterward, the MAC controller decides on which packet to send. (More discussion of packet scheduling is presented in Subsection 3.B.)

Four notable features characterize CSMA/ID. First, it is a fully distributed and asynchronous protocol. CSMA/ID does not need a centralized controller or a separate control channel to harmonize and synchronize the operation of nodes. Second, the FDL opens enough to transmit the MTU, so the transmitting packet and an incoming packet do not occur on the same wavelength. Third, CSMA/ID supports handling variable-length IP packets without using complicated segmentation and reassembly procedures that becomes more difficult as the line speed of optical wavelengths increases. Fourth, CSMA/ID maintains an all-optical transmission from source node to destination node on the metro ring network.

3.B. Packet Scheduling

The AP node has a FDL long enough to delay incoming packets so as to avoid packet collision when the AP node is transmitting a MTU into the optical fiber. Because the

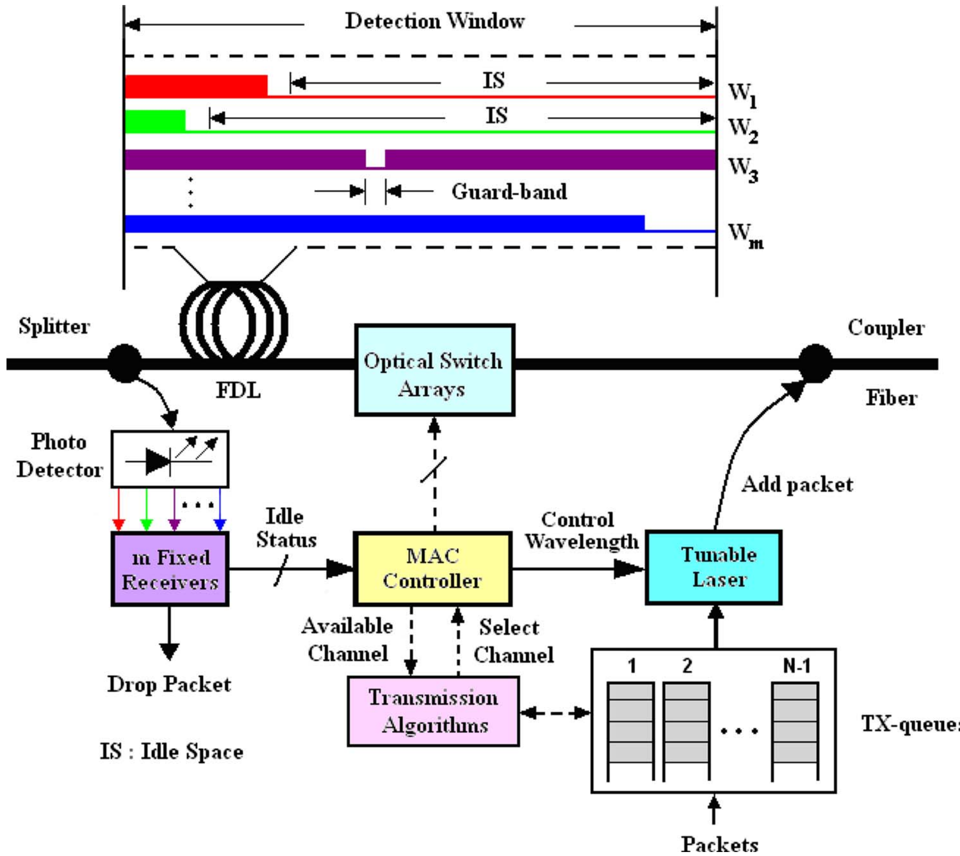


Fig. 3. CSMA/ID protocol.

IP-based packets are of variable length, the candidate packet is never a MTU packet. There can be many different idle spaces of different lengths in different channels that can be selected in a detection window for transmitting variable packets. Hence, for best bandwidth utilization, it is very important to know how to select a candidate packet from the transmitting queue and select an optimal idle space of channel to put it in. Therefore, this paper studies packet scheduling with various TQD architectures and ISA algorithms in order to evaluate their comparative performance and choose the best for the proposed protocol. The considered algorithms and architectures are described as follows.

3.B.1. Idle Space Allocation Algorithms

The idle space length in window T_{FDL} must be large enough for allocation to the candidate packet in TX queues. However, it is possible for many channels to have eligible idle space at the same time. The problem is how to select a suitable idle space efficiently to increase the system throughput. We propose four ISA algorithms as follows:

(1) *Random (RND)*. The MAC controller selects at random one of the eligible idle spaces for allocation to the candidate packet.

(2) *Longest idle space (LIS)*. The MAC controller selects the longest idle space for allocation to the candidate packet.

(3) *Best fit space (BFS)*. The MAC controller selects the smallest eligible idle space for allocation to the candidate packet.

(4) *First fit space (FFS)*. The MAC controller scans channels upwards to find an eligible idle space (from W_1 to W_m). Whenever it meets an eligible idle space, it immediately allocates it to the candidate packet and stops the scan. The candidate packets will be transmitted on the channels in ascending order. It also presents more opportunity for bigger packets to obtain a longer idle space on posterior channels ($W_m, W_{m-1}, \ldots$) than other algorithms. The FFS algorithm will be more effective than others in the metro ring networks with CSMA/ID. Because the longer idle space is generally difficultly obtained when the AP has a bigger packet to transmit and the

system is operating under high load conditions. As the result, almost all channels are put on smaller packets, but the bigger packets are blocked in the transmission queue.

These algorithms are illustrated in Fig. 4, where the candidate packet at the head of a transmitting queue has a length of 500 bytes. The candidate packet will be allocated W_3 by the LIS algorithm, W_2 by the FFS algorithm, and W_4 by the BFS algorithm. As regards the RND algorithm, the packet will be allocated randomly by any one of the channels W_2–W_4.

3.B.2. Transmitting Queue Discipline Architectures

Understanding the impact of TQD on system performance is an important issue. Here, we consider three TQD architectures: FIFO, PCQ, and PCQ_WRR scheduler.

(1) *First-in-first-out*. In the FIFO queue, the packet that arrives first will be serviced first. The MAC controller continuously monitors all detection windows T_{FDL} of the channels to make an ISA decision for the candidate packet, which is the packet at the head of the queue. Whenever the MAC controller finds eligible idle space, it uses the candidate packet size to select an idle space according to the selected ISA algorithm for the packet. In this way, packets are fairly transmitted, but system throughput is slowed because of the head-of-line (HoL) effect.

(2) *Preclassification queuing*. The majority of packets in the Internet are from the Ethernet, which standardly allows packet lengths between 40 and 64 Kbytes. However, the measurement traced from MCI's backbone OC-3 links [32] shows a discrete packet-size distribution from 40 to 1500 bytes [7–9,12,13] only, with most packet lengths being 40, 552, or 1500 bytes. Hence, according to the IP packet-size distribution, an IP packet that arrives from the LAN is preclassified into one of three class queues: 40 bytes (Q_3), 41–552 bytes (Q_2), or 553–1500 bytes (Q_1). The MAC control-

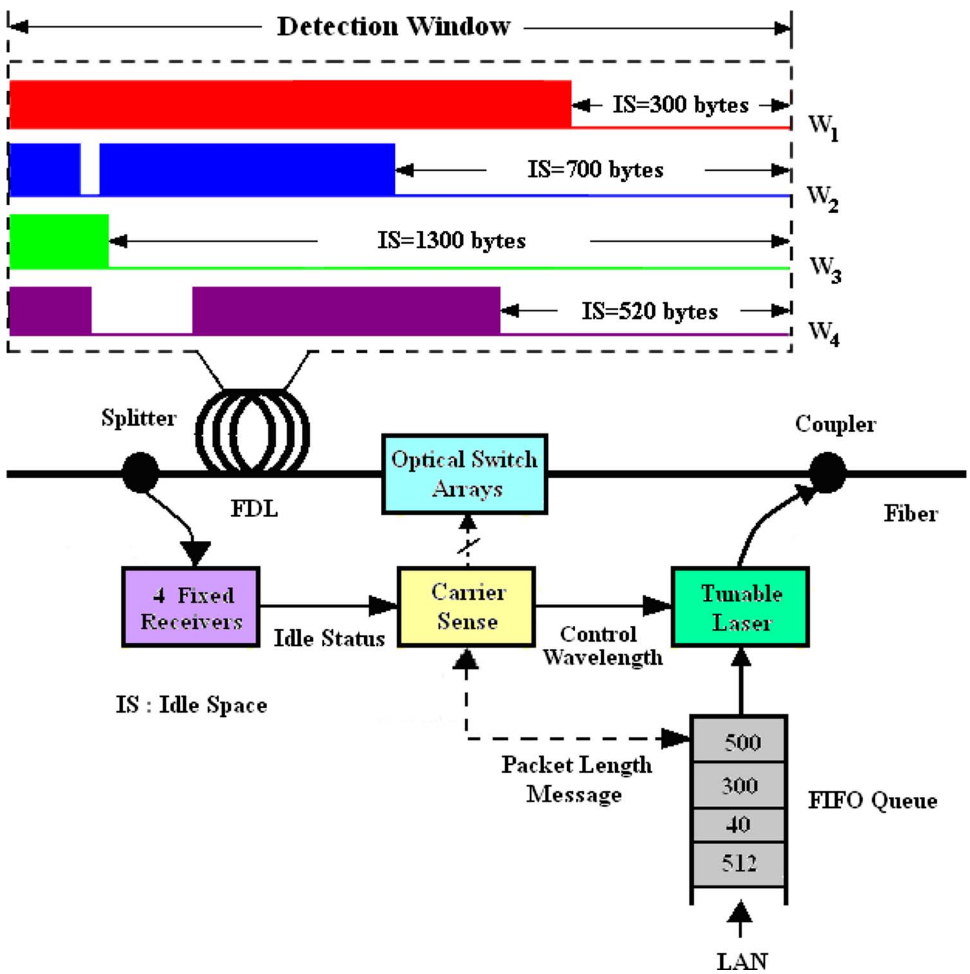


Fig. 4. Example of ISA algorithms.

ler always knows the size of the packet at the head of these queues and allocates idle space in the priority order Q_1, Q_2, and Q_3 when the eligible idle space is long enough. In other words, it transmits the largest packet first to solve the HoL problem.

(3) *Preclassification queue with weighted round-robin scheduler*. Although the PCQ discipline solves the HoL problem, it creates an unfairness issue because of queue priority; that is to say, longer packets have a greater probability of being sent. Besides, MCI's backbone OC-3 links cannot really represent the distribution of the packet length in a real-world network. Fortunately, the above issues can be mitigated by use of the WRR discipline [33] with a timer calculated. In the proposed PCQ_WRR, the weighted value of each transmission queue must be calculated after each unit time (t) according to the total length of the incoming packets from the LAN. The higher arrival rate (λ) generates the higher weighted value. The packets are transmitted in WRR fashion. For example, the weighted value and the number of packets (α,β) for three transmission queues Q_1, Q_2, and Q_3 after the time unit are (3800, 3), (2512, 5), and (840, 21). As a result, the scheduling is given as $Q_1 \to Q_2 \to Q_3 \to Q_1 \to Q_2 \to Q_3 \to Q_1 \to Q_2 \to Q_3 \to Q_2 \to Q_3 \to Q_2 \to Q_3 \to Q_3 \to Q_3 \to \ldots \to Q_3$. Therefore, the dynamic weighted value of PCQ_WRR can be applied to work in a more realistic and dynamic network environment.

Another problem of the PCQ discipline is the out-of-order delivery issue. However, it is a common problem, particularly in multipath routing networks (e.g., mesh networks or wireless networks). Fortunately, because the ring has only a single path, the out-of-order delivery problem with PCQ_WRR is slight and can be resolved by installing a resequencing queue (RSQ) at the AP node [34], as shown in Fig. 5. Packets arriving at the receive terminal of the destination AP are rearranged in position according to their sequence numbers.

PCQ_WRR has better throughput than PCQ because each queue is allowed to transmit candidate packets by its WRR scheduler. Under heavy loads, the idle spaces in the channels are usually short. Thus, even though Q_1 has the highest priority, Q_2 and Q_3 have more opportunities to use the available idle space. Figure 6 shows an example of four channels containing five idle segments ($\mathrm{IS}_1, \ldots, \mathrm{IS}_5$) in the FDL window. Without the WRR scheduler, IS_1, IS_2, and IS_3 will be selected in that order to transmit candidata packets. Although IS_3 is the longest, it retains less available space after IS_2 has been used to transmit a packet. The result is that the large packet in Q_1 always misses its opportunity to become a candidate packet for transmission. Under heavy loads, it is especially difficult for Q_1 to become a candidate packet. On the other hand, in the WRR scheduler, all transmitting queues have a fixed transmission order according to the statistic of (α,β) at time unit. Hence, there is a good chance Q_1 has already been selected to fill the large packet so that IS_1 and IS_2 cannot be used. The large packet in Q_1 can then be sent out in IS_3. This is why the WRR scheduler improves bandwidth utilization and lowers transmission delays.

4. Performance Analysis

In the following, packet scheduling adopts the PCQ_WRR discipline with the RND algorithm to analyze the average transfer delay in the single-ring and multiring cases under the following assumptions.

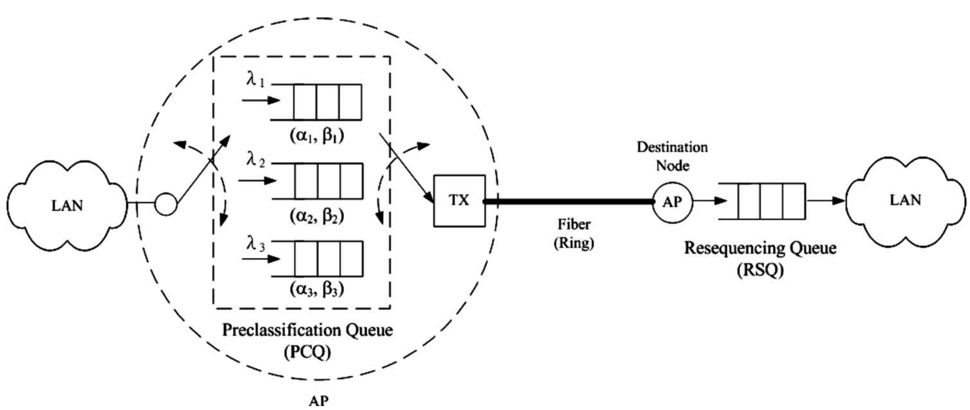


Fig. 5. Resequencing queue architecture for correcting out-of-order packet delivery.

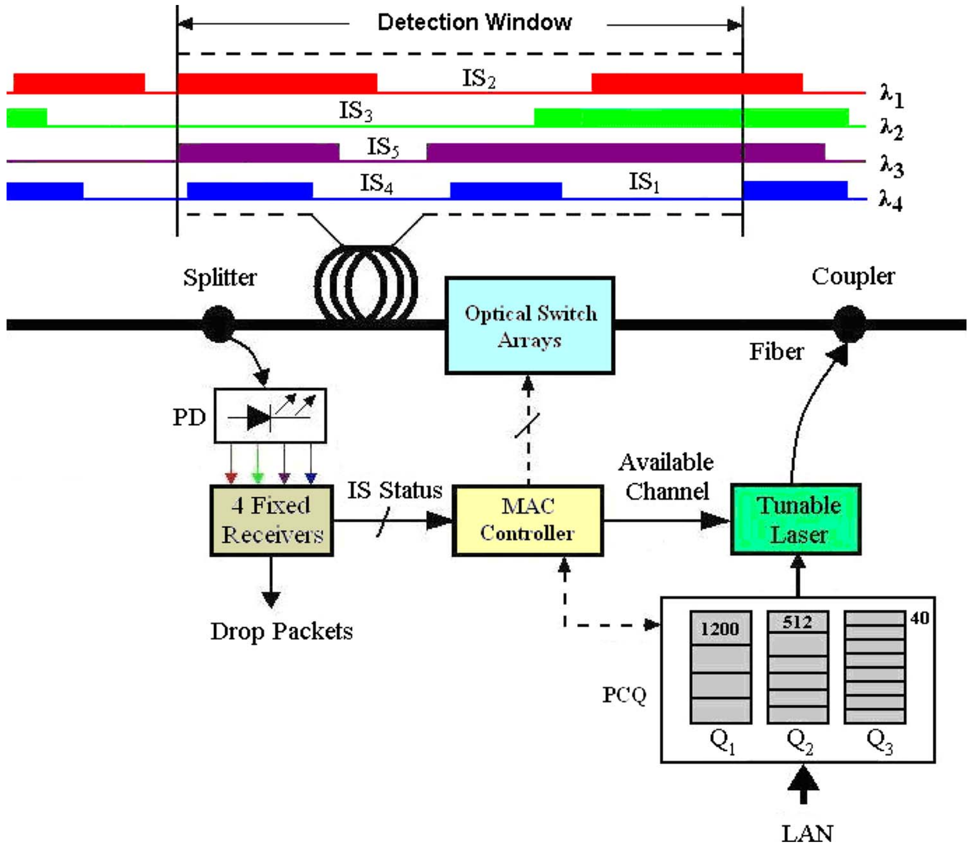


Fig. 6. Example of idle spaces in the FDL for PCQ architecture.

4.A. Assumptions and Notations

(1) The number of WDM channels (or rings) is m; each channel has N nodes.

(2) The total propagation delay of the WDM ring is τ seconds; the distances between the nodes are equal.

(3) Packets arrive as an independent identically distributed (*i.i.d.*) Poisson process with rate λ_i at each node such that $\lambda_{i,1}$, $\lambda_{i,2},\ldots,$ and $\lambda_{i,q}$ are in q different class queues as determined by packet length; hence, the aggregate arrival rate of the network is

$$\lambda = \sum_{i,1=1}^{N-1} \lambda_{i,1} + \sum_{i,2=1}^{N-1} \lambda_{i,2} + \cdots + \sum_{i,q=1}^{N-1} \lambda_{i,q}.$$

(4) The arrival stream of packets at node i destined for node $i \oplus j$ is a Poisson process with a rate of $\lambda_{i,i\oplus j}$, where $\oplus$ indicates the addition modulo N; thus, $\lambda_i = \Sigma_{j=1}^{N-1} \lambda_{i,i\oplus j}$.

(5) The mean packet generation for nodes is equal; each sends equal traffic to its destinations (uniform and symmetric traffic): $\lambda_i = \lambda/N$, $\lambda_{i,i\oplus j} = \lambda_i/(N-1) = \lambda/[N(N-1)]$ and $\lambda_{i,i}=0$, for $0 \leqslant i \leqslant (N-1)$, $1 \leqslant j \leqslant (N-1)$. The mean packet generation of PCQ_WRR for the nodes is equal. The number of packets generated in each queue of PCQ_WRR in a unit time t is n_1, n_2, and n_q, and the mean packet length is L_1, $L_2,\ldots,$ and L_q. According to the packet distribution of the MCI's backbone OC-3 links, the weight of the packet length in a class r queue is $w_r = n_r L_r / (\Sigma_{k=1}^{q} n_k L_k)$ (where $\Sigma_{r=1}^{q} w_r = 1$). The ratio of the arrival rate in each queue is $\lambda_{i,1}:\lambda_{i,2}:\ldots:\lambda_{i,q} = n_1 L_1/t : n_2 L_2/t : \ldots : n_q L_q/t = w_1 : w_2 : \ldots : w_q$, and the mean total packet length in PCQ is $\bar{L} = \Sigma_{r=1}^{q} n_r L_r / (\Sigma_{r=1}^{q} n_r)$. Therefore, the arrival rate $\lambda_{i,k}$ in node i can be given by $\lambda_{i,k} = w_k \lambda_i = w_k \times \lambda/[N(N-1)]$, for $k=1,2,\ldots,q$.

(6) The length of the packets is randomly determined by independent, identically and geometrically distributed random variables (denoted by *r.v.* M bits) with mean $E[M]$ and probability mass function [35] $P_r(M=c) = \beta(1-\beta)^c$, $c=0,1,2,\ldots,$ where $\beta = 1/(1+E[M])$.

(7) The channel bit rate of the WDM ring is R (bits/s) and the packet transmission time is $X(=M/R)$ seconds.

(8) The MTU of the network is equal to the delay line ($L=1500$ bytes) with $T=L/R$ seconds to transmit (neglecting the guard-band distance).

(9) The length of a transmit queue is infinite; no packets are lost.

The following notation is used in the analytical formulas: T_D is the average packet transfer delay, T_{Qj} is the queuing delay of packet j, T_Q is the average packet queue-waiting delay, α_j is the residual time of packet j, α is the packet residual time, V_j is the duration of all whole vacation intervals for which packet j in each queue must wait before being transmitted, V is the steady-state duration of all whole vacation intervals, T_S is the average transmission time, N_r is the average packet numbers in a class r queue of PCQ, and N_j is the average packet numbers in PCQ.

4.B. Analysis of the Single-Ring Case

In CSMA/ID node architecture a TT transmits only one packet at a time. Packets arrive with Poisson distribution at rate λ_i at each node. An $M/G/1$ queue with vacations can model our system. Given the above assumptions, the average queuing delay for the jth packet is given by

$$E[T_{Qj}] = E[\alpha_j] + \sum_{r=1}^{q} E[N_r] \times E[X_r] + E[V_j], \tag{3}$$

where

$$\begin{aligned}\sum_{r=1}^{q} E[N_r] \times E[X_r] &= w_1 E[N_j] \times E[X_1] + w_2 E[N_j] \times E[X_2] \\ &\quad + \cdots + w_q E[N_j] \times E[X_q] \\ &= (w_1 E[X_1] + w_2 E[X_2] \\ &\quad + \cdots + w_q E[X_q]) E[N_j] = E[\bar{X}] E[N_j]. \end{aligned} \tag{4}$$

By Little's formula, the average number of the packets in the PCQ_WRR, $E[N_j]$, can be obtained as $E[N_j]=\lambda_j \times T_{Qj}$. Therefore, $E[N_j]\times E[\bar{X}]$ can be substituted as $\lambda_j T_Q \times E[\bar{X}]$ in the steady state. Then, letting $V=\lim_{j\to\infty} E[V_j]$, the steady-state version of Eq. (3) can be rewritten as

$$T_Q = E[\alpha] + \lambda_i T_Q E[\bar{X}] + V. \tag{5}$$

Since the queuing delay T_Q may be affected by conflict with upstream traffic, we use the length of the delay line (or MTU) L as a slot unit to analyze the ring system. The arrival process is assumed to be a Poisson process, so the mean packet residual time α is simply

$$E[\alpha] = \frac{w_1 L_1 + w_2 L_2 + \cdots + w_q L_q}{R} = \frac{\bar{L}}{R}. \tag{6}$$

To approximate V for multichannel slotted ring networks, we use an upstream source to send packets through node i while, at the same time, node i also has traffic with average load $\rho_{Bi}=\Sigma_{k=2}^{N-1}\Sigma_{j=N-k+1}^{N-1}\lambda_{i\oplus k,i\oplus k\oplus j}E[X_j]$ to be sent. In this case, the upstream traffic will block the packet at the queue head of node i. Therefore, ρ_{Bi} can be derived from

$$\begin{aligned}\rho_{Bi} &= \sum_{k=2}^{N-1} \sum_{j=N-k+1}^{N-1} \lambda_{i\oplus k, i\oplus k\oplus j} E[X_j] = \frac{(N-1)\times(N-2)}{2} \times \frac{\lambda}{N(N-1)} \times \frac{E[M]}{R} \\ &= \frac{(N-2)\times\lambda\times E[M]}{2\times N\times R}. \end{aligned} \tag{7}$$

By this assumption, the average density ρ_{Bi} can be treated as a probability that the slot unit is fully occupied, and packets in node i have the probability ${\rho_{Bi}}^{i}(1-\rho_{Bi})$ of waiting to transmit. The mean waiting time $E[d]$ of finding an empty slot can be expressed as

$$E[d]=\sum_{i=0}^{\infty} i\frac{\bar{L}}{R}\rho_{Bi}{}^{i}(1-\rho_{Bi})=\frac{\bar{L}}{R}\times\frac{\rho_{Bi}}{1-\rho_{Bi}}=E[\bar{X}]\times\frac{\rho_{Bi}}{1-\rho_{Bi}}. \tag{8}$$

The steady-state duration of all vacation intervals V is equal to $\lambda_i T_Q \mathrm{E}[d]$. Equations (5) and (6) can be combined to obtain the average queuing delay:

$$\begin{aligned} T_Q &= E[\alpha]+\lambda_i T_Q E[X]+\lambda_i T_Q E[d] \\ &= \frac{\bar{L}}{R}+\lambda_i T_Q E[\bar{X}]+\frac{\rho_{Bi}(\lambda_i T_Q E[\bar{X}])}{1-\rho_{Bi}}. \end{aligned} \tag{9}$$

Therefore

$$T_Q=\frac{\dfrac{\bar{L}}{R}}{1-E[\bar{X}]\times\left(\dfrac{\lambda_i}{1-\rho_{Bi}}\right)}. \tag{10}$$

Because the packet transfer delay is composed of queuing, transmission, and propagation delays, the average packet transfer delay is

$$T_D=T_Q+T_S+\tau', \tag{11}$$

where τ' is the average propagation delay from a source node to a destination node, which is often expressed as $\tau/2$. The average transmission delay is

$$T_S=E[X]+E[d]=E[X]+E[\bar{X}]\times\frac{\rho_{Bi}}{1-\rho_{Bi}}. \tag{12}$$

Thus, the average transfer delay is given by

$$T_D=T_Q+S+\frac{\tau}{2}. \tag{13}$$

4.C. Analysis of the Multiring (WDM Ring) Case

To analyze the multiple WDM ring network, we assume that the traffic load from the upstream source is distributed equally among m rings. To simplify the analysis, let the circulation of slots on m rings be synchronized [36,37]; that is, a node can observe m slot units of different rings at the same time. Since the traffic load from the upstream source is distributed uniformly among the m rings, the average bridge traffic load of each ring, ρ_B, can be expressed as

$$\rho_B=\frac{\rho_{Bi}}{m}. \tag{14}$$

The probability that the packet is at the queue head and without an empty slot is $(\rho_B)^m$. Therefore, the probability that the packet has to wait i slot units before being transmitted is $(\rho_B)^{m\times i}(1-(\rho_B)^m)$. Similar to Eq. (8), let $E[d_B]$ be the average length of time to find an empty slot:

$$E[d_B]=\sum_{i=0}^{\infty} i\frac{\bar{L}}{R}(\rho_B)^{m-i}[1-(\rho_B)^m]=E[\bar{X}]\times\frac{(\rho_B)^m}{1-(\rho_B)^m}. \tag{15}$$

Since each packet in the queue has to wait for L/R, the average queuing delay is

$$T_Q=E[\alpha]+\lambda_i T_Q E[\bar{X}]+\lambda_i T_Q E[d_B]=\frac{\bar{L}}{R}+\lambda_i T_Q E[\bar{X}]+\lambda_i T_Q E[\bar{X}]\times\frac{(\rho_B)^m}{1-(\rho_B)^m}. \tag{16}$$

Therefore, the average queuing delay is given by

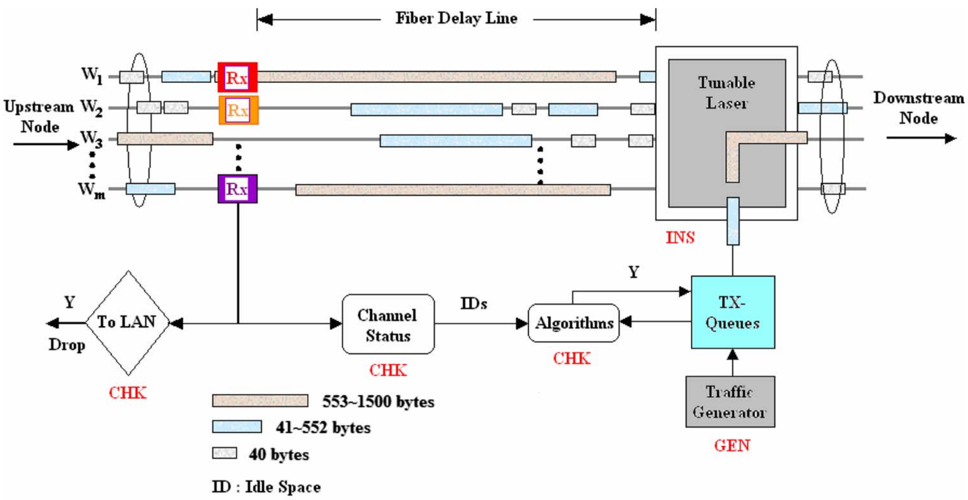
$$T_Q = \frac{\dfrac{\bar{L}}{R}}{1 - E[\bar{X}] \times \dfrac{\lambda_i}{1-(\rho_B)^m}}. \tag{17}$$

The average transmission delay is

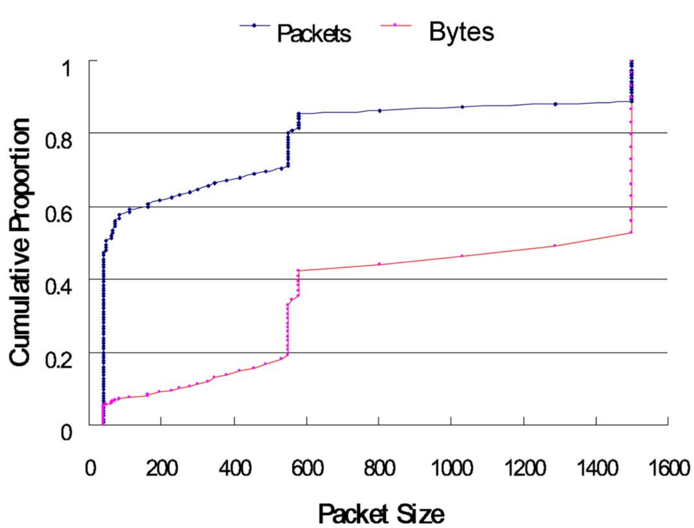
$$T_S = E[X] + E[d_B] = E[X] + E[\bar{X}] \times \frac{(\rho_B)^m}{1-(\rho_B)^m}. \tag{18}$$

Thus, the average transfer delay is obtained as

$$T_D = T_Q + T_S + \frac{\tau}{2}. \tag{19}$$



(a)



(b)

Fig. 7. Simulation models: (a) system simulation model; (b) OC3 traffic distribution model.

Table 1. Simulation Parameters for OPS Ring Networks

Protocols	Multitoken	CSMA/CA	CSMA/CP	CSMA/ID
Node architecture	TT−FR^m+FT−FR	TT–FR	TT−FR^m	TT−FR^m
Node numbers (N)	16	16	16	16
Wavelength numbers (m)	8	2, 4, 8	2, 4, 8	2, 4, 8
Token holding schemes	Exhaust, token holding time	X	X	X
Idle space allocation	X	X	RND	RND, LIS, BFS, FFS
Queue discipline	FIFO	VOQ_LQF	FIFO	FIFO, PCQ, PCQ_WRR (unit time=250 μs)
Length of ring network (R)	100 km			
Channel speed (C)	10 Gbits/s			

5. Simulation and Discussion

For evaluation of the theoretical analysis, simulation programs are produced to model the metro ring network. It is important to note that all simulations are run for sufficient time to obtain steady-state results. In general, 10×10^6 time units are simulated per point in each curve. Simulations are conducted by SIMSCRIPT II code, and each experimental value is calculated by the variance reduction technique with 40 replicated simulations using different seeds. The results are obtained with 95% confidence.

5.A. Simulation Models

The CSMA/ID simulation model is shown in Fig. 7(a). Four processes are used in the model: GEN, INS, RX, and CHK. GEN is responsible for generating IP packets with a size distribution that matches the trace from MCI's backbone OC-3 links [32] in Fig. 7(b). Thus, the mean packet size and the MTU can be calculated as 353.8 and 1500 bytes. INS is responsible for coordinating the transmission of packets in the TX queue and bypassing packets from FDL. CHK is responsible for checking the destination addresses of incoming packets, detecting the idle space status, and deciding which channel a packet should be put in by the ISA algorithms.

5.B. Performance Evaluation of OPS Ring Networks

Three MAC protocols are studied and compared with the CSMA/ID protocol on OPS metro ring networks. The MAC protocols and their general descriptions with relation to performance are as follows:

(1) CSMA/CP: the length of most packets in MCI's backbone OC-3 links is 40 bytes. Therefore, for reducing redundant fragments per packet and achieving good bandwidth utilization, the minimal idle space for fragmented packets should be designed equal to 40 bytes [13]. Hence, the length of the FDL is equal to 40+overheader length. Besides, a fragmental penalty has an overhead of packet header (20 bytes). However, the MAC controller never knows which idle space is large enough to reduce the number of fragmentations per packet when a large packet is being transmitted. Because the length of the FDL is insufficient, TQD and ISA adopt only the FIFO queue and the RND algorithm.

(2) CSMA/CA: the queue discipline is the virtual output queue (VOQ) with the longest queue first (LQF) [8] to guarantee fairness.

(3) Multitoken ring: this paper examines the effect of two multitoken ring parameters, exhaust and token holding time, on the performance of the network.

Following from the above description, the simulation parameters of the four protocols are given in Table 1. The throughput per node is the sum of bits sent from a node divided by the simulation time. Figure 8(a) shows the average transfer delay per packet versus the offered load under the multitoken ring protocol. The results show that both of the token holding schemes have low performance with $N=16$, $m=8$, and $R=100$ km. Even under light loading, both transmission delays exceed 18 ms because the token packet stays in a node's token queue for a long time. The maximum throughput is also relatively low in both cases, between 2.5 and 3 Gbits/s. Therefore, the multitoken protocol is not a good candidate for supporting metro ring networks. Figure 8(b) shows the average transfer delay per packet versus the offered load under

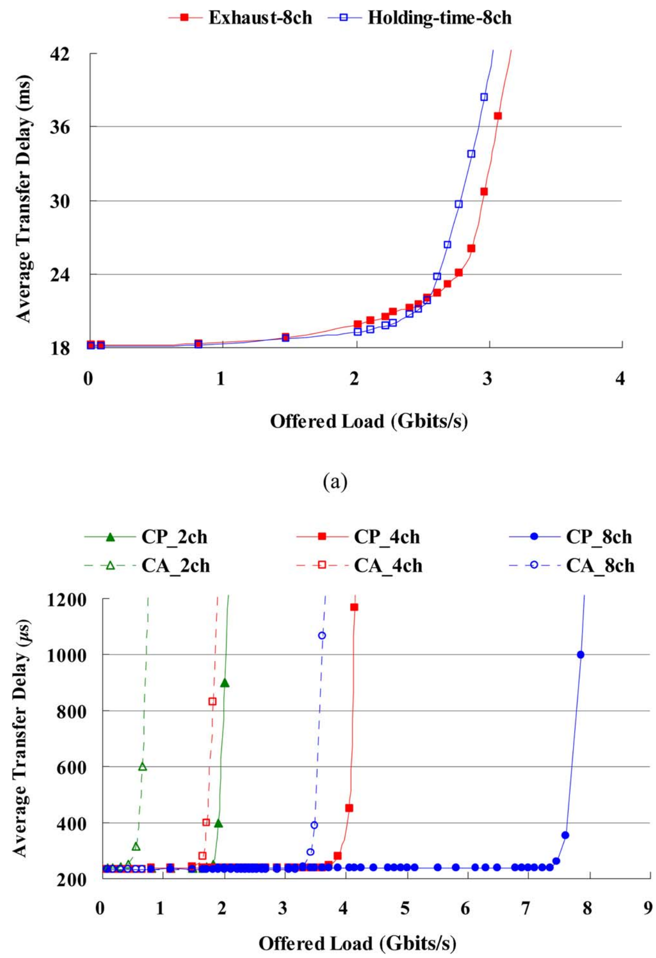


Fig. 8. Average transfer delay per packet with $N=16$ and $R=100$ km for (a) multitoken ring protocol with $m=8$; (b) CSMA/CA and CSMA/CP protocols with $m=2,4,8$.

the CSMA/CA and CSMA/CP protocols for various numbers of channels. The figure shows that the maximum throughput of each node at steady state by the CSMA/CP protocol is about 7.4 Gbits/s with $N=16$, $m=8$, and $R=100$ km, while for the same conditions the CSMA/CA protocol achieves only 3.25 Gbits/s. From the above it is seen that the performance of the CSMA/CP protocol is much better than the others.

The proposed system was simulated under three TQD architectures and four ISA algorithms for comparative evaluation of the average transfer delay per packet per the 12 TQD–ISA combinations. Figure 9 shows the relationship between the 12 combinations with $N=16$, $m=8$, and $R=100$ km. Of the individual TQD architectures, the simulation results show PCQ_WRR has the best performance, PCQ intermediate, with the FIFO queue showing the worst performance. Of the individual ISA algorithms, FFS shows the best performance followed by BFS, then LIS, and finally the RND algorithm. The reason is that FFS always uses the former channels, so the latter channels have more opportunity to retain large idle space for transmitting longer packets under heavy load. However, of the 12 TQD–ISA combinations, the best performance is provided by PCQ_WRR with FFS; its performance attains 7.5 Gbits/s at $N=16$, $m=8$, and $R=100$ km. This result is slightly higher than the CSMA/CP protocol. On the other hand, the worst performance of the combinations is FIFO with the RND algorithm, with a performance of about 6.1 Gbits/s. The difference between them, ~1.4 Gbits/s, is significant.

The CSMA/ID protocol based on PCQ_WRR and FFS at $N=16$, $m=8$, and $R=100$ km attains a total offered load capacity of 10 Gbits/s $\times$ 8 = 80 Gbits/s. The total received load of all the nodes equals 16×7.5 Gbits/s = 120 Gbits/s, so a 1.5-fold increase in channel capacity could be achieved since the spatial wavelength reuse property facilitates 1.5-fold channel utilization under balanced traffic conditions. Note, however, that the simulation results here depend strongly on the packet size distribution and the operational assumptions.

Next, in-depth simulation compares the diversity of the CSMA/ID and CSMA/CP protocols with regard to the impact of overhead, which includes a fixed packet header length and various guard-band distances. Here, the header length is equal to 20 octets and the guard-band distance is between 0 to 40 octets (0–32 ns at channel speed =10 Gbits/s). Figure 10(a) shows the impact of overhead using CSMA/ID and CSMA/CP protocols with $N=16$ and various numbers of channels, with CSMA/ID using the PCQ_WRR architecture and the FFS algorithm. The figure shows that CSMA/CP performance becomes increasingly weaker than CSMA/ID as the overhead increases. The reason is that, although the fragmentation scheme increases the utilization of idle space, it also requires more overhead when the guard-band distance or header length is large. In fact, almost half the packets of MCI's backbone OC-3 links are 40 bytes. Otherwise the number of fragmentations per packet will increase significantly, resulting in a serious difference of performance between CSMA/CP and CSMA/ID. The CSMA/ID has more propagation delay than CSMA/CP under steady-state conditions. However, it is only around 10 μs shown in Fig. 10(a). Figure 10(b) shows the average transfer delay of CSMA/ID and CSMA/CA under an asymmetric traffic load for various numbers of server nodes. In the simulations, CSMA/ID is using PCQ_WRR with the FFS algorithm, whereas CSMA/CA is using the VOQ discipline. The asymmetric traffic load is generated by the client–server model [38]. In the model, each client node generates a 1/3 traffic load for the server nodes and a 2/3 traffic load for the client nodes. Each server node generates a uniform traffic load for all nodes. According to the simulation model, reducing the number of server nodes in the network increases the phenomenon of asymmetric traffic load. From the simulation results, the average transfer delay for a symmetric traffic load is shorter than that for an asymmetric traffic load. The results also show that CSMA/ID performance is obviously greater than that of CSMA/CA. The reason is the CSMA/CA node has only one fixed receiver but the CSMA/ID node provides a dedicated receiver for each channel. Therefore CSMA/ID has more wavelength reuse than CSMA/CA. In addition, the wavelength reuse of a symmetric traffic load is always better than that of an asymmetric traffic load in a ring network with a destination removal policy.

Finally, Fig. 11 shows CSMA/ID with analytical values and simulation results for the average transfer delay per packet in the PCQ_WRR architecture with the RND algorithm for various numbers of channels. The figure shows the analytical values approach, the simulation results for CSMA/ID based on PCQ_WRR with the RND algorithm. Note that the analytical results are shown with a slightly higher load than the simulation results. This is caused by the assumption that analysis is adopted in a

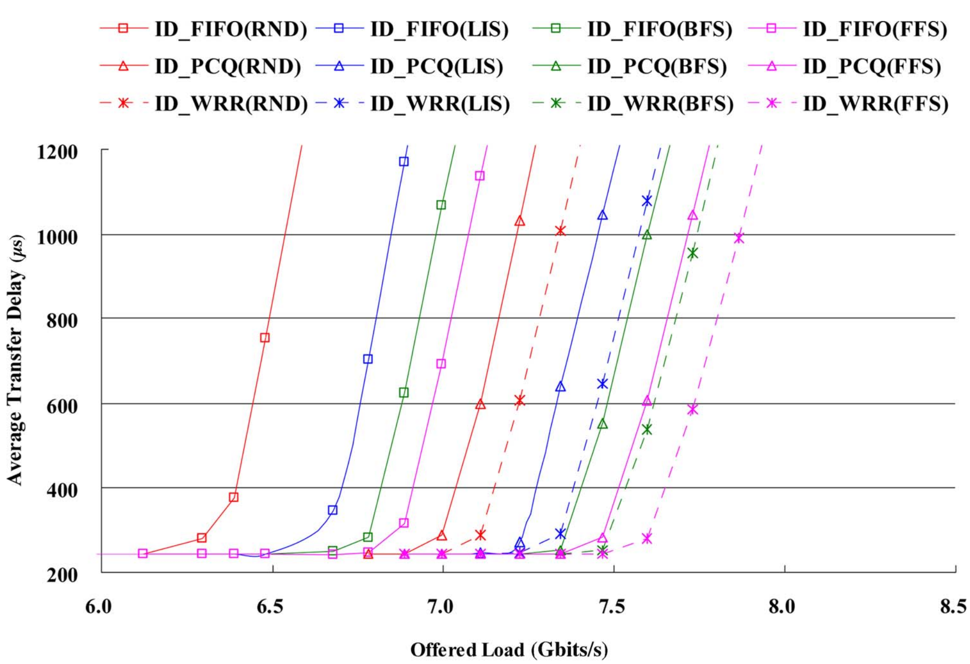
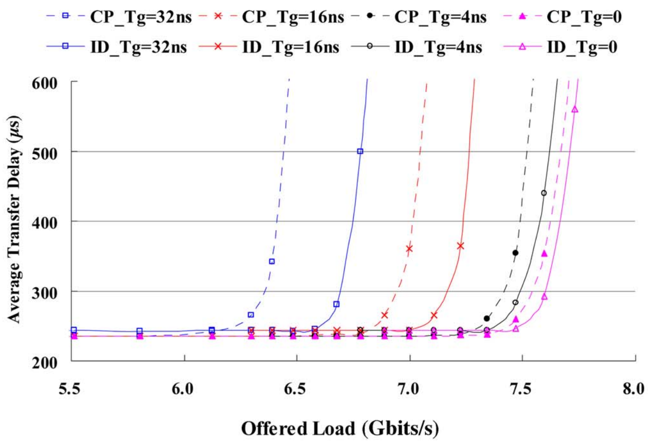
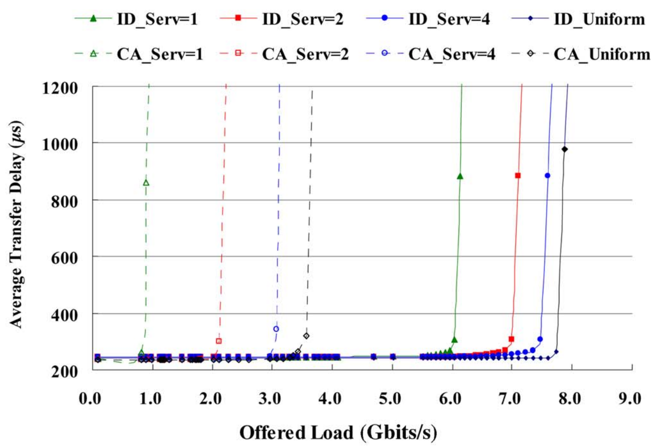


Fig. 9. Average transfer delay of the CSMA/ID ring with $N=16$, $m=8$, and $R=100$ km. Curves are shown for 12 combinations of TQD architectures and ISA algorithms.



(a)



(b)

Fig. 10. Average transfer delay versus offered load at $N=16$, $m=8$, and $R=100$ km for (a) overhead impact of CSMA/ID and CSMA/CP under symmetric traffic load; (b) CSMA/ID and CSMA/CA under asymmetric traffic load.

min-slot viewpoint to achieve an approximate performance of the CSMA/ID model. In the approach model, the idle length of minimum slots is usually slightly larger than the candidate packet, so the analytical results have better performance than the simulation results. From the above simulation and analysis, it is shown that the proposed protocol exhibits excellent characteristics, achieving high throughput and short delays for all-optical communication in the OPS metro ring network.

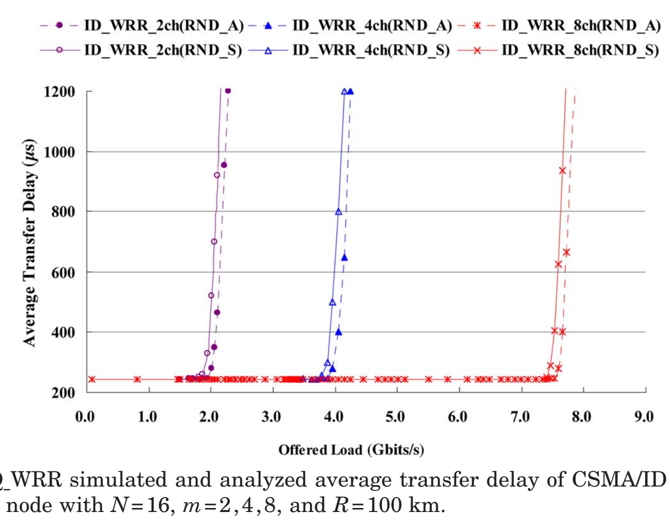


Fig. 11. PCQ_WRR simulated and analyzed average transfer delay of CSMA/ID ring versus offered load per node with $N=16$, $m=2,4,8$, and $R=100$ km.

6. Conclusion and Future Work

A novel MAC protocol has been proposed for all-optical OPS ring networks. The protocol supports IP packet transmission directly over WDM from LAN to MAN. We have investigated how to merge and collapse the middle layers between IP and WDM for the next-generation of optical LANs and MANs. This presented protocol avoids packet collision and packet fragmentation in the optical domain. For increasing channel utilization, packet scheduling with regard to ISA and queue discipline has been studied to understand their impact on performance. Simulation results for queue disciplines have shown that PCQ_WRR achieves better performance than FIFO and PCQ. These results are compared to results for PCQ_WRR with the RND algorithm for verification. Among the compared ISA algorithms, FFS shows the best performance, but the network system using the PCQ_WRR discipline together with the FFS algorithm gives the best bandwidth utilization among the 12 tested architecture–algorithm combinations. Its performance can approach 7.5 Gbits/s in the case of $N=16$ and $m=8$, so the spatial wavelength reuse property facilitates 1.5-fold channel utilization under balanced traffic conditions. Three related MAC protocols (multitoken ring, CSMA/CA, and CSMA/CP) are also studied and compared with CSMA/ID under symmetric and asymmetric traffic loads. Simulation results show the proposed protocol provides superior throughput relative to multitoken ring, CSMA/CA, and CSMA/CP protocols for all-optical communication.

There remains work to be done in the future. First, how to support a quality of service (QoS) scheme in CSMA/ID with PCQ_WRR architecture is being investigated. Second, more simulation cases about the asymmetric traffic load are being studied. Finally, we plan to propose a novel packet scheduling of CSMA/ID based on a mixed burst-length–timer scheme to decrease the influence of the guard-band time on the slow optical switch architecture of which the switching time belongs to the millisecond or microsecond level.

Acknowledgments

The authors acknowledge the National Science Council of Taiwan (NSCT) for supporting this research under project number NSC 96-2221-E-151-029.

References and Links

1. S. V. Kartalopoulos, "Elastic bandwidth, optical-fiber communication," IEEE Circuits Devices Mag. **18**(1), 8–13 (2002).
2. Y. Frignac, G. Charlet, W. Idler, R. Dischler, P. Tran, S. Lanne, S. Borne, C. Martinelli, G. Veith, A. Jourdan, J.-P. Hamaide, and S. Bigo, "Transmission of 256 wavelength-division and polarization-division-multiplexed channels at 42.7 Gb/s (10.2 Tb/s capacity) over 3 $\times 100$ km of Teralight fiber," in *Optical Fiber Communication Conference and Exhibit* (Optical Society of America, 2002), pp. FC5-1–FC5-3.
3. N. Ghani, S. Dixit, and T. S. Wang, "On IP-over-WDM Integration," IEEE Commun. Mag. **38**(3), 72–84 (2000).
4. C. S. Jeiger and J. M. H. Elmirghani, "Photonic packet WDM ring networks architecture and performance," IEEE Commun. Mag. **40**(11), 110–115 (2002).
5. R. Srivastava, R. K. Singh, and Y. N. Singh, "WDM-based optical packet switch architecture," J. Opt. Netw. **7**, 94–105 (2008).
6. M. A. Marsan, A. Bianco, E. Leonardi, M. Meo, and F. Neri, "MAC protocols and fairness control in WDM multirings with tunable transmitters and fixed receivers," J. Lightwave Technol. **14**, 1230–1244 (1996).
7. I. M. White, M. S. Rogge, K. Shrikhande, and L. G. Kazovsky, "A summary of the HORNET project: a next-generation metropolitan area network," IEEE J. Sel. Areas Commun. **21**, 1478–1494 (2003).
8. I. M. White, M. S. Rogge, K. Shrikhande, and L. G. Kazovsky, "Design of a control-channel media-access-control protocol for HORNET," J. Opt. Netw. **1**, 460–473 (2002).
9. K. S. Kim and L. G. Kazovsky, "Design and performance evaluation of scheduling algorithms for unslotted CSMA/CA with backoff MAC protocol in multiple-access WDM ring networks," Inf. Sci. (N.Y.) **149**, 135–149 (2003).
10. K. Bengi, "Access protocols for an efficient and fair packet-switched IP-over-WDM metro network," Comput. Netw. **44**, 247–265 (2004).
11. M. Chaitou, G. He'buterne, and H. Castel, "Performance of multicast in WDM slotted ring networks," Comput. Commun. **30**, 219–232 (2007).
12. N. Bouabdallah, A. L. Beylot, E. Dotaro, and G. Pujolle, "Resolving the fairness issue in bus-based optical access networks," IEEE J. Sel. Areas Commun. **23**, S12–S18 (2005).

13. J. H. Ho, W. S. Hwang, and C. K. Shieh, "A priority-aware CSMA/CP MAC protocol for the all-optical IP-over-WDM metropolitan area network," J. High Speed Netw. **16**, 157–173 (2007).
14. A. K. Turuk, R. Kumar, and R. Badrinath, "A token-based distributed algorithm for medium access in an optical ring network," Opt. Commun. **231**, 199–212 (2004).
15. C. James, A. Fumagalli, and I. Chlamtac, "The multitoken interarrival time (MTIT) access protocol for supporting variable size packets over WDM ring network," IEEE J. Sel. Areas Commun. **18**, 2094–2104 (2000).
16. F. Davik, M. Yilmaz, S. Gjessing, and N. Uzun, "IEEE 802.17 resilient packet ring tutorial," IEEE Commun. Mag. **42**(3), 112–118 (2004).
17. V. Ramnath, H. S. Cheng, and L. H. Ngoh, "Experimental and trial-based study of resilient packet ring," J. Opt. Netw. **1**, 258–269 (2008).
18. W. P. Chen, W. F. Wang, and W. S. Hwang, "A novel and simple beforehand bandwidth reservation (BRR) MAC protocol for OBS metro ring networks," J. High Speed Netw. **17**, 59–72 (2008).
19. W. P. Chen, B. Y. Liao, and W. S. Hwang, "Adaptive DBA: an algorithm for EPON based on all ONU REPORT messages," Int. J. Innov. Comput. I. **4**, 2171–2185 (2008).
20. G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: a dynamic protocol for an Ethernet PON (EPON)," IEEE Commun. Mag. **40**(2), 74–80 (2002).
21. H. Rongqing, Z. Benyuan, H. Renxiang, C. T. Allen, K. R. Demarest, and D. Richards, "Subcarrier multiplexing for high-speed optical transmission," J. Lightwave Technol. **20**, 417–427 (2002).
22. W. Wong and M. A. Summerfield, "Collision avoidance in all-optical wavelength division multiplexed packet networks," J. Lightwave Technol. **21**, 2652–2660 (2003).
23. S. M. Gemelos, I. M. White, D. Wonglumsom, K. Shrikhande, T. Ono, and L. G. Kazovsky, "WDM metropolitan area network based on CSMA/CA packet switching," IEEE Photon. Technol. Lett. **11**, 1512–1514 (1999).
24. S. Matsuo, S. H. Jeong, T. Segawa, H. Okamoto, Y. Kawaguchi, Y. Kondo, H. Suzuki, and Y. Yoshikuni, "Stable and fast wavelength switching in digitally tunable laser using chirped ladder filter," IEEE J. Sel. Top. Quantum Electron. **13**, 1122–1128 (2007).
25. C. M. Gallep and E. Conforti, "Reduction of semiconductor optical amplifier switching times by preimpulse step-injected current technique," IEEE Photon. Technol. Lett. **14**, 902–904 (2002).
26. G. I. Papadimitriou, C. Papazoglou, and A. S. Pomportsis, "Optical switching: switch fabrics, techniques, and architectures," J. Lightwave Technol. **21**, 384–405 (2003).
27. L. Y. Lin and E. L. Goldstein, "Opportunities and challenges for MEMS in lightwave communications," IEEE J. Sel. Top. Quantum Electron. **8**, 163–172 (2002).
28. T. Rogowski, S. Faralli, G. Bolognini, F. Di Pasquale, R. Di Muro, and B. Nayar, "SOA-based WDM metro ring networks with link control technologies," IEEE Photon. Technol. Lett. **19**, 1670–1672 (2007).
29. J. Cheyns, C. Develder, E. V. Breusegem, D. Colle, F. D. Turck, P. Lagasse, M. Pickavet, and P. Demeester, "Clos lives on in optical packet switching," IEEE Commun. Mag. **42**(2), 114–121 (2004).
30. A. Sharaiha, J. Topomondzo, and P. Morel, "All-optical logic AND–NOR gate with three inputs based on cross-gain modulation in a semiconductor optical amplifier," Opt. Commun. **265**, 322–325 (2006).
31. Q. Xin, P. Hartmann, R. V. Penty, I. H. White, and W. P. Krug, "A novel scalable photonic analog switch architecture based on semiconductor optical amplifiers," in *Avionics Fiber-Optics and Photonics* (IEEE, 2005), pp. 23–24.
32. WAN packet size distribution, available at http://www.nlanr.net/NA/Learn/packetsizes.html.
33. A. Francini, F. M. Chiussi, R. T. Clancy, K. D. Dtucker, and N. E. Idirene, "Enhanced weighted round robin schedulers for accurate bandwidth distribution in packet networks," Comput. Netw. **37**, 561–578 (2001).
34. X. Ye and T. David, "On the large deviation of resequencing queue size: $2-M/M/1$ case," in *Proceedings of IEEE INFOCOM* (IEEE, 2004), pp. 2374–2382.
35. H. M. Ghafir and C. B. Silio, "Performance analysis of a multiple-access ring network," IEEE Trans. Commun. **41**, 1494–1506 (1993).
36. C. S. Kang, "A broadband ring network: multichannel optical slotted ring," Comput. Netw. ISDN Syst. **27**, 1387–1398 (1995).
37. L. N. Bhuyan, D. Ghosal, and Q. Yang, "Approximate analysis of single and multiple ring networks," IEEE Trans. Commun. **38**, 1027–1040 (1989).
38. W. S. Hwang, W. F. Wang, J. Y. Wang, and C. C. Li, "A carrier preemption access control protocol for supporting IP packets over WDM ring networks," in *International Symposium on Communications Conference* (2001), session 19.2.