A Collision Avoidance Protocol for Supporting All-optical Ring Network

Wen-Ping Chen, Shih-Wei Kau, Wen-Shyang Hwang, Bin-Yih Liao+, and Wen-Tsung Hu
Department of Electrical Engineering, Department of Electronic Engineering+
National Kaohsiung University of Applied Science
pen@mail.ee.kuas.edu.tw, shihwei@mail.ee.kuas.edu.tw, wshwang@mail.ee.kuas.edu.tw, byliao@cc.kuas.edu.tw,
huwt@wshlab2.ee.kuas.edu.tw

Abstract

Recently, the number of multimedia applications in the Internet has grown enormously, and massive IP packets of variable-length are being made. There is a vast bandwidth requirement at the backbone network, such as wavelength division multiplexing (WDM) metro ring networks, to satisfy the demand; however, most of the current WDM metro ring networks are ineffective or complex. We propose a simple and effective MAC protocol named carrier senses multiple access with idle detection (CSMA/ID) to handle variable-length packets over an all-optical metro ring network. For numerically evaluating performance, an analytical model was developed by pre-classification queue with weighted round robin (PCQ_WRR) architecture and random algorithm. Further, there are two kinds of different MAC protocols: CSMA/CP and CSMA/ID are also examined and compared. By the simulation results show that our protocol has better throughput than CSMA/CA and CSMA/CP protocols under symmetric load on all-optical ring networks.

Keywords: *All-optical, Ring, Protocol*

1. Introduction

There has been an enormous increase in Internet bandwidth requirements because of exploding levels of information traffic from applications such as electronic commerce, multimedia and voice-over Internet protocol (VoIP). 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 ever-increasing demand, copper networks have been upgraded and replaced to great extent 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 Gb/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 limitation. DWDM technology now offers a solution for bandwidth insatiability, promising a backbone area network that can support more than 10.2 Tb/s.

Due to the many services and tremendous user population on the Internet, IP packet traffic dominates utilization of the Internet. However, IP packets now are transferred, switched and manipulated through complex protocol stacks such as IP/ATM/SONET/WDM and IP/HDLC/SONET/WDM. These extended network stacks result in high overhead, complicate system infrastructure and increases cost; the redundancy has become an important research issue. In addition, many wave division multiplexing (WDM) systems have been deployed in wide area networks (WANs), so the bottleneck of communications will be pushed from backbone networks to local access networks (LANs). In these circumstances, the metro network plays a critical role in the overall expansion of network services [4]. Metropolitan area networks (MANs) provide services within individual metropolitan areas, and they 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.

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 [1], HORNET [2] and Slot-ring [3]. Asynchronous transmission is for variable-length packets such as token-based [4] rings, HORNET (CSMA/CA with Backoff) [5], and CSMA/CP [6]. The latter is much better suited to IP traffic and considered in this study. In 1999 Stanford University proposed a HORNET test-bed, which uses hybrid optical-electronic technology to transmit packets and therefore the transmission rate is constrained. It uses TT-FR architecture and the CSMA/CA protocol [5] with a backoff process to handle variable-size packets. The architecture has poor performance, however, because the 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

978-0-7695-3382-7/08 $25.00 © 2008 IEEE
DOI 10.1109/ISDA.2008.323

same wavelength at its input. In 2007, [6] proposed CSMA/CP which uses packet fragmentation to solve the problem of packet collision in WDM ring networks. CSMA/CP is more efficient and has lower packet latency, but has two drawbacks. First, nodes must monitor the channel status all the time, regardless of whether searching idle space or transmitting packets, including the time required. Second, the CSMA/CP fragmentation scheme requires packet header overhead and guard-band distance, with further processing overhead from fragmentation and reassembly.

Therefore, to overcome the above issues, this study proposes a novel 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. To verify and demonstrate that the proposed CSMA/ID protocol offers high performance, this paper also considers two different MAC protocols compares them with CSMA/CP.

The WDM metro ring network architecture for the proposed protocol is presented in Section 2. A novel MAC protocol is proposed on the OPS metro ring networks with PCQ structure with WRR 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 network architecture presented here is a single and unidirectional fiber ring network which connects N nodes. Each optical fiber between the nodes is composed of m channels (W_1, W_2, W_3,…, W_m). 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, the node named access point (AP) is used for connecting LAN to MAN, while a node named present-of-point (PoP) is used for connecting MAN to 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 multi-ring network.

In the network, each AP is equipped with a tunable transmitter and m fixed receivers (TT-FR^m), where every receiver is dedicated to a particular channel for eliminating unfair access to node position and to increase the chance of packet transmission. The packets in an AP may come/pass from/to the attached LAN, gigabit Ethernet (GbE), passive optical networks (PONs) or wireless LAN. The packets are transmitted/received on/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.

The node architecture of a WDM metro ring network is shown in Fig. 1, 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 adopts a pre-classification queue with weighted round-robin (PCQ_WRR) scheduler. Each AP node listens to all wavelengths by continuously monitoring the subcarrier or baseband optical signals. It is similar to OFC’2000 and HORNET 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 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. The protocol processing time of the MAC controller is long enough because the fiber delay line (FDL) installed on the transmission path delays the packets from the upstream AP nodes for a fixed time T_{FDL} (for example, maximum transfer unit (MTU)). Thus there is no fragmentation and spatial wavelength reuse is improved. 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.

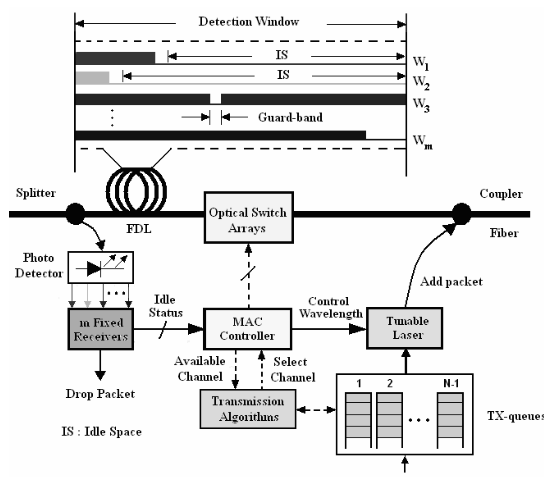

Figure 1. The node architecture of CSMA/ID.

3. MAC 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. 1. 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. After the decision had been made, the transmission of an AP in the CSMA/ID network can be divided into add/drop packet models. In the drop packet model, each AP listens to all wavelengths by monitoring either sub-carriers 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 for removing the optical signal part of the received packet in the fiber. If the address does not match the AP address, the optical signal in FDL is bypassed to the next AP node; in the add packet model, the MAC controller measures all idle spaces in every channel during the window T_{FDL} for any 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.

- *Transmission Queue Discipline (TDQ)*

(1) First-in-first-out (FIFO): 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 spaces, it uses the candidate packet size to select an idle space according to the selected ISA algorithm for the packet. In this way, packet collision is avoided, but the system throughput is slowed because of the effect of head-of-line (HoL).

(2) Pre-classification queuing (PCQ): the majority of packets in the Internet are from the Ethernet, which standard allows packet length between 40 and 64K bytes. However the measurement traced from MCI's backbone OC-3 links [25] shows a discrete packet-size distribution from 40 to 1500 bytes only, with most packet lengths being 40, 552 or 1500 bytes. Hence, according to the IP packet-size distribution, the IP packet that arrives from the LAN is pre-classified into one of three class queues: 40 bytes (Q_3), 41–552 bytes (Q_2) and 553–1500 bytes (Q_1). The MAC controller 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. In other words, it transmits the largest packet first to solve the HoL problem.

(3) Pre-classification queuing with weighted round-robin (PCQ_WRR): 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. Fortunately, the fairness issue can be mitigated by use of the weighted round-robin (WRR) discipline. In the proposed PCQ_WRR, the weighted value of each transmission queue must be calculated after each unit-time according to the total length of the incoming packets from the LAN. The higher arrival rate (λ) generates the higher weighted value. The packets will be transmitted in weighted round-robin 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 (1500, 1), (1512, 3) and (240, 6). As a result, the scheduling will be given as $Q_1 \rightarrow Q_2 \rightarrow Q_3 \rightarrow Q_2 \rightarrow Q_3 \rightarrow Q_2 \rightarrow Q_3 \rightarrow Q_3 \rightarrow Q_3 \rightarrow Q_3$.

- *Traffic-flow fairness index (TFI)*

In order to examine the fairness of the PCQ discipline, we define a fairness index is named traffic-flow fairness index (TFI). Therefore, the TFI of node i can be defined as

$$TFI = \sum_{j=1}^{q} \left(T_{ij} - \overline{T_i}\right)^2 / \overline{T_i}^{\,2} \tag{1}$$

Where q is the number of queues, T_{ij} is the ratio between output and input traffic flow for the j^{th} queue in node i, $\overline{T_i}$ is the mean ratio between output and input traffic flow for all queues in node i.

In unfair queue architecture, the highest priority queue is always given the best opportunities to transmit packets so each queue will have very different values of T_{ij} and Q_{ij}. Under these definitions, the *TFI* can become very large. On the other hand, in a fair architecture each queue will have similar values of T_{ij} and Q_{ij}. The *TFI* will then be very small or even approach zero. In particular, if the delay fairness index is zero then the queue architecture is perfectly fair. This means that the queuing delay and traffic-flow are the same in all queues.

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 multi-ring cases under the following assumptions.

- **Assumptions**

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

(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 $\lambda_{i,1}$, $\lambda_{i,1}, \cdots$, and $\lambda_{i,q}$ into q different class queues as determined by packet length; hence, the aggregate arrival rate of the network is $\lambda = \sum_{i,1=1}^{N-1} \lambda_{i,1} + \sum_{i,2=1}^{N-1} \lambda_{i,2} + \ldots + \sum_{i,3=1}^{N-1} \lambda_{i,q}$.

(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 ⊕ indicates the addition modulo N; thus, $\lambda_i = \sum_{j=1}^{N-1} \lambda_{i,i\oplus j}$.

(5) The mean packet generation for nodes is equal; each sends equal traffic to its destinations (uniform and symmetric traffic):

$$\lambda_i = \lambda / N, \quad \lambda_{i,i\oplus j} = \frac{\lambda_i}{N-1} = \frac{\lambda}{N(N-1)} \text{ and } \lambda_{i,i} = 0, \text{ for}$$
$$0 \le i \le N-1, \; 1 \le j \le N-1 \tag{2}$$

(6) The mean packet generation of PCQ_WRR for the nodes is equal. The number of packets generated in each queue of PCQ_WRR in an interval time t is n_1, n_2 and n_q, and the

mean packet length is L_1, L_2,…, and L_q. According to the packet distribution of MCI's backbone OC-3 links, the weight of the packet length in class r queue is $w_r = n_r L_r / \sum_{k=1}^{q} n_k L_k$ (r=1, 2,…, or q). The ratio of the arrival rate in each queue is $\lambda_{i,1} : \lambda_{i,2} : ... : \lambda_{i,q} = \frac{n_1 L_1}{t} : \frac{n_2 L_2}{t} : ... : \frac{n_q L_q}{t} = w_1 : w_2 : ... : w_q$, (where $\sum_{r=1}^{q} w_r = 1$), and the mean total packet length in PCQ is $\bar{L} = \sum_{r=1}^{q} n_r L_r / \sum_{r=1}^{q} n_r$. Therefore, the arrival rate $\lambda_{i,k}$ in node i can be given by $\lambda_{i,k} = w_k \lambda_i = w_k \times \frac{\lambda}{N(N-1)}$ and $\lambda_{i,k}$=0, for k=1, 2, …, q.

(7) The length of the packets is randomly determined by the independent, identically and geometrically distributed random variables (denoted by *r.v.* M bits) with mean $E[M]$ and probability mass function [8] $P_r(M=k)=\beta\times(1-\beta)^k$, k=0,1,…, where $\beta = \frac{1}{1 + E[M]}$.
(8) The channel bit rate of the WDM ring is R (b/s) and the packet transmission time is X (= M/R) seconds.
(9) The maximum transfer unit (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).
(10) The length of a transmit queue is infinite, and no packets are lost.

● *Notation:*

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

● *Analysis of the single-ring case*

Packets arrive with Poisson distribution at rate λ_i at each node. An M/G/1 queue with vacations can model our system. With the above assumptions, the average queuing delay for the j^{th} packet is given by

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

where

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

By Little's formula, the average number of packets in the PCQ_WRR, $E[N_j]$, can be obtained as $E[N_j] = \lambda_i \times T_{Qj}$. 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 equation (3) can be rewritten as

$$T_Q = E[\alpha] + \lambda_i T_Q E[\bar{X}] + V \quad (5)$$

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

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

To approximate V for the 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 $\rho_{Bi} = \sum_{k=2}^{N-1} \sum_{j=N-k+1}^{N-1} \lambda_{i\oplus k, i\oplus k\oplus j} E[X_j]$ 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

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

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

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

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

$$T_Q = E[\alpha] + \lambda_i T_Q E[X] + \lambda_i T_Q E[d] = \frac{\bar{L}}{R} + \lambda_i T_Q E[\bar{X}] + \frac{\rho_{Bi}(\lambda_i T_Q E[\bar{X}])}{1-\rho_{Bi}} \quad (9)$$

Therefore

$$T_Q = \frac{\bar{L}/R}{1 - \lambda_i E[\bar{X}] - \lambda_i E[\bar{X}] \times \frac{\rho_{Bi}}{1-\rho_{Bi}}} = \frac{\bar{L}/R}{1 - E[\bar{X}] \times (\frac{\lambda_i}{1-\rho_{Bi}})} \quad (10)$$

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

$$T_D = T_Q + T_S + \tau' \quad (11)$$

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

$$S = E[X] + E[d] = E[X] + E[\bar{X}] \times \frac{\rho_{Bi}}{1-\rho_{Bi}} \quad (12)$$

Thus, the average transfer delay is given by

$$T_D = T_Q + T_S + \tau / 2 \quad (13)$$

● *Analysis of the multi-ring (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 [9], 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 $\rho_B^{mi} \times (1-(\rho_B)^m)$. Similar to the subsection above, let $E[d_B]$ be the average length of time to find an empty slot unit

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

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

$$\begin{aligned} T_Q &= E[\alpha] + \lambda_i T_Q E[\overline{X}] + \lambda_i T_Q E[d_B] \\ &= \frac{\overline{L}}{R} + \lambda_i T_Q E[\overline{X}] + \lambda_i T_Q E[\overline{X}] \times \frac{(\rho_B)^m}{1-(\rho_B)^m} \end{aligned} \tag{16}$$

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

The average transmission delay is

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

Thus, the average transfer delay is obtained as

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

Table 1. Simulation parameters for OPS ring networks.

Protocols	CSMA/CA	CSMA/CP	CSMA/ID
Node Architecture	TT-FR	TT-FR^m	TT-FR^m
Fragment Length	X	40 bytes	X
TQD	VOQ_LQF	FIFO_RND	PCQ_WRR
Ring Length	100 km	100 km	100 km
Channel Speed	10 Gb/s	10 Gb/s	10 Gb/s

5. Simulation Results and Discussion

For evaluation of the theoretical analysis, simulation programs have been produced to model the metro ring network. It is important to note all simulations were run for sufficient time to obtain steady-state results. In general, ten million time units were simulated per point in each curve. The simulations are carried out by the SIMSCRIPT II codes. The traffic model is responsible for generating IP packets with a size distribution that matches the trace from the MCI's backbone OC-3 links. Thus, the mean packet size, the maximum transfer unit (MTU) and the ratio of the WRR in the queues of Q_1, Q_2 and Q_3 can be calculated as 353.8 bytes, 1500 bytes and 6:5:1. The simulation parameters of four protocols are given in Table 1.

We now compare the fairness of the queuing system under three scenarios: FIFO, PCQ, and PCQ_WRR. The TFI is plotted as a function of bandwidth in Fig. 2. The results show that FIFO is the fairest scheme, even though it is less efficient in other aspects. Fortunately, the TFI the PCQ_WRR is still quite fair; the indices in this scenario are small and nearly constant. The PCQ discipline, on the other hand, displays exponential behavior. These results demonstrate that the PCQ_WRR is quite fair.

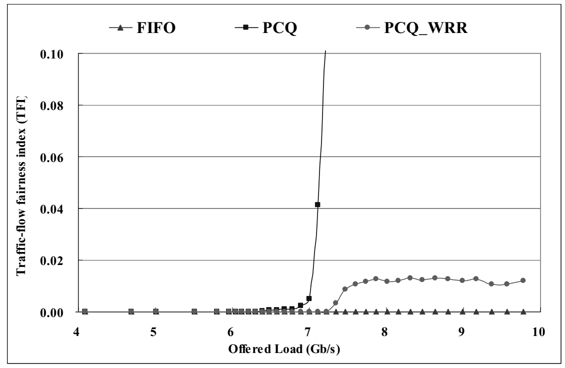

Figure 2. TFI for three different TDQ as a function of offered load.

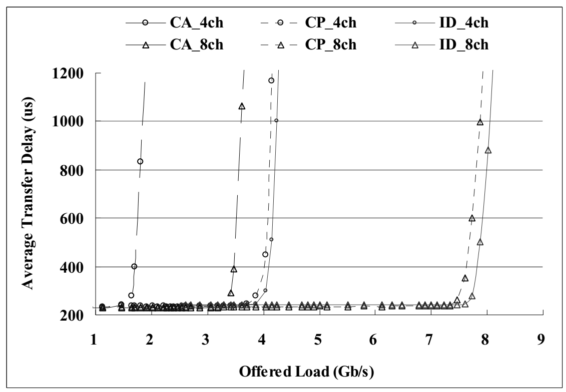

Figure 3. The average transfer delay with N=16 and m=4, 8 for CSMA/CA, CSMA/CP and CSMA/ID protocols.

Fig. 3 shows the average transfer delay per packet versus the offered load under three difference protocols for various numbers of channels, here, the CSMA/ID is adopts PCQ_WRR with the FFS algorithm. The figure shows the maximum throughput of each node at steady-state by the CSMA/ID protocol is about 7.6 Gb/s with N=16 and m=8, while for the same conditions the CSMA/CP protocol achieves 7.4 Gb/s and the CSMA/CA protocol achieves only 3.25 Gb/s. The results show that the performance of CSMA/ID protocol is much better than of CSMA/CA and CSMA/CP protocols.

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 a 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 Gb/s). Fig. 4 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 a 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, the almost half the packets of the MCI's backbone OC-3 links are 40bytes. Otherwise the fragment times per packet will increase significantly, resulting in a serious difference of performance between CSMA/CP and CSMA/ID.

Finally, Fig. 5 shows CSMA/ID with analytical values and simulation results of 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.

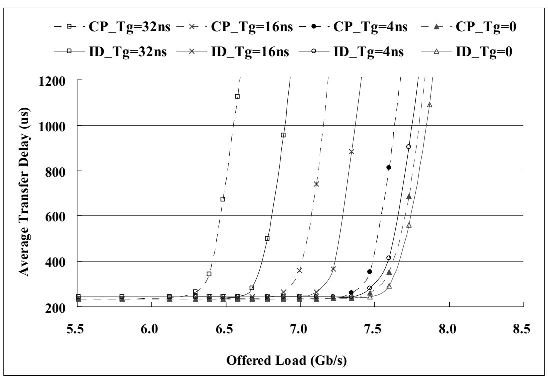

Figure 4. Average transfer delay versus offered load at N=16, m=8 for overhead impact of CSMA/ID and CSMA/CP under symmetric traffic load.

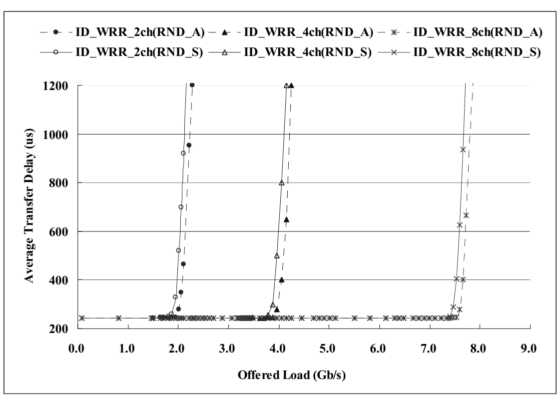

Figure 5. The average transfer delay of CSMA/ID simulated and analyzed average with N=16 and m=2, 4, 8.

6. Conclusions

A novel MAC protocol has been proposed for optical WDM 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 next-generation optical LANs/MANs. This presented protocol avoids packet collision and packet fragmentation in the optical domain. For increasing channel utilization, packet scheduling with regard to idle space allocation and queue discipline have been studied to understand their impact on performance. The simulation results for queue disciplines have shown that PCQ_WRR achieves better performance than FIFO and PCQ. These results were compared to results for PCQ_WRR with the RND algorithm for verification. Besides, the CSMA/ID using the PCQ_WRR discipline together with the FFS algorithm gives the best bandwidth utilization than that of CSMA/CA and CSMA/CP protocols. Its performance can approach 7.6 Gb/s in the case of $N = 16$ and $m = 8$. Simulation results showed the proposed protocol displays superior throughput relative to CSMA/CA and CSMA/CP protocols for all-optical communications.

Acknowledgment

The authors would like to thank National Science Council (NSC) in Taiwan R.O.C. for supporting this research; the project number is NSC 96-2221-E-115-017.

References

[1] M.A. Marsan, A. Bianco, E. Leonardi, M. Meo, F. Neri, "MAC protocols and fairness control in WDM multirings with tunable transmitters and fixed receivers", *IEEE Journal of Lightwave Technology*, vol. 14, no. 6, Jun. 1996, pp.1230 –1244.

[2] I.M. White, M.S. Rogge, K. Shrikhande, L.G. Kazovsky, "A summary of the HORNET project: a next-generation metropolitan area network", *IEEE Journal of Selected Areas in Communications*, vol. 21, no. 9, Nov. 2003, pp.1478–1494.

[3] H.T. Lin, W.R. Chang, H.T. Wu, "FARE: An efficient integrated MAC protocol for differentiated services in WDM metro rings", *Computer Communications*, vol. 30, no. 6, Mar. 2007, pp.1315–1330.

[4] A.K. Turuk, R. Kumar, and R. Badrinath, "A token-based distributed algorithm for medium access in an optical ring network", *Optics Communications*, vol. 231, no. 1-6, Feb. 2004, pp.199–212.

[5]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", *Information Sciences*, vol. 149, no. 1-3, Jan. 2003, pp.135–149.

[6] J.H. Ho, W.S. Hwang, C.K. Shieh, "A priority-aware CSMA/CP MAC protocol for the all-optical IP-over-WDM metropolitan area network", Journal of High Speed Networks, vol.16, no.2, 2007, pp.157–173.

[7] http://www.nlanr.net/NA/Learn/packetsizes.html.

[8] H.M. Ghafir, C.B., Jr. Silio, "Performance analysis of a multiple-access ring network", *IEEE Transaction on Communications*, vol.41, no.10, Oct. 1993, pp.1494–1506.

[9] C.S. Kang, "A broadband ring network: multichannel optical slotted ring", *Computer Network and ISDN Systems*, vol. 27, no.9, Aug. 1995, pp.1387–1398.