INTERNATIONAL JOURNAL OF COMMUNICATION SYSTEMS
Int. J. Commun. Syst. 2008; **21**:1155–1170
Published online 18 June 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/dac.944

Performance evaluation of CSMA/ID MAC protocol for IP over WDM ring networks

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

[1]*Department of Computer Science and Information Engineering, Diwan University, Taiwan*
[2]*Department of Electrical Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan*
[3]*Department of Electrical Engineering, National Cheng Kung University, Tainan, Taiwan*

SUMMARY

In this paper, a packet pre-classification media access control protocol based on a carrier sense multiple access with idle detection (CSMA/ID) scheme is investigated for supporting IP packets over all-optical WDM ring networks. The purpose of the protocol is to increase throughput and to decrease the packet transmission delay of IP packets over optical networks in a metropolitan area network. This protocol avoids both packet collision and packet fragmentation. In order to improve the utilization of the network, the packets transmitted from a local area network are first pre-classified into various class queues of an access point (AP) according to their length. After checking the available space based on the wavelength received by the receivers of the AP, the packets in the queues are transmitted. An analytical model is developed to evaluate the performance of the protocol, with simulation results showing good network efficiency. The proposed network has short-term variations that introduce unfairness conditions. This problem could be overcome by assigning a quota on individual queues to allow all queues fair access.

Received 1 April 2006; Revised 5 March 2008; Accepted 29 March 2008

KEY WORDS: IP over WDM; CSMA/ID; packet pre-classification; analytical model; fairness

1. INTRODUCTION

With the explosion of information traffic due to the rise of the Internet, electronic commerce, computer networks, voice, data, and video transmission, the need for a medium with the bandwidth capabilities for handling a vast amount of information is paramount. Recent advances in solid-state and photonic technologies have delivered bit wavelengths of 2.5, 10, and 40 Gbp/s. The data can be sent over optical fibers, a transmission medium that permits light to travel through it without amplification for hundreds of kilometers. Currently, the total bandwidth of an optical fiber

*Correspondence to: Jih-Hsin Ho, Department of Computer Science and Information Engineering, Diwan University, Taiwan.
†E-mail: hjsin@hpds.ee.ncku.edu.tw, hjsin@dwu.edu.tw

exceeds 2 Tbit/s (200 × 10 Gbit/s), 2.4 Tbit/s (120 × 20 Gbit/s), 3 Tbit/s (300 × 11.6 Gbit/s), and 3.2 Tbit/s (80 × 40 Gbit/s) [1–3]. Research has demonstrated that the number of wavelengths per fiber could increase to more than 1 000 [4]. This indicates that WDM can be a solution for the ever-growing bandwidth demand.

Owing to the growing number of services and users on the Internet, IP packets dominate data networks. These packets are transferred, switched, and manipulated through complex protocol stacks, such as IP/ATM/SONET/WDM and IP/HDLC/SONET/WDM. How to merge and collapse the middle layers to reduce cost, complexity, and redundancy is an important research issue [5–7]. Additionally, since many WDM systems have been deployed in wide area networks (WANs), the bottleneck of communications will be pushed from the backbone networks to local access networks. As a result, how to apply WDM to local and metropolitan area networks has attracted a lot of research interest [4–8].

A number of papers have examined WDM ring networks. Cai *et al.* proposed the MTIT access protocol for supporting variable size packets over WDM ring networks based on a fixed-transmitters-and-fixed- receivers architecture [8]. To achieve all-optical communications, MTIT adopts the source removal policy [6] for dropping packets from networks to prevent packet re-circulation. Shrikhande *et al.* developed HORNET as a testbed for a packet-over-WDM ring metropolitan area network (MAN) [7]. To facilitate signal regeneration and destination removal, HORNET utilizes opto-electronic and electro-optic conversion, which may constrain the transmission rate of the network. Although the IP standard allows a packet length of between 40 and 64 kbytes, a measurement trace from one of the MCIs backbone OC-3 links shows a discrete packet-size distribution, from 40 to 1500 bytes (see Figure 1) [8]. The smallest packet of 40 byte corresponds to TCP ACK packets and the 1500-byte packets are Ethernet's maximum transfer unit (MTU). Figure 1 shows that almost 5.34% (byte volume) of the packets are 40 bytes long, 27.58% (byte volume) of the packets are 41–552 bytes long, and 67.08% (byte volume) of the packets are 553–1500 bytes long. Hwang *et al.* proposed an all-optical media access control (MAC) protocol based on avoiding packet collision using a fragment packet scheme for all-optical WDM multi-rings with a tunable transmitter and fixed receiver (TT-FR) [9]. However, avoiding packet collisions using this fragment scheme creates a large number of fragmented packet and introduces

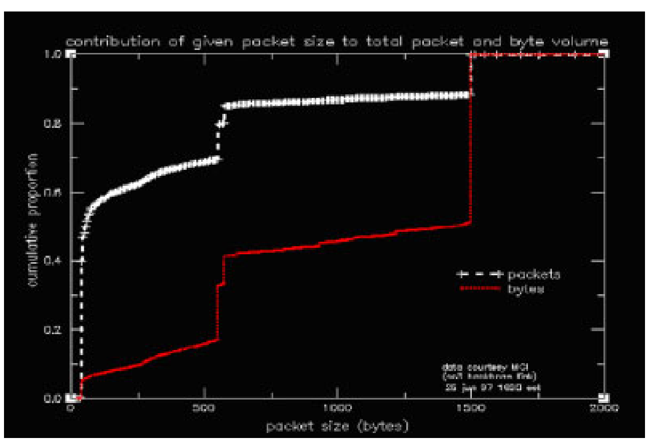

Figure 1. Cumulative distribution function (CDF) of IP packet sizes on an Internet backbone link.

DOI: 10.1002/dac

complexity. For this reason, we propose a new MAC protocol that avoids packet collision without a fragment scheme. In this paper, the WDM ring network architecture, carrier sense multiple access with idle detection (CSMA/ID) protocol, and transmission algorithms are presented in Section 2. Analytical models for evaluating the average packet delay performance are developed in Section 3. Then, Section 4 validates the accuracy of the proposed model by comparing the analytical results with those obtained using simulations. Finally, Section 5 contains the conclusion.

2. NETWORK ARCHITECTURE AND CSMA/ID MAC PROTOCOL

2.1. The network architecture

The network architecture used in this paper is a single, unidirectional fiber ring network, which connects an N number of nodes. The optical fiber is composed of W data channels $(\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_w)$, as shown in Figure 2. The network scope is assumed to cover a metropolitan area (i.e. a ring circumference of about 100 km); therefore, the system is referred to as a WDM metro ring. The access points (APs) connect local area networks (LANs) to the MAN ring network, while PoP connects the MAN to the WAN. Each data channel makes use of one specific wavelength to convey the optical signal. Therefore, using WDM technology, channels can work independently without interfering with each other. Logically, the network can be treated as a multi-ring network.

The node structure of the network is shown in Figure 3. Each node has one tunable transmitter and W fixed receivers, one for each data channel. For the optical signal sent from upstream nodes, a splitter is used to tap off a small portion of the optical power from the ring to the receivers. Every receiver detects the optical signal carried in its corresponding wavelength within the output

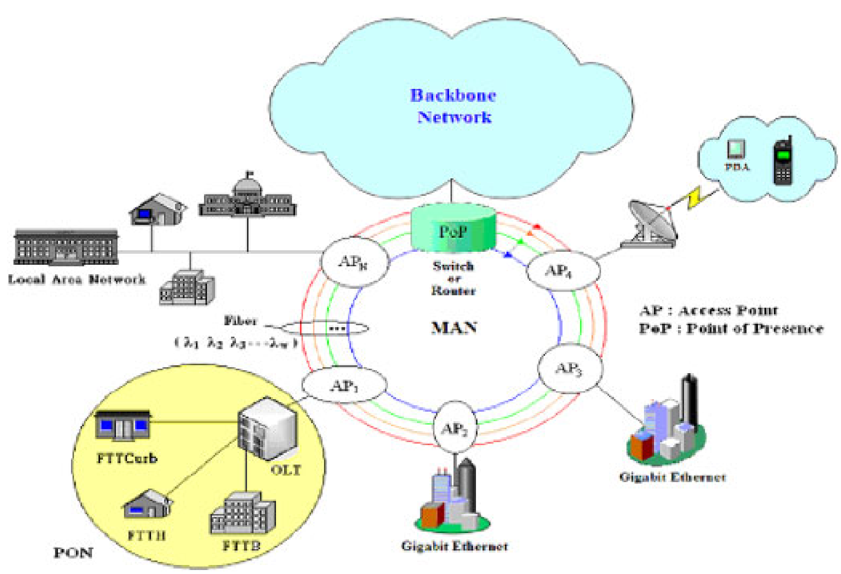

Figure 2. Architecture of a metro WDM ring.

DOI: 10.1002/dac

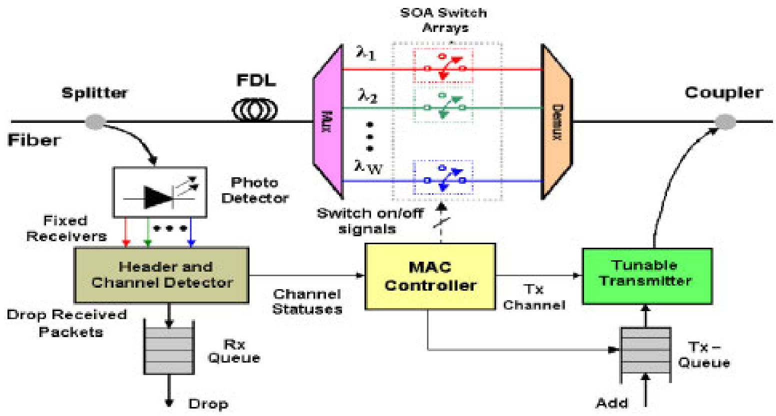

Figure 3. The node structure of the network.

branch from the splitter for node address identification. If the destination address in the incoming packet header matches the node address, the packet data are sent to the host. Meanwhile, the MAC control scheme is signaled to activate the opening of the on–off switch for the corresponding data channel in order to remove the received packet carried in the major portion of the optical signal through the delay line. If the node is not the packet's destination, the detected packet is ignored and the process of scanning the next packet is started.

2.2. CSMA/ID MAC protocol

The downstream AP recognizes an incomplete IP packet by the presence of the sub-carrier signal and pulls it off the ring. The carrier sense can check the available channel length (ACL) to notify the Tx to transmit the packet to the queue. Based on the protocol, each node monitors the wavelengths and detects the corresponding ACL, provided that there are IP packets for transmission. If an IP packet is being transmitted to a target channel while the node is detecting another IP packet arriving on the same channel at its input, a ring access problem (an access collision) will occur. Such collisions are due to the fact that the node cannot know if the opening is long enough to accommodate the packet. With the carrier access scheme, to guarantee the correctness of the protocol operations, the delay line inside the nodes must be used to delay the incoming packet. In addition, the delay line should be long enough to cover the maximum IP packet length (1500 bytes) so that unnecessary fragmentation can be avoided along with packet collision, thus improving the utilization of the bandwidth. The fiber delay line inside the AP is also responsible for processing IP packet time. Figure 4 shows the CSMA/ID MAC protocol flowchart. The MAC protocol decides whether the packet in the queue can be transmitted or not according to idle channel messages, transmit packet lengths, and the transmission algorithms. This protocol has three features: (1) it is a fully distributed, asynchronous protocol that does not need a centralized controller or a separate control channel to harmonize and synchronize the operations of nodes; (2) a transmitting packet will not collide with an incoming packet on the same wavelength because the FDL length (1500 bytes) is long enough to avoid this; and (3) it supports variable-length IP packets without complicated segmentation and reassembly, which becomes harder as the line speed of optical wavelengths increases.

DOI: 10.1002/dac

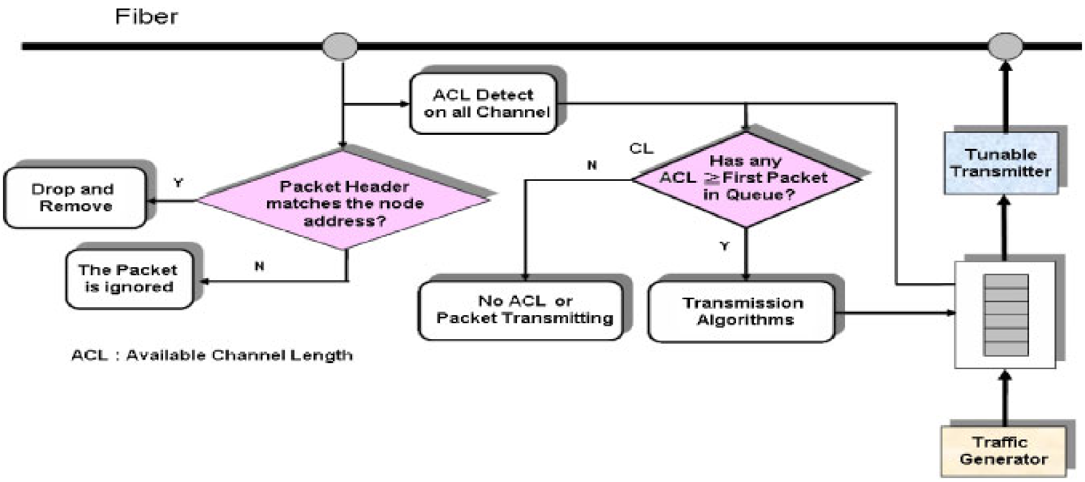

Figure 4. The CSMA/ID MAC protocol flowchart.

2.3. Transmission algorithms for CSMA/ID MAC protocol

2.3.1. First in first out (FIFO) queueing scheduling. The packet that first arrived will be the first to be served in different Tx-queues (Q1, Q2, and Q3). The MAC controller monitors all channels' ACL in the delay line, which is larger than or equal to the length of the packet in the FIFO queue. We select one queue based on the ACL algorithms and transmit the packet to the corresponding delay line; otherwise, the packet has to remain in the buffer (electronic memory) of the TX-queue until sufficient ACL is available. In this manner, the packet collision problem can be avoided; however, the head-of-line will decrease the throughput performance.

2.3.2. Pre_classification queueing (PCQ) scheduling. The process is as follows:

1. IP packets are pre-classified into three kinds of queues (Q_1, Q_2, and Q_3) by the buffer selector. The three kinds of queues are for storage of 553–1500, 41–552, and 40 byte packets, respectively. The MAC controller reads the IP packet size storage message and sends it to the appropriate queue.
2. Since each node is equipped with a receiver for its corresponding data channel, an IP packet can be transmitted via a corresponding available data channel to its destination node. The receiver is responsible for checking the destination address of incoming IP packets and detecting available channels to notify the MAC controller.
3. Using the information from (1) and (2), the above MAC controller delivers a message to the active buffer selector to transmit Q_1, Q_2, or Q_3 buffer packets. Figure 5 illustrates the MAC controller model flowchart and Figure 6(a) illustrates an example of the MAC controller scheme. If the maximum ACL is 552 bytes and the three kinds of queue storage for the first IP packet are 1200, 512, and 40 bytes, respectively, then the MAC controller transmits a message to a tunable transmitter, which transmits Q_2 buffer's packets. Figure 6(b) illustrates another example of the MAC controller scheme. If the receivers detect that the ACL is smaller than the first packet for the buffer selector, then the MAC controller does not transmit a message to the TX, and the TX will not transmit any packets to the fiber.

DOI: 10.1002/dac

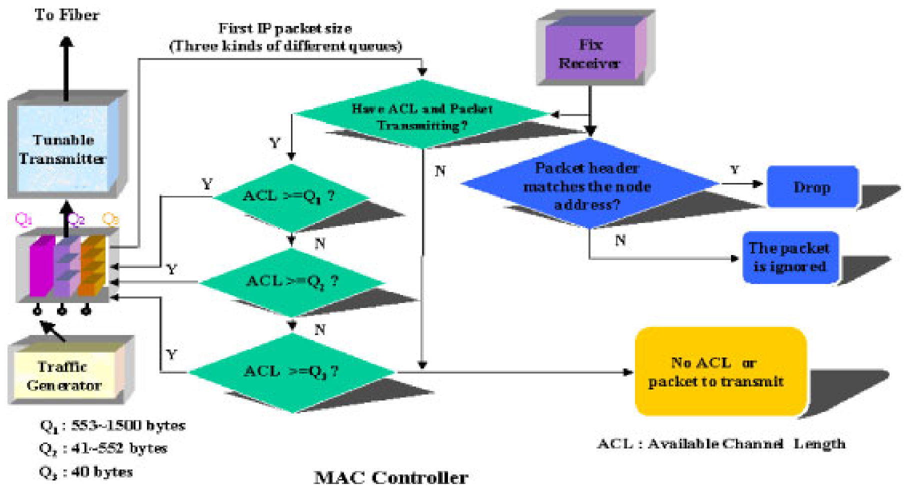

Figure 5. The MAC controller model flowchart (with IP packets pre-classified into three kinds of queues).

The downstream AP recognizes an incomplete IP packet by the presence of the sub-carrier signal and pulls it off the ring. The carrier sense can check the ACL to notify the TX to transmit the packet from the Q_1, Q_2, or Q_3 buffers.

2.3.3. Pre_classification queueing with quotas (PCQ_quota) scheduling. IP packets are pre-classified into three kinds of queues (Q_1, Q_2, and Q_3) by the buffer selector. The three kinds of queues are for storage of 553–1500, 41–552, and 40 byte packets, respectively. The classification is based on the previous long-term measurement, as in PCQ scheduling. The PCQ_quota scheduling is designed specifically for short-term variations that may cause fairness problems. Each queue is assigned a quota (in bytes), which it is allowed to transmit. If the queue is still transmitting at the end of the quota, the MAC controller is preempted and given to another queue.

2.4. *The frame format*

To support the carrier access scheme, a frame format is adopted, as shown in Figure 7. The carrier sensing mechanism for finding transmitted packets in an optical fiber can be based on sub-carrier signaling [10] or receiver monitoring. For sub-carrier signaling, each wavelength is associated with a sub-carrier frequency. When a node transmits a packet, it multiplexes the corresponding sub-carrier frequency. The nodes determine the occupancy of all wavelengths in parallel by monitoring the sub-carriers in the RF domain. In addition, since each receiver extracts the optical signals from the corresponding data channel, receiver monitoring can be used to determine the occupancy of all wavelengths. It seems natural that the receivers should be associated with an auxiliary function to monitor the status of the optical ring network. Today, the cost of such receivers is still too high for them to be economical to manufacture, but a cheaper process may be developed in the future. The start delimiter (SD) and the end delimiter (ED) mark a physical data frame conveyed in data channels for packets. The source address (SA) and the destination address (DA) serve as the address information in the network. To prevent possible errors midway through the transmission, the cyclic redundancy check (CRC) is employed. The flag (FG) field is reserved for extended protocol functions.

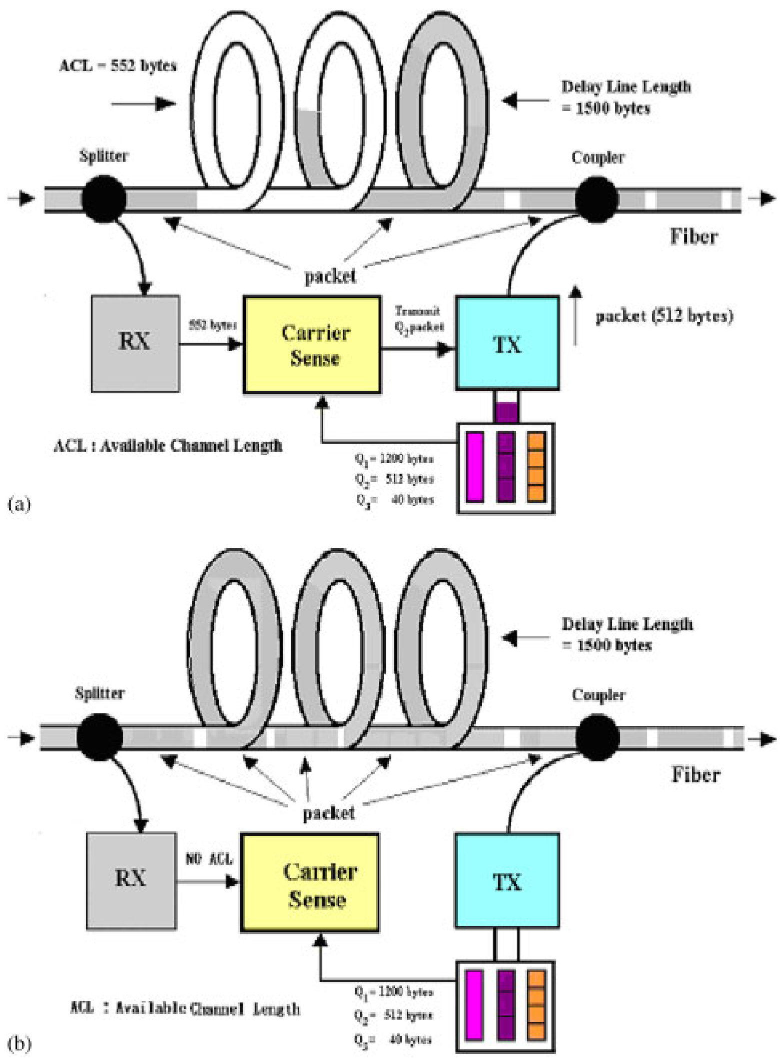

Figure 6. (a) and (b) Examples of the MAC control model with the pre-classification scheme.

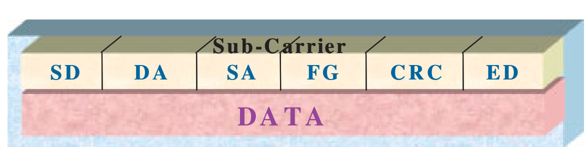

Figure 7. The frame format.

DOI: 10.1002/dac

3. PERFORMANCE EVALUATION

The transfer delay of a packet is measured from when the packet is completely stored in the source node queue until that packet has been completely received by the destination node. This delay consists of the queuing delay, the transmission delay, and the propagation delay. The queuing delay of a packet is measured from when a packet is fully stored in a queue of the source node to the time the source node was last selected by the queue before successful transmission. In this investigation, the transmission delay is defined as the interval between the source node selecting the queue to transmit the packet successfully and the time the source node last selected the queue before transmitting the packet successfully. Finally, the propagation delay of a packet is the interval between the time that the last bit of the packet reaches the destination and the moment that the last bit of the packet was transmitted. From the behavior of the expected queuing delay for the ith packet, the model can be categorized as an M/G/1 queue with vacations model [11]. Clearly, the queuing delay captures the effect of contention and is dependent on traffic density. In order to express the packet transfer delay at a node on multi-rings using an M/G/1 vacation model, we first present some assumptions and the general notations used in the subsections below.

3.1. Assumptions

For simplicity, the following assumptions are made:

1. The number of WDM channels is W.
2. The total propagation delay of the WDM ring is τ seconds, and the distances between the nodes are equal.
3. Packets that arrive are an independent, identically distributed (i.i.d.) Poisson process with rate λ_i (packets/second) at each of the N nodes on the ring, with an aggregate arrival rate for the network of $\lambda=\sum_{i=0}^{N-1}\lambda_i$.
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 addition modulo N; thus $\lambda_i=\sum_{j=1}^{N-1}\lambda_{i,i\oplus j}$. In the case of uniform and symmetric traffic on the ring, the mean packet generation for all nodes is equal and each source sends equal traffic to all destinations:

$$\lambda_i=\lambda/N, \quad \lambda_{i,i\oplus j}=\frac{\lambda_i}{N-1}=\frac{\lambda}{N(N-1)} \tag{1}$$

   and

$$\lambda_{i,i}=0 \quad \text{for } 0\leqslant i\leqslant N-1, \ 1\leqslant j\leqslant N-1$$

5. The packets have random lengths determined at each node as independent, identically and geometrically distributed random variables (denoted by $r.v.$ M (bits)) with mean $E[M]$ and probability mass function [11] $P_r(M=k)=\beta\cdot(1-\beta)^k, k=0,1,2,\ldots,$ where $\beta=1/(1+E[M])$.
6. The WDM ring channel bit rate is R (bps) and the packet transmission time is $X(=M/R)$ seconds.
7. Define MTU as equal to the delay line ($L=1500$ bytes) with $T_i=L/R$ seconds to transmit the MTU.

DOI: 10.1002/dac

3.2. Notations

The following notations are used in the analytical formulas below: D, average packet transfer delay; X, packet transmission time; TQ_i, queue-waiting delay of packet i; TQ, average packet queue-waiting delay; α_i, residual time of packet i; α, packet residual time; V_i, duration of all the whole vacation intervals for which packet i must wait before being transmitted; V, steady-state duration of all whole vacation intervals; and S, average transmission delay.

3.3. Analysis of the single-ring case

With the above assumptions, we model the queuing and transmission delay using an M/G/1 queue with vacations, as illustrated in Figure 8. The average queuing delay, TQ_i, for the ith packet is given by

$$E[\mathrm{TQ}_i]=E[\alpha_i]+E[M_i]E[X]+E[V_i] \tag{2}$$

The queuing delay and transmission delay capture the effect of contention and upstream traffic dependence. Thus, we consider the delay line (or MTU) as a slot unit. Therefore, the dependence occurs when the full slots are uniformly and independently distributed on a single ring. Since the arrival process is assumed to be Poisson, this residual time α can be considered to be uniformly distributed between 0 and L/R. Therefore, the mean packet residual time is simply

$$E[\alpha]=\frac{L}{2\times R} \tag{3}$$

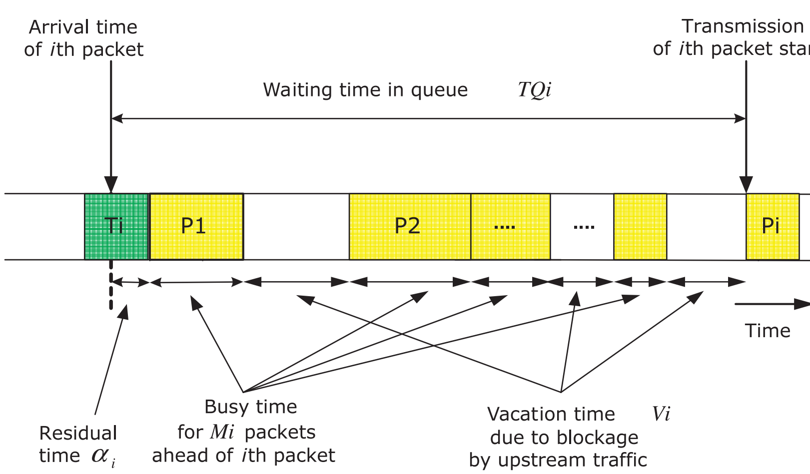

Figure 8. Calculation of the average waiting time in the M/G/1 system with vacations. The average waiting time $E[\mathrm{TQ}_i]$ of the ith packet is $E[\mathrm{TQ}_i]=E[\alpha_i]+E[M_i]E[X]+E[V_i]$.

DOI: 10.1002/dac

Using Little's formula, the value of $\lim_{i\to\infty} E[M_i]E[X]$ is $\lambda_i \mathrm{TQ}E[X]$. Letting $V=\lim_{i\to\infty} E[V_i]$, we can thus express the steady-state version of Equation (2) as

$$\mathrm{TQ}=E[\alpha]+\lambda_i \mathrm{TQ}E[X]+V \tag{4}$$

Next, we calculate approximation V for multi-channel slotted ring networks. Packets sent by an upstream source use node i as a bridge to reach their destinations. This bridge has an average traffic load of $\rho_{\mathrm{B}i}=\sum_{k=2}^{N-1}\sum_{j=N-k+1}^{N-1}\lambda_{i\oplus k,i\oplus k\oplus j}E[X_j]$.

This upstream traffic blocks the head of the queue packet at node i. Substituting the above assumptions into $\rho_{\mathrm{B}i}$ gives

$$\begin{aligned}\rho_{\mathrm{B}i}&=\sum_{k=2}^{N-1}\sum_{j=N-k+1}^{N-1}\lambda_{i\oplus k,i\oplus k\oplus j}E[X_j]\\&=\frac{(N-1)(N-2)}{2}\times\frac{\lambda}{N(N-1)}\times\frac{E[M]}{R}\\&=\frac{(N-2)\times\lambda_i\times E[M]}{2\times R}\end{aligned} \tag{5}$$

With this assumption, the average density $\rho_{\mathrm{B}i}$ can be viewed as the probability that MTU is full on the ring. The probability that a packet has to wait i more MTUs before it can be transmitted is $\rho_{\mathrm{B}i}^{i}(1-\rho_{\mathrm{B}i})$. The mean waiting time $E[d]$ to find an empty MTU can be expressed as

$$E[d]=\sum_{i=0}^{\infty} i\frac{L}{R}\rho_{\mathrm{B}i}^{i}(1-\rho_{\mathrm{B}i})=\frac{L\cdot\rho_{\mathrm{B}i}}{R(1-\rho_{\mathrm{B}i})} \tag{6}$$

The steady-state duration of all the whole vacation intervals V is equal to $\lambda_i \mathrm{TQ}\cdot E[d]$. Combining Equations (3) and (6), we obtain the average queuing delay

$$\begin{aligned}\mathrm{TQ}&=E[\alpha]+\lambda_i \mathrm{TQ}E[X]+\lambda_i \mathrm{TQ}E[d]\\&=\frac{L}{2\cdot R}+\lambda_i \mathrm{TQ}E[X]+\lambda_i \mathrm{TQ}\frac{L\cdot\rho_{\mathrm{B}i}}{R\cdot(1-\rho_{\mathrm{B}i})}\end{aligned} \tag{7}$$

which can be reduced to

$$\mathrm{TQ}=\frac{L}{2\cdot R\cdot\left(1-\lambda_i E[X]-\lambda_i\dfrac{L\cdot\rho_{\mathrm{B}i}}{R\cdot(1-\rho_{\mathrm{B}i})}\right)} \tag{8}$$

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

$$D=\mathrm{TQ}+S+\tau' \tag{9}$$

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

$$\begin{aligned}S&=E[X]+E[d]\\&=E[X]+\frac{L\cdot\rho_{\mathrm{B}i}}{R\cdot(1-\rho_{\mathrm{B}i})}\end{aligned} \tag{10}$$

DOI: 10.1002/dac

Thus, the average transfer delay is given by

$$D = \mathrm{TQ} + S + \tau/2 \tag{11}$$

3.4. Analysis of the multi-ring (WDM ring) case

In order to analyze multiple WDM ring networks, it is assumed that the bridge traffic load from the upstream source is equally distributed among W rings. To simplify the analysis, let the circulation of slots on W rings be synchronized [12, 13]. That is, a node can observe W MTUs on different rings at the same time. Since the bridge traffic load from the upstream source is uniformly distributed among the W rings, the average bridge traffic load of each ring, ρ_{B}, can be expressed as

$$\rho_{\mathrm{B}} = \rho_{\mathrm{B}i}/W \tag{12}$$

The probability that the packet at the head of a queue cannot get an empty MTU among the currently passing W MTUs is $(\rho_{\mathrm{B}})^W$. Therefore, the probability that the packet has to wait i MTUs before it can be sent out is $(\rho_{\mathrm{B}})^{W\cdot i}(1-(\rho_{\mathrm{B}})^W)$.

As in Section 3.3, let $E[d_{\mathrm{B}}]$ be the average time required to find the arrival of an empty MTU. Then we have

$$E[d_{\mathrm{B}}] = \sum_{i=0}^{\infty} i\frac{L}{R}(\rho_{\mathrm{B}})^{W\cdot i}(1-(\rho_{\mathrm{B}})^W) = \frac{L\cdot(\rho_{\mathrm{B}})^W}{R\cdot(1-(\rho_{\mathrm{B}})^W)} \tag{13}$$

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

$$\mathrm{TQ} = E[\alpha] + \lambda_i \mathrm{TQ} E[X] + \lambda_i \mathrm{TQ} E[d_{\mathrm{B}}] \tag{14}$$

Therefore, we have

$$\mathrm{TQ} = \frac{E[\alpha]}{1-\lambda_i E[X] - \lambda_i E[d_{\mathrm{B}}]} \tag{15}$$

The average transmission delay is

$$\begin{aligned} S &= E[X] + E[d_{\mathrm{B}}] \\ &= E[X] + \frac{L\cdot(\rho_{\mathrm{B}})^W}{R\cdot(1-(\rho_{\mathrm{B}})^W)} \end{aligned} \tag{16}$$

Thus, the average transfer delay is given by

$$D = \mathrm{TQ} + S + \tau/2 \tag{17}$$

4. SIMULATIONS AND RESULTS

The simulation model is shown in Figure 9. Four processes were used in the model: traffic generator, Rx, INS, and CHK. Traffic generator is responsible for generating IP packets based

DOI: 10.1002/dac

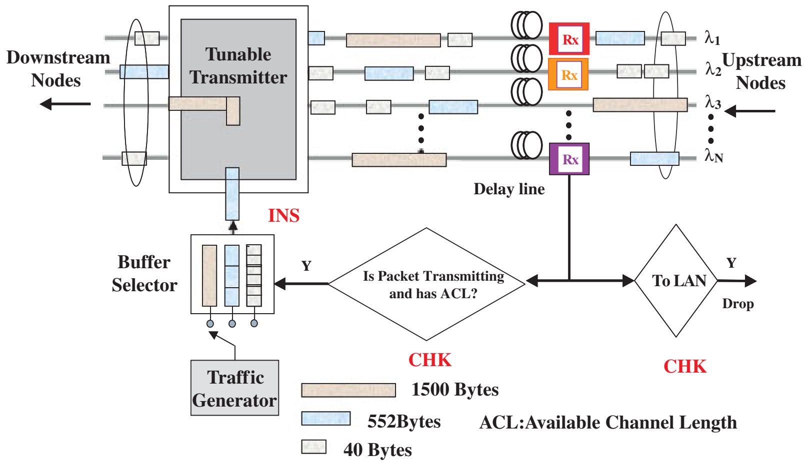

Figure 9. The simulation model.

Table I. The simulation parameters.

Number of nodes (N)	16
Number of channels (W)	8, 4, 2, 1
Ring network length	50 km
Channel speed	10 Gbps (OC-192)
Size of the delay line	1500 bytes
IP packet size	OC-3 Traffic: 40–1500 bytes
Average packet size	512 bytes

on 40 bytes = 50%, 41–551 bytes = 15%, 552 bytes = 17%, 553–1500 bytes = 3%, and 1500 bytes = 15% to generate the traffic load in Figure 1. INS is responsible for coordinating the transmission of packet sizes in the transmission queue and the shift of packets from the delay line. Rx is the receiver process that receives packets and checks ACL. CHK is responsible for checking the destination address of incoming packets. To simulate the delay line and the input fiber link of nodes under the condition of multi-channels, two sets of N queues are used.

The simulation experiments are based on the codes by SIMSCRIPT II and are replicated corresponding to using the variance reduction technique with different sequences for pseudo-random numbers. The results were obtained at a 95% confidence level.

The parameters of the network are shown in Table I. Figure 10 presents the simulated and analytical results of the average packet transfer delay in this network. The curves demonstrate that a high node offer load can be achieved with low transfer delay when the number of channels is large. The agreement between the simulation results and the analytical results is excellent. Figure 11 plots the packet transmission delay *versus* the offered load per node. PCQ_quota has the least delay, followed by PCQ and FIFO. We observe that the delay characteristics of the designed scheduling algorithms are better than those of the Hornet (TT-FR architecture) protocol [7] when the number

DOI: 10.1002/dac

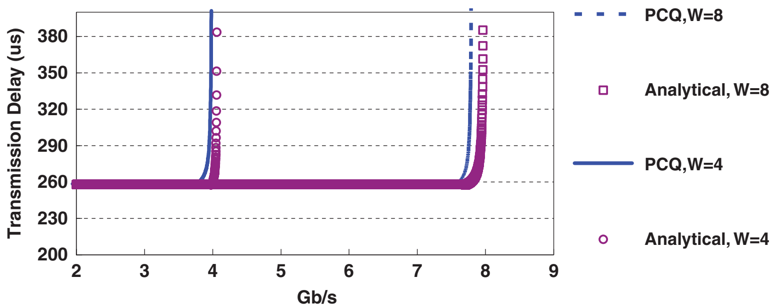

Figure 10. Average transfer delay *versus* offered load per node, when the number of channels equals 4 and 8.

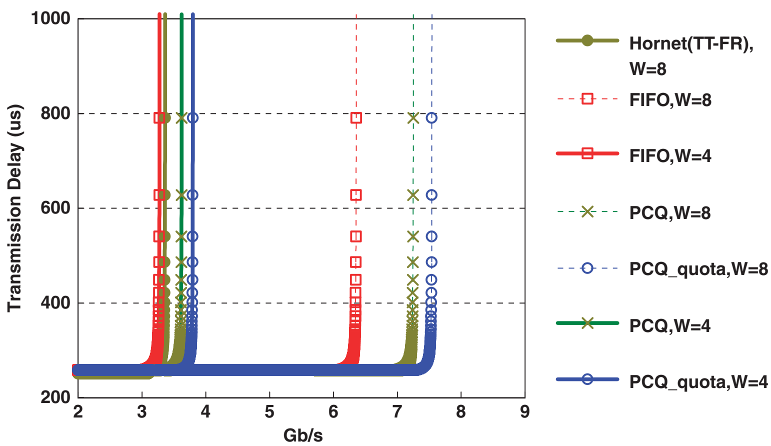

Figure 11. Average transfer delay *versus* offered load per node of transmission algorithms, when the number of channels equals 4 and 8.

of channels is equal to 8. Figures 12 and 13 show the delay fairness index and throughput fairness index [14] of the three transmission algorithms *versus* the offered load per node. Let us now compare the three TX-queues in terms of fairness. We define the throughput fairness index of a node i as c^2 (second moment) of the throughput from node i to all other nodes:

Throughput fairness index of node i is

$$f_p = \sum_{i=1}^{N} \sum_{j=1}^{q} \frac{(P_{ij} - \overline{P})^2}{\overline{P}^2} \tag{18}$$

where P_{ij} is the throughput of queue j in node i, i.e. the average number of bits transmitted by node i in queue j in a unit time, and $\overline{P} = \sum_{i=1}^{N} \sum_{j=1}^{q} P_{ij}/N \times q$. According to this definition, the smaller the throughput fairness index of a TX-queue, the better the throughput fairness. Here we use the individual quotas of Q1 quota (=7416 bytes), Q2 quota (=3050 bytes), and Q3 quota

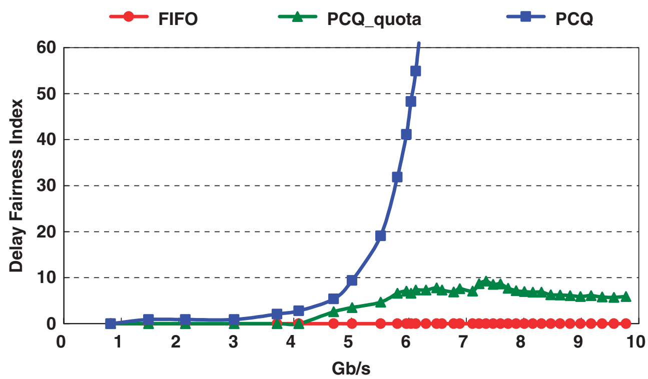

Figure 12. Delay fairness index of the three transmission algorithms *versus* offered load per node.

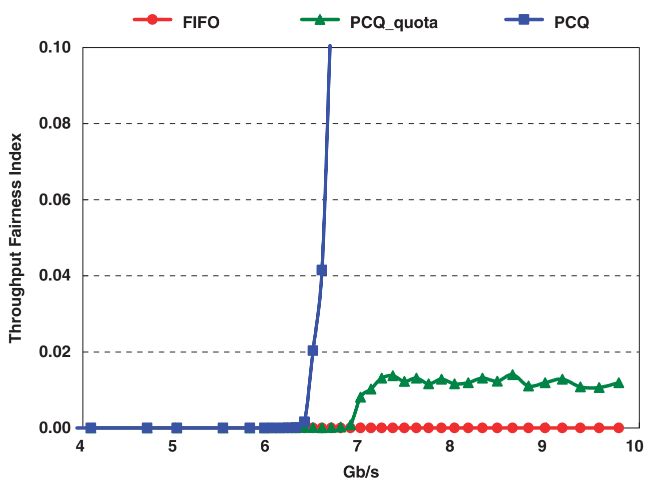

Figure 13. Throughput fairness index of the three transmission algorithms *versus* offered load per node.

(=590 bytes) in the PCQ_quota transmission algorithm. We observe that FIFO and PCQ_quota have delay and throughput fairness index values very close to zero, meaning that they are fair transmission algorithms in terms of delay and throughput. We observe that PCQ_quota has better delay and throughput fairness index values than those of the PCQ transmission algorithm. Without a quota scheme, it is difficult to find the suitable idle spaces for Q_1 (553–1500 bytes) to become a candidate packet because the total packet number of Q_1 is less than Q_2 (41–512 bytes) and Q_3 (40 bytes) under heavy load. On the other hand, using the quota scheme, there is a good chance that Q_2 and Q_3 have already filled their quotas so that idle spaces cannot be used. The packet in Q_1 can then be sent out in the suitable idle spaces. This is why the quota scheme improves bandwidth utilization and lowers transmission delays.

DOI: 10.1002/dac

5. CONCLUSION

In summary, in this paper we proposed a novel MAC protocol for all-optical WDM ring networks. The protocol supports the transmission of IP packets directly over WDM from LAN to MAN. How to merge and collapse the middle layers between IP and WDM for the next generation optical LANs/MANs was investigated. This protocol can avoid packet collision, reuse wavelength, and lacks a fragment packet scheme. In the verification, the simulated results closely match the analytical values, demonstrating the performance of the network. The throughput characteristic of the network is almost proportional to the number of channels in the network. With regard to the utilization of bandwidth of all the optical ring networks, our protocol displays the excellent characteristics of high throughput, low delay, and a good fairness index for all-optical communications.

REFERENCES

1. Yamada Y, Nakagawa SL, Kurosawa Y, Kawazawa T, Taga H, Goto K. 2 Tbit/s (200 × 10 Gbit/s) over 9240 km transmission experiment with 0.15 nm channel spacing using VSB format. *IEEE Electronics Letters* 2002; **38**(7):328–330.
2. Shimojoh N, Naito T, Tanaka T, Nakamoto H, Ueki T, Sugiyama A, Torii K, Suyama M. 2.4-Tbit/s WDM transmission over 7400 km using all Raman amplifier repeaters with 74-nm continuous single band. *Twenty-seventh European Conference on Optical Communication*, *ECOC'01*, Amsterdam, Netherlands, vol. 6, September 2001; 8–9.
3. Bissessur H, Charlet G, Idler W, Simonneau C, Borne S, Pierre L, Dischler R, De Barros C, Tran P. 3.2 Tbit/s (80/spl times/40 Gbit/s) phase-shaped binary transmission over 3/spl times/100 km with 0.8 bit/s/Hz efficiency. *Electronics Letters* 2002; **38**(8):377–379.
4. Kartalopoulos SV. Elastic bandwidth [optical-fiber communication]. *IEEE Circuits and Devices Magazine* 2002; **18**(1):8–13.
5. Ghani N, Dixit S, Wang TS. On IP-over-WDM Integration. *IEEE Communication Magazine* 2000; **38**(3):72–84.
6. Cai J, Fumagalli A, Chlamtac I. The multitoken interarrival time (MTIT) access protocol for supporting variable size packets over WDM ring network. *IEEE Journal on Selected Areas in Communication* 2000; **18**(10): 2094–2104.
7. Shrikhande KV *et al.* HORNET: a packet-over-WDM multiple access metropolitan area ring network. *IEEE Journal on Selected Areas in Communication* 2000; **18**(10):2004–2016.
8. Xu L, Perros HG, Rouskas GN. Access protocols for optical burst-switched ring networks. *Journal of Information Sciences* 2003; **149**(1–3):75–81.
9. Hwang W-S, Wang W-F, Wang J-Y, Li C-C. A carrier preemption access control protocol for supporting IP packets over WDM ring networks. *International Symposium on Communications* (*ISCOM*), Tainan, Taiwan, November 2001.
10. Hui R, Zhu B, Huang R, Allen CT, Demarest KR, Richards D. Subcarrier multiplexing for high-speed optical transmission. *IEEE Lightwave Technology* 2002; **20**(3):417–427.
11. Ghafir HM. Performance analysis of a multiple-access ring network. *IEEE Transactions on Communications* 1993; **41**(10):1494–1506.
12. Kang CS. A broadband ring network: multichannel optical slotted ring. *Computer Network and ISDN Systems* 1995; **27**(9):1387–1398.
13. Bhuyan LN. Approximate analysis of single and multiple ring networks. *IEEE Transactions on Computers* 1989; **38**(7):1027–1040.
14. Xu L, Perros HG, Rouskas GN. A simulation study of optical burst switching and access protocols for WDM ring networks. *Computer Networks* 2003; **41**:143–160.

DOI: 10.1002/dac

AUTHORS' BIOGRAPHIES

Jih-Hsin Ho received his BS degree in computer science and information engineering from Tatung University, Taipei, Taiwan, in 1993, and the MS and PhD degrees in Electrical Engineering from National Cheng Kung University, Taiwan, in 1998 and 2007, respectively. He is currently an assistant professor teaching at the Department of Computer Science and Information Engineering, Diwan University, Tainan, Taiwan. His current research interests include performance evaluation, WDM networks, Internet QoS.

Wen-Ping Chen received the BS degree in electrical engineering from National Taiwan Institute of Technology, Taiwan in 1992, MS degree in electrical engineering from National Sun Yat-sen University, Taiwan in 2000, and the PhD researching in computer network from National Kaohsiung University of Applied Sciences, in 2003. From 1993 to 2003, he worked for National Kaohsiung University of Applied Sciences as a teaching assistant in the department of electrical engineering.

Wen-Shyang Hwang received BS, MS and PhD degrees in Electrical Engineering from National Cheng Kung University, Taiwan, in 1984, 1990 and 1996, respectively. He is currently a professor of Electrical Engineering, and the chairman of department of computer science and information engineering in National Kaohsiung University of Applied Sciences, Taiwan, Republic of China. His current research interests are in the fields of storage area networks, WDM Metro-ring networks, performance evaluation, multimedia wireless communications, software design on embedded system, Internet QoS and Internet applications.

Ce-Kuen Shieh is currently a professor teaching at the Department of Electrical Engineering, National Cheng Kung University. He received his BS, MS and PhD degrees from Electrical Engineering Department of National Cheng Kung University, Tainan, Taiwan. His current research interests include distributed and parallel processing systems, computer networking and operating systems.

DOI: 10.1002/dac