Design and Performance Analysis of a Contention-Based Reservation Protocol for a Local Area Optical Internet

Jin Seek Choi, Jang Won Lee, and Minho Kang

In this paper, we propose and analyze a new multiple access protocol for a local area optical Internet based on a wavelength division multiplexing technique which uses a passive star coupler. The proposed contention-based reservation protocol can support variable-length as well as fixed-length messages for transporting Internet packets with one reservation of a minislot at the beginning of a packet transmission. The minislot is used to reserve the data channel on the basis of the slotted ALOHA protocol and the control node ensures subsequent message transmission on the same wavelength. Thus, all messages need not be broken down to many fixed-length packets, and consecutive messages are transmitted through the same wavelength. Moreover, the proposed protocol reduces the collision probability of minislots and improves wavelength utilization. We determine the maximum throughput and verify the results with simulation.

I. INTRODUCTION

The advancement of semiconductors and optical technologies in the past decade has led to the evolution of communication networks. Particularly, optical fiber is replacing metallic cable in telecommunication networks because of its well known superior characteristics, such as large bandwidth, very low error rates, and low-cost [1]-[3]. The large bandwidth of optical fiber can be used by a set of parallel channels, each operating at different wavelengths. This is called wavelength division multiplexing (WDM) technology.

IP routers have the ability to directly access optical networks by eliminating the protocol overhead between the IP and the optical networks, thus forming an optical Internet. An optical Internet handles a super frame as an aggregate of several short IP packets at the network edges, to be transmitted as a whole [5]. An optical Internet can transmit a fully optical data path without electric conversion and throughputs in the hundreds of terabits/second range. The optical Internet can support almost all services which have different transfer capabilities [4] (Fig. 1).

For implementing an optical Internet, there are two major approaches: broadcast-and-select and wavelength routing networks [2]. Wavelength routing networks generally prevail among wide area networks because of their good scalability and wavelength re-usability [6], [7]. However, current optical switching technologies are still expensive for this type of network [1], [3]. On the other hand, a broadcast-and-select network is advantageous in a local or limited area due to the advantages of its simplicity and multicasting capability [8], [9]. The broadcast-and-select network is especially attractive in a

---

Manuscript received May 16, 2001; revised Dec. 24, 2001.
This work was supported in part by the ICU and OIRC-ERC program funded by KOSEF.
Jin Seek Choi (phone : +82 42 866 6144, e-mail: jin@icu.ac.kr) is with School of Engineering, Information and Communications University, Daejeon, Korea.
Jang Won Lee (phone : +1 765 494 3519, e-mail: lee46@ecn.purdue. edu) is with the Electrical and Computer Engineering Department, Purdue University, West Lafayette, IN, USA.
Minho Kang (e-mail: mhkang@icu.ac.kr) is with School of Engineering, Information and Communications University, Daejeon, Korea.

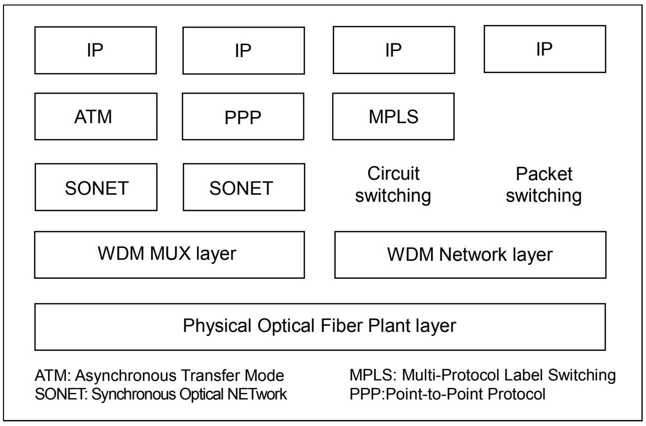

Fig. 1. WDM overlay data protocol models.

local area environment where all nodes are connected to a single broadcast facility (e.g., a passive star coupler).

A multiple access control (MAC) protocol is important in developing a local area optical Internet because it can transmit packets over broadcast-and-select WDM networks [2]. Until now, many MAC protocols have been proposed [10]-[15]. Depending on the usage of the tunable transmitter and the tunable receiver, most MAC protocols must deal with a large number of stations on ultra high speed channels; this makes reservation-based MAC protocols especially attractive. A separate channel (called a control channel) is also used in a number of protocols for transmitting control packets because a node has to wait for its control packet to return before sending a data packet; this is called pre-transmission coordination. It uses a set of distributed channel controllers. While most protocols can support only fixed-length data packet transmission, Internet packets tend to be widely variable in length. Few protocols have been proposed for variable-length packets, and protocols for fixed-length packets lead to a large transmission overhead for transmitting variable-length packets, such as warm-up time, tuning time, and wavelength conversion time [16]-[23].

Jia and Mukherjee proposed a scheduling algorithm based on a time-division multiple access reservation protocol [14]. It can reduce the impact of tuning overhead by tuning the transmitter immediately following the transmission of a control packet. However, the scheduling protocol has to maintain global information. Maintaining such global information has the serious drawback of a large warm-up time in the network operation for adding or deleting nodes. Lee and Un proposed a synchronous reservation protocol which can overcome the need for global information [15]. In this scheme, however, the limited use of minislots can increase control channel conflict, lead to a large overhead for control channel use, and make the system unstable when the input traffic is bursty (e.g., mixed with fixed- and variable-length messages). Moreover, every message has to be broken down to many fixed-length packets and each packet can be transmitted through different wavelengths. In addition, this protocol did not consider tuning overhead for consecutive message transmission over the different wavelengths [15]. Although Lee and Choi also proposed dynamic reservation protocols which slightly improve channel utilization by eliminating data channel conflict, they still have the same problems [18], [19].

In this paper, we propose a new MAC protocol for a broadcast-and-select network based on star topology. The design objective is to transmit variable length packets over the same data channel without packet fragmentation. The proposed protocol can use the same wavelength continuously while transmitting a long message. The protocol reduces the probability of control channel conflict and improves channel utilization. Furthermore, the protocol does not need global information.

The remainder of this paper is organized as follows. In section II, we describe a single-hop WDM local area network architecture and propose a new MAC protocol. In section III, we analyze the maximum throughput. In section IV, we verify the analytical results with simulation.

II. THE PROPOSED PROTOCOL

1. Network Architecture

The network architecture considered in this paper is a single-hop broadcast-and-select WDM Local Area Network (LAN) (Fig. 2). There are M+1 nodes with M access nodes and one control node. Every node connects with a passive star coupler of a star topology. The usable bandwidth is divided into N+1 wavelengths, $\lambda_0, \lambda_1, \lambda_2, \ldots, \lambda_N$. The wavelength λ_0 is dedicated to the transmission of the control information. The other N wavelengths are dedicated to the transmission of actual data traffic. In the design, we assume that the number of nodes would typically be larger than the number of available channels but we do not limit the number. In what follows, a wavelength defines a channel.

Each node has a transceiver with a fixed-transmitter, a fixed-receiver, a tunable-transmitter, and a tunable-receiver structure which is called an FT-FR-TT-TR transceiver. Fixed transceivers (transmitters and receivers) are locked onto the control channel λ_0, but tunable transceivers can tune to any wavelength λ_i within $i \in N$ data channels. Both transmitters and receivers can simultaneously send and receive information. Each node also has one register, R_RI (Register for Reservation Information). The register consists of N entries. Each entry includes the information of the source address, the destination address, and the message length for currently transmitting nodes. The entry is indexed by the wavelength i. At every slot,

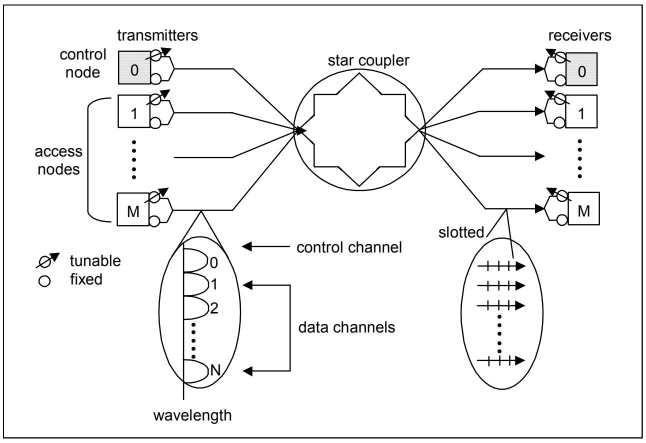

Fig. 2. A single-hop WDM local area network architecture with a passive star coupler.

wavelength data channel	source address	destination address	message length
1	Null	Null	Null
2	12	3	2
3	1	7	1
	:	:	:
N-1	Null	Null	Null
N	8	4	4

Register for reservation information

Fig. 3. The data structure of the register for reservation information (R_RI).

the values of each entry are updated on the basis of the reservation information. Figure 3 shows the data structure of the register.

Figure 4 shows a typical control and data channel structure over time. The first horizontal line represents the control channel λ_0. The other N lines represent the data channels. All optical links are synchronized at the beginning of each slot. Figure 5 shows a typical slot structure of the data channel. A data packet, called a message, may be composed of a sequence of data slots, $T_d=n\times T_X$, where n is an arbitrary integer. On the other hand, the control channel is time slotted with a fixed slot length T_X. Each slot contains a control frame which consists of X minislots with a length of T_m each and a reservation information region of length T_r. The number of minislots, X, is defined by

$$X = \left\lfloor \frac{T_X - T_r}{T_m} \right\rfloor \qquad (1)$$

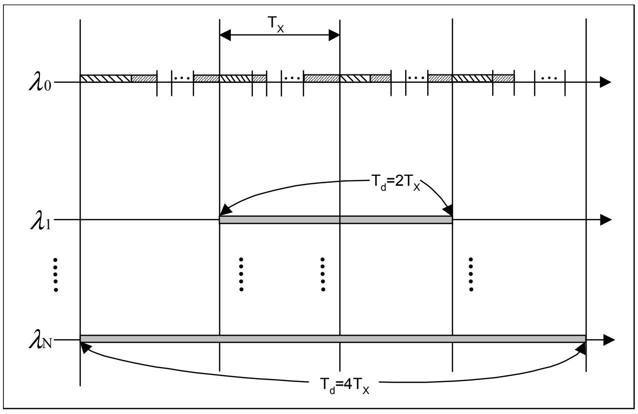

Fig. 4. Structure of control and data channels.

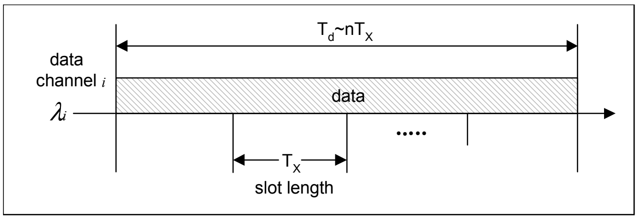

Fig. 5. Slot structure of a data channel.

where $\lfloor z \rfloor$ means the largest integer not greater than z. Each minislot contains a request packet which informs all nodes that the source node has a message to be sent to a destination node. The reservation information consists of N bits. Each bit indicates the reservation status of the corresponding data channel. Each bit is dedicated to a data channel. We assume that the size of the reservation information can be very small compared with the slot length. This is because the number of data channels is usually smaller than that of nodes. Figure 6 shows the detail of the slot structure of a control channel.

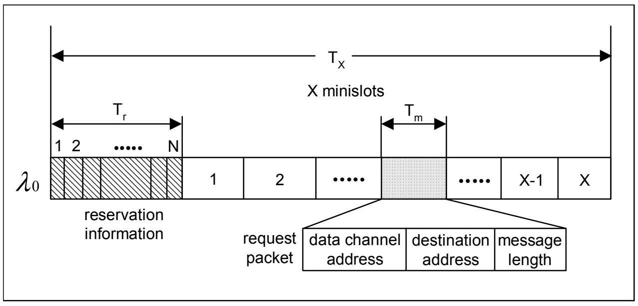

Fig. 6. Slot structure of a control channel.

2. Access Protocol Based on Reservation Information

The access protocol exploits a distributed assignment scheme for the access of data channels. Each node is assumed to be in one of two states: idle or backlogged. We define an idle node as a node having an empty buffer. A backlogged node is one with a message for transmission. The backlogged state is further divided into two sub-states: transmitting and contending. A transmitting node is a node that has a message and is currently transmitting its message through the data channel. A contending node has a message and is trying to reserve a data channel but is not currently transmitting its message. Figure 7 shows the state transition diagram.

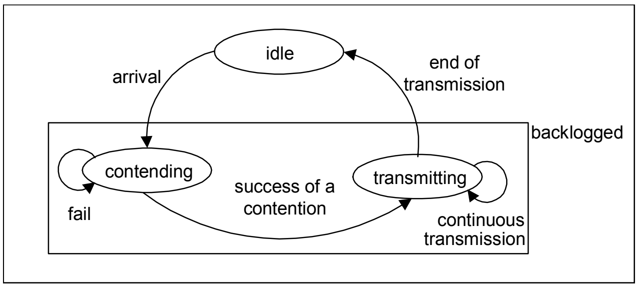

Fig. 7. State transition diagram of a node.

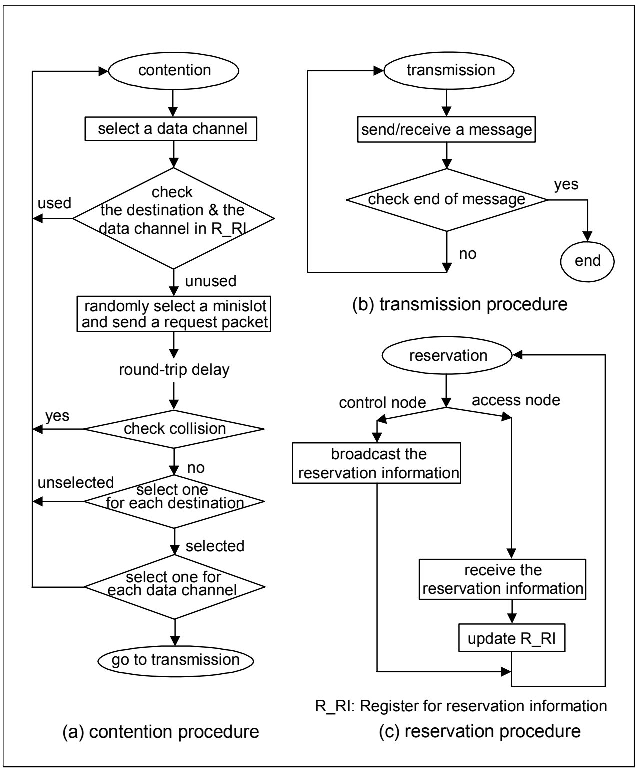

Fig. 8. Block diagram of contention-based reservation protocol.

Figure 8 shows the block diagram of the access protocol, which consists of three parts: contention, transmission, and reservation procedures. Every node acts according to its state. At the end of a slot, an idle node receives a new message and changes its state to contending. At the beginning of a slot, all contending nodes, including those that received a message in the previous slot, start the contention procedure as follows. First, the node randomly selects a data channel, checks whether it is in use, and checks the destination address in the register R_RI. Here, the destination address is included in an incoming message. The R_RI includes the source and destination addresses as well as the data channel numbers for currently transmitting nodes. Once the data channel or the destination address is registered in the R_RI, the node aborts the contention procedure since it may conflict with the continuously transmitting nodes. Only when the data channel and the destination address are not registered in the R_RI, does the node continue the contention procedure. Thus, the node preferentially excludes the conflicts of the data channels and the destinations that are already reserved by the transmitting nodes.

Next, the node randomly selects a minislot and sends a request packet to the selected minislot on the basis of the slotted ALOHA protocol. After a delay for the round-trip, the node receives the returned minislots and checks for collision of the request packets. If the request packet collides with other packets, then the node aborts the transmission and reschedules the transmission of the request packet at a later time chosen at random [26]. If the request packet does not collide with others, then the node selects one packet for each data channel and each destination among the request packets that are successfully returned without collision. In this case, the request packet is not blocked if it does not have the same data channel and destination address as those of earlier successful request packets and if the number of successful request packets within a slot does not exceed that of the data channels, *N*. Then, the node saves the results in the R_RI and changes its state to transmitting. Since every node receives all request packets, it is possible for every node to get and store the same result locally. They keep the result in the R_RI.

In the transmitting state, source/destination nodes start the transmission procedure as follows. They tune their tunable transmitters/receivers to the reserved data channel. All source/ destination pairs send/receive a message on the selected data channel at the beginning of the next slot. If the node has a remaining message, it continuously sends the message on the same data channel. Otherwise, the transmitting node quits the transmission and changes its state to idle at the next slot.

Since the control node receives all request packets, it is possible to get the same results at every slot. The control node also keeps track of the time at which each transmitter and receiver is tuned in. The control node broadcasts the results as reservation information to every slot (we call this the reservation procedure). Meanwhile, access nodes receive the reservation information as well as the request packets and keep track of the time at which transmitters, receivers, and channels will be free. Thus, access nodes infer the same results from the reservation information within a few slot times, even if they are newly added to the network.

Figure 9 shows an example of the protocol operation. This figure shows the time line that marks a message flow from left to right. The above two lines refer to the time lines on the transmitter side. The bottom two lines refer to the time lines on the receiver side. The boxes on the top of the figure show the contents of the R_RI register. In this example, we assume that M=20, N=5, X=5, and T_X=1. We also assume that wavelength 5 and destination node 2 are initially unused (i.e., wavelength 5 and destination node 2 are not marked in R_RI).

At the end of slot k-2, node 3 receives a message of length 4 which is destined to node 2. At the beginning of slot k-1, node 3 starts the contention procedure as follows. Node 3 checks destination address 2 and data channel number 5 in the R_RI. Since the data channel and the destination address are not registered in the R_RI, node 3 randomly selects the minislot of 2 and transmits a request packet to the minislot (#1 in Fig.9). The request packet indicates that node 3 has a message of length 4 destined to node 2 to be sent through data channel 5. In this contention, the node selects minislot 2. After a round-trip delay, node 3, node 2, and the control node successfully receive the returned request packets without collision, and they know that there is no earlier request packet with the same destination and that the same data channel and the number of successfully returned request packets does not exceed that of the data channels, N (#2). Each node saves the result in the R_RI (#3). At the beginning of slot k, node 3 starts the transmission procedure and continuously sends a message destined to node 2 through data channel 5 during the four consecutive slots (#4). At the beginning of slot k, the control node starts the reservation procedure as follows. The control node generates the control frame that consists of a set of blank minislots followed by reservation information. The control node broadcasts the reservation information marked at the corresponding 5th bit (#5). This indicates that node 3 has a remaining message greater than one slot. After a round-trip delay, each node receives the reservation information and updates the message length, subtracting one at the fifth row in R_RI (#6). The above reservation procedure is repeated at k+1 (#7 and #8). At k+2, the control node broadcasts the reservation information without marking the corresponding 5th bit (#9). After a round-trip delay, each node recognizes that node 3 has no more messages to be sent at the next slot k+3 and saves the result in the R_RI (#10). At k+3, node 3 sends the final portion of the remaining message on data channel 5 and then releases data channel 5 (#11). At the end of the slot k+3, node 3 changes its state into idle. Consequently, node 3 sends a message on data channel 5 during the consecutive four slots and the message can be sent on the same data channel 5 without packet fragmentation. This is accomplished with one reservation of minislot 2 at $k-1$. At the end of slot k+3, each node resets the source address, the destination address, and the message length in the fifth row in the R_RI.

The access protocol is independent of the tuning time. The only difference is the time shift of the data transmission as much as the tuning time. Figure 10 shows an example of the same access operation mentioned above when the tuning time is one slot time. Since the tuning time may be from a few nanoseconds to a few milliseconds depending on the device technology, the assumption could be reasonable.

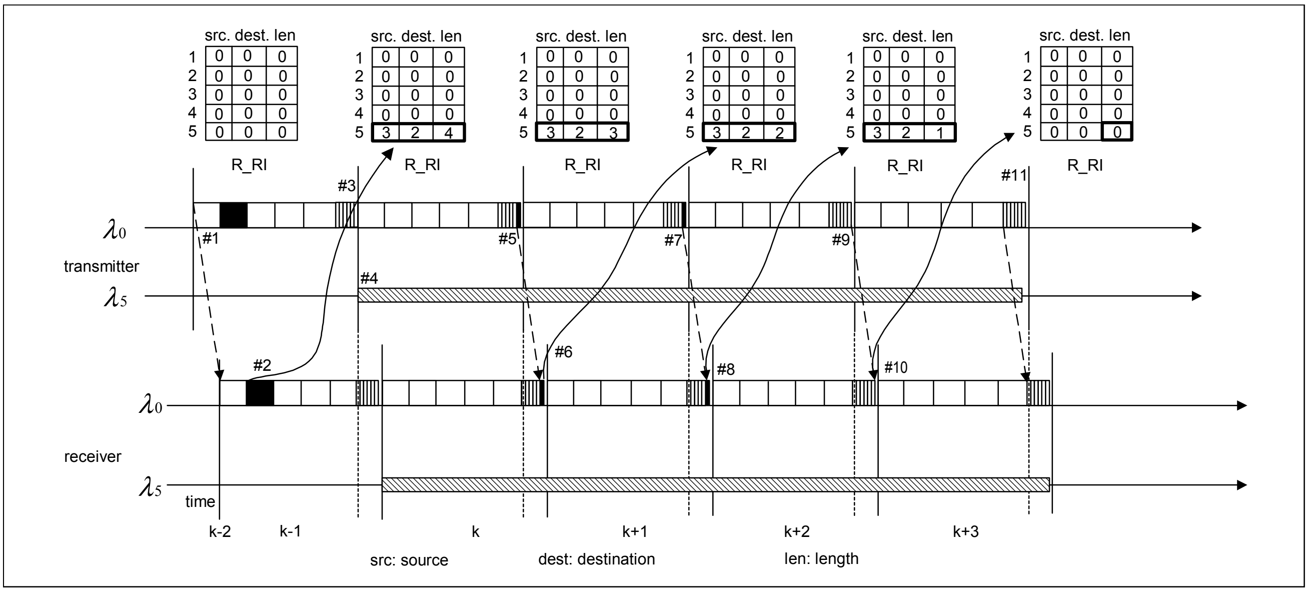

Fig. 9. An example of the protocol operation for a message transmission.

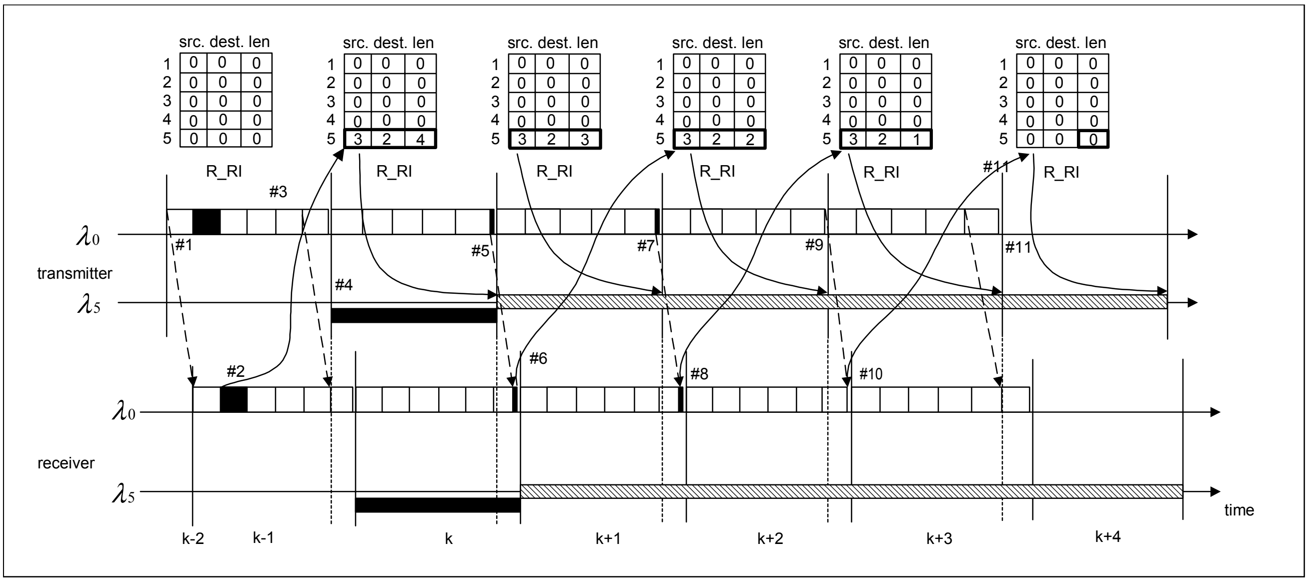

Fig. 10. An example of the protocol operation for a message transmission when the tuning time is one slot time.

3. Node Initialization Protocol Based on Minislots

When a new node is added to the network, it begins a node initialization procedure until it gets the synchronization from the R_RI. The synchronization state is accomplished by interpreting the minislots on the control channel as follows. At first, the node checks the minislots and updates the R_RI. The R_RI is repeatedly updated by the minislots that include which nodes start to send a packet through specified data channels. When the reservation information corresponds to the R_RI, the node knows the source and the destination address pairs of all the data channels and sets its state as the synchronized state. Then, the node performs the initialization protocol and sends messages through the proposed access protocol. Thus, a node easily joins or leaves the network at anytime without maintaining global information on connection histories. Moreover, the initialization procedure will be accomplished within a finite time since there is no node sending infinite-length messages.

When the control node fails or the control node is disconnected from the network, the proposed protocol can be run as the conventional multi-channel slotted ALOHA protocol mentioned in [10]. In this case, the protocol cannot handle variable-length messages since the control node cannot broadcast the reservation information and the access nodes cannot use it. Moreover, the protocol can be run regardless of the propagation delay since the proposed protocol is a reservation-based protocol.

III. PERFORMANCE ANALYSIS

We analyze the system as follows. There are M nodes, N data channels, and X minislots in the system. The number of nodes is independent of the number of data channels or the number of minislots. We assume that $M \geq X \geq N$ [13], [15], [23]. In addition, message length is assumed to have a geometric distribution with an average of $1/L$. The basis of the message length is the length of a slot time, T_X. Each node is in one of three states: a transmitting state in which the node is transmitting a data packet, an idle state in which the node does not have a data packet to transmit, and a contending state in which the node is trying to reserve a data channel but is not transmitting at this slot. When a node is in either contending or transmitting, we call the state a backlogged state. Only the idle node can generate a message, and the destination of the message is assumed to be uniformly distributed among M nodes. According to the network diameter and data rate, the round-trip delay, τ, may have various values. In our analysis, the round-trip delay is assumed to be negligible compared to the slot length ($\tau << T_X$) [15]. Although tuning time depends on the device technology, we assume that it is included in the message format and is negligible compared with the slot length ($T_X >> T_t$). In our analysis, we follow the assumption of ($\tau << T_X$).

With the above assumptions, we analyze the proposed protocol using a discrete time Markov chain. First of all, let state variable (k_1,k_2) describe the state in which the number of nodes in the transmitting state is k_1 and the number of nodes in the backlogged state is k_2. To obtain the balanced equation of the state variable, we define some probabilities.

Let $PC(c|t,b)$ be the probability that c nodes among t nodes continue to transmit their messages in the next slot, given that the number of nodes in the transmitting state is t and the number of nodes in the backlogged state is b. Here, c is the number of continuously transmitting nodes and $t-c$ is the number of nodes finishing the transmission at the end of the slot. $t-c$ nodes change the state into idle. Since the length of a packet is geometrically distributed with parameter L, $PC(a|b,c)$ is given by [25]

$$PC(c|t,b) = {}_tC_c(1-L)^c L^{t-c} \text{ for } c \le t \le b. \tag{2}$$

Let $PA(a|c,t,b)$ be the probability that a nodes among ($M-b+t-c$) idle nodes generate a message to be sent. Here, c is the number of continuously transmitting nodes, t is the number of nodes in the transmitting state, and b is the number of nodes in the backlogged state. In the idle state, each node is assumed to generate a message according to the Bernoulli process with parameter λ. So, $PA(a|c,t,b)$ is given by [25]

$$PA(a|c,t,b) = {}_{M-b+t-c}C_a \lambda^a (1-\lambda)^{M-b+t-c-a}. \tag{3}$$

Also, let $PS(s|a,c,t,b)$ be the probability that s nodes among $a+b-t$ contending nodes successfully reserve data channels, given that there are a nodes with new arriving and $b-t$ failing contending nodes in the previous slot. To get $PS(s|a,c,t,b)$, we define the following two probabilities: $PX(g|a,c,t,b)$ and $PT(s|g,a,c,t,b)$. $PX(g|a,c,t,b)$ is the probability that g nodes among $a+b-t$ nodes have no conflict on the data channel and the destination address with those of the continuously transmitting nodes, given that c is the number of continuously transmitting nodes. Here, g is the number of nodes which can transmit a request packet in the minislot, and $a+b-t$ is the number of contending nodes. $PT(s|g,a,c,t,b)$ denotes the probability that s nodes among g nodes reserve the data channel successfully, given that there are c continuously transmitting nodes and there are $a+b-t$ contending nodes. This probability includes three blocking probabilities: the blocking probability caused by a collision of minislots, the blocking probability caused by the other request packets using the same data channel, and the blocking probability caused by the same destination.

To get $PX(g|a,c,t,b)$, we define the following two probabilities: $PD(r|a,c,t,b)$ and $PW(g|r,a,c,t,b)$. $PD(r|a,c,t,b)$ defines the probability that r nodes among $a+b-t$ nodes choose destinations which are not in conflict with that of the continuously transmitting nodes. Here, $a+b-t-r$ nodes are blocked because of a data channel conflict with the continuously transmitting nodes. $PW(g|r,a,c,t,b)$ is the probability that g nodes among r nodes have no data channel conflict with the continuously transmitting nodes, that is, $PW()$ means that g contending nodes choose unused data channels as well as unused destination addresses. They are given by

$$PD(r|a,c,t,b) = {}_{a+b-t}C_r \left(\frac{M-c}{M}\right)^r \left(1-\frac{M-c}{M}\right)^{a+b-t-c} \tag{4}$$
$$\text{for } g \le r \le a+b-t$$

and

$$PW(g|r,a,c,t,b) = {}_rC_g \left(\frac{N-c}{N}\right)^g \left(1-\frac{N-c}{N}\right)^{r-g} \tag{5}$$
$$\text{for } 0 \le r \le a+b-t$$

respectively. Using $PD(r|a,c,t,b)$ and $PW(g|r,a,c,t,b)$, we calculate the probability, $PX(g|a,c,t,b)$, that g nodes among $a+b-t$ contending nodes can transmit their reservation packets, given that c nodes are in a continuously transmitting state. It is given by

$$PX(g|a,c,t,b) = \sum_{r=g}^{a+b-t} PW(g|r,a,c,t,b) PD(r|a,c,t,b) \tag{6}$$

To get $PT(s|g,a,c,t,b)$, we define the following three probabilities: $Q(x,i,g)$, $Prd(m,j,i)$, and $Prw(n,l,i)$. $Q(x,i,g)$ denotes the probability that each of i minislots among x minislots is chosen by only one request packet, when g nodes are trying to choose a minislot to transmit a request packet (i.e., the probability that i nodes among g nodes can transmit their request packets successfully). Then,

$$\begin{aligned} Q(x,i,g) &= \frac{{}_xC_{ig}C_i i!}{xg} \\ &\times \sum_{v=0}^{\min(x-i,g-i)} (-1)^v {}_{x-i}C_{v\,g-i}C_v v!(x-i-v)^{g-i-v}. \end{aligned} \tag{7}$$

$Prd(m,j,i)$ denotes the probability that j nodes among i nodes can successfully reserve their destination nodes given that there are m available destinations. Then, we get

$$Prd(m,j,i) = \frac{{}_mC_{m-j}\sum_{y=o}^{j}(-1)^y {}_jC_y(j-y)^i}{m^j}. \tag{8}$$

Also, $Prw(n,l,j)$ denotes the probability that l nodes among j nodes can successfully reserve their data channels, given that there are n available data channels. Then, we get

$$Prw(n,l,j) = \frac{{}_nC_{n-l}\sum_{z=0}^{l}(-1)^z {}_lC_z(l-z)^j}{n^j}. \tag{9}$$

From $Q(,,)$, $Prd(,,)$ and $Prw(,,)$,we can get $PT(s|g,a,c,t,b)$ as

$$\begin{aligned} PT(s|g,a,c,t,b) &= \sum_{i=s}^{g} Q(x,i,g) \\ &\times \sum_{j=s}^{i} Prd(M-c,j,i) \\ &\times \sum_{l=s}^{j} Prw(N-c,l,j) \end{aligned} \tag{10}$$
$$\text{for } s \le g \le a+b-t,$$

where M is the number of destination nodes, X is the number of minislots and N is the number of data channels. Compared to the analysis in [18] in which the number of available slots is $M-c$, in our analysis, the number of available minislots is always M. From (6) and (10), we can obtain $PS(s|a,c,t,b)$ as follows:

$$PS(s|a,c,t,b)=\sum_{g=s}^{a+b-t} PT(s|g,a,c,t,b)\times PX(g|a,c,t,b). \quad (11)$$

Let us define π_{k_1,k_2} as the Markov state that the system is in, where k_1 is the number of successfully transmitting nodes and k_2 is the number of backlogged nodes in the system. And let $P_{(k_1,k_2),(u_1,u_2)}$ denote the state transition probability that the system in state (k_1,k_2) goes to state (u_1,u_2). Then, π_{k_1,k_2} is obtained from the two following balanced equations [27]

$$\Pi \leftarrow \Pi P \quad (12)$$

and

$$\sum_{k_1=0}^{N}\sum_{k_2=0}^{M-k_1} \pi_{k_1,k_2}=1, \quad (13)$$

where Π is the state matrix of $\{\pi_{k_1,k_2}\}$ and P is the transition matrix of $\{P_{(k_1,k_2),(u_1,u_2)}\}$. For $0 \le k_1 \le min(N, k_2)$, $0 \le k_2 \le M$, $0 \le u_1 \le min(N, u_2)$, $0 \le u_2 \le M$, the transition matrix becomes

$$\begin{aligned}P(k_1,k_2),(u_1,u_2) = &\sum_{k_3=0}^{\min(k_1,u_1)} PS(u_1-k_3|u_2-k_2+k_1-k_3,k_3,k_1,k_2)\\ &\times PA(u_2-k_2+k_1-k_3|k_3,k_1,k_2)\\ &\times PC(k_3|k_1,k_2).\end{aligned} \quad (14)$$

Therefore, the throughput of the network system, S_T, is obtained by

$$S_T=\sum_{k_1=1}^{N}\sum_{k_2=1}^{M-k_1} k_1\pi_{k_1,k_2}. \quad (15)$$

IV. PERFORMANCE RESULTS AND DISCUSSION

In this section, we show the analytical results for the maximum throughput and verify them with simulation. For simplicity, we assume the simulation model as follows. Several short packets will be gathered into a single data unit. The super frame should reduce the traffic self-similarity. This allows us to consider that the empty nodes can generate a new message with probability ρ using a simple random traffic model [5]. The message length is geometrically distributed with a mean value of $1/L$. The destination address is uniformly distributed among M destination nodes. The system is assumed to be operating at an equilibrium point. Each node contains a single buffer and belongs to one of two states [24]: the idle state in which the node does not have a data message to be sent, and the backlogged state in which the node has a message to be sent. Only when a node is in the idle state, can it generate a new message. When a reservation packet collides with other reservation packets in a time slot, it is retransmitted at the next time slot with probability 1. At that time, the addresses about destination, wavelength, and control minislots are newly assigned. The simulation is running for 10^6 slot times. The model does not count for back-off delay due to collisions. In the simulation, we assume that M=20, N=5, and X=5, 7 and 10. In the result, the offered load is defined as ρ/L. Throughput is defined as the average number of occupied data channels per time slot. In the figures that follow, the analysis results are represented by a solid line and the simulation results by a plus.

Figure 11 shows the throughput versus the offered load for various message lengths, L=0.1 − 1.0 with X=7. We see that the close match between the analytical and simulation results indicates that the analysis is adequate for obtaining the performance of the proposed protocol, in which there is significant relation between the message length and the maximum throughput. We observe that the maximum throughput proportionally increases as the average message length increases. This is because messages can be transmitted with one reservation of a minislot at the beginning of a message transmission. This reduces the probability of a minislot conflict, and channel utilization is significantly improved. Thus, the proposed protocol is more useful when the message length is longer.

We also observe that the throughput is stable under heavy loads, unlike that of L=1 when L<1. In case of L=1, the proposed protocol is the same as the contention-based reservation protocol using the slotted ALOHA protocol [17]. The throughput of a minislot becomes that of the slotted ALOHA protocol, Ge^{-G}, where G is the total number of retransmissions and new transmissions in a given slot and e is the base of the natural logarithm. Under heavy loads, S_T may be considerably reduced if the number of minislots is insufficient. That leads to a significant throughput degradation when $X<N\cdot e$ (the solid line in Fig. 11). On the other hand, in the case of L<1, the throughput can be maintained at a high level under heavy loads because the proposed protocol can reduce the probability of minislot conflict proportional to the message length. From the results, we observe that the proposed protocol is more useful when the

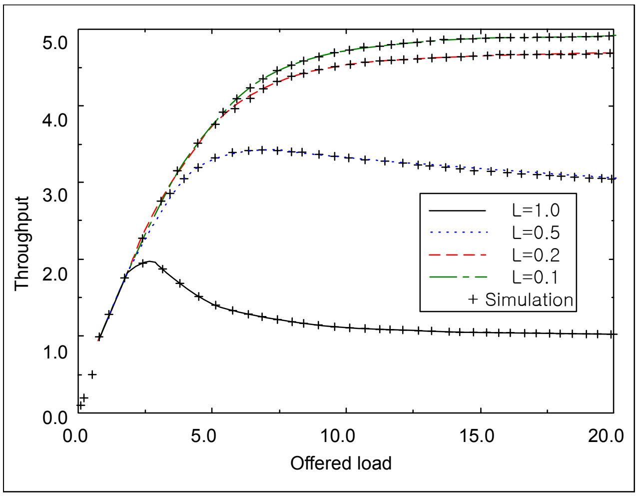

Fig. 11. Throughput for various message lengths.

message length is long and the number of minislots is large. For example, the number of minislots, X=7, is enough to maintain the maximum throughput under a heavy load when L=0.2.

Figure 12 shows the throughput against the offered load for various numbers of minislots, X=5, 7, and 10 at M=20, N=5 and L=0.2. The maximum throughput for X minislots is

$$S_T = min\lfloor N, X/e \rfloor. \tag{16}$$

As the offered load increases, the throughput will rapidly increase as more messages become available for transmission. The through put will eventually decrease when the number of nodes in the contending state begins to run over that of the nodes to be transmitted. When the value of X is controlled such that $N{\cdot}e \geq X$, the throughput can be increased up to N and cannot be decreased when the number of contending nodes runs over that of the nodes to be transmitted. The main cause of this improvement is that the probability of the minislot conflict is considerably reduced when the number of minislots becomes large.

Figure 13 plots the analytical results of the throughput for the proposed protocol and compares them to the result of Lee and Un's scheme when X =7 and 10 [15]. The result shows that our scheme outperforms Lee and Un's scheme in throughput, because the proposed protocol significantly reduces the probability of minislot conflicts by using one minislot to transmit a whole message. Even if we consider the overhead of the reservation information, the performance gain is still considerable. Moreover, the proposed protocol improves wavelength use by allowing consecutive message transmission from different sources.

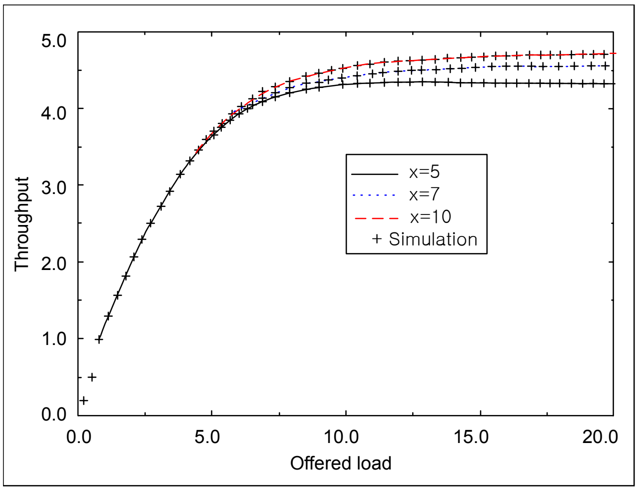

Fig. 12. Throughput for various numbers of minislots for M=20, N=5, and L=0.2.

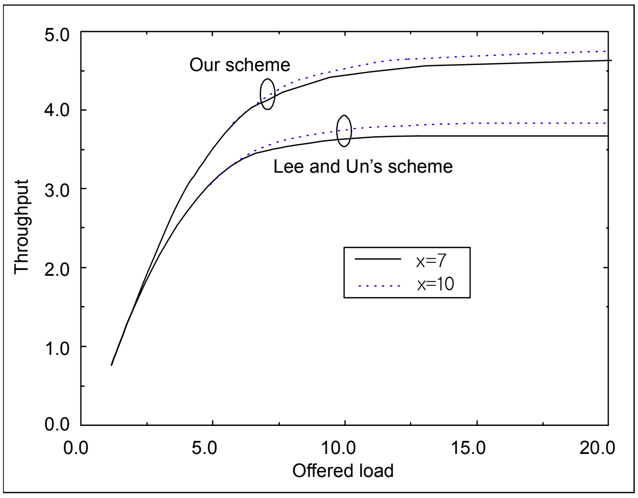

Fig. 13. Performance comparison at M=20, N=5, and L=0.2.

V. CONCLUSION

In this paper, we proposed a new contention-based reservation protocol to support variable-length messages. The proposed protocol uses reservation information as well as minislots. The minislot is used to reserve the data channel on the basis of the slotted ALOHA protocol. The reservation information generated by the control node ensures subsequent message transmission on the data channel without using the minislots. A whole message can be scheduled by reserving one minislot. Thus, the protocol significantly reduces the probability of minislot conflict and improves data channel utilization. Moreover, the protocol can use the same data channel while transmitting a long packet, so a packet need not be broken down to many fixed-length packets. The proposed protocol has little overhead since the reservation information has a form of bitmap address which is considerably shorter than the slot length. It is not necessary to maintain global information, so any node can join or leave the network at anytime without network re-initialization. Therefore, we conclude that the proposed protocol is suitable for a local area optical Internet which accommodates multimedia Internet traffic over a WDM network.

REFERENCES

[1] R. Ramaswami and K.N. Sivarajan, *Optical Networks a Practical Perspective*, Morgan Kaufmann, 1998.

[2] B. Mukherjee, *Optical Communication Networks*, McGraw-Hill, 1997.

[3] B.S. Ham, “A Novel Method of All-Optical Switching Quantum Router,” *ETRI J.*, vol. 23, no. 3, Sept. 2001.

[4] M. Veeraraghavan and M. Karol, “Using WDM Technology to Carry IP Traffic,” *34th Conf. on Information Sciences and System*, Princton, Mar. 2000.

[5] L. Tancevski, A. Ge, G. Castanon, and L.S. Tamil, “A New Scheduling Algorithm for Asynchronous, Variable Length IP Traffic Incorporating Void Filling,”*OFC’99*, San Diego, CA, Feb. 1999.

[6] J.Y. Wei, C.D. Liu, S.Y. Park, K.H. Liu, R.S. Ramamurthy, H.Kim, and M.W. Maeda, “Network Control and Management for the Next Generation Internet,” *IEICE Trans. Comm.*, vol. E83-B, no. 10, Oct. 2000.

[7] J.S. Choi, N. Glomie, F. Lapeyrere, F. Mouveaux, and D. Su, “A Classification of Routing and Wavelength Assignment Algorithm in WDM Networks,” *Proc. of OPNET’2000*, Paris, France, Feb. 2000.

[8] K. Shrikhande, A. Srivatsa, I.M. White, M.S. Rogge, D. Wonglumson, S.M. Gemelos, and L.G. Kazovksy,“CSMA/CD MAC Protocols for IP-HORNET: An IP over WDM Metropolitan Area Ring Network,” *Proc. of GLOBECOM'00*, Nov. 2000, pp. 1303-1307.

[9] E. Modiano, R. Barry, and E. Swanson, “A Novel Architecture and Medium Access Control (MAC) Protocol for WDM Networks,” *Proc.of OFC’98*, 1998, pp. 90-91.

[10] B. Mukherjee, “WDM-Based Local Lightwave Networks Part I: Single-Hop Systems,” *IEEE Network Mag.*, May 1992, pp. 12-27.

[11] I.M.I. Harbab, M. Kavehrad, and E.W. Sundberg, “Protocols for Very High-Speed Optical Fiber Local Area Networks Using a Passive Star Topology,” *IEEE J. Lightwave Technol.*, Dec. 1987, pp. 1782-1794.

[12] N. Bambos, J. Bannister, L.A. Bergman, J. Cong, E. Gafni, M. Gerla, L. Kleinrock, S. Monacos, P. Hu, B. Kannan, B. Kwan, P. Palnati, J. Peck, and S. Walton, *The Super Computer Supernet (SSN): A High-Speed Electro-Optic Campus and Metropolitan Network*, ISBN: 0-8194-2066-2, SPIE’96, 1996.

[13] H.B. Jeon and C.K. Un, “Contention-Based Reservation Protocols in Fiber Optic Local Area Networks with Passive Star Topology,” *Electronic Lett.*, vol. 26, no. 12, 1990, pp. 780-781.

[14] F. Jia, “The Receiver Collision Avoidance (RCA) Protocols for a Single-Hop WDM Lightwave Networks,” *Proc. of ICC*, 1992, pp. 6-12.

[15] J.H. Lee and C.K. Un, “Performance of a Synchronous Reservation Protocol for Variable-Sized Messages in a WDM-Based Local Networks,” *Proc. of GLOBECOM’95*, 1995, pp. 2210-2214.

[16] F. Jia and B. Mukherjee, “Variable-Length Message Scheduling Algorithms for a WDM Based Local Lightwave Network,” *Proc. of INFOCOM*, 1994, pp. 1362-1369.

[17] H.B. Jeon and C.K. Un, “Contention-Based Reservation Protocols in Multiwavelength Optical Networks with a Passive Star Topology,” *Proc. of ICC’92*, vol. 3, Nov. 1992, pp. 1473-1477.

[18] J.H. Lee and C.K. Un, “Dynamic Scheduling Protocol for Variable-Sized Messages in a WDM-Based Local Networks,” *IEEE J. Lightwave Technol.*, vol. 14, no. 7, July 1996, pp. 1595-1600.

[19] J.S. Choi and H.H. Lee, “A Dynamic Wavelength Allocation Scheme with Status Information for Fixed-and Variable-Length Message,” *GLOBECOM’98*, Australia, 1998.

[20] T. Strosslin and M. Ganaire, “A Flexible MAC Protocol for All-Optical WDM Metropolitan Area,” *Proc. of IPCCC’00*, 2000, pp. 567-572.

[21] M. Mishra, E.L. Johnson, and K.M. Sivalingam, “Scheduling in Optical WDM Networks Using Hidden Markov Chain-Based Traffic Predicators,” *Proc. of ICON’00*, 2000, pp. 380-384.

[22] M.W. Janoska and T.D. Todd, “A Simplified Optical Star Network Using Distributed Channel Controllers,” *IEEE J. Lightwave Technol.*, vol. 12, no. 11, Nov. 1994, pp. 2011-2022.

[23] J.S. Choi, J.K. Choi, and K.S. Lee, “Contention-Based Reservation Protocol in Multiwavelength Local Network with Fixed-and Variable-Length Message,” *ICT’97*, Australia, Apr. 1997.

[24] I. Chlamtac and A. Ganz, “Channel Allocation Protocols in Frequency-Time Controlled High Speed Networks,” *IEEE Trans. on Comm.*, vol. 36, Apr. 1988, pp. 430-440.

[25] W. Feller, *An Introduction to Probability Theory and Its Applications*, 3rd ed., Wiley, New York, 1968.

[26] J.L. Hammond and P.J.P. O'Reilly, *Performance Analysis of Local Computer Networks*, Addison Wesley, 1986.

[27] R.W. Wolff, *Stochastic Modelling and the Theory of Queues*, Prentice-Hall, International Ed., 1989.

[28] L. Wang and M. Hamdi, “Efficient Protocol for Multimedia Streams on WDMA Networks,” *Int’l Conf. on Information Networking (ICOIN12)*, 1998, pp. 241-246.

Jin Seek Choi has been working for School of Engineering at Information and Communications University since March 2001. He graduated from Sogang University with BS degree in electronic engineering in 1985. He received MS and PhD degrees in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), Korea, in 1987 and 1995, respectively. He has worked for National Institute of Science and Technology (NIST), Washington DC, USA as Visiting Researcher from September 1998 to August 2000. He has worked at Kongju National University from 1995 to 2001, where he was a faculty member in the Information and Communication Department. He had worked at LG Information and Communication Co. from 1987 to 1991 where he worked on the development of Etherent, FDDI bridge and ISDN systems. His current research interest includes QOS guaranteed optical Internet, and QOS guaranteed high-speed switching and routing protocols.

Jang Won Lee received his BS degree in electronic engineering from Yonsei University, Seoul, Korea, in 1994 and his MS degree in electrical engineering from KAIST (Korea Advanced Institute of Science and Technology), Daejeon, Korea, in 1996. He is currently working toward his PhD degree in the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA. His research interests include pricing and resource allocation issues in communication networks.

Minho Kang is Director of the Optical Internet Research Center at the Information and Communications University, Daejeon, Korea. He received BSEE, MSEE, and PhD degrees from Seoul National University, University of Missouri-Rolla, and The University of Texas at Austin in 1969, 1973, and 1977, respectively. From 1977 to 1978, he was with AT&T Bell Laboratories, Holmdel, NJ. During 1978 and 1989, he was a Department Head and a Vice President at Electronics and Telecommunications Research Institute. Also during 1985 and 1988, he was the Electrical and Electronics Research Coordinator at the Korean Ministry of Science and Technology. During 1990-1998, we was an Executive Vice President at Korea Telecom in charge of R&D, quality assurance, and overseas business development groups. In 1999, he joined the Information and Communications University, where he is a Professor of Engineering. Dr. Kang served as the Chairman at the Asia Pacific Telecommunity Study Group of Bangkok for 1996-1999 and the General Assembly of Korean Telecommunications Technology Association for 1995-1997. He is a member of National Academy of Engineering in Korea, a Vice Chairman of Optical Society of Korea, and a Co-Chairman of the first International Conference on Optical Internet 2002. He was awarded the Order of Merit-DongbaekJang by Korean Government and Grand Technology Medal by 21st Century Management Club in 1983 and 1991, respectively, for the contribution to optical communications development.