Hierarchical Dynamic Bandwidth Allocation Algorithm for Multimedia Services over Ethernet PONs

Kye-Hyun Ahn, Kyeong-Eun Han, and Young-Chon Kim

In this paper, we propose a new dynamic bandwidth allocation (DBA) algorithm for multimedia services over Ethernet PONs (passive optical networks). The proposed algorithm is composed of a low-level scheduler in the optical network unit (ONU) and a high-level scheduler in the optical line terminal (OLT). The hierarchical DBA algorithm can provide expansibility and efficient resource allocation in an Ethernet PON system in which the packet scheduler is separated from the queues. In the proposed DBA algorithm, the OLT allocates bandwidth to the ONUs in proportion to the weight associated with their class and queue length, while the ONU preferentially allocates its bandwidth to queues with a static priority order. The proposed algorithm provides an efficient resource utilization by reducing the unused remaining bandwidth caused by the variable length of the packets. We also define the service classes and present control message formats conforming to the multi-point control protocol (MPCP) over an Ethernet PON. In order to evaluate the performance, we designed an Ethernet PON system on the basis of IEEE 802.3ah "Ethernet in the first mile" (EFM) using OPNET and carried out simulations. The results are analyzed in terms of the channel utilization, queuing delay, and ratio of the unused remaining bandwidth.

Keywords: Dynamic bandwidth allocation algorithm, MAC protocol, Ethernet PON.

Manuscript received Dec. 9, 2003; revised Feb. 19, 2004.

Kye-Hyun Ahn (phone: +82 42 860 1861, email: khyun@etri.re.kr) is with Broadband Convergence Network Research Division, ETRI, Daejeon, Korea.

Kyeong-Eun Han (email: spring32@dreamwiz.com) and Young-Chon Kim (email: yckim@moak.chonbuk.ac.kr) are with the Department of Computer Engineering, Chonbuk National University, Jeonju, Korea.

I. Introduction

While access networks have experienced little change in recent years, long-haul networks have changed dramatically due to the emergence of wavelength division multiplexing technologies. Local area networks have scaled up in speed from 10 to 100 Mbps, and subsequently to 1 Gbps. Even 10 Gbps Ethernet products have started to emerge. As a result, there is a growing gulf between these high-capacity LANs and backbone networks, with the low-speed access networks constituting the bottleneck. There is an urgent need for a new technology for access networks which would provide the necessary improvements, one which is inexpensive, simple, scalable, and capable of delivering integrated voice, data, and video services to subscribers [1]-[7]. The passive optical network (PON) is one such technology and is considered an attractive solution for access networks [1]-[11].

There are two prevalent data-link layer protocols for the PON architecture, namely the asynchronous transfer mode (ATM) PON and Ethernet PON. The ATM PON was initially developed and defined as a PON-based optical access network that uses ATM as its layer-2 protocol. There are a number of proposals which allow for ATM PONs to share upstream bandwidth [8]-[11]. However, the expense and complexity associated with the ATM PON have caused it to be of declining interest to industry [1], [2].

An Ethernet PON is a PON that carries all the data encapsulated in Ethernet frames and is backward compatible with existing IEEE 802.3 Ethernet standards as well as other relevant IEEE 802 standards [12]. Since Ethernet is an inexpensive technology that is ubiquitous and interoperable with a variety of legacy equipment, it represents the perfect choice for delivering IP packets over PONs and for supporting multimedia traffic efficiently [1]-[6]. The IEEE 802.3ah "Ethernet in the First Mile (EFM)" Task Force is carrying out the standardization of Ethernet PON as a solution for access networks [12].

The Ethernet PON should be able to deliver multiple services such as voice communications, high-definition video, video conferencing, real-time transactions, and data traffic. To support these applications with their diverse requirements, Ethernet PONs need to have class-of-services mechanisms built in [3]-[7]. However, several features of an Ethernet PON don't allow for easy adaptation of the existing algorithms. These features include the distributed nature of the scheduling domain, with the queues and the scheduler being located at a large distance from each other, a limited control-plane bandwidth, and a significant queue switchover overhead [3], [4].

In this paper, we propose a hierarchical dynamic bandwidth allocation (DBA) algorithm to provide multiple services and to resolve the problems associated with the existing allocation algorithms in an Ethernet PON. We adopt a hierarchical scheme in order to solve the scalability issues by eliminating the need for separate GATE and REPORT messages to each queue. Our scheme also solves the switchover overhead issues, which arise from the fact that all of the queues in one optical network unit (ONU) are served consecutively, with no guard times between their transmissions. Additionally, the proposed DBA algorithm uses a proportional allocation algorithm applying differential weights to service queues for the optical line terminal's (OLT's) scheduler and a preferential allocation based on a strict priority for the ONU's scheduler. Repeatedly performing a proportional rule results in a fragmentation of resources, with each fragment potentially having insufficient bandwidth to transmit a single Ethernet packet. Since the occurrence of such fragments is quite frequent, the resource utilization is decreased. In order to implement the DBA algorithm, we classify the different services into four priority categories and present the control message formats in the multi-point control protocol (MPCP) used by Ethernet PONs.

This paper is organized as follows. Section II presents the control message formats and the dynamic bandwidth allocation procedure based on the MPCP protocol. Section III describes the proposed DBA algorithms. Section IV presents the OPNET simulation model environment and the results obtained from this simulation under various traffic environments. Finally, section V concludes the paper.

II. Multi-Point Control Protocol

A PON is a high bandwidth point-to-multipoint optical fiber network. PONs generally consist of an OLT, which is connected to optical network units (ONUs) using only fiber cables, optical splitters, and other passive components. The OLT is located at a local exchange, and the ONU is located either on the street, in a building, or even in a user's home [1]-[12].

As shown in Fig. 1, in the downstream direction (from OLT to ONUs), the traffic goes from one point to multiple points. In the upstream direction (from ONUs to OLT), the traffic from multiple points are directed toward a single point (OLT). Because Ethernet is by nature a broadcast-based protocol, in the downstream direction it fits perfectly with the PON architecture: packets are broadcast by the OLT and are extracted by their destination ONU based on their media access control (MAC) address. In the upstream direction, however, the ONUs should share the channel capacity and resources [1]-[12]. Therefore, a MAC protocol which is able to allocate the bandwidth dynamically and to guarantee fairness between the ONUs is needed [1]-[7].

In the IEEE 802.3ah EFM, the MPCP is actively being discussed as a suitable MAC protocol for Ethernet PONs. The MPCP is being developed by this task force in order to support

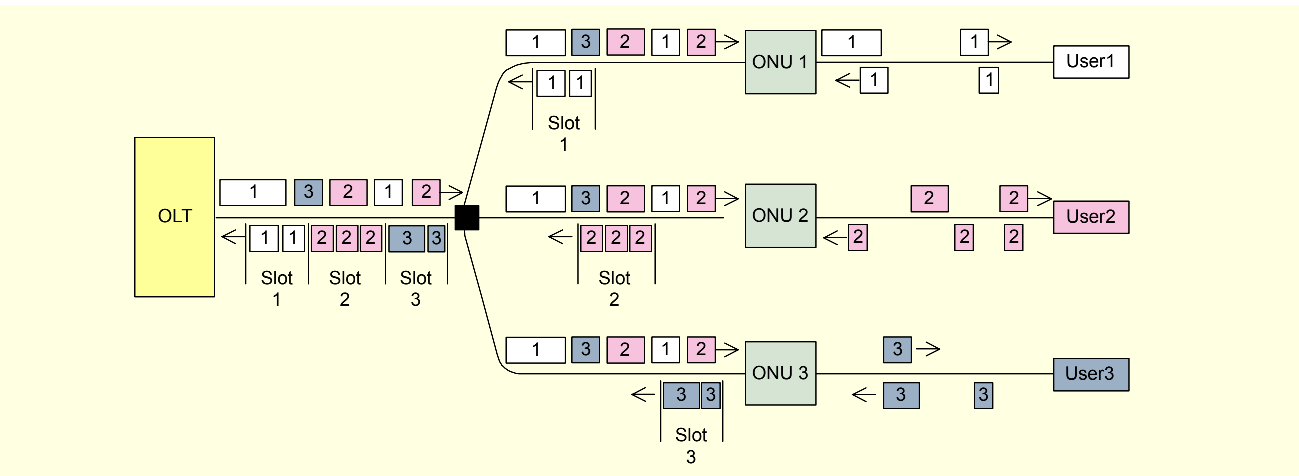

Fig. 1. Packet transmission in an Ethernet PON [3], [4].

timeslot allocation by the OLT [12]. The MPCP does not concern itself with a particular bandwidth-allocation scheme, but rather is a supporting mechanism that can facilitate the implementation of various bandwidth allocation algorithms in an Ethernet PON.

The procedure for bandwidth allocation relies on two control messages: GATE and REPORT. Both GATE and REPORT messages are MAC control frames (type 88-08) and are processed by the MAC control sub-layer [12].

The purpose of the GATE message is to grant transmission windows to ONUs for both discovery messages and normal transmission. Up to four grants can be included in a single gate message. The number of grants can also be set to zero, in order for the GATE message to be used as an MPCP keep-alive mechanism from the OLT to the ONU. The first three bits in the *number of grants* field, *b*0, *b*1, *and b*2, contain the number of grants and are composed of the valid length and start time pairs in this message, which is a number between 0 and 4. The *b*3 bit is used as the Discovery flag. The Force Report flag bits from *b*4 to *b*7 are used to ask the ONU to issue a REPORT message related to the corresponding grant number at the corresponding transmission opportunity indicated in this GATE.

The GATE and REPORT message formats are shown in Figs. 2 and 3, respectively [12].

In the REPORT messages, the ONUs indicate the upstream bandwidth they need per 802.1Q priority queue. REPORT messages are also used as keep-alives from the ONU to the OLT. The *number of queue sets* field specifies the number of requests in the REPORT message. A REPORT message may hold multiple sets of *report bitmap* and *queue #n* fields, as specified in the *number of queue sets* field. The report bitmap is an 8-bit bitfield flag register that indicates which queues are represented in this REPORT message.

Field (b0 … b7)	Octets
Destination address	6
Source address	6
Length/Type = 88-08	2
Opcode = 00-02	2
Timestamp	4
Number of grants/Flags	1
Grant #1 start time	0/4
Grant #1 length	0/2
Grant #2 start time	0/4
Grant #2 length	0/2
Grant #3 start time	0/4
Grant #3 length	0/2
Grant #4 start time	0/4
Grant #4 length	0/2
Sync time	0/2
Pad/Reserved	13–39
FCS	4

Fig. 2. GATE message format.

Field (b0 … b7)	Octets
Destination address	6
Source address	6
Length/Type = 88-08	2
Opcode = 00-03	2
Timestamp	4
Number of queue sets	1
Report bitmap	1
Queue #0 report	0/2
Queue #1 report	0/2
Queue #2 report	0/2
Queue #3 report	0/2
Queue #4 report	0/2
Queue #5 report	0/2
Queue #6 report	0/2
Queue #7 report	0/2
Pad/Reserved	0–39
FCS	4

Repeated n times as Indicated by number of queue sets (Report bitmap through Queue #7 report)

Fig. 3. REPORT message format.

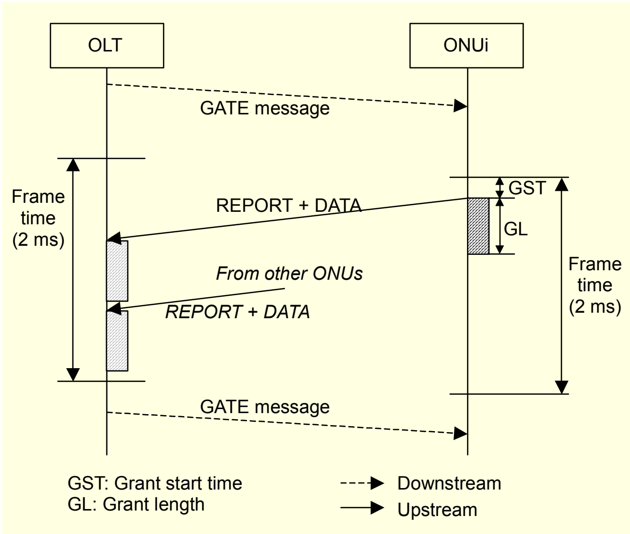

Fig. 4. Timing diagram of GATE and REPORT messages.

A timing diagram of the GATE and REPORT messages between the OLT and ONUs is shown in Fig. 4.

We assume that the OLT and ONUs are already synchronized through a ranging procedure. Synchronization is making data consistent with data that is distributed in various devices [14]. The ONU*i* detects its grant information from the GATE message through a downlink. As the grant information, the ONU*i* transmits its REPORT message and data packets without collisions through an uplink channel. The OLT collects report information from all ONUs and then performs a dynamic bandwidth allocation algorithm considering the reports. The results are broadcasted for the ONUs through a GATE message. The request information with the REPORT message sent in frame *i* is actually reflected in the uplink access in frame *i*+2.

III. Proposed Algorithm

1. Classification of Service Classes

The performance of a packet-based network can be conveniently characterized by several parameters: bandwidth, packet delay, delay variation, and packet-loss ratio. Quality of service refers to a network's ability to provide bounds on some or all of these parameters on a per-connection basis. Not all networks, however, can maintain a per-connection state or even identify connections. To support diverse application requirements, networks separate all traffic into a limited number of classes and provide differentiated services for each class. Such networks are said to maintain classes of service (CoS) [3], [4].

In order to support CoS, an Ethernet PON must classify traffic into CoS and provide differentiated treatment for each class. IEEE 802.1p distinguishes and illustrates the seven traffic classes as shown in Table 1 [3], [4], [13].

Table 1. Mapping of traffic classes into priority queues [13].

Number of queues	Traffic types queue assignments
1	{*Best effort*, Excellent effort, Background, Voice, Controlled load, Video, Network control}
2	{*Best effort*, Excellent effort, Background}, {*Voice*, Controlled load, Video, Network control}
3	{*Best effort*, Excellent effort, Background}, {*Controlled load*, Video}, {*Voice*, Network control}
4	{*Background*}, {*Best effort*, Excellent effort}, {*Controlled load*, Video}, {*Voice*, Network control}
5	{*Background*}, {*Best effort*, Excellent effort}, {*Controlled load*}, {*Video*}, {*Voice*, Network control}
6	{*Background*}, {*Best effort*}, {*Excellent effort*}, {*Controlled load*}, {*Video*}, {*Voice*, Network control}
7	{*Background*}, {*Best effort*}, {*Excellent effort*}, {*Controlled load*}, {*Video*}, {*Voice*}, {*Network control*}

In this paper, we consider that each ONU is equipped with four queues serving four priority classes, denoted *P*0, *P*1, *P*2, and *P*3. First, the highest priority service, the *P*0 class, supports applications that require a bounded end-to-end delay and generate a constant bit rate, e.g. voice and network control services. The other three classes support traffic with a variable bit rate. The *P*1 class serves applications that have a variable bit rate and require low delay and a guaranteed bandwidth such as real-time video services. The *P*2 class is used for applications that are not delay sensitive but that require bandwidth guarantees. The lowest priority service, the *P*3 class, is not sensitive to end-to-end delay or jitter. Under a very heavy network load this queue may receive no service. This queue will only be an allocated bandwidth if surplus bandwidth is available.

2. Hierarchical Scheduling

The Ethernet PON needs to have class-of-service mechanisms built in. However, several features unique to this type of network don't allow for the easy adaptation of existing algorithms to the Ethernet PON. These features include the distributed nature of the scheduling domain, with the queues and scheduler being located at a large distance from each other, a limited control-plane bandwidth, and a significant queue switchover overhead [3], [4].

We can consider two cases. These cases are introduced in [4]. The first case is where the scheduler is located solely in the OLT. In this case, the ONUs do not have any scheduler themselves. We call this flat scheduling, and a schematic diagram is provided in Fig. 5.

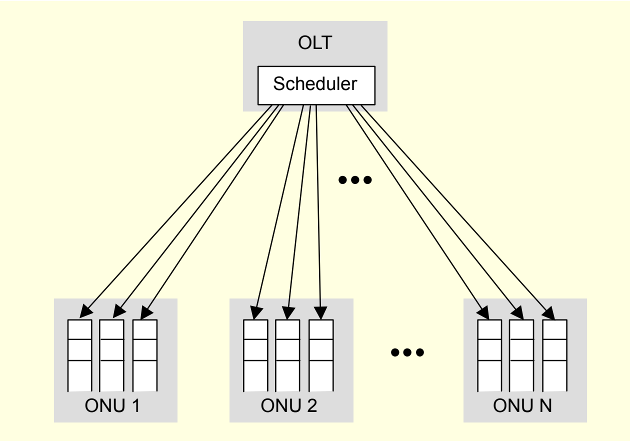

Fig. 5. Flat scheduling.

This sole scheduler with the overall information allocates bandwidth to each queue fairly and efficiently. However, flat scheduling cannot satisfy the system scalability and efficiency requirement of an Ethernet PON system. An Ethernet PON is a distributed system, so that the propagation delay between the queues and the scheduler often exceeds the transmission time of a packet. This is because the distance between the OLT and ONUs in an Ethernet PON system is quite long, with the maximum distance being 20 km. Therefore, in the case of multiple queues, a separate GATE message needs to be sent to each one. Also, a guard time is requested between packets transmitted from other queues located in another ONU within the Ethernet PON.

As [4] shows, if the average Ethernet packet size is 500 bytes, the Ethernet PON line rate is 1 Gbps, and the guard time is 1 ms, the overhead will be about 20 percent. However, for the smallest packet sizes (64 bytes), this overhead can reach 66% [4].

The other case is referred to as hierarchical scheduling. In this case, the schedulers are located at the ONUs as well as at the OLT. The high-level scheduler, situated at the OLT, allocates bandwidth to the ONUs, while the low-level scheduler, situated at the ONUs, distributes the bandwidth allocated by the OLT among its own queues. The GATE messages between the OLT and the ONUs would grant an aggregated bandwidth per ONU. Therefore, a hierarchical scheme solves the scalability issues by eliminating the need for a separate GATE in each queue. It also solves the switchover overhead issues which arise from the fact that all of the queues in one ONU are served consecutively with no guard times between their transmissions [4]. Figure 6 shows a hierarchical scheduling model.

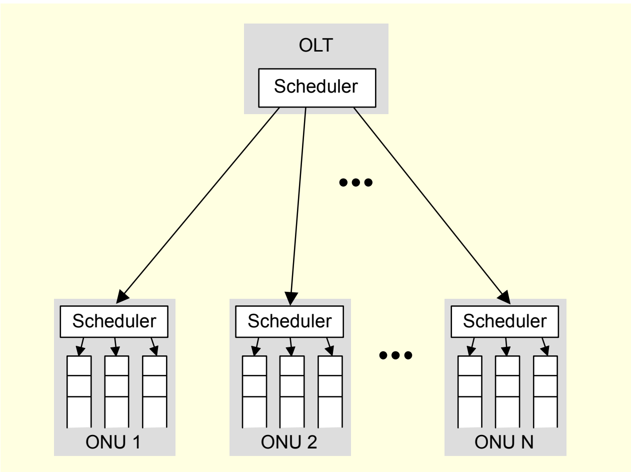

Fig. 6. Hierarchical scheduling.

3. Hierarchical DBA Algorithm

In this section, we describe the proposed algorithm in detail. If a packet is received from a user, the ONU identifies its class and places it in the corresponding queue. Between slots, an ONU stores all the packets received from the user in their respective queues. When a slot arrives, the ONU transmits a REPORT message, which contains multiple queue lengths. Thus, the high-level scheduler in the OLT determines the queue length per class for each ONU. Considering the overall information about the network, the high-level scheduler determines the amount of bandwidth to allocate to each queue. GATE messages are then sent to inform each ONU of the aggregated bandwidth which they have been allocated. At each ONU, the low-level scheduler distributes the allocated aggregated bandwidth among the priority queues.

A. Algorithm for OLT's Scheduler

First of all, we describe the DBA algorithm for the high-level scheduler in the OLT. It adopts a proportional method. The queue length of the *Pm* class of ONU_i is denoted as QL_m^i . This is obtained from the REPORT message, and the length indicates the amount of requirement transmission capacity for a frame time. In this paper, the queue length and bandwidth are a frame time scale. Let BW_{total} be the total transmission capacity within a frame time and $BW_{G,m}^i$ be the guaranteed capacity per frame time for the *Pm* class of ONU_i.

Step 1. Bandwidth guarantees for *P*0 class

The high-level scheduler begins by allocating the guaranteed transmission capacity for the *P*0 classes. Let BW_{fixed}^i be the capacity per frame time for the *P*0 class of ONU_i. The predetermined amount of bandwidth is assigned regardless of whether or not there are frames to send. Although it may waste resources, it provides a fast and reliable transmission service.

$$BW_{alloc}^i = BW_{fixed}^i \tag{1}$$

When this step is implemented, BW_{fixed}^i deserves careful attention. Consider the following problem for example.

There is a *P*0 service which has a 4-Mbps guaranteed bandwidth. The fixed bandwidth with a frame of 2 ms for the service is about 1,000 bytes. So this protocol allocates about 1,000 bytes for every frame. If the service has a constant data arrival rate, then the protocol can provide a successful transmission for data packets which are constant in size. However, if the service temporarily has a packet larger than 1000 bytes, this queue is blocked forever. Therefore, if there is a possibility of this problem occurring, we need to make a restriction that the minimum guaranteed bandwidth of a service cannot be less than 6.07 Mbps.

Step 2. Bandwidth guarantees for *P*1and *P*2 classes

The scheduler also allocates the guaranteed bandwidths to the *P*1 and *P*2 classes. However, the total guaranteed bandwidth for each priority queue can not exceed the total bandwidth. If the queue length is smaller than the guaranteed bandwidth, then the amount of bandwidth allocated is limited to the queue length. The available bandwidth for the following steps will therefore be equal to the total bandwidth minus the sum of the bandwidths allocated to each ONU, as in (3).

$$BW_{alloc}^i = BW_{alloc}^i + \min(QL_1^i, BW_{G,1}^i) + \min(QL_2^i, BW_{G,2}^i) \tag{2}$$

$$BW_{avail} = BW_{total} - \sum_i BW_{alloc}^i \tag{3}$$

Step 3. Dynamic allocation for $P1$, $P2$ and $P3$ classes

After dealing with the issue of the guaranteed bandwidths, the scheduler performs dynamic allocation by considering the queue lengths and priorities of the different classes. To do this, the requested bandwidth has to be calculated as follows:

$$QL_m^i = \max(0, QL_m^i - BW_{G,m}^i), \tag{4}$$

where the class number m can be 1, 2, or 3, and $BW_{G,3}^i$ is zero. $BW_{G,1}^i$ and $BW_{G,2}^i$ are used for the purpose of subtracting the bandwidth already allocated from the queue length before performing step 3.

The scheduler can allocate as much resources as requested, as long as the available bandwidth is not smaller than the sum of the required queue lengths. The aggregated allocated bandwidth for ONU_i is determined as shown in (5)

$$BW_{alloc}^i = BW_{alloc}^i + QL_1^i + QL_2^i + QL_3^i. \tag{5}$$

Otherwise, it proportionally allocates bandwidth to the ONUs based on the queue length and the weights of the classes. Let W_m be the weight for the Pm class. We assume that the W_1, W_2, W_3 are 3, 2 and 1, respectively. The amount of requested bandwidth for ONU_i is reevaluated as shown in (6). The QL_{req}^i can reflect differential priorities for classes as well as the queue length.

$$QL_{req}^i = \sum_{m=1,2,3} (QL_m^i \times W_m), \tag{6}$$

$$QL_{total} = \sum_i QL_{req}^i \ . \tag{7}$$

The aggregated allocated bandwidth for ONU_i is determined as follows:

$$BW_{alloc}^i = BW_{alloc}^i + \left(\frac{QL_{req}^i}{QL_{total}} \right) \times BW_{avail} \tag{8}$$

B. Algorithm for the ONU's Scheduler

While an ONU transmits the length of each of its queues, the OLT only responds by sending the aggregated bandwidth through the GATE message. The low-level scheduler decides which packets are to be transmitted among the various priority queues within the assigned slot. The scheduling algorithm at the ONU, which is shown in Fig. 7, can be similar to that used in the OLT except that the scheduling object consists of multiple service queues.

Steps 1 and 2, which are used to manage the guaranteed bandwidth, remain unchanged but are performed after the processing at the OLT scheduler. We assume that ONU_i is allocated the amount of bandwidth BW_{alloc}^i by the OLT. A queue for the Pm class of ONU_i can serve its data packets during the BW_m^i. The BW_m^i is determined as shown in the following algorithms.

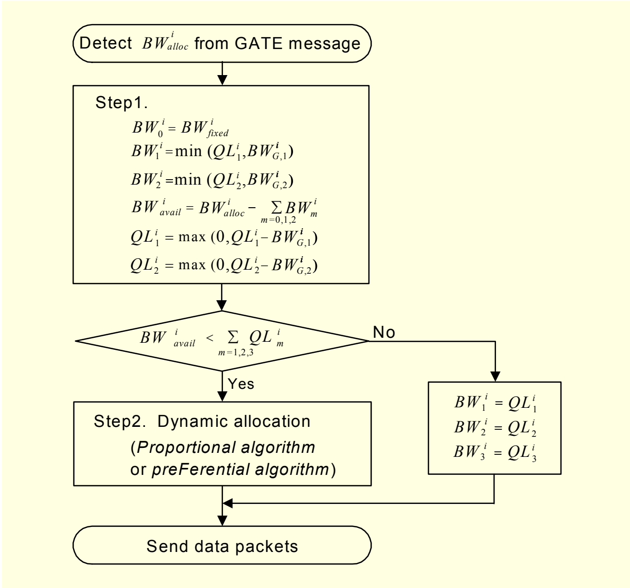

Fig. 7. Algorithm for ONU_i's scheduler.

Step 1. Bandwidth guarantees

First, the algorithm assigns the promised bandwidths for the $P0$, $P1$, and $P2$ classes.

$$\begin{aligned} &BW_0^i = BW_{fixed}^i, \ BW_1^i = \min(QL_1^i, BW_{G,1}^i), \\ &BW_2^i = \min(QL_2^i, BW_{G,2}^i). \end{aligned} \tag{9}$$

The available bandwidth (BW_{avail}^i) is determined as follows:

$$BW_{avail}^i = BW_{alloc}^i - BW_{fixed}^i - \min(QL_1^i, BW_{G,1}^i) - \min(QL_2^i, BW_{G,2}^i),$$

where QL is not the current queue length as determined by the newly arrived packets, but rather the queue length communicated to the OLT through the REPORT contained in the previous frame. The previously assigned bandwidths are subtracted from QL_1^i and QL_2^i, respectively.

$$QL_1^i = \max(0, QL_1^i - BW_{G,1}^i), \quad QL_2^i = \max(0, QL_2^i - BW_{G,2}^i)$$

Step 2. Dynamic allocation for $P1$, $P2$ and $P3$ classes

Next, the proposed algorithm determines whether or not the available bandwidth is equal to the sum of the queue lengths. If it is, then each of the queues will have a reasonable chance to transmit its own packets. Otherwise, the DBA algorithm has to distribute the available resources among the various queues. In this case, we consider two algorithms, proportional-proportional (P-P) and proportional-preferential (P-F). The difference between them is the allocation algorithm for ONU's scheduler.

■ Proportional-proportional algorithm

The ONU's scheduler performs proportional allocation by considering the queue lengths and priorities of the different classes such as the OLT's scheduler.

The proportional allocation algorithm considers the length and order of the priority of each queue, and proportionally allocates its assigned slot among the different queues. Each queue gets a portion of the slot and serves several packets. The portions are determined as follows:

$$BW_m^i = BW_m^i + \frac{(QL_m^i \times W_m)}{\sum_{m=1,2,3}(QL_m^i \times W_m)} \times BW_{avail}^i. \qquad (10)$$

Since an Ethernet packet has a variable size, there may be a remainder bandwidth. Figure 8 is a good example to illustrate the remainder.

This example does not involve any guaranteed bandwidth. The second queue has three packets whose sizes are 10, 7, and 8, and the queue length (QL_2^i) is 25. According to the result of the scheduling algorithm, the queue receives 20 resources and can serve two packets whose sizes are 10 and 7. Since there are insufficient resources, the last packet cannot be transmitted, so a remainder occurs.

Each queue has its own remainder which is not used for packet transmission. However, if they were to be combined to form a single fragment with a continuous bandwidth, then these remainders could be useful. Only a single remainder occurs per ONU.

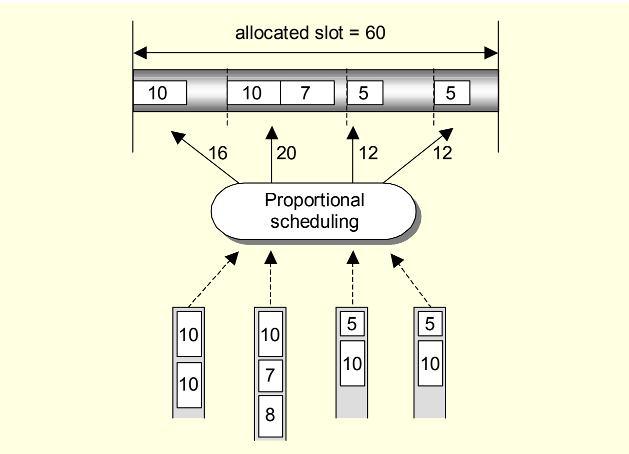

Fig. 8. An example using proportional scheduling.

In order to solve this remainder problem, we suggest that a preferential scheduling algorithm, rather than a proportional scheme, be used at the low-level scheduler.

■ Proportional-preferential algorithm

This algorithm allocates bandwidth to queues in order of priority. A higher priority queue that contains packets will be serviced before a lower priority queue. Thus, a portion of class 1 is first determined by performing (11) and (12) with m=1. The QL_m^i is the queue length included in the last REPORT message. It is updated when an ONU sends a REPORT message. Although the queue received many new packets after the last REPORT was sent, the value of QL_m^i is not affected.

$$BW_m^i = BW_m^i + \min(BW_{avail}^i, QL_m^i) \qquad (11)$$

$$BW_{avail}^i = \max(0, BW_{avail}^i - QL_m^i) \qquad (12)$$

If there is any available bandwidth, the scheduler allocates resources for the next class, and this procedure repeatedly executes until the available bandwidth has dried up, or the allocation for the lowest class has finished.

With the same environment as that shown in Fig. 8, this algorithm assigns data packets, as shown in Fig. 9. There is no remainder in this example, but a single remainder per ONU may occur because of the variable packet size and proportional scheduling at the OLT.

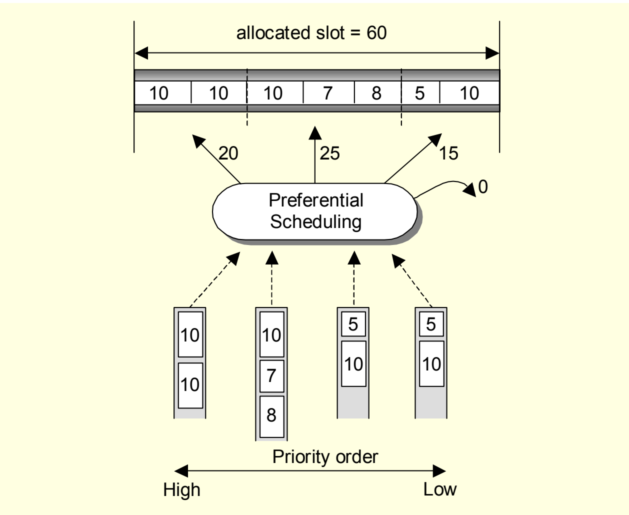

Fig. 9. An example using preferential scheduling.

Figure 10 shows a conceptual model of the proposed hierarchical DBA algorithm. The OLT's scheduler uses a proportional scheduling, and the ONU's scheduler uses a preferential scheduling. Since the OLT allocates an aggregate bandwidth among the ONUs, it is reasonable to adopt a proportional scheduling considering weights and queue lengths.

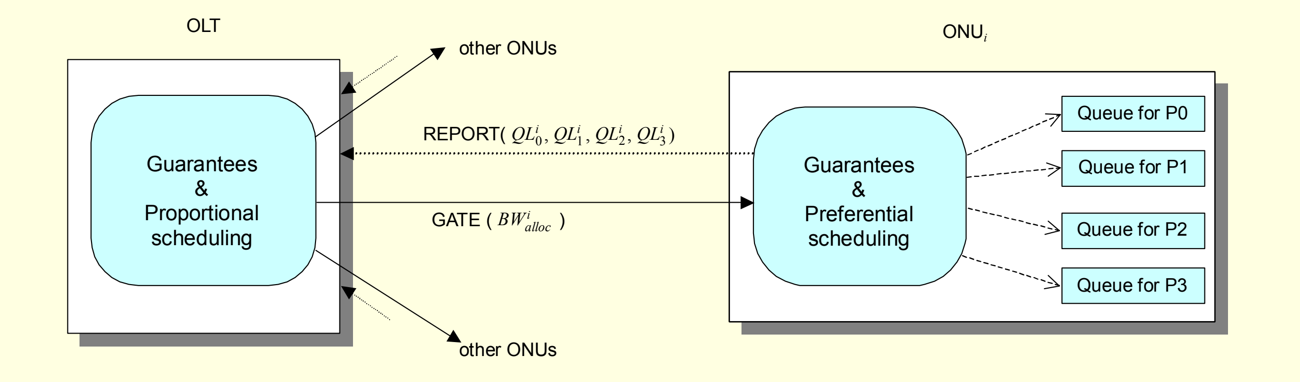

Fig. 10. The proposed hierarchical DBA algorithm.

In this paper, we do not consider that the OLT employs the preferential scheduling. If the OLT's scheduler uses the preferential approach in Step 3, it effectively decreases the delay for the high-class traffic. However, it creates a lack of allocatable resources for the low-class traffic. In particular, P1, P2 and P3 classes have variable bit rates. If any queue with a high class requires a burst transmission to the OLT, the OLT's scheduler allocates the sufficient resources to the queue. Thus, queues with a low class may have to wait for another chance. It is an over-preference that the scheduler tries to completely allocate the requirements for the burst traffic. Moreover, since the OLT's scheduler considers all of the queues for the Ethernet PON, a delay for the low class traffic may frequently occur.

Although we do not consider the preferential algorithm for the OLT's scheduler, the approach can be adopted according to the network traffic environments.

We can also modify this algorithm so that it takes into consideration the lengths and weights of the queues. However, if we do this, there is a risk that an enormous amount of non-real-time data will be transmitted faster than a small amount of real-time data. We evaluate this risk using a simulation in the next section.

IV. Performance Evaluation

We implement the simulation model for the proposed algorithms using OPNET, and the performances of the proposed algorithms are analyzed and evaluated in terms of the channel utilization, queuing delay, and ratio of the remainder.

1. Simulation Environments

In our simulations, we consider an Ethernet PON consisting of an OLT and 16 ONUs and assume that the upstream and downstream transmission speeds of the Ethernet PON are the same. We also assume that the data rate values (between the OLT and coupler and between the coupler and ONUs) are both 1Gbps. Also, the frame time is fixed at 2 ms, and the guard time is taken to be 1μs [12].

The formats of the control messages are based on the MPCP. To support multiple services, we assume that there are four kinds of queues for each service class in the ONUs. Additionally, we assume that each queue has an almost infinite capacity. The simulation parameters are summarized in Table 2.

Table 2. Simulation parameters.

Parameters	Value
Number of ONUs	16
Link capacity	1 Gbps
Frame length	2 ms
Number of classes	4
Guard time	1 μs
Inter packet gap	96 bits
Length of GATE/REPORT	64 bytes
Guaranteed bandwidth	12,320 bits
Weight of classes for DBA	3 (=P1), 2 (=P2), 1 (=P3)

Table 3. Distribution of packet length.

Packet length	Probability
$l = 64$	0.03
$64 < l < 580$	0.17
$l = 580$	0.18
$580 < l < 1518$	0.12
$l = 1518$	0.50

Each service traffic with a variable bit rate is represented as an ON-OFF model. The duration lengths of an ON and OFF period are generated according to an exponential distribution with a mean of 1 ms.

The distribution of the IP packet length is shown in Table 3. Let *l* be the packet length in bytes.

The offered load to the Ethernet PON is assumed to be the same as the total amount of the generated traffic loads, and each ONU generates an equal amount of traffic. When we vary the offered load in our simulation experiments, we maintain the same load for *P*0 and split the remaining load among *P*1, *P*2, and *P*3 equally.

2. Simulation Results

The proposed algorithm, which is denoted by P-F (proportional-preferential) DBA, is compared with the P-P (proportional-proportional) DBA in terms of the channel utilization, queuing delay, and ratio of the unused remainder. We verified that both algorithms provide differential service qualities among the classes.

Figures 11 and 12 show the delay performances obtained using P-P DBA and P-F DBA, respectively. Since it has a fixed amount of bandwidth periodically, the *P*0 class maintains a very

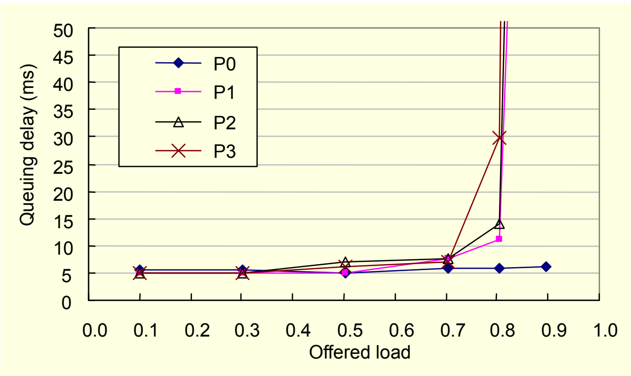

Fig. 11. Queuing delay for classes under P-P DBA.

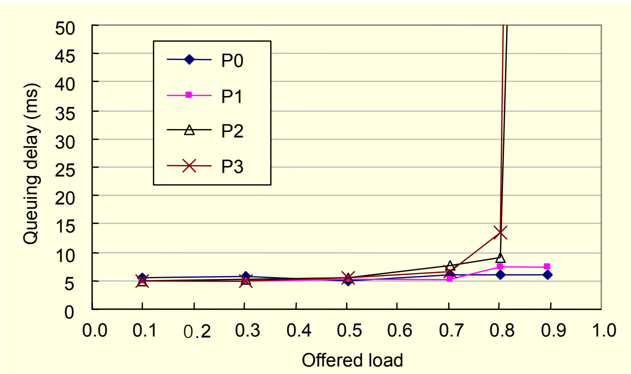

Fig. 12. Queuing delay for classes under P-F DBA.

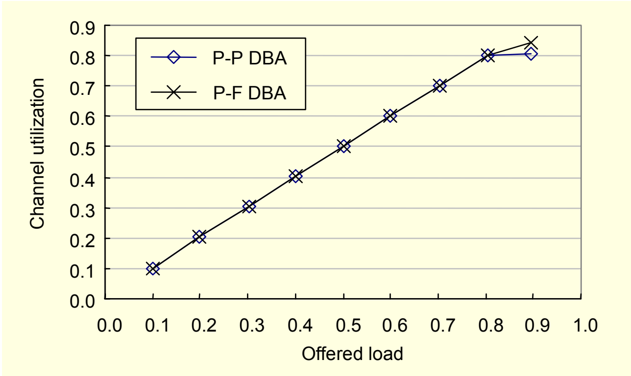

Fig. 13. Channel utilization.

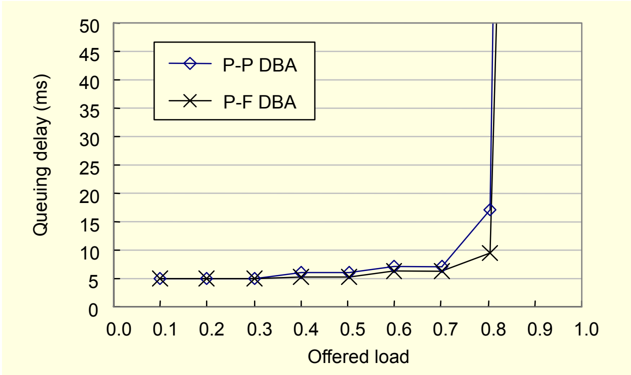

Fig. 14. Queuing delay.

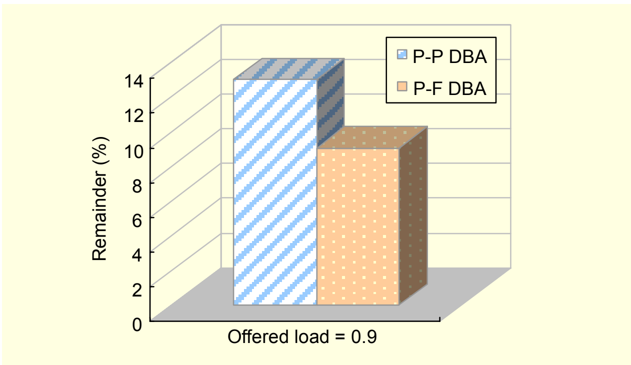

Fig. 15. Ratio of remainder bandwidth (at load 0.9).

low delay, regardless of the load. On the other hand, the delays for the other classes increase according to the load because they are served by a dynamic bandwidth allocation.

Under P-F DBA, in particular, *P*1 obtains a low delay even if the offered load is high. Moreover, the results for *P*2 and *P*3 are lower than those obtained under P-P DBA. This means that P-F DBA can provide an efficient bandwidth allocation, as well as taking CoS into consideration.

We can also see in Figs. 13 and 14 that the proposed algorithm is more efficient than P-P DBA.

In Fig. 15, the results are shown in terms of a ratio of the unused remainder for the total allocated bandwidth, where the load is 0.9. The amount of unused bandwidth for P-P DBA is larger than that for P-F DBA.

V. Conclusion

In this paper, we propose the hierarchical proportional-preferential dynamic bandwidth allocation algorithms, involving multiple services in an Ethernet PON. The use of a hierarchical scheme can solve the scalability issues by eliminating the need for separate GATE and REPORT messages to be sent to each queue over an Ethernet PON system. The use of this scheme also solves the switchover overhead issues due to the fact that all of the queues in one ONU are served consecutively, with no guard times between their transmissions. Therefore, the proposed algorithm is composed of the OLT's scheduler and the ONU's scheduler, which use a proportional allocation scheme and a preferential allocation scheme, respectively. Since in the proposed algorithm, the ONU's scheduler uses a preferential scheme rather than a proportional scheme, the amount of unused remainder bandwidth can be decreased.

We classified the services into four priority categories and designed the hierarchical DBA. In order to evaluate its performance, the algorithm is also implemented over an Ethernet PON system using OPNET. The proposed algorithm is compared with the overall proportional algorithm. The simulation results are analyzed in terms of the queuing delay, channel utilization, and the ratio of the unused remainder. The proposed hierarchical P-F DBA algorithm can achieve an efficient resource utilization and can provide a differential class of services.

References

[1] Glen Kramer and Gerry Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next Generation Optical Access Network," *IEEE Comm. Magazine*, Feb. 2002, pp.66-73.

[2] Glen Kramer, Biswanath Mukherjee, and Gerry Pesavento, "IPACT: A Dynamic Protocol for an Ethernet PON (EPON)," *IEEE Comm. Magazine*, Feb. 2002, pp. 74-80.

[3] Glen Kramer, Biswanath Mukherjee, S.Dixit, Yingjua Ye, and Ryan Hirth, "Supporting Differentiated Classes of Service in Ethernet Passive Optical Networks," *J. of Optical Networking*, vol. 1, no. 8/9, Aug. /Sept. 2002.

[4] G. Kramer et al., "Fair Queuing with Service Envelopes (FQSE): A Cousin Fair Hierarchical Scheduler and it's Application in Ethernet PON," *Technical Report No. CSE 2003-6*, Univ. of California, Davis, CA, Mar. 2003.

[5] Su-il Choi and Jae-doo Huh, "Dynamic Bandwidth Allocation Algorithm for Multimedia Services over Ethernet PONs," *ETRI J.*, vol. 24, no. 6, Dec. 2002, pp. 465-468.

[6] Maode Ma, Yongqing Zhu, and Tee Hiang Cheng, "Bandwidth Guaranteed Polling MAC Protocol for Ethernet Passive Optical Networks," *Proc. INFOCOM'2003*, vol. 1, Mar. 2003.

[7] Ho-Sook Lee et al., "A Two-Step Scheduling Algorithm to Support Dual Bandwidth Allocation Policies in an Ethernet Passive Optical Network," *ETRI J.*, vol. 26, no.2, Apr. 2004, pp. 185-188.

[8] J.D. Angelopoulos, I.S.Venieris, and G.I. Stassinopoulos, "A TDMA Based Access Control Scheme for APONs," *J. of Lightwave Technology,* vol.11, no.5/6, May/June 1993, pp. 1095-1103.

[9] J.D. Angelopoulos, I.S. Venieris, and E.N.Protonotarios, "A Transparent MAC Method for Bandwidth Sharing and CDV Control at the ATM Layer of Passive Optical Networks," *J. of Lightwave Technology,* vol. 14, no. 12, Dec. 1996, pp. 2625-2635.

[10] Youngjin M, Changhwan O, and Kiseon K, "A MAC Scheme Based the Cell Arrival Timing Information for Multimedia Services over ATM-Based PON," *IEICE Transaction Comm.*, vol. E82-B, no. 10, Oct. 1999, pp.1702-1705.

[11] Jongwook Jang and E.K. Park, "Dynamic Resource Allocation for Quality of Service on a PON with Home Networks," *IEEE Comm. Magazine*, vol. 38, Issue: 6, June 2000, pp.184-190.

[12] ANSI/IEEE Standard 802.3ah, 2002 ed., "*Media Access Control Parameters, Physical Layers and Management Parameters for subscriber access networks*," IEEE 802.3ah, 2002.

[13] IEEE P802.1D Draft 3, "Draft Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges," June 2003, http://www.ieee802.org/1/files/private/d-rev-drafts/d3/802-1D-D3.pdf.

[14] Byung Yun Lee, Gil Haeng Lee, and Hoon Choi, "Performance Analysis of SyncML Server System Using Stochastic Petri Nets," *ETRI J.*, vol. 26, no. 4, Aug. 2004, pp. 360-366.

Kye-Hyun Ahn has been an Engineer Staff of Electronics and Telecommunications Research Institute (ETRI) since August 2003. She received her PhD degree from Chonbuk National University, Jeonju, Korea, in 2003. Her current research interests include broadband communications in general and quality of service related architectures and protocols in particular.

Kyeong-Eun Han received the BS and MS degrees in computer engineering from Chonbuk National University, Jeonju, Korea, in 2001 and 2003. She is working toward the PhD degree in computer engineering at Chonbuk National University. She focuses on the design of network protocols and architecture for optical access networks such as Ethernet PON and WDM-based PON.

Young-Chon Kim received the BS, MS, and PhD degrees from Korea University, Seoul, Korea, in 1980, 1982, and 1987. He joined the Department of Computer Engineering at Chonbuk National University, Jeonju, Korea, in 1986, and is now a Professor at the School of Electronics and Information Engineering of Chonbuk National University. From 1989 to 1990, he was a Post-Doctorate Fellow at the University of California, Irvine. Also, from 1998 to 1999, he worked as a Visiting Researcher at the Network Research Laboratory at the University of California, Davis. He works with the Research Institute of Industrial Technology at Chonbuk National University. His research interests include high-speed optical communication networks, next generation IP networks, and broadband wireless communication networks.