Inter-ONU Bandwidth Scheduling by Using Threshold Reporting and Adaptive Polling for QoS in EPONs

Yeon-Mo Yang, Sang-Ook Lee, Hae-Won Jung, Kiseon Kim, and Byung-Ha Ahn

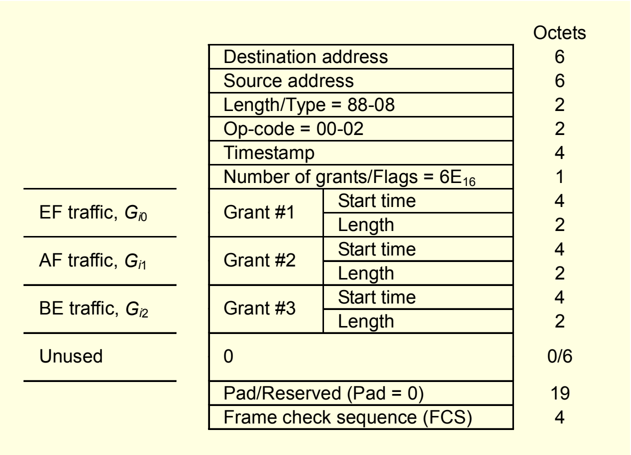
ABSTRACT—A dynamic bandwidth allocation (DBA) scheme, an inter-optical network unit (ONU) bandwidth scheduling, is presented to provide quality of service (QoS) to different classes of packets in Ethernet passive optical networks (EPONs). This scheme, referred to as TADBA, is based on efficient threshold reporting from, and adaptive polling order rearranging of, ONUs. It has been shown that the network resources are efficiently allocated among the three traffic classes by guaranteeing the requested QoS, adaptively rearranging the polling orders, and avoiding nearly all fragmentation losses. Simulation results using an OPNET network simulator show that TADBA performs well in comparison to the available allocation scheme for the given parameters, such as packet delay and channel utilization.

Keywords—Dynamic bandwidth allocation (DBA), Ethernet passive optical networks (EPONs), QoS, threshold reporting, adaptive polling.

I. Introduction

Passive optical network (PON) technology is considered as an efficient solution for deploying an access network because it can offer scalable, high-bandwidth, and cost-effective services. An Ethernet PON (EPON), an extended platform, basically preserves the advantages of Ethernet networks while reducing the cost and improving the quality of services (QoS) to a requested level [1], [2]. Practically, EPONs consist of one optical line terminal (OLT) situated at a central office and multiple optical network units (ONUs) located at the equipment at a customer's premises. In the upstream direction (from ONUs to the OLT), as it is a shared medium by ONUs, scheduling is required to prevent a data collision from different ONUs. A primitive solution is to assign a fixed time slot to each ONU. While assigning a fixed time slot regardless of its demand is simple, this scheme cannot adapt to bursty heterogeneous traffic and may waste bandwidth considerably. Consequently, a dynamic bandwidth allocation (DBA) scheme is required to support better QoS in EPONs.

Recent developments on DBA in EPONs in achieving QoS to a desired level are in progress [1]-[4]. In this letter, we propose a DBA scheme, referred to as TADBA, which is based on efficient threshold reporting and adaptive polling order rearranging, an inter-ONU bandwidth scheduling. The scheme provides better QoS (packet delay and channel utilization) under unbalanced traffic conditions by minimizing fragmentation losses, adaptively rearranging polling orders, and fairly utilizing the surplus bandwidth.



			Octets
	Destination address		6
	Source address		6
	Length/Type = 88-08		2
	Op-code = 00-02		2
	Timestamp		4
	Number of grants/Flags = $6E_{16}$		1
EF traffic, G_{i0}	Grant #1	Start time	4
		Length	2
AF traffic, G_{i1}	Grant #2	Start time	4
		Length	2
BE traffic, G_{i2}	Grant #3	Start time	4
		Length	2
Unused	0		0/6
	Pad/Reserved (Pad = 0)		19
	Frame check sequence (FCS)		4

Fig. 1. Format of a GATE message for the *i*-th ONU (64 bytes).

Manuscript received Oct. 14, 2005; revised Oct. 24, 2005.

Yeon-Mo Yang (phone: +82 62 970 2426, email: yangym@gist.ac.kr), Sang-Ook Lee (email: yashin96@gist.ac.kr), and Byung-Ha Ahn (email: bayhay@gist.ac.kr) are with the Department of Mechatronics, GIST, Gwangju, Korea.

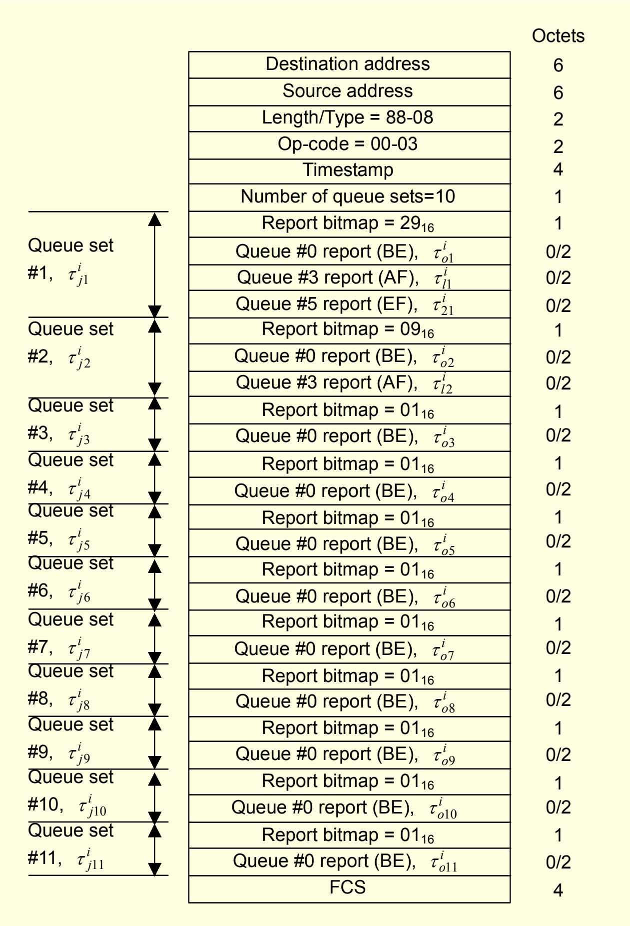
Kiseon Kim (email: kskim@gist.ac.kr) is with the Department of Information and Communications, GIST, Gwangju, Korea.

Hae-Won Jung (email: hw-jung@etri.re.kr) is with Broadband Convergence Network Research Division, ETRI, Daejeon, Korea.

II. Multiple Thresholds in a REPORT Message

The multi-point control protocol (MPCP) [6] provides services in regard to bandwidth requests and permissions with no collisions in the upstream direction. The protocol relies on two Ethernet control messages (GATE and REPORT) in its normal operation. GATE messages are transmitted by the OLT at fixed or variable cycle times to the ONUs; whereas each ONU reports its own queue status to the OLT by using REPORT messages. A GATE message can contain up to four grants with the start time and length of the granted period, and a REPORT message can include up to multiple sets of eight queue-length reports. GATE and REPORT messages are transmitted as Ethernet frames of 64 bytes. Figure 1 shows the proposed GATE message, while Fig. 2 depicts a threshold reporting-based REPORT message.

As the size of a REPORT message is limited to 39 bytes among 64 bytes, we have proposed eleven levels of threshold reporting that is different on each traffic class. For the given constraint, this method uses 39 bytes for reporting (7 bytes × 1 + 5 bytes × 1 + 3 bytes × 9 = 39 bytes). Threshold values (τ, see section III) are considered to minimize the remainder slot caused by packet fragmentation. When a report size at an ONU is less than B^{MIN}, defined as the minimum guaranteed transmission window in (1), only the largest threshold $\tau_{i,1}^{C}$ will be sent to the OLT. This will considerably reduce the report size compared to sending eleven threshold values. When the report size is greater than B^{MIN}, it first checks the number of frames exceeding B^{MIN}, that is, the threshold number. In case the threshold number is less than eleven, the threshold value is its corresponding packet size, and all the threshold values are sent to the OLT. When the threshold number is greater than eleven, it calculates the eleven limits ($v_{i,l}^{C}$) that are assigned and distributed evenly at eleven levels of reporting. Then, the newly obtained threshold is the nearest value to each one of the eleventh limits. A detailed threshold map is shown in Fig. 3.



		Octets
	Destination address	6
	Source address	6
	Length/Type = 88-08	2
	Op-code = 00-03	2
	Timestamp	4
	Number of queue sets=10	1
Queue set #1, τ_{j1}^{i}	Report bitmap = 29_{16}	1
	Queue #0 report (BE), τ_{o1}^{i}	0/2
	Queue #3 report (AF), τ_{l1}^{i}	0/2
	Queue #5 report (EF), τ_{21}^{i}	0/2
Queue set #2, τ_{j2}^{i}	Report bitmap = 09_{16}	1
	Queue #0 report (BE), τ_{o2}^{i}	0/2
	Queue #3 report (AF), τ_{l2}^{i}	0/2
Queue set #3, τ_{j3}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o3}^{i}	0/2
Queue set #4, τ_{j4}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o4}^{i}	0/2
Queue set #5, τ_{j5}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o5}^{i}	0/2
Queue set #6, τ_{j6}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o6}^{i}	0/2
Queue set #7, τ_{j7}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o7}^{i}	0/2
Queue set #8, τ_{j8}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o8}^{i}	0/2
Queue set #9, τ_{j9}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o9}^{i}	0/2
Queue set #10, τ_{j10}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o10}^{i}	0/2
Queue set #11, τ_{j11}^{i}	Report bitmap = 01_{16}	1
	Queue #0 report (BE), τ_{o11}^{i}	0/2
	FCS	4

Fig. 2. Format of a REPORT in the i-th ONU (64 bytes).

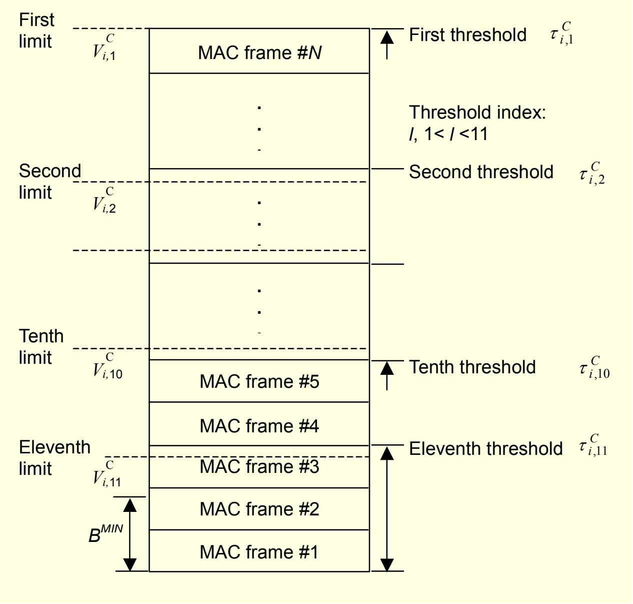


Fig. 3. Threshold values at threshold index l in the i-th ONU: the threshold values are used by the ONU to choose the non-fragmented queue size.

III. Dynamic Bandwidth Allocation

Basically, DBA can be described as follows: Each ONU sends its own threshold information of the three prioritized queues for best effort (BE), assured forwarding (AF), and expedited forwarding (EF) traffics using the same REPORT message. The OLT calculates the grants for the three prioritized service classes per ONU using the reported threshold information and delta of previous information, and transmits the grants for an ONU in a GATE message at the defined cycle time. Then, each ONU sends its traffic based on the received grant size from the given start time.

The proposed algorithm can be described for an EPON with N ONUs as follows: The transmission speed of the EPON is R_{bps}, guard time is T_g, and the cycle time is T_{cycle}. Concerning

T_{cycle}, for better QoS, there are two limit points such as T_{MIN} and T_{MAX}. When the requested size is less than one fourth of T_{MAX}, T_{cycle} will be T_{MIN}; otherwise, it is T_{MAX}. Additionally, the OLT will rearrange the polling order of ONUs adaptively depending on the sorting result of the report queue size from the ONUs. In other words, the ONU with the largest queue length will be polled first in the next polling cycle.

Let B^{MIN} be the minimum guaranteed transmission window (TW) in bits, per ONU. Then, B^{MIN} can be expressed as

$$B^{MIN} = \frac{(T_{cycle} - N \times T_g) \times R_{\text{bps}}}{N}. \tag{1}$$

The OLT and each ONU maintain state variables as follows.

- R_i^C (R_i^{EF}, R_i^{AF} and R_i^{BE}): the requested TW for each traffic class (in bits),
- D_i^C (D_i^{EF}, D_i^{AF} and D_i^{BE}): the difference between $R_i^C(t)$ at the current time and $R_i^C(t-1)$ at the previous time at the i-th ONU (in bits), and
- $\tau_{i,l}^C$ ($\tau_{i,l}^{EF}$, $\tau_{i,l}^{AF}$ and $\tau_{i,l}^{BE}$): the threshold value for each queue where l is an index of the threshold and $1 \le l \le 5$, that is, the number of threshold levels is five (in bits).

Here, C denotes one of the service classes such as EF, AF, or BE. When $R_i^C(t) > R_i^C(t-1)$, $D_i^C(t)$, called a delta, is obtained as follows:

$$D_i^C(t) = R_i^C(1) - R_i^C(t-1). \tag{2}$$

Now, let R_i be the sum of the requested TW (R_i^C) for each traffic class, as in $R_i = R_i^{EF} + R_i^{AF} + R_i^{BE}$. Then, we will define $R^{SUM} = \sum_i^N R_i$ as the sum of the requested TW, and $B^{Total} = T_{cycle} \times R_{\text{bps}}$ as the cycle TW at each cycle. When R^{SUM} is smaller than B^{Total}, the requested TWs of EF and AF, R_i^{EF} and R_i^{AF}, will be obtained by adding each delta value. When R^{SUM} is greater than B^{Total}, since a limited time slot size is available, only R_i^{EF} will consider the delta (D_i^C). The details of the delta scheduling are explained in burst polling-based delta-DBA (BP-DDBA), or DDBA for simplicity [4].

Let B_i^g be the granted TW for each i-th ONU. After the delta operation to the requested TW, we will investigate the bandwidth at the i-th ONU as follows:

Step 1. At an instant, some ONUs might have less traffic, light-loaded; while the other ONUs require more than the minimum bandwidth, heavily-loaded. When $R_i < B^{MIN}$,

$$B_i^g = R_i, \tag{3}$$

which results in a total surplus bandwidth: $B_{total}^{surplus} = \sum_i^M (B^{MIN} - R_i)$, where M is the set of light-loaded ONUs.

Step 2. When $R_i \ge B^{MIN}$, let $T_{i,l}$ be the sum of the l-th threshold reporting ($\tau_{i,l}^C$) for each traffic class as in $T_{i.l} = \tau_{i,l}^{EF} + \tau_{i,l}^{AF} + \tau_{i,l}^{BE}$. Then, a total matching threshold is given by $T_l^{matching} = \sum_{k \in K} T_{k,l}$, where K is the set of heavily loaded ONUs. Here, $N = M + K$, and we assume that $\tau_{i,l}^C > \tau_{i,l+1}^C$, that is, $\tau_{i,x}^C > \tau_{i,y}^C$ when $y > x$. If there is the smallest l such that $T_l^{matching} \le (B_{total}^{surplus} + K \cdot B^{MIN})$ for any l-th threshold, then

$$B_i^g = T_{i.l}. \tag{4}$$

Otherwise, B_i^g is as follows:

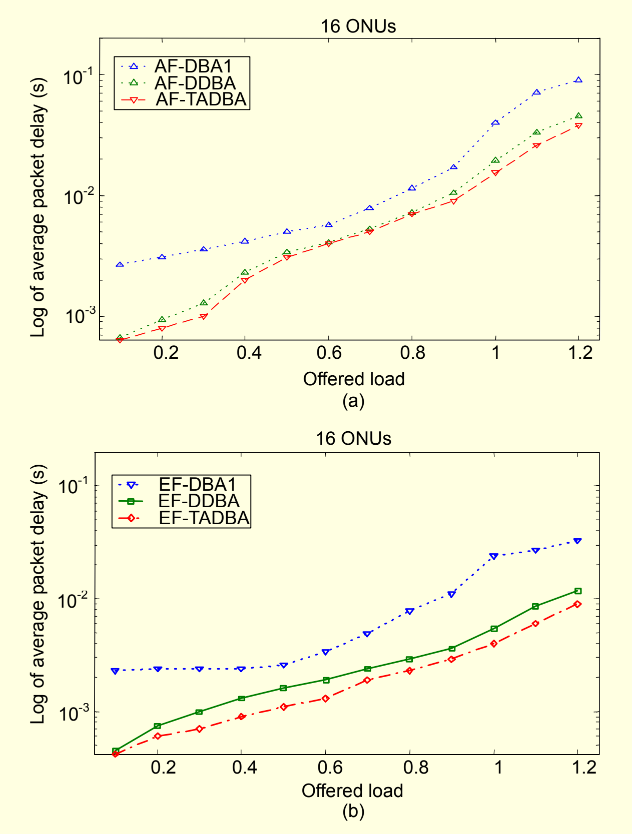
$$B_i^g = B^{MIN} + \frac{B_{total}^{surplus} \times (R_i - B^{MIN})}{\sum_{k \in K} (R_k - B^{MIN})}. \tag{5}$$

Then, the calculated bandwidth B_i^g among (3), (4), and (5) will be the grant sizes, G_{i0}, G_{i1} and G_{i2} for BE, AF, and EF at the i-th ONU, respectively.

IV. Experimental Results

We performed a simulation study by using an OPNET network simulator. It consists of a 1 by 16 EPON system with the star topology connected by full duplex 1 Gbps links whose center node is 1 by 16 PSC (passive star coupler) and periphery nodes are 16 ONUs and one ONT. The distance between the OLT and ONUs are 20 km. Also, each ONU supports three priority queues whose size is 5 Mbytes. The guard time between two consecutive TWs is 1.6 µs. The cycle times, T_{MIN} and T_{MAX}, are 0.4 and 1.6 ms, respectively. While simulating, we considered two groups of ONUs: one low-traffic group light-loaded and one high-traffic group heavily-loaded, that is, unbalanced traffic conditions (the numbers of heavily-loaded ONUs are 13, 14, 15, and 16).

Figures 4(a) and 4(b) illustrate a comparison of average packet delay of EF and AF classes with respect to competitive DBA1 of [3], DDBA of [4], and TADBA. The figures show the relationship between the offered load (independent parameter) and the logarithm of the average packet delay (dependent parameter). As shown Fig. 4(a), when the offered load is around 0.8, the average delay of AF packets with DBA1 rises to almost 0.07 s; however, the packet delay increased under TADBA goes to 0.04 s at an offered load of 1.2, while still being lower than those of DBA1. This is due to the unwanted fragmentation losses of DBA1, where the allocation time per ONU is based on a single threshold, and the bandwidths of heavily-loaded ONUs are vulnerable to fragmentation and prolonged delay caused by an unbalanced traffic load. When the incoming packet has a considerable increase, TADBA considers the remaining time slot existing between ONUs and OLT during inter-ONU scheduling.



(a)

(b)

Fig. 4. Comparison of average packet delay with three schemes: DBA1, DDBA, and TADBA for each traffic class, (a) AF and (b) EF.

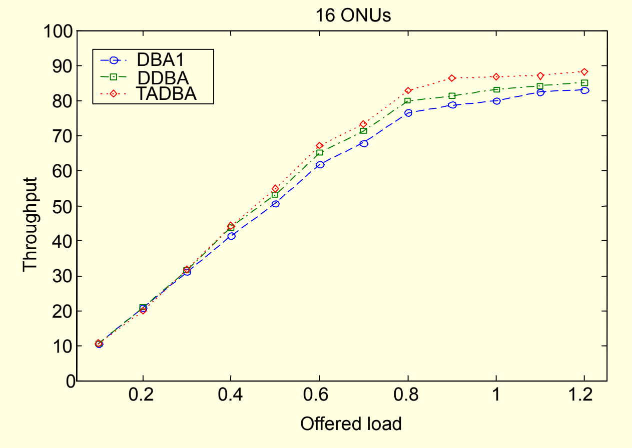


Fig. 5. Channel bandwidth utilization.

Also, as shown in Fig. 4(b), it provides a prioritized service to EF traffic as in [4], resulting in a lesser average packet delay. As a consequence, TADBA shows a lot of improvement in handling the packet delay.

Figure 5 shows the utilization of the channel under DBA1, DDBA, and TADBA. Although there is no constraint on packet delay for the BE traffic class, it is important to maximize the utilization. When the offered load is light, such as less than 0.4, differentiating T_{cycle} between T_{MIN} and T_{MAX} affects the channel utilization in TADBA, that is, it increases the utilization. With the increase in the network load, the channel utilization of the TADBA group exceeds that of DBA1 at the offered load of 0.6. After that point, it shows a different trend between DBA1 and TADBA. The channel utilization of all traffic classes in TADBA is very similar to that in DDBA when the offered load is relatively low. Beyond the load of 1.0, it increases to a value which is higher than that of all traffic classes in the TADBA scheme. More specifically, TADBA shows a utilization of 88% compared with DDBA and DBA1, which show 85 and 83% utilization, respectively.

The improvements can be explained as follows. TADBA utilizes the surplus bandwidth more effectively and efficiently by taking account of the threshold reporting and polling order rearranging than DDBA does when the offered load exceeds 1.

V. Conclusions

In this letter, we have presented a TADBA scheme for EPONs. This scheme can effectively allow all ONUs to fairly share the uplink bandwidth according to their threshold reports. That is, TADBA ensures that all service classes proportionally share the bandwidth on the ratio of the demand of a single class to total demand. By avoiding nearly all fragmentation losses through effective threshold reporting and adaptively rearranging polling orders as to the report sizes from ONUs, TADBA results in better performance in terms of packet delay and utilization compared with other DBA algorithms under the given unbalanced load environment.

References

[1] G. Kramer et al., “Supporting Differentiated Classes of Service in Ethernet Passive Optical Networks,” *J. Opt. Networks*, vol. 1, no.8/9, Aug. 2002, pp. 280-98.

[2] G. Kramer, “Design of Next-Generation Subscriber Access Systems Based on Ethernet Passive Optical Networks (EPON),” Doctoral Dissertation, Univ. of California, Davis, CA, Jun. 2003.

[3] C.-M. Assi, Y. Ye, S. Dixit, and M.- A. Ali, “Dynamic Bandwidth Allocation for Quality-of-Service over Ethernet PONs,” *IEEE J. Selected Areas in Communications*, vol. 21, no. 9, Nov. 2003, pp. 1467-1477.

[4] Y.-M. Yang, J.-M. Nho, and B.-H. Ahn, “Supporting Quality of Service by Using Delta Dynamic Bandwidth Allocations in Ethernet Passive Optical Networks,” *J. Opt. Networks*, vol. 4, no. 2, Feb. 2005, pp. 68-81.

[5] Y. Koh and K. Kim, “Loss Probability Behavior of Pareto/M/1/K Queue,” *IEEE Communications Letters*, Jan. 2003, pp. 39-41.

[6] IEEE Std. 802.3ah, *IEEE Std for Information Technology - Telecommunications and Information Exchange between Systems - Local and Metropolitan Area Networks - Specific Requirements*, IEEE, 2004.