A New Prioritized Request-Based Dynamic Bandwidth Allocation for Maintaining Fairness in PON with Multiple Heavy Users

YU, Zhenfei
Graduate School of Information Science and Electrical Engineering, Kyushu University

KARIYAWASAM, Amila
Graduate School of Information Science and Electrical Engineering, Kyushu University

Kimura, Shunji
Graduate School of Information Science and Electrical Engineering, Kyushu University

https://hdl.handle.net/2324/7361946

出版情報：IEICE Proceeding Series, 2024-11-25. 電子情報通信学会
バージョン：

A New Prioritized Request-Based Dynamic Bandwidth Allocation for Maintaining Fairness in PON with Multiple Heavy Users

Zhenfei YU[a], ***Student Member,*** **Amila KARIYAWASAM,** ***Member, and*** **Shunji KIMURA,** ***Member***

SUMMARY This study proposes a DBA that can reduce the maximum delay by 32.74% and 18.68% compared to the *LS-ED* and *WDBA*, respectively and maintain the Maximum Delay Fairness Index at a high level when overloaded ONUs exist.
keywords: *PON system, DBA, fairness, heavy users, maximum delay*

1. Introduction

Passive Optical Network (PON) has become a leading technology in broadband access networks due to its high bandwidth capacity and cost-effectiveness [1]. At the same time the variability in network usage is changing dramatically. This is because of discrepancies in volume of traffic and delay requirements depending on applications such as online gaming, online trading to IoT (internet of things) [2]. On the other hand, deployment of NG-PON2 has been reported giving operators the freedom to control wavelengths of transmission over current time-division multiple access (TDMA) [3],[4]. This also opens opportunity to serve business users through existing networks originally deployed for mass users. Against this background, it is of great interest to understand how to maintain fairness among users when allocating bandwidth as networks become more dynamic. Therefore, in NG-PON, the challenges are still concentrated on Dynamic Bandwidth Allocation (DBA).

For TDMA based PON, various approaches have been studied to maintain fairness when allocating bandwidth among optical network units (ONUs) while maintaining other parameters such as jitter, uplink utilization and computational complexity in acceptable values [5],[6]. *Proportional dynamic bandwidth allocation* was proposed to guarantee high fairness among ONUs where bandwidth proportional to the requested bandwidth is allocated to each ONU [7]. However, this approach has drawbacks of allowing overloaded ONUs to invade the bandwidth of underloaded ONUs. To avoid overloading the network by heavy users, *Limited DBA* was proposed where bandwidth allocation is kept under the limit value [5]. *Limited DBA* algorithm has drawbacks of causing some ONUs to consistently be in a state of insufficient bandwidth, while other ONUs have excess bandwidth that is not fully utilized.

In this study, we propose a novel DBA to handle heavy users in PON while maintaining fairness. In the proposed DBA, network overloading by the heavy users is avoided through Max-Min strategy and fairness is maintained by giving higher priority to ONUs that were treated unfairly in previous cycle. We analyze performance of the proposed DBA using computer simulations and compare results with *Limited scheme with excess distribution allocation* (*LS-ED*) and *Weight-based DBA* (*WDBA*).

The rest of this paper is organized as follows: We compare previously reported algorithms regarding fairness of DBA in section 2. Details of proposed algorithm is given in section 3. Section 4 presents performance comparison of the proposed DBA algorithm against *LS-ED* and *WDBA* followed by a Conclusion in section 5.

2. Related work

Various DBA algorithms have been proposed to maintain fairness when allocating bandwidth to subscribers of PON. Here, we review some of those proposals and compare basic ideas for maintaining fairness in DBA.

R. Kaur et al. proposed a *Game Theoretic Bandwidth Allocation Scheme*, which is based on the principle of proportionality [7]. In an ONU, there are four types of transmission containers (TCONTs) with different priorities. Here, we take TCONT Type 2 as an example for explanation. In the inter-ONU fairness scheme, the bandwidth allocated to the *i*-th ONU's TCONT Type 2 is proportional to the ratio of the requested bandwidth of that ONU's TCONT Type 2 to the sum of all ONUs' TCONT Type 2 requested bandwidth. This algorithm guarantees fairness in TCONT type 2 bandwidth allocation. On the other hand, this algorithm ensures that all ONUs' TCONT type 2 experience similar delays, but with a higher average delay.

Similarly based on the principle of proportional allocation, K. Kantelis et al. proposed the *Adaptive Protocol with Weights for Fair Distribution of Resources*, which differs by allocating bandwidth according to the weight proportion of each ONU [8]. With the aim of providing fairness, the algorithm supervises the buffer occupancy of each ONU and in case this exceeds 50%, then the assigned weight of the ONU is updated by adding 0.25 units for every 10% of increase. First, the minimum guaranteed bandwidth is calculated based on the weights, and then the excess bandwidth is distributed to the unsatisfied ONUs according

The authors are with the Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka-shi, 819-0395 Japan.
[a] E-mail: yu.zhenfei.358@s.kyushu-u.ac.jp

to their weights. It lowers the chance of packet loss caused by buffer occupancy.

To prevent overloaded ONUs from monopolizing uplink bandwidth, the *Limited DBA* was proposed. By setting an upper limit on the grant size, this approach prevents unfair bandwidth allocation while reducing latency. The *LS-ED* enhances the *Limited* DBA to boost statistical multiplexing while continuing to prevent ONUs from monopolizing the uplink. In the excess distribution schemes, the overloaded ONUs evenly utilize the unused or excess bandwidth left over by the underloaded ONUs [9]. Both the *Adaptive Protocol with Weights for Fair Distribution of Resources* and the *LS-ED* achieve relative fair bandwidth allocation, but the unused bandwidth is not efficiently utilized.

To maintain a high fairness index of bandwidth allocation throughout the DBA operation, X. Bai et al. proposed the *WDBA* [10]. During the operation of *WDBA*, ONUs' requests that are below the threshold are fully satisfied, and the threshold is updated accordingly. When ONUs' requests exceed the threshold, the remaining bandwidth is equally distributed among the remaining ONUs. As a result, satisfied ONUs experience minimal delay, whereas unsatisfied ONUs experience significant delays.

Considering drawbacks of previously proposed DBAs, we propose a novel DBA that focuses on current network conditions and reduces network delay while maintaining fairness by considering the data from the previous DBA cycle (request and allocation data). The proposed DBA gives higher priority to ONUs that were previously treated unfairly, thereby maintaining the overall fairness of bandwidth allocation, and reducing the significant delay experienced by unsatisfied ONUs.

3. Proposed DBA algorithm

The basic idea of the proposed DBA is to prioritize ONUs that were treated unfairly in the previous cycle, granting them higher priority in the current cycle. This is implemented through *prioritized requests*. We refer to the proposed DBA algorithm as PR-DBA.

A_i and R_i of Eq. 1 represent the uplink bandwidth resources allocated to the i-th ONU and the uplink bandwidth resources requested by the i-th ONU, respectively. Here x_i is called as relative allocation. n is the number of ONUs participating in bandwidth allocation. x_f represents the fair allocation mark, shown in Eq. 2. Dividing by n serves the purpose of calculating the average, to provide a benchmark for fair distribution during the bandwidth allocation process. ONU's perception of fairness is defined in Eq. 3. For the i-th ONU, the bandwidth allocation fairness is p_i. By setting the upper limit of p_i to 1, it prevents severely overloaded ONUs from encroaching on the bandwidth of underloaded ONUs, which would result in unfair bandwidth allocation. The *prioritized request* is defined as in Eq. 4, which is the real request multiplied by the ONU's perception of fairness. The *prioritized requests* contain two pieces of information: the ONUs' bandwidth requests and the degree of unfair treatment they received previously. These help reduce their delay and ensure their bandwidth requirements are met. The *prioritized requests* are sorted in ascending order. The sorting ensures that ONUs with smaller bandwidth requests or those who experienced unfair treatment in the previous cycle are given higher priority.

$$x_i = \frac{A_i}{R_i} \tag{1}$$

$$x_f = \frac{\sum_{i=1}^{n} x_i}{n} \tag{2}$$

$$p_i = min\left(1, \frac{x_i}{x_f}\right) \tag{3}$$

$$prioritized\ R_i = R_i * p_i \tag{4}$$

The specific implementation process is as follows:

Step 1: Calculate perception of fairness for each ONU according to the previous cycle information.

Step 2: Calculate *prioritized requests*.

Step 3: Sort the *prioritized requests* in ascending order.

Step 4: Allocate the real requested bandwidth to each ONU according to the sorted order, until the remaining bandwidth divided by the remaining unallocated ONUs is less than the current last allocation. Then, distribute the remaining bandwidth among the remaining ONUs according to the proportion of their real requests.

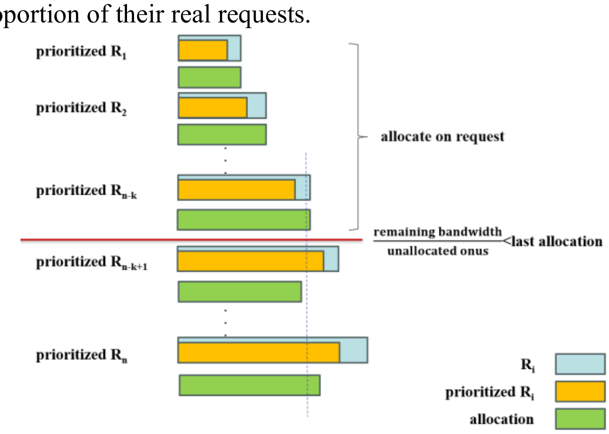

Fig. 1 Illustration of the PR-DBA operation steps.

Fig. 1 illustrates the operation process of the PR-DBA. The lengths of the blue, yellow, and green rectangles represent the ONUs' real requests, the ONUs' *prioritized requests*, and the bandwidth allocated to ONUs, respectively. The last allocation refers to the current final bandwidth allocation.

4. Performance Evaluation and Discussion

Table 1 Simulation parameters.

Parameter	Value
simulation length	1.5 s
DBA cycle	1.5 ms
DBA guard time	1 µs
uplink transmission rate	1 Gbps
number of ONUs	32
packet size	500 bytes
traffic generator	Poisson

We study the performance of the PR-DBA by simulating an EPON access network model. Table 1 lists the common simulation parameters used in our simulation.

In this study, traffic delay is defined as the difference between the time packets arrive at the ONU buffer and the time they arrive at the OLT buffer. Simulation results are compared against the *LS-ED* and *WDBA* under 2 conditions: with overloaded ONUs and without overloaded ONUs.

In this simulation, overloaded ONUs are defined as ONUs attempting to use more than 1.5-times the average uplink capacity. As the traffic load increases, we randomly increase the number of overloaded ONUs. Fig. 2 shows the number of overloaded ONUs under different traffic loads. The choice to add overloaded ONUs starting from a traffic load of 0.8 is due to the fact that, under lighter traffic loads, the uplink bandwidth is sufficient, and the impact of overloaded ONUs on bandwidth allocation is minimal. Overloaded ONUs usually experience significant delays. Therefore, the presence or absence of overloaded ONUs is of great importance for evaluating DBA algorithms.

By setting the traffic load to 1.2 in the simulation, it is possible to observe how the delay in the uplink increases and how the system stability changes under heavy and uneven traffic loads conditions.

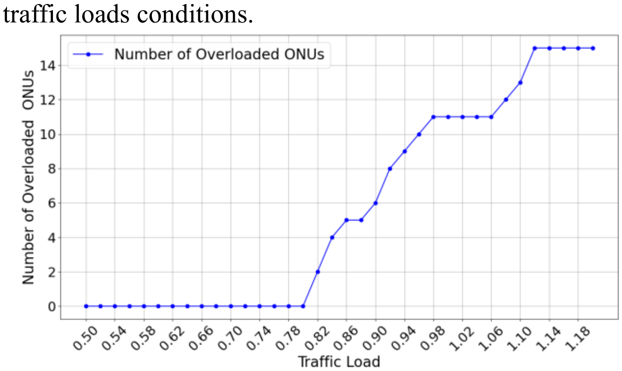

Fig. 2 Number of overloaded ONUs vs traffic load.

We compared maximum delay of the proposed algorithm with that of *LS-ED* and *WDBA* and results are presented in Fig. 3(a) and Fig. 3(b). Under the conditions without and with overloaded ONUs, the maximum delay for *LS-ED* starts increasing first and grows the fastest with traffic load. This is because *LS-ED* sets a bandwidth allocation upper limitation for each ONU and evenly distributes the unused bandwidth, resulting in some bandwidth being idle. Although *WDBA* fully utilizes all the bandwidth, it evenly distributes the unused bandwidth among the unsatisfied ONUs without considering the buffer occupancy of each ONU, causing the most overloaded ONU to experience the maximum delay.

In both scenarios, the PR-DBA performs the best. This is because it first sorts the ONUs based on *prioritized requests*, giving higher priority to those that have been treated unfairly. It then uses a Max-Min strategy to accommodate as many satisfied ONUs as possible, reducing the delay for these ONUs. For the unsatisfied ONUs, it employs proportional allocation to distribute the remaining bandwidth, ensuring that all unsatisfied ONUs experience the same level of starvation. This approach prevents the most severely starved ONUs from emerging, thereby reducing the maximum delay.

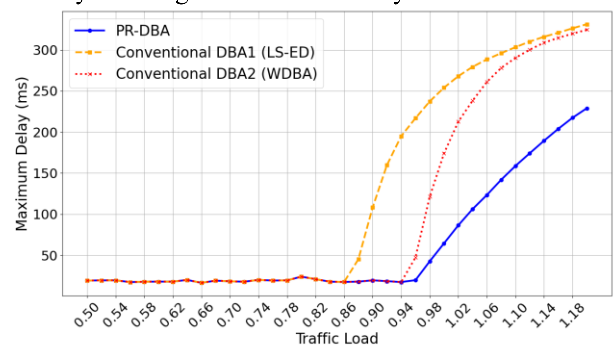

Fig. 3(a) Maximum Delay vs traffic load for 3 DBA algorithms without overloaded ONUs.

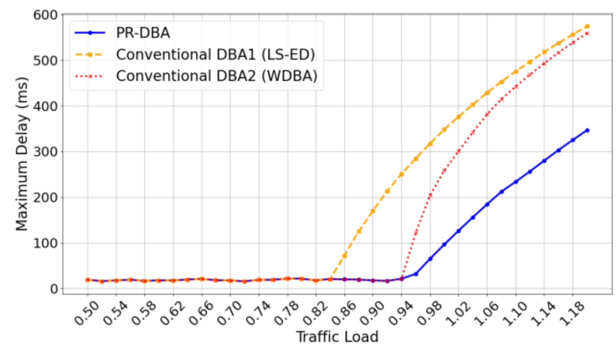

Fig. 3(b) Maximum Delay vs traffic load for 3 DBA algorithms with overloaded ONUs.

We studied differences in maximum delay to understand the reason for the lowest maximum delay of the proposed algorithm. Maximum delays of ONUs when traffic load is 1.0 are depicted in Fig. 3(c).

In Fig.3(c), the traffic generation rate for ONUs #1 to #32 gradually increases. ONUs #1 to #21 are underloaded ONUs, and ONUs #22 to #32 are overloaded ONUs. The largest maximum delays for *LS-ED* and *WDBA* were observed in the most overloaded ONUs, while the PR-DBA keeps the maximum delays experienced by all ONUs within a certain range. In the PR-DBA, ONUs #13 to #21 experience higher maximum delays. This is because, when allocating bandwidth based on the order of *prioritized requests*, previously unfairly treated overloaded ONUs are given higher priority and thus receive more bandwidth. As a result, the bandwidth allocated to ONUs #13 to #21 is reduced, leading to increased delays. By proportionally distributing the remaining bandwidth among the unsatisfied ONUs #22 to #32, the unsatisfied ONUs experience the same level of starvation, resulting in the unsatisfied ONUs experiencing almost the same maximum delay. The relatively large maximum delay for ONU #1 in all 3 DBA algorithms is due to the traffic generation rate of ONU #1 being too low, resulting in a longer time to generate a 500 bytes packet.

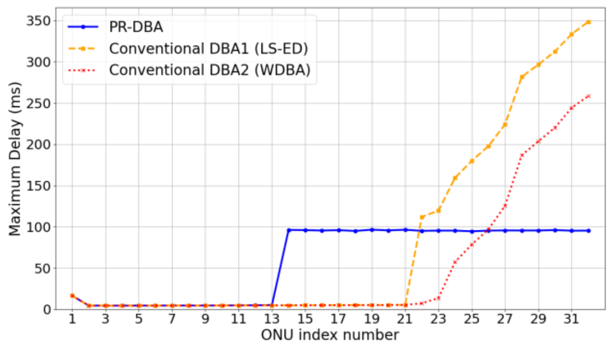

Fig. 3(c) Maximum Delay for 3 DBA algorithms with overloaded ONUs under 1.0 traffic load.

Fig. 4 illustrates the relationship between the Maximum Delay Fairness Index and traffic load for the 3 DBA algorithms with overloaded ONUs. The Maximum Delay Fairness Index *f(d)* $(0 < f(d) \leq 1)$ is defined in Eq. 5 [7],[11]. d_i is the ratio of maximum delay observed by *i*-th ONU to the average of maximum delay observed by all ONUs. Here, the Maximum Delay Fairness Index states how fairly the bandwidth is allocated with respect to the maximum delay measured in each ONU. Totally fair allocation should have an index of 1. As the disparity increases, a scheme which favors only a selected few ONUs has a fairness index near 0.

$$f(d) = \frac{\left(\sum_{i=1}^{n} d_i\right)^2}{n\sum_{i=1}^{n} d_i^2} \tag{5}$$

As the traffic load increases, the difference in maximum delay experienced by satisfied ONUs and unsatisfied ONUs becomes significant, resulting in a decrease in the maximum delay fairness index for all 3 DBA algorithms. However, compared to *LS-ED* and *WDBA*, the PR-DBA maintains the Maximum Delay Fairness Index at a higher level, meaning that the bandwidth allocation of the PR-DBA does not cause some ONUs to experience significant delays, resulting in poor user experience and decreased service quality. The PR-DBA improves Maximum Delay Fairness Index by 57.10% compared to *LS-ED* and by 32.69% compared to *WDBA*, with overloaded ONUs.

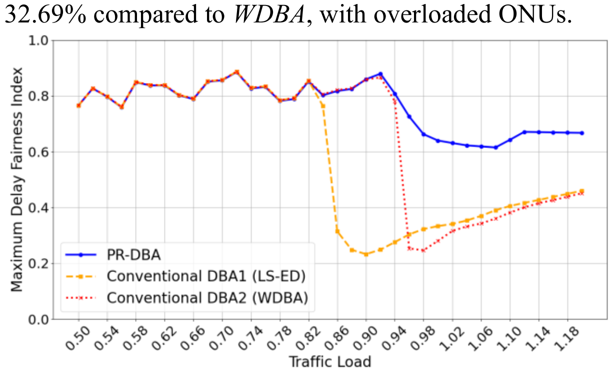

Fig. 4 Maximum Delay Fairness Index vs traffic load for 3 DBA algorithms with overloaded ONUs.

5. Conclusion

In this paper, the proposed DBA is compared to the *LS-ED* and *WDBA* in terms of maximum delay and Maximum Delay Fairness Index to demonstrate its advantages under heavy and uneven traffic load conditions. The results show that the proposed DBA reduces maximum delay by 32.74% and 18.68% compared to the *LS-ED* and *WDBA*, respectively when overloaded ONUs exist. Moreover, compared to *LS-ED* and *WDBA*, the proposed DBA maintains steady Maximum Delay Fairness regardless of heavy traffic loads.

References

[1] Pfeiffer, Thomas, et al, "PON going beyond FTTH [Invited Tutorial]." Journal of Optical Communications and Networking 14.1 (2022): A31-A40.
[2] Kaur, Ravneet, et al, "Performance analysis of fairness oriented dynamic bandwidth algorithm in integrated fiber-wireless architecture based on XG-PON and Wi-Fi." 2018 20th International Conference on Transparent Optical Networks (ICTON). IEEE, 2018.
[3] Aurzada, Frank, et al, "Capacity and delay analysis of next-generation passive optical networks (NG-PONs)." IEEE Transactions on Communications 59.5 (2011): 1378-1388.
[4] Li, Yifan, et al, "Fair and efficient DWBA algorithm based on SLA differentiated polling interleaved scheduling for NG-EPON." Optical Fiber Technology 61 (2021): 102451.
[5] McGarry, Michael P., Martin Reisslein, and Martin Maier, "Ethernet passive optical network architectures and dynamic bandwidth allocation algorithms." IEEE Communications Surveys & Tutorials 10.3 (2008): 46-60.
[6] Hussain, Sattar, and Xavier Fernando, "EPON: An extensive review for up-to-date dynamic bandwidth allocation schemes." 2008 Canadian Conference on Electrical and Computer Engineering. IEEE, 2008.
[7] Kaur, Damanjeet, Garima Gupta, and Vivekanand Jha, "A game theoretic bandwidth allocation scheme towards improving the fairness of XG-PON systems." 2021 International Conference on Information and Communication Technology Convergence (ICTC). IEEE, 2021.
[8] Kantelis, Konstantinos, et al, "Adaptive protocol with weights for fair distribution of resources in healthcare-oriented PONs." 2020 International Conference on Computer, Information and Telecommunication Systems (CITS). IEEE, 2020.
[9] Dhaini, Ahmad R., et al. "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks." Journal of lightwave technology 25.1 (2007): 277-286.
[10] Bai, Xiaofeng, Abdallah Shami, and Chadi Assi, "On the fairness of dynamic bandwidth allocation schemes in Ethernet passive optical networks." Computer Communications 29.11 (2006): 2123-2135.
[11] Jain, Rajendra K., Dah-Ming W. Chiu, and William R. Hawe, "A quantitative measure of fairness and discrimination." Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA 21 (1984): 1.