A Modified Max-Min Fair Dynamic Bandwidth Allocation Algorithm for XG-PONs

Ilias Gravalos[a,b], Konstantinos Yiannopoulos[a,c], Georgios Papadimitriou[d] and Emmanouel A. Varvarigos[a,b]
(a): Computer Technology Institute and Press "Diophantus," Rio 26500, Greece.
(b): Computer Engineering and Informatics Department, University of Patras, Rio 26500, Greece.
(c): University of Peloponnese, Department of Informatics and Telecommunications, Tripoli 22100, Greece.
(d) Department of Informatics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece.

***Abstract*— Passive Optical Networks (PONs) are a promising architecture for providing broadband access to the end user and XG-PON is the latest ITU-T PON standard, which provides line rates of up to 10 Gb/s. In XG-PON systems, the shared access of the upstream bandwidth necessitates the use of media access control (MAC) protocols that efficiently allocate the available capacity among multiple subscribers. Within this context, we present in this work an XG-PON oriented dynamic bandwidth allocation (DBA) algorithm that distributes the available upstream bandwidth by taking into account delay limits and fairness considerations. To this end, the proposed algorithm implements a max-min fair allocation scheme in order to distribute the available bandwidth to all the connected Optical Network Units (ONUs) at the PON. In the same time, the algorithm is engineered to always provide a minimum level of service on every XG-PON frame and this is achieved by a combination of status reporting and traffic monitoring techniques. Our simulation results show that the utilization of the proposed algorithm accomplishes low and bounded latency and jitter values over a range of traffic loads, and that it also results in a fair bandwidth distribution amongst ONUs irrespective of their spatial separation.**

Keywords—XG-PON; dynamic bandwidth allocation; max-min fairness.

I. Introduction

The expansion of the networked services that are offered to end users gives rise to increased bandwidth requirements. Additionally, there is a constant need to develop access networks that are inexpensive, scalable and capable of delivering integrated voice, video and data services to the subscribers. The Fiber-To-The-x (FTTx) access models, namely FTTH (Home), FTTC (Curb) and FTTB (Building) offer direct fiber connection close to the user premises and are being envisaged as a suitable solution for meeting the growing bandwidth demands. PONs in particular are a perfectly suitable candidate technology for implementing FTTx, since they are capable of providing high capacity and reliable access, and are also built in cost-effective fashion that fully exploits the passive and shared nature of the optical infrastructure. Currently, two major research domains for PON standardization exist, the first being the IEEE EPON (802.3av) [1][2] that is mainly designed for non-time critical data transfers, and the second being the ITU XG-PON (G.987) [2][3][4] that also enables the transfer of time critical data.

All transmissions in the standardized PONs are performed between the Optical Line Terminal (OLT) and multiple ONUs. The ONUs are connected via a passive optical splitter with the OLT in a tree-based architecture in which the OLT resides at the root. PONs are also implemented over a single downstream wavelength that communicates the OLT data to all ONUs, and a separate upstream wavelength that carries the transmissions from the ONUs to the OLT. Communication at the downstream direction is straightforward and the OLT broadcasts data frames to all connected ONUs. Special header fields are used to address a specific ONU and traffic bearing entities in it (traffic containers). In the upstream direction, however, the transmissions from the ONUs are combined by a power splitter and an arbitration mechanism is required so as to schedule transmissions and avoid collisions at the splitter output. To this end, upstream transmissions are co-ordinated by the OLT, which grants each ONU a suitable start time for its transmission as well as a transmission duration that is adapted to the ONU bandwidth needs.

In order to efficiently utilize the available upstream bandwidth, a DBA algorithm is required to distribute bandwidth in a flexible fashion that matches the time-varying traffic demands of ONUs. Furthermore, after taking into consideration that an XG-PON must accommodate a large number of users, it becomes evident that the DBA must be fair, with respect to bandwidth sharing, and simultaneously keep the average delay and jitter within acceptable levels. However, only a limited number of DBA algorithms that are specifically designed for XG-PONs have been previously reported [8], [9] [10], [11]. Significantly more effort has been put on EPONs [6], [7], [12], [13], still the results from previous works suggest that EPON-oriented DBA algorithms exhibit high delays when applied to XG-PONs [5]. This owes to the fact that XG-PONs employ very brief transmission scales (125 μsec long) to ensure that time-critical data are delivered in time, and, as a result, novel DBA algorithms are required to meet the stringent timing requirements in XG-PONs.

Within this context we propose a DBA algorithm that relies on a modified max-min fair scheme [16]-[19] to share the available upstream bandwidth among ONUs. The max-min fair mechanism has been extensively studied in E-PON, but to the best of our knowledge there is no corresponding study for G-PON architecture. Though, max-min constitutes a suitable method for XG-PONs, due to the constant upstream frame. The

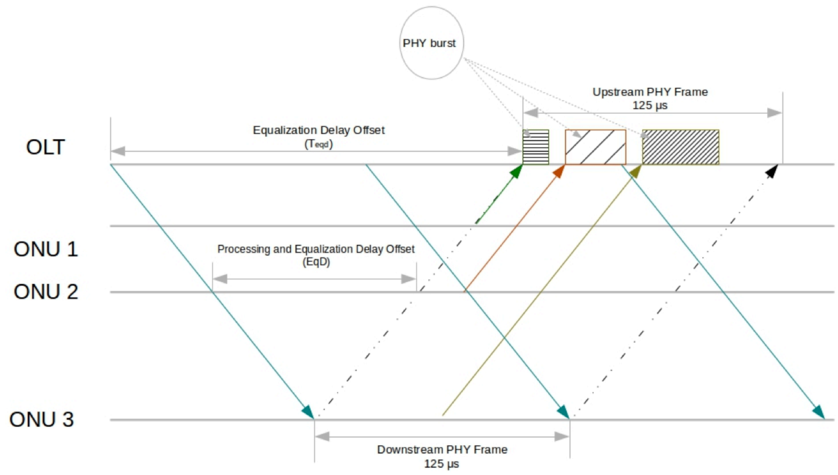


Figure 1: Frame transmission over the XG-PON

key idea behind the proposed algorithm is to redistribute any upstream capacity that remains unutilized within the upcoming 125 μsec and provide overloaded ONUs with an excess bandwidth that corresponds to their incoming traffic. Moreover, a hybrid status reporting and traffic monitoring mechanism is also proposed to ensure that even the most distant ONUs receive their fair share of bandwidth every 125 μsec. In our approach, a virtual buffer occupancy report is estimated at the OLT for all ONUs that have not been able to report their real occupancy prior to the execution of the max-min fair DBA. Our results show that the combination of the max-min fair DBA with the virtual reporting mechanism contributes to a significant decrease on the average delay that frames experience at the PON in comparison with two well-known variations of the interleaved-polling-with-adaptive-cycle-time (IPACT) DBA. The results also show that the attained latency, jitter and throughput are almost identical for all ONUs irrespective of their distance from the OLT.

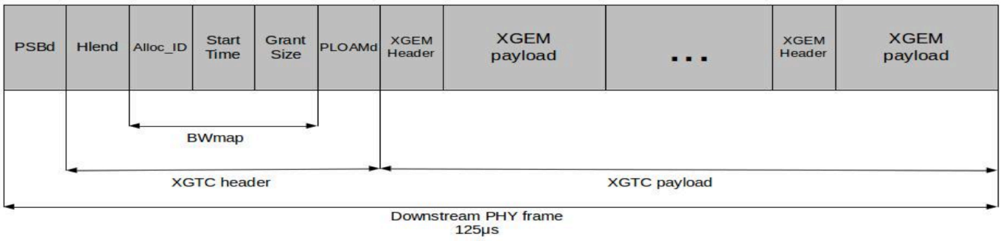
The remainder of this paper is organized as follows. In Section II we present fundamental operations of the XG-PON standard while in Section III we describe the proposed dynamic bandwidth allocation algorithm, as well as the accompanying prediction process that facilitates the estimation of bandwidth requests from seemingly non-reporting ONUs. We then present in Sections IV and V, respectively, the simulation setup and the results that have been obtained for the proposed algorithm. We also compare the proposed algorithm with limited and gated IPACT variants in Section V. Finally, Section VI summarizes the presentation and concludes the paper.

II. XG-PON PROCEDURES

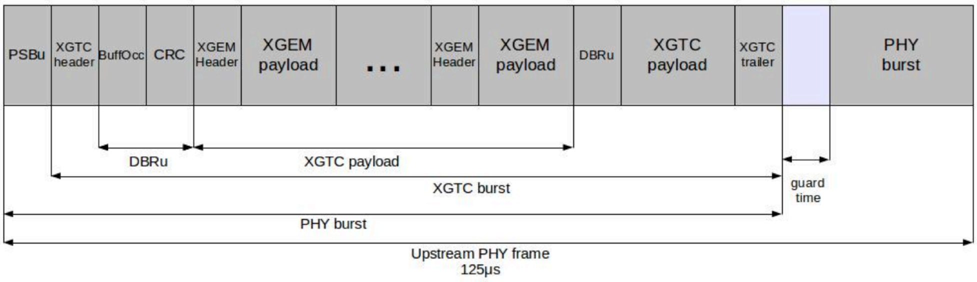
In the current Section we provide a description of the key XG-PON procedures, namely synchronization and framing, which affect the proposed DBA algorithm and the hybrid reporting mechanism. With respect to synchronization, Figure 1 depicts the communication basics over the XG-PON system. XG-PONs utilize frames of a fixed duration (125 μsec) both in the downstream and the upstream direction. Downstream synchronization is straightforward, since the OLT directly controls the beginning and the end of each 125 μsec-long frame in this direction. Upstream synchronization, however, is more challenging since ONUs transmit their data in an asynchronous fashion (bursts). Following the XG-PON standard conventions, upstream bursts (a) must arrive within the limits of a single upstream frame, and (b) must not collide at the common segment of the PON (splitter and feeder fiber). This can be achieved only if all bursts are synchronized to a common reference time, i.e. the start time of the upstream frame. To accomplish this, the OLT calculates an equalization delay offset T_{eqd} that corresponds to the temporal difference between the downstream and upstream frame start times. T_{eqd} equals the maximum ONU response time (RTD), which is calculated by adding the ONU processing time (typically 35$\pm$1 μsec) with the round-trip delay of the most distant ONU (typically 10 μsec/km). By subtracting the RTD of each respective ONU from T_{eqd}, the OLT also determines an ONU-specific static equalization delay EqD_i which is observed in all upstream transmissions. This procedure ultimately aligns upstream bursts that belong in a common upstream frame with the frame beginning.

In contrast with the downstream/upstream frame alignment, which only takes place once in the PON lifetime, burst collision avoidance is performed at the OLT on a per-frame basis. Prior to the beginning of a downstream frame, the OLT executes the DBA algorithm that takes into account the capacity requirements (occupancy reports) of the ONUs and assigns them with burst lengths (*GrantSizes*) in the next upstream frame. The OLT also calculates a *StartTime* for each respective burst, which is measured from the beginning of the upstream frame. *StartTimes* are calculated in a straightforward fashion by noting that a burst may only commence after the previous burst has finished (see for example Eq. (6) in Section IV). Typically, guard times (typically 8 bytes) are also inserted between bursts to allow for transmission drifts.

GrantSizes and *StartTimes* are communicated on the BWmap field of the downstream frame, whose structure is illustrated in Figure 2(a). The BWmap contains multiple allocation structures in order to serve a number of individual logical buffers (Alloc-IDs) within each ONU. In a similar fashion, occupancy reports from ONUs are carried in the DBRu field of



(a)



(b)

Figure 2: Structure of (a) downstream (b) upstream PHY frames.

upstream bursts, which are detailed in Figure 2(b). Additional fields are also present on the downstream frame and the upstream bursts (i.e. for physical layer synchronization, management and data transfer purposes), however they are not relevant to the purposes of this work and are not further discussed.

III. Traffic Aware Max-Min DBA

Bandwidth allocation is the process by which the OLT calculates the *GrantSizes* and the corresponding *StartTimes* for the ONUs. A dynamic allocation mechanism is generally preferable, since bandwidth assignments can be dynamically adapted to the instantaneous ONU traffic load, thus improving the upstream bandwidth utilization. Hence, the DBA mechanism at the OLT must satisfy the traffic demands of all ONUs and allocate the available bandwidth in an efficient and fair manner. In this work, we exploit the advantages of a well-known policy for resource sharing, namely the max-min fair algorithm, in order to allocate the available upstream bandwidth. The max-min fair algorithm is particularly appealing for DBA purposes in a PON, since the upstream bandwidth is shared under the following rules [18]:

- The upstream bandwidth is allocated in order of increasing demand.
- An ONU is never allocated with a GrantSize larger than its occupancy report.
- ONUs that cannot be fully served get an equal share of the remaining bandwidth.

Moreover, the complexity of the max-min fair algorithm is low (polynomial), which can be of practical importance, since the OLT has a very limited time to execute the DBA between successive downstream frames.

Given the max-min fair algorithm, each ONU n is allocated during downstream frame t with a *GrantSize* W_t^n that is calculated from the following steps:

- *Initialization*: Sort the ONU capacity demands R_t^n in an increasing order within a demand set S. Initialize the *GrantSizes* W_t^n to zero.
- *Step 1*: Calculate the available excess bandwidth by

  $$B_L = \frac{BW - \sum_{k=1}^{m} W_t^k}{|S|}, \tag{1}$$

  where (a) BW is the total bandwidth that is allocated for upstream bursts (38880 bytes minus framing overheads and guard times), (b) m is the total number of ONUs, and (c) $|S|$ equals the number of elements in S.
- *Step 2*: For all ONUs in S update the *GrantSizes* W_t^n as

  $$W_t^k = \min\{R_t^k, W_t^k + B_L\}. \tag{2}$$

- *Step 3*: Remove from S all ONUs that have been fully served and thus satisfy

  $$W_t^k = R_t^k. \tag{3}$$

- *Step 4*: If $|S| \geq 0$ then repeat the procedure from Step 1, else end.

The proposed dynamic allocation algorithm requires the constant OLT awareness of the traffic demands R_t^n of all ONUs at the PON. This can be achieved in two distinct ways: (a) the ONUs explicitly report their buffer status to the OLT (status reporting - SR), or (b) the OLT implicitly estimates the per ONU traffic requirements by monitoring the transmission of idle frames in upstream transmissions (traffic monitoring – TM). Our approach is to utilize a hybrid reporting mechanism bandwidth allocation. In more detail, the OLT solicits buffer occupancy reports from the ONUs, so as to be constantly aware of instantaneous traffic changes. Following Figure 1, however, the next downstream frame occasionally starts while some

ONUs' bursts have not arrived at the OLT. As a result, the OLT has inconsistent buffer status for these ONUs and this can lead to inefficient bandwidth utilization.

In order to improve utilization, we apply a traffic estimation mechanism to ONUs that have not explicitly reported their buffer occupancy C_t^n. Following this approach, the OLT assigns the aforementioned ONUs with a virtual capacity demand R_t^n that is calculated either from previous *GrantSizes* (Max-Min Based on Grant - MMBoG)

$$R_t^n = \begin{cases} C_t^n\,, \ if\ C_t^n \neq null \\ \qquad\quad and\ C_t^n > 0 \\ \frac{\sum_{i=1}^{t-1} w_i^n}{t-1}, \qquad else \end{cases}, \tag{4}$$

or from previous occupancy reports (Max-Min Based on Report - MMBoR)

$$R_t^n = \begin{cases} C_t^n\,, \ if\ C_t^n \neq null \\ \qquad\quad and\ C_t^n > 0 \\ \frac{\sum_{i=1}^{t-1} c_i^n}{t-1}, \qquad else \end{cases}, \tag{5}$$

Hence, during every BWmap construction, the OLT has explicit demands about the buffer occupancy from ONUs that reported their status and implicit demands by traffic monitoring the rest. Moreover, zero buffer reported ONUs are also provided with a virtual request, so that (a) any traffic that arrives to them prior to the next upstream frame can be served immediately, and (b) they are forced to report their occupancy.

Finally, after the appropriate bandwidth assignment has been made to each ONU, the OLT must determine the transmission *StartTime* (*ST*) of upstream bursts. Assuming that the processing delay is the same for all ONUs, the best utilization is achieved by serving ONUs in ascending order according to their distance from the OLT (or equivalently their upstream propagation time t_{prop}). Thus, the upstream burst *StartTimes* are calculated in a straightforward fashion as

$$ST_t^k = ST_t^{k-1} + W_t^{k-1} + GuardBand. \tag{6}$$

IV. Simulation Model

A simulation model was developed in OMNET++ [19] to evaluate the performance of the proposed algorithm and compare it with IPACT variants. The simulation setup consists of the OLT, 10 and 32 ONUs placed at random distances from 1 to 20 km, and an optical splitter (OS). The downstream and upstream channel rates were set to 9.985 Gb/s and 2.488 Gb/s respectively, following the recommendations of the XG-PON standard. Moreover, all communications abide to the frame formats of Figure 2, which were fully developed for both the downstream and the upstrea m. The basic OLT functionalities are summarized as follows:

- The OLT generated a downstream frame every 125 µsec. All downstream frames carried a valid BWmap field with *StartTimes* and *GrantSizes* for all ONUs.
- The OLT received the asynchronous upstream bursts from the ONUs and isolated the respective DBRu fields.

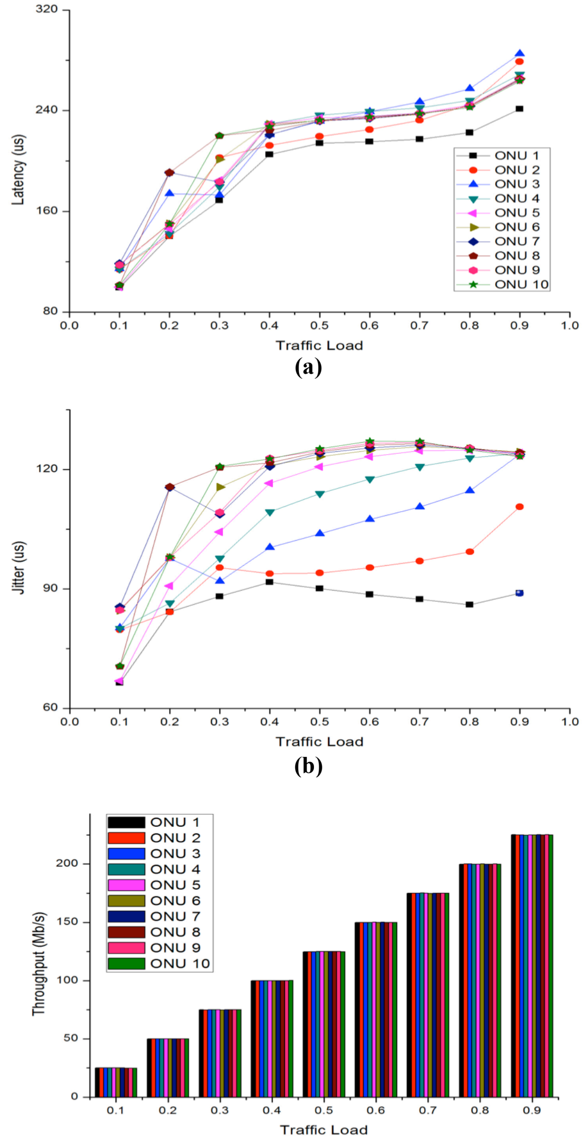
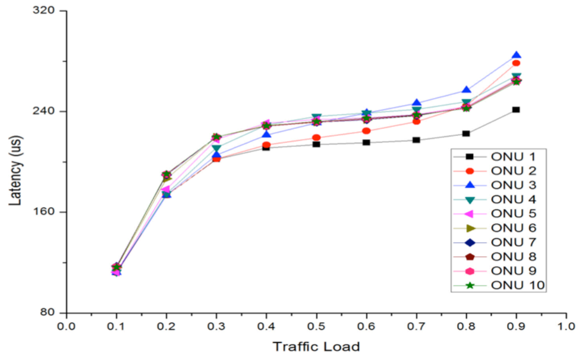


Figure 4: a) Average packet latency, b) Average packet's latency jitter, c) Average Throughput for MMBoG.

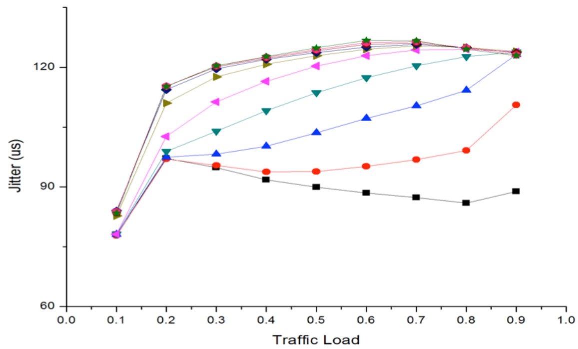
- Exactly before the generation of the downstream frame, the OLT executed the max-min fair DBA algorithm, calculated the *StartTimes* and *GrantSizes* and constructed the BWmap.

As far as the optical splitter is concerned, the module was implemented as a frame replicator in the downstream direction. The splitter created a number of replicas of the downstream frame at its input and each replica was placed at a different port, thus feeding the connected ONUs. In the upstream direction, the optical splitter simply forwarded the received frames from the ONUs to the OLT. The OS does not implement any buffering. Finally, all ONUs were implemented with the following functionalities:

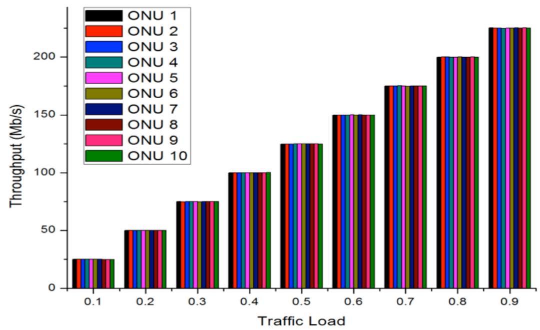
- ONUs received the synchronous downstream frames and isolated their own *StartTimes* and *GrantSizes*.



(a)



(b)



(c)

Figure 5: a) Average packet latency, b) Average packet's latency jitter, c) Average throuput for MMBoR.

- ONUs framed the data from their buffers in upstream bursts with a length equal to *GrantSize* and then waited for a period equal to EqD_i+*StartTime* to transmit.
- ONUs sampled their buffers, constructed their respective DBRu fields, updated the upstream burst contents and then sent the burst.

Traffic generators (TGs) were also implemented to provide the ONUs with incoming data from the end-user network. The traffic was generated in the form of fixed size frames, while the frame interarrival times followed a Poisson distribution with a mean value λ that was derived from the desired per ONU load ρ_{ONU}, following

$$\rho_{ONU} = \frac{\rho_{PON}}{m} = \frac{\lambda \cdot L}{m \cdot R_{u/s}}. \tag{7}$$

In Eq. (7) ρ_{PON} represents the total PON load, m is the number of ONUs, $R_{u/s}$ equals the upstream line rate (2.488 Gb/s) and L is the traffic generator frame size (1000 bytes).

V. Performance Evaluation

In order to evaluate the performance of the proposed algorithm we executed a number of simulation experiments with respect to the average packet latency (delay), the latency standard deviation (jitter) and the throughput. Figures 4 and 5 summarize the performance of the proposed DBA versus the traffic load. It is straightforward to verify that the experienced delay is always less than three frame durations (375 μsec) for all ONUs even when the load approaches 90%. Moreover, the algorithm achieves similar delays and identical throughput for all ONUs, irrespective of their distance from the OLT, which is a clear indication that (a) the max-min DBA serves ONUs in a fair fashion, and (b) the bandwidth estimation methods tend to eliminate the prioritization of the closest ONUs that are able to report in a more timely fashion as compared with remote ONUs. A bounded behavior is observed for the latency deviation (jitter), as well. Even though jitter cannot be fully eliminated due to the discrete-cycle PON operation, the presented results demonstrate that jitter saturates to less than a single frame duration (125 μsec) for all ONUs as the network load increases. This is of practical importance in XG-PONs, which are designed with QoS provisioning considerations in mind, since a number of delay-sensitive applications (such as IPTV and VoIP) require both a fixed delay and an upper bound on the jitter.

We further assessed the performance of the proposed DBA by comparing it with two well-known IPACT principles - the limited and the gated, since other principles perform similar to the limited IPACT in terms of latency, while the gated IPACT exhibits the lowest latency in general [20]. For the limited IPACT implementation, the maximum *GrantSize* W_{max} was set to the upstream frame nominal size (38880 bytes) divided by the number of the ONUs. The *GrantSize* was calculated from the bandwidth estimations as

$$W_t^n = \begin{cases} R_t^n\,, & if\ R_t^n \le W_{max} \\ W_{max}\,, & else \end{cases}. \tag{8}$$

For the gated IPACT implementation the *GrantSize* was set equal to the bandwidth estimation.

$$W_t^n = R_t^n\,. \tag{9}$$

Whenever the total *GrantSizes* exceeded the nominal frame size in gated IPACT, consecutive upstream frames where utilized to serve the ONUs. The results of the comparison are presented in Figure 6, which summarizes the average (over all ONUs) delay, jitter and throughput that is achieved by all four DBAs. The results show that the proposed DBA outperforms IPACT in terms of latency and jitter, and the benefit that is gained is more pronounced at increased network loads. Clearly, IPACT requires at least an additional frame duration at loads over 80%, while the limited principle does not even provide a bounded jitter value. As a final remark, the figure demonstrates that all DBAs attain identical throughputs and therefore the proposed algorithm does not underutilize the PON capacity.

VI. CONCLUSION

To better utilize the XG-PON rate capabilities and satisfy capacity demands of FTTx models we proposed a hybrid status reporting and traffic monitoring dynamic bandwidth allocation method to serve traffic among the ONUs. In our algorithm the OLT grants portions of the available bandwidth utilizing the max-min fair mechanism over the estimated buffer occupancy reports of each ONU. The OLT estimates the occupancy for the ONUs that have not reported demands, and the estimation is based either on the average size of previous grant sizes, or the average buffer occupancy of previous reports. Simulation results showed that the proposed algorithm outperforms the IPACT limited and gated methods in average latency and latency jitter. Moreover, the algorithm ensures fairness among all ONUs and does not impose any throughput loss.

ACKNOWLEDGMENT

This work has been funded by the NSRF (2007-2013) Synergasia-II/EPAN-II Program "Asymmetric Passive Optical Network for xDSL and FTTH Access," General Secretariat for Research and Technology, Ministry of Education, Religious Affairs, Culture and Sports (contract no. 09SYN-71-839).

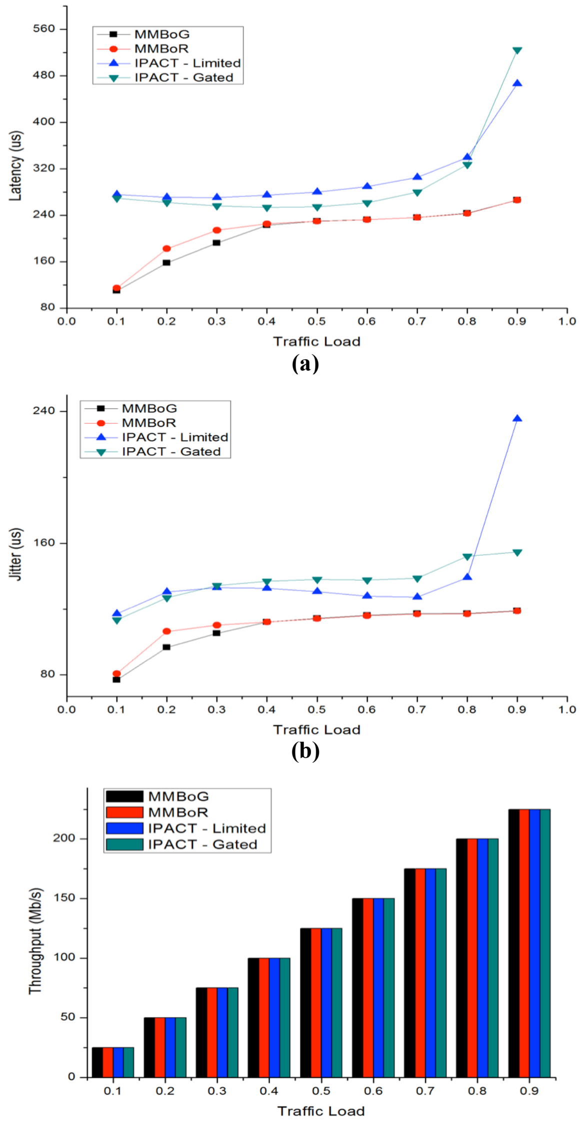


Figure 6: Comparison of DBA methods in a) Average packet latency, b) Average packet's latency jitter and c) throughput vs. traffic load.

REFERENCES

[1] IEEE Std 802.3av-2009, IEEE Standard for Ethernet - Amendment 1: Physical Layer Specifications and Management Parameters for 10Gb/s Passive Optical Networks. IEEE. 2009.

[2] Elaine Wong, "Next-Generation Broadband Access Networks And Technologies", Journal Of Lightwave Technology, 30 (4), 2012

[3] ITU-T, "10-Gigabit-capable passive optical network (XG-PON) systems", October 2010.

[4] FrankJ.Effenberger, "The XG-PON System: Cost Effective 10 Gb/s Access", Journal Of Lightwave Technology, 29 (4), 2011

[5] AURZADA et al. "Capacity And Delay Analysis Of Next-Generation Passive Optical Networks (NG-PONS)", IEEE transactions on communications, 59(5), 2011.

[6] M. Mcgarry, M. Reisslein, and M. Maier, "Ethernet Passive Optical Network Architectures and Dynamic Bandwidth Allocation algorithms," IEEE Commun. Surveys Tutorials, vol. 10, no. 3, 2008, pp. 46-60

[7] J. Zheng and H.T. Mouftah, "A Survey of Dynamic Bandwidth Allocation Algorithms for Ethernet Passive Optical Networks," Optical Switching Netw., 6(3), 2009, pp. 151-162

[8] B. Skubic et al., "A Comparison of Dynamic Bandwidth Allocation for EPON, GPON, and Next-Generation TDM PON," IEEE Comm. Mag., 47(3), 2009, pp. 40-48

[9] Siyang Yu, Xue Chen, Xintian Hu, and Caili Du, "A Dynamic Bandwidth Allocation Scheme for Nested – PON", Second International Conference on Instrumentation & Measurement, Computer, Communication and Control, 2012

[10] Man Soo Han, Hark Yoo, and Dong Soo Lee, "Development of Efficient Dynamic Bandwidth Allocation Algorithm for XG-PON", ETRI Journal, 35(1), 2013

[11] Tamal Das, Ashwin Gumaste, Akhil Lodha, Ashish Mathew, and Nasir Ghani, Generalized Framework and Analysis for Bandwidth Scheduling in GPONs and NGPONs – The K-out-of-N Approach, Journal Of Lightwave Technology, 29(19), 2011

[12] Kramer, G., Singhal, N.K., Dixit, S., Banerjee, A., Mukherjee, B., Ye, Y., Fair queueing with service envelopes (FQSE): A cousin-fair hierarchical scheduler for subscriber access networks, IEEE Journal on Selected Areas in Communications 22 (8) , pp. 1497-1513

[13] Banerjee, A., Kramer, G., Mukherjee, B., "Fair sharing using service-level agreements (SLAs) for open access in EPON", IEEE International Conference on Communications 3 , pp. 1858-1862

[14] D. Bertsekas, R. Gallager. "Data Networks", Prentice- Hall, 1992.

[15] D. Nace and M. Pi´oro. Max-min fairness and its applications to routing and load-balancing in communication networks: a tutorial. Communications Surveys & Tutorials, IEEE, 10(4):5– 17, 2008

[16] A. Tomaszewski. "A polynomial algorithm for solving a general max-min fairness problem". European transactions on telecom- munications, 16(3):233–240, 2005

[17] B. Radunovic and J.-Y. Le Boudec. A unified framework for max-min and min-max fairness with applications. Networking, IEEE/ACM Transactions on, 15(5):1073 –1083, 2007

[18] S.Keshav, An Engineering Approach to Computer Networking, Addison-Wesley, Reading, MA, 1997.

[19] http://www.omnetpp.org/

[20] G. Kramer, B. Mukherjee, and G. Pesavento: "Interleaved Polling with Adaptive Cycle Time (IPACT): A Dynamic Bandwidth Distribution Scheme in an Optical Access Network" Photonic Network Comm., 2002, 4(1), pp 89-107.