EFFICIENT DYNAMIC BANDWIDTH ALLOCATION FOR EPON

A.S.M. Delowar Hossain, M. A. Ali
Department of Electrical Engineering, Graduate School of The City University of New York
New York, NY 10016-4309, USA
ASMDHossain@gc.cuny.edu

ABSTRACT
Ethernet Passive Optical Network (EPON) has emerged as the solution of choice for the access network. In EPON, the data transfer channel is shared among users in the upstream direction. Therefore, it requires an arbitration method to allow ONUs to share the upstream bandwidth. There are many research papers on bandwidth arbitration schemes. Most of the Dynamic Bandwidth Allocation (DBA) schemes in the literatures do not consider the overall network condition to globally optimize their algorithm. Furthermore, they do not consider the ONU MAC frame boundary when granting bandwidth to an ONU. This causes mismatch between granted bandwidth and ONU frame boundary. Since fragmentation of Ethernet frame to fit the remaining ONU slot is not allowed, it causes unused ONU slot bandwidth (ONU slot remainder). In addition, number of them require DBA calculation time when no upstream traffic is allowed (idle cycle time). In this paper, we propose a DBA scheme that eliminates ONU slot remainder and idle cycle time and reuses all unused cycle bandwidth in a globally optimized fashion to maximize the upstream bandwidth utilization. Improved utilization will result in minimization of packet delay, packet loss ratio and ONU buffer occupancy. Simulation was done to demonstrate the advantages of the proposed DBA.

Keywords: EPON, DBA, MPCP, OLT, ONU.

1. INTRODUCTION
The capacity of backbone networks has recently increased drastically, but there has been little change in the access networks. With the rapid growth of the number of users and multimedia services, access networks become bottlenecked. Passive optical network (PON), a point-to-multipoint optical network composed of passive elements only, is regarded as the most likely technology to solve this problem [1-4]. EPON, a PON that carries Ethernet traffic, is the best candidate for next generation access network because it is cheap, simple and widely used. In the PON architecture, a centralized optical line terminal (OLT) is connected to dispersed optical network units (ONU) over point-to-multipoint topologies, e.g. tree, ring, bus. An OLT is located at the local exchange and ONUs are close to the end user location. In downstream transmission from OLT to ONUs, a frame is broadcasted to all ONUs through an optical splitter and each ONU filters the received frame depending on its destination address. In upstream transmission from ONUs to OLT, ONUs must share a single optical fiber trunk (Fig.1). Therefore, the Multi-Point Control Protocol (MPCP) is developed by the IEEE 802.3ah task force to support upstream bandwidth allocation without collision among ONUs [5]. The MPCP is mainly operated by two control messages, GATE and REPORT. REPORT is the ONU bandwidth request to OLT and GATE is the granted upstream bandwidth by OLT. There are many research papers on bandwidth arbitration schemes. All schemes has strong and weak points. In general, most of them grant upstream bandwidth based on a single ONU report and do not consider the overall network condition to globally optimize their algorithm. Also, most of them ignore the issue of mismatch between granted bandwidth to an ONU and frame boundary of that ONU data. Thus, causes waste of bandwidth in a granted ONU slot (fragmentation of Ethernet frame to fit the remaining slot is not allowed). In addition to that, they require DBA calculation time when no upstream traffic is allowed (idle cycle time). In this paper, we describe a Dynamic Bandwidth Allocation (DBA) scheme that eliminates ONU slot remainder and idle cycle time totally and reuses all unused cycle bandwidth dynamically to maximize the upstream bandwidth utilization. The proposed DBA is globally optimized and always considers the entire network demands before any bandwidth assignment. This very mechanism is crucial for supporting QoS requirement in a network where time sensitive traffic has to be considered over all other network demands. This DBA scheme will
Figure 1: EPON Tree Architecture

demonstrate better utilization and packet loss ratio than IPACT[6] and DBA2 [7] at network saturation, when efficiency is mostly needed to minimize packet delay and save packets from being dropped.

In Section 2 we present a short review of MPCP and DBA. In Section 3 we describe the proposed DBA algorithm. In Section 4 we present simulation results and discussion of the results. Finally the paper is concluded with some remarks.

2. MPCP AND DBA

In response to an ONU request for bandwidth (REPORT), MPCP at OLT transmits a GATE message to a particular ONU with the following information: ONU transmission start time and the length of the transmission. MPCP timestamps the message with its local time. Upon receiving a GATE message matching its MAC address, each ONU will program its local registers with “transmission start time” and “transmission length”. Also, the ONU will update its local clock to that of the timestamp in the received control message, hence avoiding any potential clock drift and maintaining synchronization with the OLT. When the transmission “start timer” expires, the ONU will start its contention-free transmission. Many DBAs emerged, based upon MPCP, to guarantee efficient collision free transmission. We will discuss IPACT for the sake of comparison. IPACT is an ONU-based polling scheme, which reduces the polling cycle with an interleaved polling approach. When $ONU_i$ is transmitting packets through the upstream channel, the OLT informs $ONU_{i+1}$ of its starting time and its transmission bandwidth size. $ONU_{i+1}$ is polled before the transmission from $ONU_i$ is completed. The first bit from $ONU_{i+1}$ arrives at the OLT only after the OLT receives the last bit from $ONU_i$(including a guard time after each ONU transmission). Because the bandwidth granted for $ONU_i$ is equal to its demand, it can be fully used. However, this scheme may cause an ONU with heavy traffic load to monopolize the upstream channel so that packets from other ONUs are delayed longer. To solve this problem, a limited service discipline is used to redefine a maximum transmission bandwidth ($B_{\text{max}}$) for all ONUs. With $B_{\text{max}}$, the maximum transmission bandwidth is confined so that the average queuing delay of packets may decrease. When ever the ONU request is more than $B_{\text{max}}$, OLT grants the ONU the prefixed bandwidth $B_{\text{max}}$. Note that $B_{\text{max}}$ is a fixed number and it is not customized to match with ONU frame boundary. So, the received maximum bandwidth $B_{\text{max}}$ may not match with an ONU frame boundary and will cause bandwidth wastage [8](MAC frame cannot be fragmented to fit the remaining of the slot). Also, IPACT issues grants based on single ONU report without global knowledge of other ONU demands; thus it is not globally optimized. To address this issue, [7] proposed globally optimized DBA, but with no consideration to ONU slot remainder[8]. Also it assumes that there is always at least an ONU demanding less than $B_{\text{max}}$ and uses that premise to eliminate idle cycle time caused by DBA calculation. But when all ONUs demand exceed $B_{\text{max}}$, then DBA in [7] will have idle cycle time. Our proposed DBA addresses all these issues.
3. PROPOSED DBA

The proposed DBA is a centralized scheme, where the OLT decides the upstream bandwidth allocation for the ONUs. In a given time slot, unlike other DBAs, an ONU first transmits its control report to the OLT followed by its data. The proposed DBA is a centralized scheme, where the OLT decides the upstream bandwidth allocation for the ONUs. In a given time slot, unlike other DBAs, an ONU first transmits its control report to the OLT followed by its data. The control report includes the ONUs' upstream bandwidth request, and the OLT allocates the bandwidth accordingly.

In cycle N, as soon as the last ONU's report is received by OLT, the DBA calculation for cycle N+1 starts.

First ONU in cycle N+1 transmits after receiving grant.

When last ONU's timeslot in cycle N is not long enough to compensate for DBA calculation, DBA calculation extends to next cycle.

First grant of cycle N+1. Grant calculation time is compensated by last ONU's data transmission time.

ONU that reported $\leq B_{\text{max}}$ in cycle N, will transmit at the start of cycle N+1 due to advanced grant (while DBA calculation extends from last cycle for the ONUs reported $>B_{\text{max}}$).

ONU reported $\leq B_{\text{max}}$ received advanced grant.

ONU reported $>B_{\text{max}}$ received delayed grant after DBA calculation.

In cycle N, as soon as the last ONU's report is received by OLT, the DBA calculation for cycle N+1 starts.

Figure 2: DBA Scheduling a)normal operation  b) when last ONU's slot size is small
control report (queue information) and then the data to OLT. All ONU reports are considered by OLT at the end of cycle N to do DBA calculation for cycle N+1. Then OLT allocates bandwidth (grant) to each ONU. Since the DBA scheduling takes into consideration all the ONU requests (unlike IPACT), it is globally optimized. It is capable of allocating bandwidth to needy ONUs beyond the normal maximum limit ($B_{\text{max}}$) by transferring bandwidth from the less needy ONUs. The DBA calculation at the end of cycle causes some idle cycle time when no upstream traffic is allowed. We introduce a mechanism to eliminate this idle cycle time and achieve better performance.

DBA module in OLT only needs the report frames to do DBA calculation. As mentioned earlier, an ONU transmits report frame before the data frames. Therefore, once the last ONU of cycle N has transmitted its report to OLT and the data is still being transmitted by that ONU to OLT, the OLT will complete the DBA calculation (including the issuance of the first grant) for cycle N+1. The data transmission time of last ONU of cycle N will compensate for DBA calculation for cycle N+1 (Fig. 2a). Note that the order of grant is based on the shortest request first and that will ensure the longest transmission slot at the last (to fully compensate for DBA time). When ONU load is low, the last ONU of cycle N may not have long enough transmission time to compensate for DBA calculation for cycle N+1. To address that circumstance, the DBA adds additional mechanism. Since an ONU request up to $B_{\text{max}}$ will be granted anyway, DBA will allow ONUs requesting up to $B_{\text{max}}$ in cycle N to transmit (in sequence of their request) at the beginning of cycle N+1 without waiting for DBA calculation. This will compensate for DBA calculation time for busy ONUs (report > $B_{\text{max}}$), which may extend from cycle N to cycle N+1 (Fig. 2b). If there is no busy ONU at all, then all the ONUs will receive grants up to $B_{\text{max}}$ in advance; eliminating the need of an end of cycle DBA calculation. Thus, for all conditions the DBA calculation time is compensated for. Assuming a 10km trunk with a maximum ONU distance of 1 km from splitter, the round trip time (RTT) from OLT to ONU is 110us. The last ONU’s data transmission slot time ($B_{\text{max}} \approx 125$us) will compensate for DBA calculation time and OLT’s initial grant. Incase the last ONU slot is less than 125us, then the low demand (report < $B_{\text{max}}$) ONU will transmit at the beginning of next cycle (advanced granted slot), while the OLT completes DBA calculation and issues first grant. In another words, the globally optimized DBA calculation is completed without any idle cycle time. The sequence of transmission for ONUs reported > $B_{\text{max}}$ is based on the shortest report first (SRF); if there is a tie, then the decision is based on the sequence of report. Note that all ONUs are synchronized to a common reference clock extracted from the OLT downstream traffic. Clocking information, in the form of a synchronization marker, is included in a frame at the beginning of each downstream cycle. The synchronization marker is a 1-B code that is transmitted every 2 ms to synchronize the ONUs with the OLT.

After the issue of grant scheduling we need to look at the bandwidth distribution. Cyclic ONU reportings provide OLT adequate buffer information so that OLT can grant bandwidth in such a way that eliminates ONU slot remainder. ONU report is formulated after comparing ONU queue size to a pre-established maximum slot bandwidth size ($B_{\text{max}}$). We define $B_{\text{max}}$ as follows:

$$B_{\text{max}} = \frac{1}{N} \left\lfloor \frac{\text{R}_{\text{onu}}(T_{\text{cyc}} - (N \cdot G))}{1} \right\rfloor = 15 \text{KB}$$  \hspace{1cm} (1)

where polling cycle $T_{\text{cyc}} = 2$ms, guard band G between two ONUs transmission is 5 microsecond, maximum upstream transmission rate $\text{R}_{\text{onu}}$ (from ONUs to OLT) is 1Gbps, and ONU count $N = 16$. ONU report structure (Fig. 3) is formulated as follows:

$$R_j = \begin{cases} Q_{j, \text{boundaries}}, & \text{if } Q_{j, \text{total}} > B_{\text{max}} \\ Q_{j, \text{total}}, & \text{if } Q_{j, \text{total}} \leq B_{\text{max}} \end{cases}$$  \hspace{1cm} (2)
where \( R_j \) is the report, \( Q^\text{total}_j \) is the queue length and \( M^\text{boundaries}_j \) is the frame boundaries of \( ONU_j \). ONU reports the whole queue length when queue size \( \leq B_{\text{max}} \), otherwise it reports an array of 8 frame boundaries starting with the first boundary which is just below the \( B_{\text{max}} \). Following the first boundary below \( B_{\text{max}} \), the next 6 boundaries will be at least 1KB apart; the last frame boundary is

\[
Q_{\text{total}} \leq B_{\text{max}} \quad : \quad \text{report queue length}
\]

\[
Q_{\text{total}} > B_{\text{max}} \quad : \quad \text{report frame boundaries}
\]

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Report frame structure}
\end{figure}

At first the DBA calculates the total unclaimed bandwidth (\( B_{\text{cycle\_remainder}} \)) of the next cycle by adding what was unclaimed by low demand ONUs from the allowed amount of bandwidth (\( B_{\text{max}} \)) as in (3):

\[
B_{\text{cycle\_remainder}} = \sum_{j} (B_{\text{max}} - Q^\text{total}_j)
\]  

(3)

where \( U \) is the total number of ONUs reported queue size (\( Q^\text{total}_j \)) less than \( B_{\text{max}} \). Then it will calculate total over demand (\( B_{\text{total\_over\_demand}} \)) of all the busy ONUs (ONU report > \( B_{\text{max}} \)). Total ONU excess demand is calculated as follows:

\[
B_{\text{total\_over\_demand}} = \sum_{j} (Q^\text{total}_j - B_{\text{max}})
\]  

(4)

where \( O \) is the total number of ONU reported queue size (\( Q^\text{total}_j \)) more than \( B_{\text{max}} \). Note that when the \( B_{\text{cycle\_remainder}} \geq B_{\text{total\_over\_demand}} \) then each busy ONU is allocated bandwidth according to their request without any further calculation. Otherwise, DBA determines the share (in addition to \( B_{\text{max}} \)) of the cycle remainder for each busy ONU according to the ratio of its excess demand and total over demand of all busy ONUs. A busy ONU’s additional share (\( B^\text{extra}_j \)) from total excess cycle bandwidth (\( B_{\text{cycle\_remainder}} \)) is as follows:

\[
B^\text{extra}_j = \frac{Q^\text{total}_j - B_{\text{max}}}{B_{\text{total\_over\_demand}}}
\]  

(5)

The final calculated bandwidth for an busy ONU is as follows:

\[
B^\text{calculated}_j = B^\text{extra}_j + B_{\text{max}}
\]  

(6)

Finalizing grant according to (6) will cause wastage of bandwidth due to ONU slot remainder. Therefore, OLT will issue grant (\( B_{\text{assigned}} \)) by finding a frame boundary equal (or less) to calculated (\( B_{\text{calculated}} \)) grant as follows:

\[
B_{\text{assigned}} = \frac{Q^\text{total}_j}{B_{\text{max}}}
\]  

5
Note that for a 1B mismatch of the last frame boundary of a 1538B frame, there will be bandwidth wastage of as much as 1537B (1538B-1B); it is 10% of maximum allowed bandwidth of an ONU slot (15kB). OLT recalculates cycle remainder and total over demand after every grant as follows:

\[ B_{\text{cycle remainder}} = B_{\text{cycle remainder}} - (B_{\text{assigned}} - B_{\text{max}}) \]  
(8)

\[ B_{\text{total over demand}} = B_{\text{total over demand}} - (Q_{\text{total}} - B_{\text{max}}) \]  
(9)

Then OLT applies (5) to (7) to determine next grants until all busy ONUs are granted bandwidth. This dynamic calculation of OLT transfers all unused bandwidth from low demand ONUs to high demand ONUs in a globally optimized and fair manner. The DBA eliminates ONU slot remainder and distributes excess cycle bandwidth to other needy ONUs without any additional DBA calculation time. Slot remainder elimination and dynamic reuse of all unused bandwidth, save packets from waiting in queue until the next cycle and make room for new packets to arrive. This results in minimization of packet delay and lower packet loss ratio (see Section 4). To achieve the above mentioned advantages, this DBA involves no additional complexities. Only the modification of OLT DBA software and ONU reporting scheme (frame boundary is reported using existing reporting format) will bring the approach to reality.

### 4. SIMULATION RESULTS

We simulated two of the best performing algorithms to compare with the proposed DBA, Limited IPACT [6] and DBA2[7]. In this simulation \( B_{\text{max}} = 15KB \). Maximum time slot for each cycle \( T_{\text{cyc}} = 2ms \) and guard band G between two ONUs transmission is 5 microsecond. The load of each 16 ONUs are almost identical and ONU buffer size is 10MB. ONUs receive synthetically generated self-similar traffic [9] from user end in a maximum rate of 100Mbps (\( R_{\text{user}} \)) and from ONUs to OLT maximum upstream transmission rate is 1Gbps(\( R_{\text{onu}} \)). In tree architecture both downstream and upstream delays are the same (110 microseconds for 10km trunk). At full network load we simulated over 50 million packets in a time bounded discrete event driven simulation. Note that the relation between total network load (TNL) and offered ONU load (OOL) is \( TNL = (OOL*16)/10 \), and around OOL=0.6 the network is saturated.

We observe that proposed DBA has better performance than IPACT and DBA2 at higher load. In Fig. 4(a)-(c), the overall performance looks almost similar for all DBAs until TNL is around 0.8 (OOL \( \approx 0.5 \)). This is due to the following reasons: at 0.5 ONU load, ONUs receives in average 50Mbps, which is 12.5kB per cycle (maximum duration 2ms). When ONU transmits that 12.5kB out of its buffer, it will have to add 20 bytes of overhead per frame. We found in (1) that \( B_{\text{max}} \) for an ONU per cycle is 15kB and that normally is an ONUs cycle bandwidth. So, generally around this load each ONU’s bandwidth demand per cycle is close to \( B_{\text{max}} \). OLT running any of the algorithms will assign grants equivalent to entire queue length as reported by ONU (since it is less than \( B_{\text{max}} \) - see Section 3). This full queue length is the frame boundary of the last frame of ONU and grant will be on that frame boundary for all algorithms. Thus, none of the DBAs should have any slot remainder and generally there will be no high demand ONUs requiring redistribution of unclaimed cycle bandwidth. So, the proposed DBA will work same as others, because proposed DBA's advantages are more effective at higher load when network resources are scarce. Note that due to bursty traffic, even in low load suddenly excess traffic may arrive in buffer raising the ONU demand over \( B_{\text{max}} \). In that case, IPACT and DBA2 may have some ONU slot remainder. The performance improvement of proposed DBA over others appears when ONU load is greater than 0.5 (TNL\( \approx 0.8 \)). Then ONUs queue sizes start to get larger and ONUs bandwidth demand
exceed $B_{\text{max}}$. Under this circumstance, IPACT and DBA2 are forced to grant all ONU's the same maximum bandwidth ($B_{\text{max}}$) for that cycle. Thus, wastes bandwidth due to ONU slot remainder; but the proposed DBA wastes no bandwidth (due to elimination of ONU slot remainder).

**Figure 4**: Proposed DBA versus other DBAs
Furthermore, proposed DBA assigns that saved slot remainder bandwidth along with other unclaimed cycle bandwidth of low demand ONUs to the needy ONUs. That allows proposed DBA to serve the needy ONUs earlier than IPACT and DBA2. That enhances the performance of proposed DBA over others. In Fig. 4(b), we see better utilization of proposed DBA due to ONU slot remainder savings and reuse of unused cycle bandwidth. That in turn allows more packets to be served in the network than IPACT and DBA2. It causes shorter buffer stay (packet delay - Fig. 4a) for packets and allows more packets to enter the queue, resulting lower packet loss ratio (Fig 4c).

At the rate of 100Mbps ($R_{user}$) with 16 ONUs, the OLT-ONU line (1Gbps) is saturated when each ONUs average load is around 0.6 (62.5% ~ 60Mbps) of its full load. The delay, packet loss and utilization reaches the maximum at the network saturation.

CONCLUSION

We demonstrated that the proposed DBA performs better than IPACT and DBA2. It eliminates the bandwidth wastage caused by ONU slot remainder in IPACT and DBA2 and distributes that bandwidth among needy ONUs without any idle cycle time. This results in better utilization, minimized delay and lower packet loss ratio especially at higher network load when network resources are very scarce. The proposed DBA is globally optimized and always considers overall network demands before any bandwidth assignment. This very mechanism is crucial for supporting QoS requirement in a network where time sensitive traffic has to be considered first over all other network demands. Therefore, this algorithm can be extended to QoS aware functions with promising performance.

REFERENCES