Online excess bandwidth distribution for Ethernet passive optical networks

Jason R. Ferguson,[1] Martin Reisslein,[2,*] and Michael P. McGarry[3]

[1]*ADTRAN, 9201 North 25th Avenue, Suite 200, Phoenix, Arizona 85021, USA*
[2]*Department of Electrical Engineering, Arizona State University, Goldwater Center, Tempe, Arizona 85287, USA*
[3]*Department of Electrical and Computer Engineering, University of Akron, 301 Buchtel Common, Akron, Ohio 44325, USA*
[*]Corresponding author: reisslein@asu.edu

Received November 24, 2008; revised January 16, 2009; accepted February 11, 2009; published March 24, 2009 (Doc. ID 104379)

Excess bandwidth distribution techniques have recently been proposed to improve the dynamic bandwidth allocation in Ethernet passive optical networks (EPONs). We compare existing offline excess bandwidth distribution with conventional limited interleaved polling with adaptive cycle time (IPACT-limited) in terms of packet delay performance. We identify the factors that result in packet delay reduction with excess bandwidth distribution compared to IPACT-limited and discover that existing offline excess distribution mechanisms become unstable at moderate to high loads in long-range EPONs with large round-trip propagation delays. We propose a novel online excess bandwidth distribution (OEBD) mechanism to provide stable excess bandwidth distribution even at high loads in long-range EPONs. We demonstrate how OEBD can be tuned via parameters to provide grant sizing between limited and gated service. © 2009 Optical Society of America

OCIS codes: 060.2330, 060.4510, 060.4250, 060.4251.

1. Introduction

Excess bandwidth distribution for Ethernet passive optical networks (EPONs) was originally proposed in [1] as an improvement over the limited allocation approach of interleaved polling with adaptive cycle time (IPACT) [2,3], which we refer to as IPACT-limited. IPACT-limited is characterized by a maximum grant size $G_i^{\max}$ [bytes], also frequently referred to as minimum guaranteed bandwidth, for each optical network unit (ONU) i, $i=1,\ldots,M$. Let R_i [bytes] denote the size of the upstream transmission request from ONU i. If ONU i requests less than $G_i^{\max}$, i.e., $R_i \leq G_i^{\max}$, then IPACT-limited grants the full request; i.e., the size of the grant to ONU i is $G_i=R_i$. If the request exceeds $G_i^{\max}$, i.e., $R_i > G_i^{\max}$, then IPACT-limited grants an upstream transmission window for $G_i=G_i^{\max}$ bytes.

The intuitive reasoning behind excess bandwidth distribution is to declare the unused portions $(G_i^{\max}-G_i)$ as excess bandwidth and distribute the total excess bandwidth $\sum_{i=1}^{M}(G_i^{\max}-G_i)$ among the ONUs with $R_i > G_i^{\max}$. Thus, the basic trade-off made with excess bandwidth distribution is to extend the cycle within which all ONUs are served once (to no more than approximately $\sum_{i=1}^{M} G_i^{\max}/C$, with C denoting the upstream transmission rate in bytes/s) so that heavily loaded ONUs can transmit more than $G_i^{\max}$ in a cycle. In contrast, IPACT-limited enforces shorter cycles by strictly limiting upstream transmissions to at most $G_i^{\max}$ in a cycle, resulting in heavily loaded ONUs having to use more cycles to clear backlogs.

To the best of our knowledge, IPACT with limited allocation has not been, in detail, quantitatively compared with the various proposed excess bandwidth distribution mechanisms. That is, the outlined basic trade-off between clearing traffic backlogs with more shorter cycles with IPACT-limited versus fewer longer cycles with excess distribution techniques has not been quantitatively investigated to identify the factors leading to improvements with excess bandwidth distribution. Oftentimes, the excess bandwidth distribution research has focused on quantitatively comparing different excess bandwidth distribution techniques against each other, as reviewed in Section 2.

In this paper, we conduct for the first time, to the best of our knowledge, a detailed quantitative comparison between IPACT-limited and excess bandwidth distribution techniques (a shorter preliminary version of this work was presented in [4]). We conduct extensive simulations to assess the packet delays. We identify the factors that influence the relative performance differences between IPACT-limited and excess bandwidth distribution. We find that existing offline excess bandwidth distribution mechanisms achieve significant delay reductions compared to IPACT-limited for traffic with large bursts in mid-range EPONs. However, we also find that the existing offline excess bandwidth distribution mechanisms suffer from instability problems in long-range EPONs. The architectural and photonics level aspects of long-range EPONs have been explored in several studies (see, for instance, [5–8]), while the examination of the implications of the long reach for the medium access control has just recently begun (see, e.g., [9]).

We propose a novel online excess bandwidth distribution (OEBD) mechanism to overcome the stability problems of existing offline excess bandwidth distribution mechanisms while still maintaining fairness. OEBD accumulates unused bandwidth portions $(G_i^{\max}-G_i)$ in an excess credit pool. Heavily loaded ONUs with $R_i > G_i^{\max}$ are assigned *online* a fair share of the excess credit pool. Through extensive simulation experiments we demonstrate that OEBD has favorable delay performance while avoiding the stability problems of offline excess bandwidth distribution. Further, we demonstrate how OEBD provides tunable performance between IPACT-limited and IPACT gated service, whereby IPACT gated grant sizing always sets $G_i = R_i$.

This paper is structured as follows. In Section 2, we review related work. In Section 3, we compare IPACT-limited with the existing offline excess bandwidth distribution techniques. We first present our simulation setup and then examine the impact of the traffic and network parameters on the performance of IPACT-limited and offline excess bandwidth distribution. In Section 4, we introduce OEBD and examine its performance through simulations. We summarize our conclusions in Section 5.

2. Related Work

Excess bandwidth distribution as part of the dynamic bandwidth allocation process in EPONs [10] has received significant interest in recent years. Following the seminal work by Assi *et al.* [1], several refinements and a range of different mechanisms have been proposed for excess bandwidth distribution. Bai *et al.* [11], for instance, significantly advanced excess bandwidth allocation by introducing a weighted excess division technique that enforces fair division of the total excess bandwidth among the overloaded ONUs requesting more than $G_i^{\max}$. Bai *et al.* [11] also proposed an iterative procedure for allocating excess bandwidth that minimizes wasted bandwidth capacity thereby greatly improving the efficiency of excess bandwidth distribution.

Zheng [12] proposes a mechanism in which the upstream transmission channel is not idle between granting cycles. This is accomplished by always scheduling underloaded ONUs before overloaded ONUs when possible. When necessary to prevent the channel from remaining idle, an overloaded ONU is scheduled with no granted excess, thereby resorting to limited service in an effort to increase channel utilization. OEBD differs from this scheme in that it increases channel utilization without resorting to limited service.

The rolling excess credit pool in OEBD is similar to the sliding-cycle-based excess credit pool (SLICT) proposed by Kim *et al.* [13] in that OEBD allows excess distribution to be performed without using an offline method. OEBD differs from SLICT in the following fundamental way. A maximum granting cycle length is typically bounded by the sum of the guaranteed minimum grant sizes to each ONU. The total excess bandwidth available is then typically computed to be the unused portion of the maximum granting cycle length left over from underloaded ONUs. OEBD uses this typical method of accumulating excess credits. In SLICT, a granting cycle includes an added "shared" subcycle that is used exclusively for distributing excess bandwidth. SLICT accumulates excess credits at the time of grant to ONU i by computing the excess left over in the shared subcycle from the previous $M-1$ grants, thereby facilitating online operation. With SLICT, no excess credits are accumulated from the excess of underloaded ONUs.

ONUs with high traffic variation are identified by the early dynamic bandwidth allocation (E-DBA) proposed by Hwang *et al.* [14]. These high variation ONUs are selected to send a second set of REPORT messages at the end of a granting cycle to improve the prediction accuracy. This improved accuracy comes at the expense of lower channel utilization as a result of transmissions of a second set of REPORT messages from the high variation ONUs. Kim *et al.* [15] present an analysis of the mean queueing delay with differing numbers of classes that are serviced by a traffic prediction mechanism called service quality pre-engagement (SQP).

A modification to IPACT-limited was explored by Lee *et al.* [16], whereby the excess bandwidth of a given ONU is immediately equally distributed among the other ONUs to increase their maximum grant limits. Our OEBD is fundamentally different from the distribution technique in [16] in that it maintains an excess bandwidth credit pool for carrying excess bandwidth across cycles, as explained in Section 4, and supports unequal weight-based excess distribution.

The excess bandwidth allocation refinement by Choudhury and Saengudomlert [17] included some limited performance comparisons with IPACT-limited. Our comparisons are fundamentally different from [17] in that we consider a wide range of scenarios, including a range of round-trip propagation delays, to identify the factors leading to improvements with excess bandwidth distribution over IPACT-limited.

More recently, some excess bandwidth distribution techniques originally proposed for single-channel EPONs have been investigated in the context of wavelength division multiplexing (WDM) EPONs [18,19] with several upstream transmission channels [20]. The performance evaluation in [20] includes comparisons of excess bandwidth distribution with some form of IPACT, namely, the single table extension of IPACT to WDM [21]. Specifically, the gated allocation of IPACT, where the grant is set equal to the ONU request, without any upper limit, was considered. Gated allocation can lead to unfairness because grant sizes are determined solely as a function of reported queue depth. IPACT with limited allocation, which we consider in this study, avoids this fairness problem by strictly limiting the size of a granted upstream transmission.

3. Comparison of IPACT-Limited and Offline Excess Bandwidth Distribution

We initially consider IPACT-limited bandwidth allocation in conjunction with (a) the *offline* scheduling framework, where all ONU reports must be received before allocating bandwidth, which idles the upstream channel for one round-trip time (RTT); (b) the ONU load status *hybrid* scheduling framework, where underloaded ONUs with $R_i \leqslant G_i^{\max}$ are immediately granted bandwidth, and overloaded ONUs with $R_i > G_i^{\max}$ are granted bandwidth once all ONU reports have been received; and (c) the *online* scheduling framework, where all ONUs are immediately granted bandwidth.

Among the many different excess distribution schemes, we focus on the weighted excess division dynamic bandwidth allocation scheme from [11], which enforces fair distribution of the excess bandwidth by dividing the excess according to the weights of the ONUs. We combine this with the iterative excess allocation [11], which iteratively allocates excess bandwidth to ONUs in an effort to maximize the number of satisfied ONUs. By maximizing the number of satisfied ONUs, unused slot remainders are minimized. We refer to this combined excess bandwidth distribution scheme as *iterative*. We consider this iterative scheme in conjunction with (a) the *offline* scheduling framework, whereby all ONU reports need to be received at the optical line terminal (OLT) before commencing the dynamic bandwidth allocation, as well as (b) the ONU load status *hybrid* scheduling framework (referred to as DWBA-2 in [20]), whereby underloaded ONUs receive their grant (online) immediately and the excess bandwidth distribution is executed for the overloaded ONUs (offline), once all the report messages have been received. Note that in both cases, the excess bandwidth distribution operates in offline fashion, in that the excess is allocated only after all ONU reports for the present cycle have been received. We note that we do not consider DBA-3 from [20], which allocates overloaded ONUs (online) immediately the maximum grant size, and then additional *excess bandwidth in offline fashion*, because of its increased complexity, limited delay reductions, and increased wasted bandwidth, as evaluated in [20].

3.A. Simulation Setup

We developed an EPON simulation engine using the CSIM simulation library. We set the upstream transmission bit rate to $C=1$ Gbits/s. We initially consider an EPON with $M=16$ ONUs, each with a 10 Mbyte buffer. We consider round-trip propagation delays RTT (ONU to OLT and back to ONU) for short-range ($[8,10]$ μs corresponding to OLT-to-ONU distances up to 1 km), mid-range ($[13.36,100]$ μs corresponding to OLT-to-ONU distances up to 10 km), long-range ($[0.8,1]$ ms corresponding to OLT-to-ONU distances up to 100 km), and extra-long-range ($[1.6,2]$ ms corresponding to OLT-to-ONU distances up to 200 km) EPONs; for each range, the different ONUs draw their RTT independently at random from a uniform distribution over the respective intervals.

Each ONU independently generates self-similar traffic with a Hurst parameter of 0.75 [22] using 32 traffic sources. The burst size (number of data packets) and time between bursts were independently randomly drawn from Pareto distributions. Following common packet size models, 60% of the packets have 64 bytes, 4% have 300 bytes, 11% have 580 bytes, and 25% have 1518 bytes. Each of the 32 sources of a given ONU has initially a maximum burst size of $B=10$ Mbytes, which is achieved by truncating the Pareto distribution to produce a maximum burst size no greater than $B/1518$ bytes$=6907$ packets. Each ONU contributed equally to the overall traffic load.

For upstream transmission, each data packet is sent with a preamble of 8 bytes and an interpacket gap of 12 bytes (which count toward the upstream transmission grant). Gate and report messages each have 64 bytes and there is a $t_{\text{guard}}=1$ μs guard time between upstream transmissions. For the IPACT-limited and hybrid-iterative bandwidth allocation schemes, we initially set the maximum grant size to $G_{\max}=G_i^{\max}$ $=15{,}500$ bytes for each ONU; the corresponding cycle time is $MG_{\max}/C+Mt_{\text{guard}}$ $=2$ ms.

As a primary performance metric we consider in this paper the average packet queuing delay, defined as the average time interval from the generation of a packet at an ONU to the instant the upstream transmission of the packet commences. We note that the total packet delay in the network would be obtained by adding the packet transmission delay and the one-way propagation delay to the packet queuing delay. Whenever several grants are considered simultaneously in the scheduling, as can arise with the considered offline and hybrid approaches, we employ the largest processing time first (LPT) scheduling policy. The LPT scheduling policy is generally a good policy for minimizing the makespan (total length) of a schedule [23]. We leave more specific delay evaluations, e.g., evaluating the fraction of packets meeting a prescribed deadline, as well as the evaluation of other scheduling policies for future research.

3.B. Impact of Offline, Hybrid, and Online Scheduling

As expected, we observe from Table 1 that online-limited substantially reduces the delay compared to offline-limited since the extra RTT (between the last ONU completing the upstream transmission of a cycle and the first ONU commencing the upstream transmission of the next cycle) is eliminated for *all* grants. (Hybrid-limited, where only overloaded ONUs are scheduled in offline fashion, performed very similarly to online-limited.) With hybrid-iterative the extra RTT is avoided only for the underloaded ONUs, resulting in a relatively smaller delay reduction compared to offline-iterative.

Table 1. Packet Delay in Milliseconds as Function of Load in Mbits/s for Different Scheduling Frameworks[a]

	Load			
	200	400	600	800
Offline-limited	0.30	0.44	0.78	2.34
Offline-iterative	0.21	0.31	0.55	1.47
Online-limited	0.25	0.33	0.54	1.39
Hybrid-iterative	0.20	0.28	0.45	1.10

[a]Fixed parameters: mid-range RTT, $M=16$ ONUs, $B=10$ Mbytes maximum burst size, $G_{\max}=15{,}500$ bytes.

We further observe from Table 1 that for the considered online and hybrid scheduling, the difference between limited and iterative is relatively smaller than for offline scheduling. This is mainly because the cycles in online-limited are much shorter than in offline-limited, reducing the impact of the larger number of cycles needed to work off large traffic bursts. At the same time, it is more likely that upstream transmissions from ONUs requesting less than $G_{\max}$ mask the shorter delay between the upstream transmissions working off a large burst from a given ONU.

To further examine the impact of the larger number of cycles, we simulated a hypothetical EPON with all transmission overhead directly associated with an upstream transmission (guard time as well as report message transmission time) set to zero. For this hypothetical EPON with a load of 800 Mbits/s, online-limited achieves a delay of 1.01 ms compared to hybrid-iterative giving 0.94 ms; i.e., the gap has significantly narrowed compared to the 1.39 ms versus 1.10 ms with all the overheads. This significant narrowing of the gap indicates that the delay difference between online-limited and hybrid-iterative is to a large degree due to the upstream transmission overheads (guard time, report message transmission time), which are experienced more often when transmitting large bursts in more, but shorter cycles with online-limited. Note in particular that each cycle in a real EPON contains a guard time and a report message transmission for each ONU, even if only one or a few ONUs have data to send. These delays cannot be masked by the interleaved transmissions of several ONUs. We proceed to examine online-limited and hybrid-iterative, the best performing approaches from this section, in more detail in the subsequent section.

3.C. Impact of Burst Size B and Maximum Grant Size $G_{\max}$

We examine in Table 2 the impact of smaller burst sizes as well as larger maximum transmission grants. Recall that the results in Table 1 were obtained with a maximum burst size of 10 Mbytes for each of the 32 traffic streams producing the load at a given ONU. We observe from Table 2 that reducing the maximum burst size to 4 Mbytes and further to 65 kbytes reduces the delay difference between online-limited and hybrid-iterative, with both giving essentially the same delays for the 65 kbyte maximum burst size. This is because smaller bursts require fewer cycles for transmission, both with online-limited and hybrid-iterative.

Further, we observe from Table 2 that a larger maximum grant size of $G_{\max}=31{,}125$ bytes compared to the $G_{\max}=15{,}500$ bytes considered in Table 1 narrows the gap between online-limited and hybrid-iterative for the large $B=10$ Mbytes maximum burst size. This is again mainly because of the fewer cycles required to work off bursts, which are this time due to the larger maximum grant size and correspondingly longer cycle.

Table 2. Packet Delay in Milliseconds as a Function of Load in Mbits/s for Different Maximum Burst Sizes B[a]

	Load			
	200	400	600	800
$B=65$ k, $G_{\max}=15{,}500$ B, On.-lim.[b]	0.135	0.147	0.179	0.313
$B=65$ k, $G_{\max}=15{,}500$ B, Hyb.-it.[c]	0.135	0.147	0.179	0.309
$B=4$ M, $G_{\max}=15{,}500$ B, On.-lim.	0.206	0.271	0.421	0.974
$B=4$ M, $G_{\max}=15{,}500$ B, Hyb.-it.	0.180	0.237	0.363	0.790
$B=10$ M, $G_{\max}=15{,}500$ B, On.-lim.	0.250	0.330	0.540	1.390
$B=10$ M, $G_{\max}=15{,}500$ B, Hyb.-it.	0.200	0.280	0.450	1.100
$B=10$ M, $G_{\max}=31{,}125$ B, $M=16$, On.-lim.	0.213	0.289	0.473	1.16
$B=10$ M, $G_{\max}=31{,}125$ B, $M=16$, Hyb.-it.	0.199	0.271	0.440	1.06
$B=10$ M, $G_{\max}=15{,}500$ B, $M=16$, On.-lim.	0.247	0.327	0.534	1.38
$B=10$ M, $G_{\max}=15{,}500$ B, $M=16$, Hyb.-it.	0.202	0.276	0.448	1.10
$B=10$ M, $G_{\max}=15{,}500$ B, $M=32$, On.-lim.	0.281	0.410	0.689	1.80
$B=10$ M, $G_{\max}=15{,}500$ B, $M=32$, Hyb.-it.	0.213	0.312	0.526	1.31

[a]Fixed parameters: $M=16$ ONUs, mid-range RTT.

[b]On.-lim: online-limited.

[c]Hyb.-it.: hybrid-iterative.

Finally, we observe from Table 2 that a larger number of ONUs makes the delay differences between online-limited and hybrid-iterative more pronounced. This is primarily due to the increased upstream transmission overheads (guard time and report transmission time) that are incurred for each ONU once in each cycle.

3.D. Impact of Round-Trip Time *RTT* and Maximum Grant Size G_{max}

In this section we focus on the impact of the RTT, in conjunction with the maximum grant size, on the relative delay performance of online-limited and hybrid-iterative. (Ignore for now the OEBD results in Table 3; these are discussed in Section 4.) We first observe that for the short RTT up to 10 μs, the delays for high loads are very similar to the delays for high loads for RTTs up to 100 μs in Table 2.

Importantly, we observe from Table 3 that hybrid-iterative exhibits a pronounced threshold behavior. For loads below a critical threshold, hybrid-iterative gives substantially smaller delays than online-limited. In fact, the delay differences become more pronounced with increased RTT, with hybrid-iterative achieving delays less than half as large as online-limited for low loads and the extra long RTT up to 2 ms. However, for loads above a critical load threshold, which decreases for increasing RTT, hybrid-iterative becomes unstable and gives excessively large delays. On the other hand, online-limited robustly continues to provide low delays even at very high loads.

The explanation for this behavior is as follows. Consider an extreme scenario in hybrid-iterative where only one ONU has upstream traffic, namely, a very large burst. Then, this overloaded ONU receives all the upstream transmission bandwidth in the cycles, namely, MG_{max} per cycle (neglecting the grants for report messages to the other ONUs). A given cycle consists of an upstream transmission of MG_{max} [bytes], lasting MG_{max}/C [s], plus one RTT (*RTT*) for reporting the remaining size of the backlog and receiving the next grant. In addition, the cycle contains M guard times and the report transmission times of the other $M-1$ ONU's report messages (which we neglect in this approximate analysis). Thus, the maximum sustainable upstream transmission rate is approximately

$$\frac{MG_{max}}{RTT+\dfrac{MG_{max}}{C}+Mt_{guard}}=\frac{G_{max}}{\dfrac{RTT}{M}+\dfrac{G_{max}}{C}+t_{guard}}. \quad (1)$$

We note that this threshold is approximate in that upstream transmissions from underloaded ONUs may mask some of the RTT incurred due to offline excess bandwidth distribution, leading to a higher threshold in an actual EPON. On the other hand, the neglected overheads may slightly reduce the threshold for an actual EPON.

Table 3. Packet Delay in Milliseconds as a Function of Load in Mbits/s for Different Round-Trip Times[a]

	Load			
	200	400	600	800
G_{max}=15,500 bytes, 2 ms cycle				
Short RTT, online-limited	0.158	0.259	0.486	1.36
Short RTT, hybrid-iterative	0.123	0.209	0.398	1.08
Short RTT, OEBD	0.131	0.218	0.412	1.11
Long RTT, online-limited	2.43	2.60	2.93	3.85
Long RTT, hybrid-iterative	1.31	1.44	1.76	>2 s
Long RTT, OEBD	1.82	1.88	2.03	2.61
X-long RTT, online-limited	5.37	6.64	10.13	34.23
X-long RTT, hybrid-iterative	2.57	3.01	>2 s	>2 s
X-long RTT, OEBD	3.65	3.91	4.62	8.91
G_{max}=31,125 bytes, 4 ms cycle				
Long RTT, online-limited	1.66	1.70	1.81	2.30
Long RTT, hybrid-iterative	1.25	1.33	1.51	2.11
Long RTT, OEBD	1.68	1.73	1.87	2.43

[a]Fixed parameters: $M=16$ ONUs, $B=10$ Mbytes maximum burst size.

For the specific realizations of the randomly drawn RTT, which gave an average RTT of 44.2 μs for the mid-range EPON simulation, an average RTT of 0.871 ms for the long-range EPON simulation, and an average RTT of 1.74 ms for the extra-long-range EPON simulation, the approximate theoretical thresholds are 970.5, 691.3, and 530.5 Mbits/s, respectively. Our simulations indicate that these theoretical approximations are quite close to the actual thresholds found in simulations, which are around 912 Mbits/s for the mid-range, 690 Mbits/s for the long-range, and 513 Mbits/s for the extra-long-range scenario. For loads well below the derived threshold, hybrid-iterative is able to provide small delays and to fairly allocate excess bandwidth to overloaded ONUs in offline fashion. When the load grows well above the threshold, then waiting for the excess bandwidth distribution until all report messages are received for a cycle, reduces the capacity so much to render the network effectively unstable.

We note, however, that hybrid-iterative significantly increases the stability limit over offline-limited for the analyzed extreme scenario: when one ONU has a large burst of upstream traffic in offline-limited, the ONU can transmit $G_{\max}$ [bytes], which takes $G_{\max}/C$ [s], before it has to wait for one RTT plus the M guard times. Hence, the upstream transmission rate is limited to approximately $G_{\max}/[RTT+(G_{\max}/C)+Mt_{\text{guard}}]$, which for $M>1$ is significantly smaller than the rate in Eq. (1). To overcome the stability problems due to the offline excess bandwidth distribution in hybrid-iterative, we propose and examine a novel OEBD approach in Section 4.

4. Online Excess Bandwidth Distribution

In this section we introduce and evaluate OEBD for an *online* scheduling framework that makes grant decisions based on a single report; extensions to online just-in-time (JIT) scheduling [24] are left for future work. Recall that we let R_i denote the bandwidth requested by the considered report from a given ONU i, $i=1,\ldots,M$, whereby bandwidth is measured in units of bytes of transmitted data (i.e., corresponds to an upstream transmission window in seconds times the upstream bandwidth in bytes/s), and that we let $G_i^{\max}$ [bytes] be a constant denoting the maximum bandwidth that can be allocated to ONU i in a grant. We let w_i, $0 \leq w_i \leq 1$, $\Sigma_{i=1}^{M} w_i=1$ denote the weight of ONU i in a weighted fair excess division [11]. We define an excess bandwidth credit pool and let E_t [bytes] denote the current total amount of bandwidth credits in the excess pool. In addition, we let δ, $0 \leq \delta \leq 1$, be a constant decay factor, and denote N, $N \geq 1$, for the decay interval.

The OEBD bandwidth allocation proceeds as follows. If the considered ONU i is underloaded, i.e., requests less than the prescribed maximum allocation ($R_i \leq G_i^{\max}$), then the bandwidth R_i is granted and the excess $G_i^{\max}-R_i$ is added to the excess bandwidth pool E_t. If the considered ONU is overloaded, i.e., requests more than its prescribed maximum allocation ($R_i > G_i^{\max}$), then the ONU is allocated its prescribed maximum $G_i^{\max}$ plus up to $w_i E_t$ excess bandwidth from the pool. With a controlled excess allocation technique, the allocation is capped at R_i; i.e., the ONU is allocated $\min\{G_i^{\max}+w_i E_t, R_i\}$. Accordingly, the excess pool is reduced by $\min\{G_i^{\max}+w_i E_t, R_i\}$. In addition, after every N grants, we "decay" the pool, E_t, using the multiplicative constant δ (i.e., $E_t \leftarrow \delta E_t$).

We note that we outlined the OEBD approach for a single polling thread per ONU. Multithread polling [9] has recently been proposed to reduce the queueing delays in long-range EPONs by employing multiple polling threads per ONU. Multithread polling has been evaluated in conjunction with IPACT with online scheduling and gated grant sizing in [9]. OEBD and multithread polling are complementary in that OEBD affects the size of a given grant whereas multithread polling affects the frequency of grants to a given ONU. Consequently, OEBD and multithread polling could be combined by executing the OEBD steps for every polling thread. A quantitative study of such multithread polling with OEBD is an important direction for future research.

4.A. Simulation Results: Delay and Channel Utilization

We conducted a set of simulation experiments to understand the delay and channel utilization performance of OEBD compared to online-limited, online-gated, and hybrid-iterative. We used the same simulation parameters described in Subsection 3.A. In addition, we varied the two parameters that manage the rolling excess credit

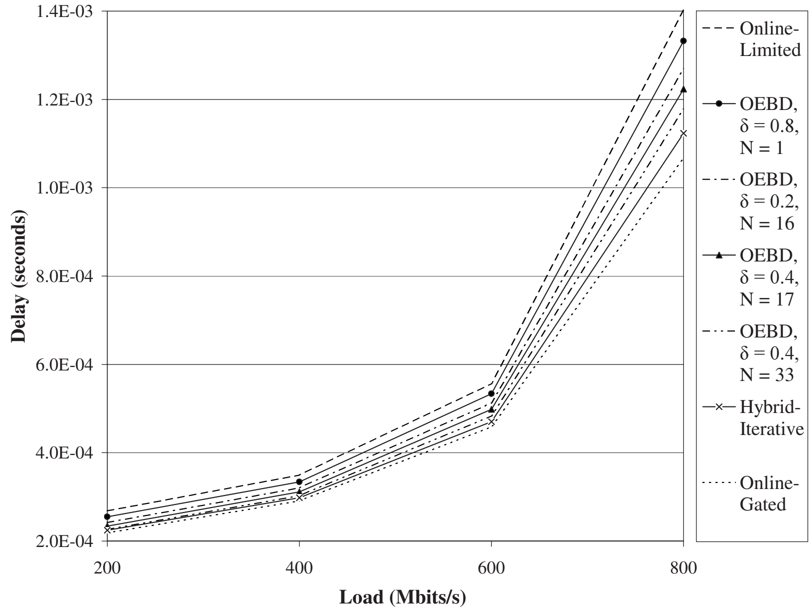

Fig. 1. Average queueing delay for mid-range RTT (up to 10 km).

pool in OEBD: (1) the multiplicative decay factor, δ, and (2) the decay interval, N. We varied δ from 0 to 1 in increments of 0.2. We selected values for N to decay the pool after every grant, $N=1$; after every M grants, $N=M$; and after every $2M$ grants, $N=2M$. We also selected N values that were one larger than M and $2M$ to remove a bias that occurs when the credit decay occurs at a multiple of M. Specifically, if the decay interval is a multiple of M then the same ONU could potentially receive a grant after every excess pool decay causing this ONU to consistently see a smaller excess pool than other ONUs. To remove this bias, the decay interval, N, must be set to a value that is not a multiple of the number of ONUs, M. This will ensure that the excess pool decay instant will rotate among all ONUs.

Figure 1 shows the average queueing delay for load values between 0.2 and 0.8 Gbits/s for an EPON with mid-range RTT (i.e., up to 10 km). Figure 2 shows the same for an EPON with a maximum range of 100 km. The figures present the parameter settings that provided the most insightful results. We notice that the performance of OEBD can be tuned, using δ and N, to provide service between limited and gated.

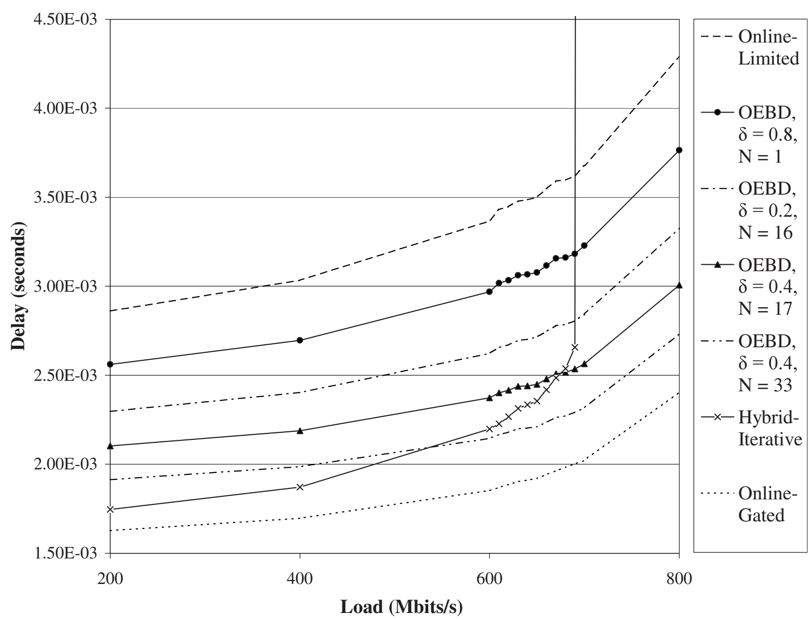

Fig. 2. Average queueing delay for long-range RTT (up to 100 km).

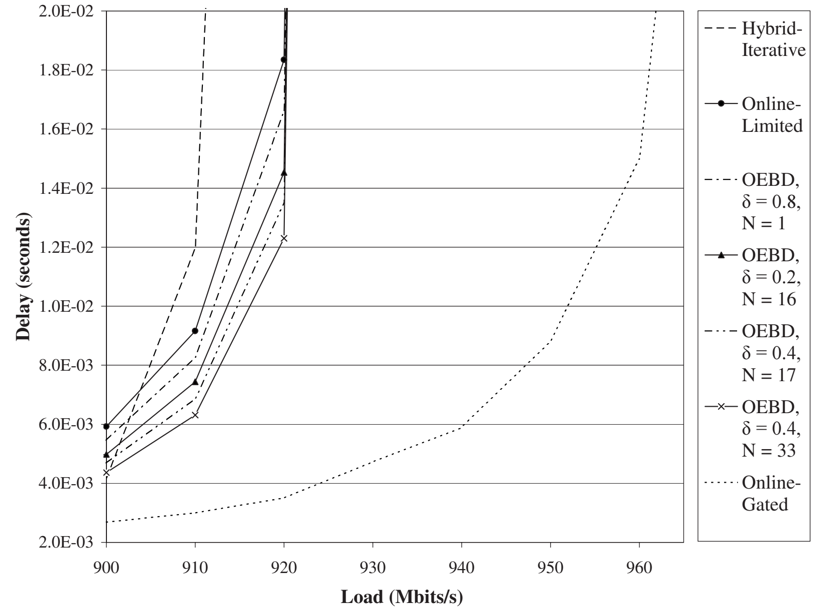

Fig. 3. Average queueing delay for mid-range RTT (up to 10 km).

With $\delta=0$ and $N=1$, the excess credit pool is set to zero after every grant and therefore excess credits never accumulate. In the absence of excess credits, limited service is provided. Therefore, OEBD with $\delta=0$ and $N=1$ provides limited service. When $\delta=1$, the excess credit pool is never decayed. As a result, an infinite number of excess credits can accumulate, leading to a gated service. When all ONUs present an equal load, as is the case in this set of simulation experiments, the average queueing delay trends lower as OEBD approaches a gated service. When there is even a single ONU that is attempting to present more than its fair share of bandwidth, in the long term, this may no longer be true.

We also notice that for EPONs with mid-range RTT, the benefit of minimizing channel idleness by using OEBD is generally not sufficient to provide better delay performance than hybrid-iterative. However, when EPONs have a longer range RTT (for example, up to 100 km) OEBD can provide lower average queueing delays with the correct parameters. Specifically, those parameter settings that decay the excess pool slower can provide lower delay (i.e., large δ and large N).

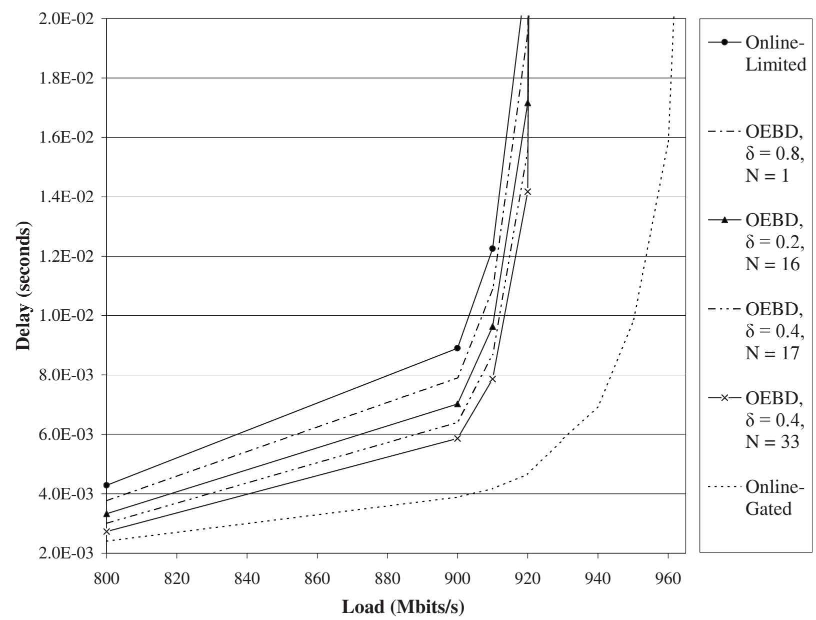

Fig. 4. Average queueing delay for long-range RTT (up to 100 km).

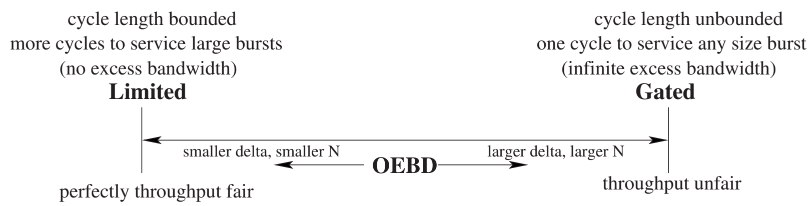

Fig. 5. Grant sizing continuum.

To examine the performance as the load approaches the channel utilization limits of each of the DBA methods we conducted simulation experiments with load values approaching 0.98 Gbits/s. Figure 3 shows the average queueing delay for an EPON with mid-range RTT (i.e., up to 10 km), while Fig. 4 shows the same for an EPON with a maximum range of 100 km. Hybrid-iterative does not show up in Fig. 4 as it becomes unstable at a much lower load level of around 690 Mbits/s. We observe from these figures that gated service provides the highest maximum channel utilization and lowest delay. We also observe that hybrid-iterative has the lowest maximum channel utilization (0.91 Gbits/s for mid-range RTT and 0.69 Gbits/s for long-range RTT) and OEBD lies between the two (0.92 Gbits/s for both RTT ranges). Hybrid-iterative idles the channel while waiting for all REPORT messages to be received in order to determine the excess bandwidth to be distributed to overloaded ONUs. This wasted channel capacity increases with increasing RTT leading to an inversely proportional relationship between the channel utilization limit and the RTT range for hybrid-iterative. OEBD avoids this channel idling by servicing ONUs in online fashion. As a result, OEBD can achieve higher channel utilization that is immune to increases in RTT ranges.

The channel utilization achieved by OEBD is still less than gated because for large bursts from ONUs, OEBD requires several granting cycles to service those bursts compared to gated, which requires only one granting cycle. The more granting cycles, the more channel capacity is consumed by upstream transmission overheads.

To illustrate the delay performance tuning capability of OEBD and associated trade-offs, in Fig. 5 we present a grant sizing continuum that exists between limited and gated service. The decay parameters (i.e., δ and N) determine where any specific implementation of OEBD lies on the continuum.

4.B. Simulation Results: Fairness

To understand how fair OEBD can be with respect to distributing excess bandwidth, we benchmarked the excess fairness of OEBD in our simulation experiments. Figures 6 and 7 display the measured fairness indices. We use the fairness metric presented in

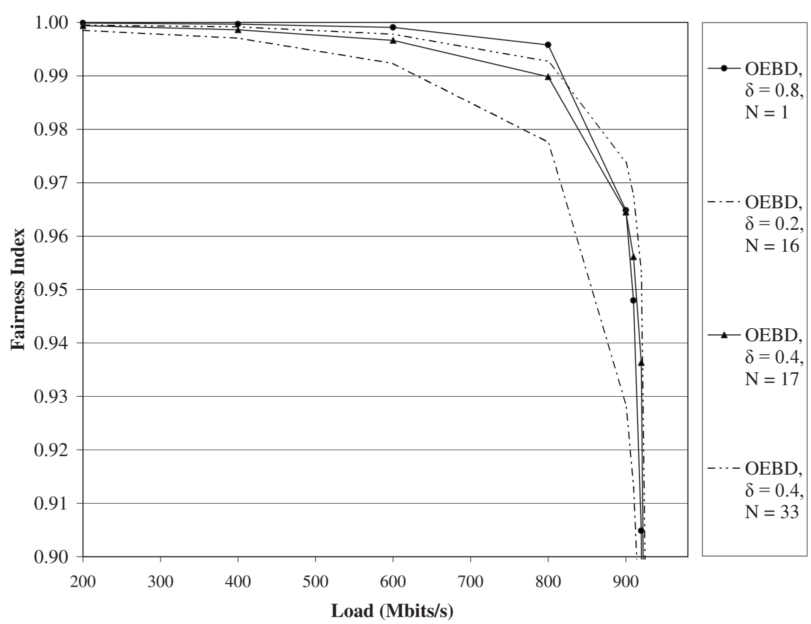

Fig. 6. Excess fairness benchmark, f_e, for mid-range RTT (up to 10 km).

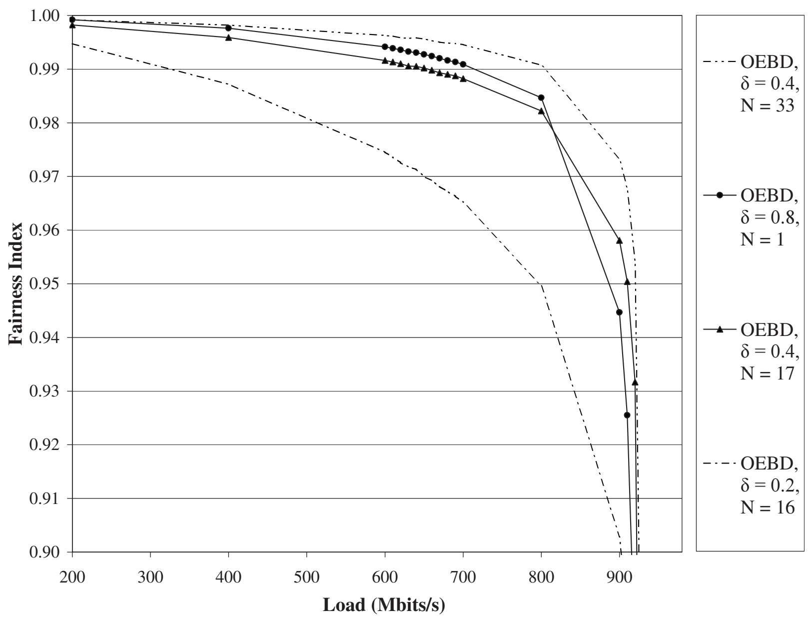

Fig. 7. Excess fairness benchmark, f_e, for longer-range RTT (up to 100 km).

[11] that is derived from Jain's fairness index [25]. Let f_e be the excess fairness index, M be the total number of ONUs, E_i be the excess bandwidth received by ONU i, and w_i be the weight of ONU i, then

$$f_e = \frac{\left(\sum_{i=1}^{M} \dfrac{E_i}{w_i}\right)^2}{M\sum_{i=1}^{M}\left(\dfrac{E_i}{w_i}\right)^2}.$$

We observe in Figs. 6 and 7 that the OEBD parameters have an impact on the fairness of excess distribution. Specifically, the parameter settings that decay the excess credits slowly (i.e., large δ, large N) exhibit improved excess fairness (i.e., $f_e \rightarrow 1$). We leave a more detailed fairness evaluation for a future study.

5. Conclusion

We have examined the delay and fairness performance of conventional IPACT with limited allocation and existing excess bandwidth allocation strategies, which allocate excess in an offline fashion. We discovered that offline excess bandwidth allocation significantly reduces the delay compared to IPACT-limited for traffic with large bursts and EPONs with mid-range RTTs. For traffic with small bursts or EPONs with short RTTs, IPACT-limited achieves delays almost as low as with offline excess bandwidth distribution.

Importantly, we found that for long-range EPONs with large RTTs, offline excess bandwidth distribution exhibits a pronounced threshold behavior: for loads below a critical threshold, offline excess bandwidth distribution provides lower delays than IPACT-limited. For loads above the threshold, offline excess bandwidth distribution becomes unstable, resulting in excessively large delays, whereas IPACT-limited continues to achieve small delays.

We introduced online excess bandwidth distribution (OEBD) to overcome the stability problems of the existing offline excess bandwidth distribution mechanisms. We found that OEBD with the correct parameter settings (i.e., δ and N) can provide lower queueing delays for long-range EPONs. Even for mid-range EPONs, we have found that OEBD provides a higher delay stability limit (i.e., higher maximum achievable channel utilization). The difference in stability limit is exacerbated for longer range EPONs.

We have also found that the grant sizing service provided by OEBD can be tuned between limited (δ=0 and N=1) and gated (δ=1). This flexibility will allow network designers to selectively trade off lower queueing delays with fairness guarantees.

There are numerous important avenues for future research on OEBD. One important direction is to further comprehensively study the parameter setting for OEBD to ensure robust, good performance across a wide range of scenarios. Another avenue is to examine the compatibility of OEBD with emerging dynamic bandwidth allocation strategies for long-range EPONs, such as multithread polling [9]. Furthermore, not only the delay performance, but also the fairness performance of OEBD requires careful evaluation.

References

1. C. Assi, Y. Ye, S. Dixit, and M. Ali, “Dynamic bandwidth allocation for quality-of-service over Ethernet PONs,” IEEE J. Sel. Areas Commun. **21**, 1467–1477 (2003).
2. G. Kramer, B. Mukherjee, and G. Pesavento, “IPACT: a dynamic protocol for an Ethernet PON (EPON),” IEEE Commun. Mag. **40**(2), 74–80 (2002).
3. G. Kramer, B. Mukherjee, S. Dixit, Y. Ye, and R. Hirth, “Supporting differentiated classes of service in ethernet passive optical networks,” J. Opt. Netw. **1**, 280–298 (2002).
4. J. Ferguson, M. Reisslein, and M. McGarry, “When are online and offline excess bandwidth distribution useful in EPONs?” in *Proceedings of the Third ICST International Conference on Access Networks* (ICST, 2008), pp. 50–59.
5. J. Lazaro, J. Prat, P. Chanclou, G. Tosi Beleffi, A. Teixeira, I. Tomkos, R. Soila, and V. Koratzinos, “Scalable extended reach PON,” in *Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference*, OSA Technical Digest (CD) (Optical Society of America, 2008), paper OThL2.
6. D. Shea and J. Mitchell, “A 10 Gb/s 1024-way split 100-km long reach optical access network,” J. Lightwave Technol. **25**, 685–693 (2007).
7. G. Talli and P. Townsend, “Hybrid DWDM-TDM long-reach PON for next-generation optical access,” J. Lightwave Technol. **24**, 2827–2834 (2006).
8. I. Van de Voorde, C. Martin, J. Vandewege, and X. Qiu, “The superPON demonstrator: an exploration of possible evolution paths for optical access networks,” IEEE Commun. Mag. **38**(2), 74–82 (2000).
9. H. Song, A. Banerjee, B.-W. Kim, and B. Mukherjee, “Multi-thread polling: a dynamic bandwidth distribution scheme in long-reach PON,” in *IEEE Global Telecommunications Conference* (IEEE, 2007), pp. 2450–2454.
10. J. Zheng and H. Mouftah, “Media access control for Ethernet passive optical networks: an overview,” IEEE Commun. Mag. **43**(2), 145–150 (2005).
11. X. Bai, A. Shami, and C. Assi, “On the fairness of dynamic bandwidth allocation schemes in Ethernet passive optical networks,” Comput. Commun. **29**, 2123–2135 (2006).
12. J. Zheng, “Efficient bandwidth allocation algorithm for Ethernet passive optical networks,” IEE Proc.-Commun. **153**, 464–468 (2006).
13. H. Kim, H. Park, D. K. Kang, C. Kim, and G. I. Yoo, “Sliding cycle time-based MAC protocol for service level agreeable Ethernet passive optical networks,” in *IEEE International Conference on Communications* (IEEE, 2005), pp. 1848–1852.
14. I.-S. Hwang, Z. Shyu, L.-Y. Ke, and C.-C. Chang, “A novel early DBA mechanism with prediction-based fair excessive bandwidth reallocation scheme in EPON,” in *Sixth International Conference on Networking* (IEEE, 2007), pp. 75–80.
15. N. Kim, H. Yun, and M. Kang, “Analysis of effect of load-based excess bandwidth reservation on performances of differentiated services in E-PON,” IET Commun. **1**, 382–390 (2007).
16. S. Lee, T.-J. Lee, M. Chung, and H. Choo, “Adaptive window-tuning algorithm for efficient bandwidth allocation on EPON,” Lect. Notes Comput. Sci. **4479**, 1217–1220 (2007).
17. P. Choudhury and P. Saengudomlert, “Efficient queue based dynamic bandwidth allocation scheme for Ethernet PONs,” in *IEEE Global Telecommunications Conference* (IEEE, 2007), pp. 2183–2187.
18. M. McGarry, M. Reisslein, M. Maier, and A. Keha, “Bandwidth management for WDM EPONs,” J. Opt. Netw. **5**, 637–654 (2006).
19. A. Banerjee, Y. Park, F. Clarke, H. Song, S. Yang, G. Kramer, K. Kim, and B. Mukherjee, “Wavelength-division-multiplexed passive optical network (WDM-PON) technologies for broadband access: a review,” J. Opt. Netw. **4**, 737–758 (2005).
20. A. Dhaini, C. Assi, M. Maier, and A. Shami, “Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks,” J. Lightwave Technol. **25**, 277–286 (2007).
21. K. Kwong, D. Harle, and I. Andonovic, “Dynamic bandwidth allocation algorithm for differentiated services over WDM EPONs,” in *The Ninth International Conference on Communications Systems* (IEEE, 2004), pp. 116–120.
22. K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation* (Wiley-Interscience, 2000).
23. M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, 3rd ed. (Springer, 2008).
24. M. McGarry, M. Reisslein, C. Colbourn, M. Maier, F. Aurzada, and M. Scheutzow, “Just-in-time scheduling for multichannel EPONs,” J. Lightwave Technol. **26**, 1204–1216 (2008).
25. R. Jain, A. Durresi, and G. Babic, “Throughput fairness index: an explanation,” ATM Forum (1999), http://www.cse.wustl.edu/jain/atmf/a99-0045.html.