Shortest Propagation Delay (SPD) First Scheduling for EPONs with Heterogeneous Propagation Delays

Michael P. McGarry, Martin Reisslein, Frank Aurzada, and Michael Scheutzow

Abstract— Due to the geographic distribution of its subscribers, Ethernet Passive Optical Networks (EPONs) have typically varying propagation delays between the Optical Network Units (ONUs) and the Optical Line Terminal (OLT). In this paper, we consider EPONs with an offline scheduling framework, which enables Quality-of-Service mechanisms by collecting bandwidth requests from all ONUs before the OLT makes dynamic bandwidth allocations for transmissions on the shared ONUs-to-OLT upstream channel. We propose and evaluate the Shortest Propagation Delay (SPD) first scheduling policy which sequences the ONUs' upstream transmissions in increasing order of the ONUs' propagation delays, i.e., the upstream transmission of the ONU with the smallest propagation delay is scheduled first. We formally analyze the competitiveness of SPD first scheduling and find that it achieves very close to optimal performance. We characterize the stability limit for Gated and Limited grant sizing in conjunction with SPD grant scheduling. We evaluate the cycle length and packet delay with SPD scheduling through probabilistic analysis and simulations and find significant reductions in packet delay with SPD first scheduling in EPONs with heterogeneous propagation delays, especially when Limited grant sizing is employed.

Index Terms— Ethernet Passive Optical Network, Grant scheduling, Packet delay, Propagation delay.

I. INTRODUCTION

Passive Optical Networks (PONs) have emerged as an attractive technology for high-speed access networks [1]–[4]. In particular, the combination of PON technologies with the ubiquitous Ethernet networking technologies has made Ethernet PON (EPON) a promising access network choice [5]–[14]. In Ethernet Passive Optical Networks (EPONs), an Optical Network Unit (ONU) provides high-speed network access to an individual subscriber or a group of subscribers. Several ONUs connect to a single Optical Line Terminal (OLT), typically in the form of a tree topology rooted at the OLT. Due to the geographic distribution of the served subscribers, the individual ONUs have typically different distances, and thus different propagation delays from the OLT. With the emergence of long reach and next-generation PONs [15]–[20] covering larger geographic areas with spans of 100 km or higher (i.e., one-way ONU-to-OLT propagation delays of

0.5 ms or higher), the disparities of the propagation delays are likely becoming more pronounced. Further, trends to consolidate central offices in fewer locations give rise to the need to serve ONUs distributed over large geographic areas [21]–[23].

EPONs avoid collisions on the shared upstream (ONUs-to-OLT) channel through a polling based medium access control protocol. The ONUs signal their bandwidth demands with REPORT messages to the OLT, while the OLT dynamically allocates bandwidth and schedules the upstream transmissions so as to avoid collisions. The OLT signals the ONUs with GATE messages their upstream transmission windows (grants). A key challenge for efficient sharing of the upstream channel is the masking of the round trip propagation delay between OLT and ONUs. One of the first approaches for masking the propagation delays has been the Interleaved Polling with Adaptive Cycle Time (IPACT) approach [9], [24] which interleaves the REPORT-GATE cycles of the individual ONUs so that they can mask each others propagation delays. The basic IPACT approach implements the online scheduling framework [25] in that the OLT considers a single ONU REPORT when making bandwidth allocation and scheduling decisions; the online scheduling framework is referred to as interleaved polling in [14], [26].

Quality of Service (QoS) control generally requires that the OLT considers and trades off requests from several ONUs when making bandwidth allocation and scheduling decisions. With the offline scheduling framework [25], which is referred to as interleaved polling with stop in [14], [26], the OLT collects REPORTs from all ONUs before making bandwidth allocation and scheduling decisions. The offline scheduling framework thus enables the wide variety of QoS mechanisms, see for instance, [27]–[35], which consider jointly all REPORTs in their bandwidth allocation and scheduling decisions. On the downside, the offline scheduling framework imposes an idle period on the upstream channel between cycles due to the OLT schedule computation time, and transmission time of the first GATE message and the round trip propagation delay to the first scheduled ONU of a new cycle, as described in more detail in Section III-B. Further idle periods are possible if an upstream transmission is not long enough to mask the propagation delay to the next ONU in the schedule.

A few studies have pursued strategies that combine online scheduling and offline scheduling. For instance, the studies [36]–[39] schedule ONUs with small bandwidth requests immediately (i.e., in online fashion), while ONUs with large bandwidth requests are only scheduled after REPORTs from all ONUs have been collected (i.e., in offline fashion) and more informed decisions are possible. When the ONU propagation

M. McGarry is with the Dept. of Electrical and Computer Eng., University of Akron, Akron, OH 44325, Email: mmcgarry@uakron.edu, Phone: (330) 972-7168, Fax: (330) 972-6487.

M. Reisslein is with the School of Electrical, Computer, and Energy Engineering, Arizona State University, Tempe, AZ 85287-5706, Email: reisslein@asu.edu, Phone: (480) 965-8593, Fax: (480) 965-8325.

F. Aurzada and M. Scheutzow are with the Institute of Mathematics, Technical University Berlin, Berlin, Germany, Email: {aurzada, ms}@tu-berlin.de.

delays are fairly homogeneous, scheduling small bandwidth request (which have a small impact on the QoS and fairness properties of the schedule) right away can indeed be a good strategy to mask propagation delays for the larger requests, which require more careful informed decisions. However, when the propagation delays of the ONUs are vastly different then scheduling a small grant for a far-away ONU can result in large idle times. Thus, for EPONs with heterogeneous propagation delays, the scheduling decisions need to take the propagation delays into consideration.

In this paper, we examine, to the best of our knowledge, for the first time the problem of efficiently masking heterogeneous propagation delays in EPONs with offline scheduling. The problem of accommodating heterogeneous EPON propagation delays has previously only been examined in [40] for an online scheduling framework. We propose and evaluate the Shortest Propagation Delay (SPD) first scheduling policy. The SPD first policy strives to mask the long round trip propagation delays to far-away ONUs by first scheduling the upstream transmissions of near-by ONUs. We prove that the SPD first policy minimizes the cycle length to within a small time period (number of ONUs times transmission delay of GATE message) of an optimal scheduling policy. We characterize the cycle length and packet delay of SPD first scheduling for Gated grant sizing in low load and high load regimes through probabilistic analysis and derive stability limits for Limited grant sizing [9]. We conduct extensive simulations to verify our analysis and to broadly assess the reductions in cycle length and packet delay as well as the increase in channel utilization achieved with SPD first scheduling.

Importantly, by including a sufficient number of close-by ONUs with small propagation an EPON using SPD first scheduling can be engineered to allow for offline scheduling with a very small imposed idle time between scheduling cycles. Further, SPD first scheduling is very simple in that a given set of served ONUs needs to be sorted only once in increasing propagation delays.

This article is organized as follows. In Section II we provide background on scheduling in EPONs and review related work. In Section III we formally model the grant scheduling problem with heterogeneous propagation delays and characterize the competitiveness of SPD first scheduling. We also derive approximations of the mean cycle length and packet delay. In Section IV we present numerical results from our analytical cycle length and delay evaluation and provide extensive simulation results for SPD first scheduling. Finally, in Section V we summarize our findings.

II. BACKGROUND AND RELATED WORK

In this section we briefly provide background on the dynamic bandwidth allocation in EPONs and review related research on scheduling in EPONs. The dynamic bandwidth allocation (DBA) in EPONs can be divided into: 1) the sizing of the upstream transmission windows (grants), and 2) the scheduling of the grants on the upstream wavelength channel [11]. Widely used basic grant sizing methods are *Gated* grant sizing, where the OLT sets the grant size equal to the ONU request and *Limited* grant sizing, where the OLT sets the grant size equal to the ONU request up to a maximum grant size; if request exceeds the maximum grant size, then the maximum grant size is allocated [9], [24]. QoS mechanisms for EPONs, such as [27]–[35] typically control the grant size to achieve specific QoS objectives. In this study, we consider the grant size as given and focus on the scheduling of the grants.

As noted in the Introduction, the basic online scheduling framework considers and schedules one ONU request at a time, whereas the offline scheduling framework collects reports from all ONUs before making scheduling decisions [14], [25], [26]. A few studies have examined a just-in-time scheduling framework, where ONU requests are collected and scheduling decisions are made when the channel is about to become idle [25], [41], [42]. In [43] the ONUs are split into two groups whereby each group is scheduled in offline fashion and the cycles of the two groups are interleaved to mask the idle time in between cycles. We also note that a few studies have sought to improve on the REPORT-GATE traffic signaling through traffic prediction, see e.g., [44]–[46]. The offline scheduling framework is important as it facilitates QoS mechanisms, such as [27]–[35], by providing requests from all ONU to be considered simultaneously in the dynamic bandwidth allocation. In this first study on grant scheduling in EPONs with heterogeneous propagation delays we focus on the offline scheduling framework. We leave the study of grant scheduling for heterogeneous propagation delays in EPONs with just-in-time scheduling, a combination of online and offline scheduling, or traffic prediction for future research.

We proceed to briefly review the research on grant scheduling policies in EPONs. The research on grant scheduling has primarily examined scheduling policies for the offline scheduling framework, where all ONU requests are considered in scheduling decisions. Scheduling policies for combinations of the online and offline scheduling frameworks and for the just-in-time scheduling framework, where a subset of the ONUs are considered, have also been studied. A number of studies have examined scheduling policies that provide pre-scribed QoS differentiation or fairness properties, e.g., [27]–[39]. We focus in this first study on heterogeneous propagation delays on minimizing the average packet delay; considering heterogeneous propagation delays in conjunction with QoS differentiation and fairness mechanisms are important directions for future research. Existing scheduling policies for minimizing the average packet delay include:

- Earliest Arrival First (EAF) scheduling [14], [47] which orders ONUs by the arrival time of the head of line packet.
- Shortest Processing Time (SPT) first scheduling [25], [41] which orders ONUs by their grant size.
- Largest Processing Time (LPT) first scheduling [14], [31] which orders ONUs by their grant size (descending order).
- Largest Number of Frames (LNF) first scheduling [25] which orders ONUs by the number of frames queued (descending order).

A recent comparison found that LNF provided slightly smaller or the same average queueing delays [25] than the other policies and we consider therefore LNF as a benchmark in our performance evaluation in Section IV.

We also note that efforts to mask propagation and other system delays (such as laser tuning times) have been examined for medium access control and scheduling in WDM star networks, e.g., [48], [49]. These WDM star networks provide all-to-all connectivity and are thus fundamentally different from the EPON tree network, where only the OLT can reach all ONUs.

III. PERFORMANCE ANALYSIS

A. Network Model and Notation

We consider the EPON reporting and granting cycle with the offline scheduling framework, which is illustrated in Fig. 1 for $N = 3$ ONUs. We denote t_{sched} for the schedule computation time, i.e., the time duration from the instant when all REPORT messages have been received at the OLT to the instant the transmission of the first GATE message commences. We denote t_G for the fixed transmission time [in seconds] of an MPCP GATE message, t_g for the fixed guard time [in seconds] required between ONU transmission windows, and t_R for the fixed transmission time of a MPCP REPORT message.

We let the constants τ_i, $i = 1, \ldots, N$, denote the one-way propagation delays [in seconds] between OLT and ONU i (which we consider to be equal to the ONU i to OLT propagation delay). We let $\tau_{(i)}$, $i = 1, \ldots, N$, denote the propagation delays sorted in ascending order, i.e., $\tau_{(1)} = \min_i \tau_i$ and $\tau_{(N)} = \max_i \tau_i$. We note that large propagation distance ranges and correspondingly large propagation delay ranges $\tau_{(N)} - \tau_{(1)}$ can result in large dynamic ranges in the signals received at the OLT. The integration research system designed in [50] has a dynamic range of 11.6 dB (see [50, Fig. 12]) accommodating a propagation distance ranges of over 40 km for 0.25 db/km fiber loss and other system parameters (number of splitters and laser power) being equal. The dynamic range of some commercially available OLT receivers is 20 dB (e.g., see [51]), allowing for distance ranges of up to 80 km. With the ongoing advances in optical receivers, see for instance [52]–[54], it is reasonable to expect that larger dynamic ranges will become available in the near future. Furthermore, reach extenders for positive power gain [23], [55]–[58] and optical attenuators for negative power gain can be used to adjust for other loss differences and/or for further extending propagation delay ranges.

For a given cycle, we let R_i be a random variable denoting the reported queue depth (in units of seconds of upstream transmission time) and G_i [in seconds] be a random variable denoting the duration of the upstream transmission window (grant) of ONU i, $i = 1, \ldots, N$. We suppose that G_i includes all "per-Ethernet frame" overheads, such as Preamble and Inter Packet Gap (IPG).

For a given cycle, we define the cycle length Γ as the time period from the instant the scheduling commences to the instant the upstream transmissions of the cycle are completely received. We define the upstream channel utilization η as the ratio of the sum of upstream transmission windows to the cycle length, i.e., $\eta = \sum_{i=1}^{N} G_i / \Gamma$

B. Problem Overview

Generally, in order to minimize packet delays and maximize the utilization on the EPON upstream channel, idle periods on the upstream channel should be minimized. In turn, with minimal idle periods, the cycle length Γ is minimized. With the offline scheduling framework, there is an idle period (stall time) between the instant the end of the last upstream transmission of the preceding cycle arrives at the OLT and the instant the beginning of the first upstream transmission of the current cycle arrives at the OLT. Clearly, this first stall time is minimized by sending the first GATE message to the ONU with the shortest propagation delay, which results in

$$t_1^{\text{stall}} = \max(t_{\text{sched}} + t_G + 2\tau_{(1)}, \ t_g). \tag{1}$$

For illustration of the problem suppose that next the GATE messages and upstream transmissions of ONUs 2 and 3 follow, see Fig. 1. If the first upstream transmission is too short to mask the round-trip propagation delay to ONU 2, a stall time t_2^{stall} occurs between the end of the first upstream transmission and the beginning of the reception of the second upstream transmission. More specifically, if

$$t_1^{\text{stall}} + G_1 + t_R + t_g < t_{\text{sched}} + 2t_G + 2\tau_2, \tag{2}$$

then a non-zero stall time t_2^{stall} occurs. On the other hand, as illustrated for ONU 3 in Fig. 1, if the round-trip propagation delay is masked by a preceding upstream transmission, then there is no stalling.

Note that in the illustration in Fig. 1, the sequence of the GATE message transmissions is equal to the sequence of upstream transmissions reaching the OLT. These two sequences do not necessarily need to be the same. In fact, one can relatively easily construct examples where first sending the GATE message to a far-away ONU, followed by sending the GATE messages and receiving the upstream transmissions of near-by ONUs, followed by the reception of the upstream transmission from the far-away ONU minimizes idle periods, and thus the cycle length. We also note that [24] briefly mentioned that GATE messages to far-away ONUs may need to be sent before GATE messages to near-by ONUs to achieve close to continuous utilization of the upstream channel, but did not analyze in any detail the scheduling of the upstream transmission windows.

For ease of notation, we include the REPORT transmission time t_R in the duration of the upstream transmission grant G_i in Sections III-C through III-E.

C. Solution Strategy

The scheduling of the upstream transmissions can be viewed as a generalized version of the scheduling problem with release times, i.e., times when a given job becomes eligible for execution. Even for fixed known release times, the problem of minimizing the total completion time is strongly NP-hard [59]. Our problem is more general in that the release times, i.e., the times when upstream transmissions can at the very earliest
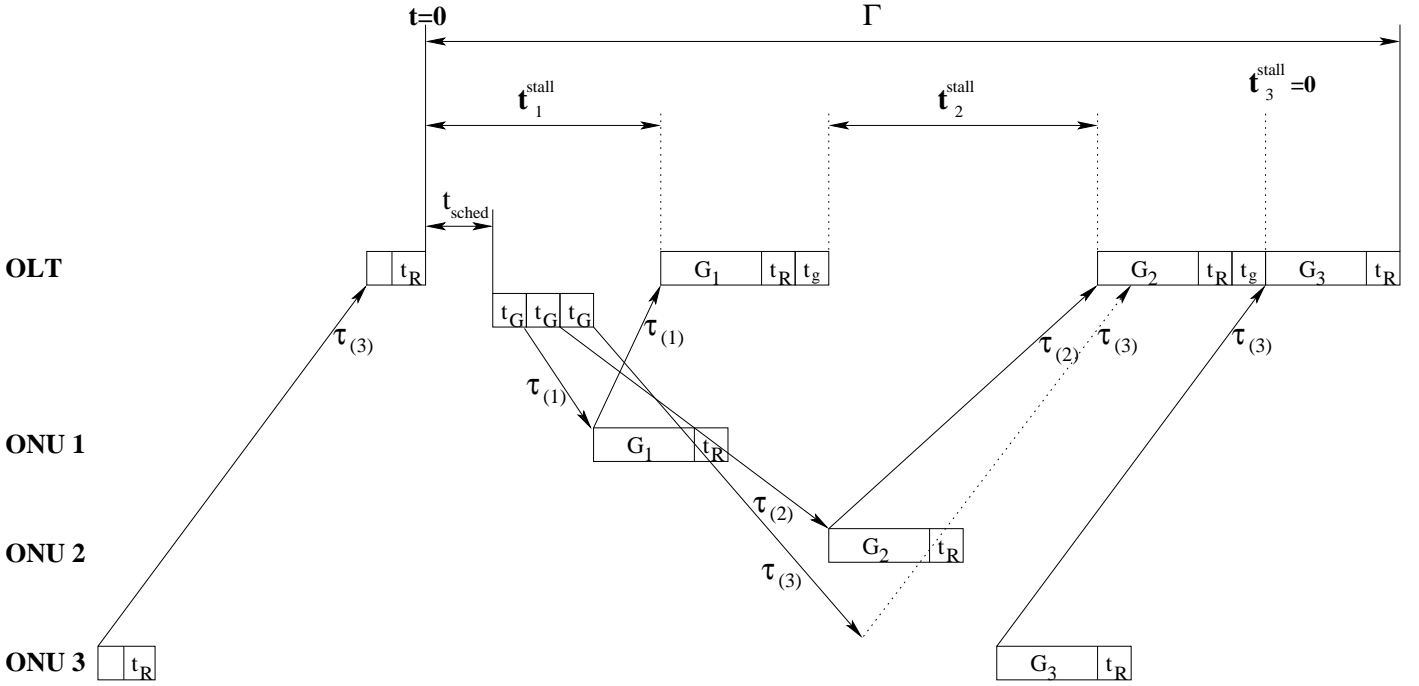
Fig. 1. Illustration of offline scheduling with three ONUs. The horizontal axis represents time and the nodes OLT, ONU 1, ONU2, and ONU 3 are illustrated on the vertical axis; the vertical axis is not scaled for propagation delay. The upstream channel experiences an idle (stall) period due to the scheduling computation time t_{sched}, transmission time of first GATE message t_G, and round trip propagation delay $2\tau_{(1)}$. There are further stall periods if the upstream transmission of an ONU does not mask the round trip propagation delay of the next ONU.

arrive at the OLT depend on the sequencing of the GATE message transmissions.

To the best of our knowledge, the combined problem of scheduling the sequence of GATE message transmissions and upstream transmissions so as to minimize the cycle time is mathematically intractable. Our solution strategy is to consider two restrictions: $(R1)$ We suppose that the GATE message transmission time t_G is equal to the REPORT message transmission time t_R, This is reasonable as both of these MPCP messages are sent in minimum-length Ethernet frames. Noting that every upstream transmission must at least contain a REPORT to obtain the current queue depth of the ONU, $t_G = t_R$ implies that $G_i \geq t_G \; \forall i$. $(R2)$ We initially suppose that the sequence of the GATE message transmissions is equal to the sequence of the upstream transmissions arriving at the OLT. Note that even with this restriction, the problem of scheduling the upstream transmissions is a generalized version of the scheduling with release times in that the release times of the jobs are not fixed, but rather depend on their position in the schedule (i.e., the number of t_G delays in the release time varies according to the sequence of the upstream transmissions). Given these two restrictions, we show in Section III-D that the shortest propagation delay (SPD) first scheduling policy minimizes the cycle time. Subsequently, in Section III-E we relax the restriction R2 on the sequence of the GATE messages and characterize the competitiveness of SPD scheduling.

D. Shortest Propagation Delay (SPD) First Optimality

Theorem 1: If $G_i \geq t_G \; \forall_i$ and the GATE message transmission sequence is equal to the sequence of the upstream

transmissions, then Shortest Propagation Delay (SPD) first scheduling of the upstream transmissions minimizes the cycle duration Γ.

Proof: Without loss of generality, we neglect the scheduling time t_{sched} and the guard times t_g in the following as they are not affected by the scheduling of the upstream transmissions.

a) Comparison of two ONUs: Consider two ONUs 1 and 2. Let $\Gamma_{1,2}$ denote the cycle length when ONU 1 is scheduled before ONU 2. As illustrated in Fig. 2, there are two cases for the evaluation of $\Gamma_{1,2}$: (A) the propagation delay τ_2 governs the cycle length, and (B) the propagation delay τ_1 governs the cycle length. Clearly, the actual cycle length $\Gamma_{1,2}$ is obtained as the maximum of the two cases:

$$\Gamma_{1,2} = \max\left(2t_G + 2\tau_2, \; t_G + 2\tau_1 + G_1\right) + G_2. \quad (3)$$

Analogously, we obtain by symmetry (which only exchanges the roles of 1 and 2) the cycle length when ONU 2 is scheduled first, followed by ONU 1:

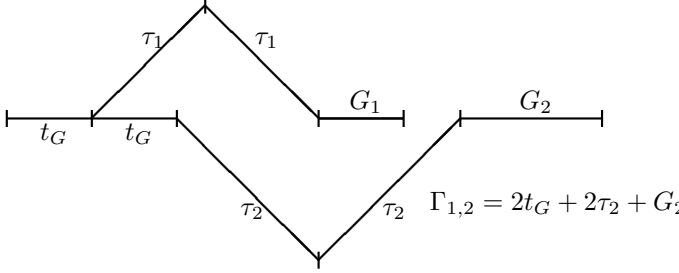$$\Gamma_{2,1} = \max\left(2t_G + 2\tau_1, \; t_G + 2\tau_2 + G_2\right) + G_1. \quad (4)$$

We want to show that if $G_i \geq t_G$ then we cannot have

$$\tau_1 > \tau_2 \quad \text{and} \quad \Gamma_{1,2} < \Gamma_{2,1}, \quad (5)$$

i.e., it cannot be that scheduling the ONU with the longer propagation delay (no. 1 in this case) leads to a shorter cycle length.

We proceed to show that (5) leads to a contradiction. We distinguish all possible cases according to where the maximum in the definition of $\Gamma_{1,2}$ and $\Gamma_{2,1}$, respectively, is attained.

Case (A): Cycle length governed by τ_2:
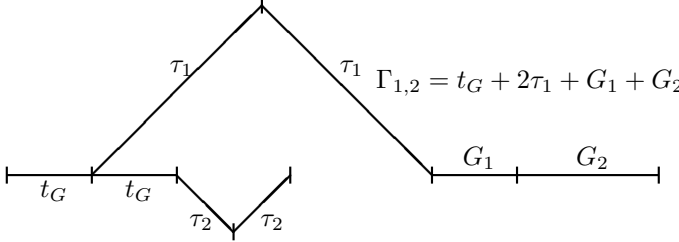

Case (B): Cycle length governed by τ_1:

Fig. 2. Illustration of cases for evaluation of cycle length $\Gamma_{1,2}$.

Case 1: $2t_G + 2\tau_2 \leq t_G + 2\tau_1 + G_1$ and $2t_G + 2\tau_1 \leq t_G + 2\tau_2 + G_2$. In this case, we have (the first condition comes from $\Gamma_{1,2} < \Gamma_{2,1}$, the other two from the condition for the case).

$$t_G + 2\tau_1 + G_1 + G_2 \quad < \quad t_G + 2\tau_2 + G_2 + G_1 \quad (6)$$
$$2t_G + 2\tau_2 \quad \leq \quad t_G + 2\tau_1 + G_1 \quad (7)$$
$$2t_G + 2\tau_1 \quad \leq \quad t_G + 2\tau_2 + G_2. \quad (8)$$

Note that equation (6) is a contradiction to $\tau_1 > \tau_2$.

Case 2: $2t_G + 2\tau_2 \geq t_G + 2\tau_1 + G_1$ and $2t_G + 2\tau_1 \leq t_G + 2\tau_2 + G_2$. Here, we get

$$2t_G + 2\tau_2 + G_2 \quad < \quad t_G + 2\tau_2 + G_2 + G_1 \quad (9)$$
$$2t_G + 2\tau_2 \quad \geq \quad t_G + 2\tau_1 + G_1 \quad (10)$$
$$2t_G + 2\tau_1 \quad \leq \quad t_G + 2\tau_2 + G_2. \quad (11)$$

Note that (10) is a contradiction to $\tau_1 > \tau_2$ and $t_G \leq G_1$.

Case 3: $2t_G + 2\tau_2 \leq t_G + 2\tau_1 + G_1$ and $2t_G + 2\tau_1 \geq t_G + 2\tau_2 + G_2$. Here, we get

$$t_G + 2\tau_1 + G_1 + G_2 \quad < \quad 2t_G + 2\tau_1 + G_1 \quad (12)$$
$$2t_G + 2\tau_2 \quad \leq \quad t_G + 2\tau_1 + G_1 \quad (13)$$
$$2t_G + 2\tau_1 \quad \geq \quad t_G + 2\tau_2 + G_2. \quad (14)$$

Note that (12) is a contradiction to $t_G \leq G_2$.

Case 4: $2t_G + 2\tau_2 \geq t_G + 2\tau_1 + G_1$ and $2t_G + 2\tau_1 \geq t_G + 2\tau_2 + G_2$. Here, we get

$$2t_G + 2\tau_2 + G_2 \quad < \quad 2t_G + 2\tau_1 + G_1 \quad (15)$$
$$2t_G + 2\tau_2 \quad \geq \quad t_G + 2\tau_1 + G_1 \quad (16)$$
$$2t_G + 2\tau_1 \quad \geq \quad t_G + 2\tau_2 + G_2. \quad (17)$$

Note that (16) is a contradiction to $\tau_1 > \tau_2$ and $t_G \leq G_1$.

b) General case: Now, consider N ONUs with propagation delays $\tau_1, \ldots, \tau_N$, respectively. We show that it is optimal to schedule them in SPD first manner. Consider any order of the ONUs 1, 2, $\ldots, N$ and suppose that we had (strictly) optimal cycle time over all orders but we did not have $\tau_1 \leq \tau_2 \leq \cdots \leq \tau_N$. If we do not have $\tau_1 \leq \tau_2 \leq \ldots \leq \tau_N$, there must be one i such that $\tau_i > \tau_{i+1}$. However, independently of the propagation delays of the preceding ONUs, it would be better (i.e., give lower cycle time) to have $\tau_i \leq \tau_{i+1}$, by the above comparison of two ONUs in part a) of this proof. (This reasoning is not influenced by the fact that the transmissions of ONUs i and $i+1$ may be delayed by preceding transmissions of other ONUs. In any case, it cannot be a loss to schedule the ONU with the shortest propagation delay first.) This is a contradiction to the optimality of the ordering. Hence, this contradiction implies the assertion, since i was arbitrary. ∎

Note that an immediate corollary of Theorem 1 is that SPD first scheduling maximizes the upstream channel utilization. As examined in detail in the next section, the bound in Theorem 1 shows that SPD first scheduling is very close to optimal, since the GATE transmission time t_G is typically small compared to traffic.

E. Competitiveness of SPD Scheduling

In this section, we examine the competitiveness of SPD first scheduling in comparison to an optimal schedule that minimizes the cycle length. The optimal schedule may have different sequences of GATE transmissions and upstream transmissions, i.e., does not need to meet restriction R2. We still require that the optimal schedule meets restriction R1 that $G_i \geq t_G$ since the EPON polling requires that each upstream transmission contains at least a REPORT. We first characterize the absolute difference of the cycle length with SPD Γ_{SPD} compared to the minimal cycle length Γ_{opt} of an optimal schedule. Next, we examine the competitive ratio of SPD scheduling, i.e., the bound on the ratio of the cycle length with SPD scheduling Γ_{SPD} to the cycle length of an optimal schedule Γ_{opt}.

Theorem 2: The cycle length with SPD first scheduling exceeds the minimal cycle length by no more than $(N-1)t_G$, i.e., $\Gamma_{\text{SPD}} \leq \Gamma_{\text{opt}} + (N-1)t_G$.

Proof: For SPD first scheduling, let t_i^{start}, $i = 1, \ldots, N$, denote the instant when the upstream transmission of ONU i begins to arrive at the OLT. We define for convenience $t_0^{\text{start}} := 0$ and $G_0 = 0$ and note that

$$t_i^{\text{start}} \quad = \quad \max(it_G + 2\tau_{(i)}, \ t_{i-1}^{\text{start}} + G_{i-1} + t_g). \quad (18)$$

The cycle length is

$$\Gamma_{\text{SPD}} = t_N^{\text{start}} + G_N. \quad (19)$$

Now, consider an imaginary EPON where the grants to all ONUs are communicated with *one* GATE message requiring only one transmission time t_G. This imaginary EPON serves as a comparison for the optimal scheduling in the real EPON. Let $\Gamma_{\text{opt},c}$ denote the cycle time in the imaginary EPON (the subscript c is for "comparison strategy"). Clearly, SPD first

scheduling of the upstream transmissions is optimal in the imaginary EPON.

We proceed to show that

$$\Gamma_{\text{SPD}} - (N-1)t_G \leq \Gamma_{\text{opt,c}} \leq \Gamma_{\text{opt}} \leq \Gamma_{\text{SPD}}, \qquad (20)$$

where the last two inequalities are trivially satisfied. We prove the first inequality by induction. In the imaginary EPON let s_i^{start} denote the instant when the upstream transmission of ONU i begins to arrive at the OLT. Denote $s_0^{\text{start}} := 0$ and note that

$$\begin{aligned} s_i^{\text{start}} &= \max(t_G + 2\tau_{(i)}, \ s_{i-1}^{\text{start}} + G_{i-1} + t_{\text{g}}), \\ &\qquad \text{for } i = 1, \ldots, N. \end{aligned} \qquad (21)$$

We obtain by induction that

$$t_i^{\text{start}} \leq s_i^{\text{start}} + (i-1)t_G. \qquad (22)$$

The case $i = 1$ is trivial. In general, we get

$$\begin{aligned} t_i^{\text{start}} &= \max(it_G + 2\tau_{(i)}, \ t_{i-1}^{\text{start}} + G_{i-1} + t_{\text{g}}) \qquad (23) \\ &\leq \max(it_G + 2\tau_{(i)}, \\ &\qquad s_{i-1}^{\text{start}} + (i-2)t_G + G_{i-1} + t_{\text{g}}) \qquad (24) \\ &\leq \max((i-1)t_G + t_G + 2\tau_{(i)}, \\ &\qquad s_{i-1}^{\text{start}} + (i-1)t_G + G_{i-1} + t_{\text{g}}) \qquad (25) \\ &= (i-1)t_G + \max(t_G + 2\tau_{(i)}, \\ &\qquad s_{i-1}^{\text{start}} + G_{i-1} + t_{\text{g}}) \qquad (26) \\ &= (i-1)t_G + s_i^{\text{start}}. \qquad (27) \end{aligned}$$

The assertion follows from (22), since

$$\begin{aligned} \Gamma_{\text{SPD}} &= t_N^{\text{start}} + G_N \qquad (28) \\ &\leq s_N^{\text{start}} + G_N + (N-1)t_G \qquad (29) \\ &= \Gamma_{\text{opt,c}} + (N-1)t_G. \qquad (30) \end{aligned}$$

$\blacksquare$

We remark that the bound in Theorem 2 is attained for an example scenario with $\tau_1 = \ldots = \tau_{N-1} = 0$, $2\tau_N = Nt_G$, and $G_1 = \ldots = G_N = t_G$. For this example, the cycle length with optimal scheduling is $\Gamma_{\text{opt}} = (N+2)t_G$, whereas the cycle length with SPD scheduling is $\Gamma_{\text{SPD}} = (2N+1)t_G$. That is, the difference $\Gamma_{\text{SPD}} - \Gamma_{\text{opt}}$ is exactly $(N-1)t_G$ in this example, and thus the bound in Theorem 2 cannot be improved.

Proposition 1: The competitive ratio for the cycle length with Smallest Propagation Delay (SPD) first scheduling is

$$\frac{\Gamma_{\text{SPD}}}{\Gamma_{\text{opt}}} \leq \min\left\{ \frac{t_G + 2\max_i \tau_i + \sum_i G_i}{t_G + 2\min_i \tau_i + \sum_i G_i}, \right. \\ \left. \frac{t_G + 2\max_i \tau_i + \sum_i G_i}{t_G + \max_i(2\tau_i + G_i)} \right\}. \qquad (31)$$

Proof: First, note that the shortest possible cycle length must satisfy

$$\Gamma_{\text{opt}} \geq t_G + 2\tau_{(1)} + \sum_{i=1}^{N} G_i \qquad (32)$$

since at least the first GATE needs to be transmitted and at least the round trip propagation delay to the nearest ONU is incurred before all the upstream transmissions (with aggregate

duration $\sum_{i=1}^{N} G_i$) can arrive at the OLT. This bound is attained when the ONU with the shortest propagation delay has a very large upstream transmission.

Second, note that cycle must be long enough to accommodate the GATE message transmission, round-trip propagation delay, and upstream transmission of each individual ONU i, i.e.,

$$\Gamma_{\text{opt}} \geq t_G + 2\tau_i + G_i \qquad (33)$$

for each ONU i, $i = 1, \ldots, N$. Since this holds for all i we can take the maximum over all ONUs yielding

$$\Gamma_{\text{opt}} \geq t_G + \max_i(2\tau_i + G_i). \qquad (34)$$

This bound is attained if one ONU with a large propagation delay has a large upstream transmission and all other ONUs have small propagation delays and upstream transmissions.

Thirdly, for the cycle time with SPD first scheduling

$$\Gamma_{\text{SPD}} \leq t_G + 2\tau_{(1)} + \sum_{i=1}^{N} G_i + \sum_{i=1}^{N-1} \delta_{i,i+1} \qquad (35)$$

where we define $\delta_{i,i+1}$ as the difference between the $(i+1)$th smallest and the ith smallest round trip propagation delay, i.e., $\delta_{i,i+1} = 2\tau_{(i+1)} - 2\tau_{(i)}$ for $i = 1, \ldots, N-1$. Note that $\sum_{i=1}^{N-1} \delta_{i,i+1}$ is the worst case for the stall times in between the upstream transmissions arriving at the OLT. This worst-case occurs if each ONU has only a REPORT message to send upstream, i.e., $G_i = t_G$. The δ-sum is telescoping and we get

$$\sum_{i=1}^{N-1} \delta_{i,i+1} = 2\tau_{(N)} - 2\tau_{(1)}. \qquad (36)$$

Thus,

$$\Gamma_{\text{SPD}} \leq t_G + 2\tau_{(N)} + \sum_{i=1}^{N-1} G_i. \qquad (37)$$

Combining (32) and (37) we get

$$\frac{\Gamma_{\text{SPD}}}{\Gamma_{\text{opt}}} \leq \frac{t_G + 2\tau_{(N)} + \sum_{i=1}^{N} G_i}{t_G + 2\tau_{(1)} + \sum_{i=1}^{N} G_i}. \qquad (38)$$

From (34) and (37) we get

$$\frac{\Gamma_{\text{SPD}}}{\Gamma_{\text{opt}}} \leq \frac{t_G + 2\tau_{(N)} + \sum_{i=1}^{N} G_i}{t_G + \max_i(2\tau_i + G_i)}. \qquad (39)$$

$\blacksquare$

The bounds in Proposition 1 show that the SPD first scheduling policy has a good competitive ratio in most cases. Indeed, the first bound is a good competitive ratio, i.e., is close to one, if $\sum_{i=1}^{N} G_i$ is large, i.e., if there is heavy traffic. The second bound is a good competitive ratio if $\tau_{(N)}$ is large compared to $\sum_{i=1}^{N} G_i$, i.e., typically if there is light traffic.

From the derivations of the preceding bounds we obtain the following insights into SPD first scheduling and potential heuristics to improve on SPD first scheduling: a) If there is heavy traffic at ONUs with short propagation delays, then it would be favorable to send them their GATE messages as soon as possible following the SPD first policy. With this strategy

the traffic from the ONUs with short propagation delays can mask the GATE transmissions and round trip propagations to the far-away ONUs. b) If there is little traffic at the ONUs with short propagation delays, then GATE messages could first be sent to the ONUs with large propagation delays, even though their upstream transmission windows should, of course, be scheduled after the windows of the ONUs with short propagation delays. With this strategy, the ONUs with short propagation delays can finish their transmissions by the time the upstream transmissions from the ONUs with long propagation delays arrive at the OLT.

F. Packet Delay Analysis for Gated Grant Sizing

In this section, we analyze the cycle length and packet delay with SPD first scheduling for Gated grant sizing. We consider approximations for light traffic and heavy traffic. We define the packet traffic load ρ_i, $i = 1, \ldots, N$ of ONU i as the ratio of the average traffic bit rate (including Ethernet frames plus "per-Ethernet frame" overhead, i.e., preamble and IPG) generated at ONU i to the upstream channel bit rate C [bit/s]. Unlike for the analysis in the preceding sections, we do not consider the REPORT message with transmission time $t_R = t_G$ as part of a grant G_i in this section, but rather account for the REPORT transmission times separately from the grant duration. We also explicitly consider the guard time t_g in this section. We denote $\rho_t = \sum_{i=1}^{N} \rho_i$ for the total load. Noting that with Gated grant sizing the "per-grant" overhead (REPORT message, guard time) becomes negligible as the grants grow very large in heavy traffic, the stability condition (maximum throughput) with gated grant sizing is $\rho_t < 1$ (which holds for arbitrary packet traffic patterns, including self-similar traffic). We let $\bar{L}$ and σ_L [in bit] denote the mean and standard deviation of the packet size (plus Preamble and IPG). In the delay analysis we consider Poisson packet traffic and define Poi$[\omega]$ for a random variable with Poisson distribution with parameter ω.

1) Low Load Scenario: In a low load situation with SPD scheduling, the cycle time is determined by the last (longest prop. delay) ONU. Namely, the cycle time in cycle $n + 1$ satisfies

$$\Gamma_{LL}(n+1) = Nt_G + 2\tau_{(N)} + \frac{\bar{L}}{C}\text{Poi}\left[\rho_{(N)}\frac{C}{\bar{L}}\Gamma_{LL}(n)\right] + t_G, \tag{40}$$

where $\rho_{(N)}$ is the load parameter of the ONU with the largest propagation delay $\tau_{(N)}$. The cycle consists of N GATE message transmission and the data packet transmissions plus REPORT message transmission of the last ONU. The other transmissions are masked by the long propagation delay of the last ONU (due to the low load assumption).

Taking expectations we get

$$\mathbb{E}\Gamma_{LL} = Nt_G + 2\tau_{(N)} + \rho_{(N)}\mathbb{E}\Gamma_{LL} + t_G. \tag{41}$$

This yields

$$\mathbb{E}\Gamma_{LL} = \frac{(N+1)t_G + 2\tau_{(N)}}{1 - \rho_{(N)}}. \tag{42}$$

Since we need the second moment of the cycle time for the delay analysis, we obtain in the same way from (40):

$$\mathbb{E}\Gamma_{LL}^2 = \mathbb{V}\Gamma_{LL} + (\mathbb{E}\Gamma_{LL})^2 \tag{43}$$

$$= \rho_{(N)}\frac{\bar{L}}{C}\mathbb{E}\Gamma_{LL} + (\mathbb{E}\Gamma_{LL})^2 \tag{44}$$

$$= \mathbb{E}\Gamma_{LL}(\rho_{(N)}\frac{\bar{L}}{C} + \mathbb{E}\Gamma_{LL}). \tag{45}$$

The delay of a packet generated at ONU i consists of the backward recurrence time of the cycle length [60, Ch. 5.5] $\frac{\mathbb{E}\Gamma_{LL}^2}{2\mathbb{E}\Gamma_{LL}}$, i.e., the time until the generated packet is included in a REPORT, the time period from the instant the transmission of the REPORT is complete to the instant the next upstream transmission commences, which is $it_G + 2\tau_{(i)}$ due to the low load assumption, and the time needed within the grant until the considered packet commences transmission $\rho_i\frac{\mathbb{E}\Gamma_{LL}^2}{2\mathbb{E}\Gamma_{LL}}$:

$$D_{LL}^{(i)} = \frac{\mathbb{E}\Gamma_{LL}^2}{2\mathbb{E}\Gamma_{LL}} + it_G + 2\tau_{(i)} + \rho_i\frac{\mathbb{E}\Gamma_{LL}^2}{2\mathbb{E}\Gamma_{LL}}$$
$$+ \tau_{(i)} + \frac{\bar{L}}{C}. \tag{46}$$

$$= (1 + \rho_i)\frac{\rho_{(N)}\frac{\bar{L}}{C} + \mathbb{E}\Gamma_{LL}}{2} + it_G$$
$$+ 3\tau_{(i)} + \frac{\bar{L}}{C}. \tag{47}$$

The overall delay is then given by

$$D_{LL} = \frac{1}{\rho_t}\sum_{i=1}^{N}\rho_{(i)}D_{LL}^{(i)}, \tag{48}$$

where $\rho_t := \sum_{i=1}^{N}\rho_i$ and $\rho_{(i)}$ is the load parameter of the ONU with the i-th smallest propagation delay.

2) Heavy Load Scenario: In a heavy load scenario, all ONUs have data to send, which masks the round-trip delays between OLT and the second to last ONUs in the SPD first schedule. In this case, the cycle length satisfies

$$\mathbb{E}\Gamma_{HL} = t_G + 2\tau_{(1)} + \sum_{i=1}^{N}\rho_i\mathbb{E}\Gamma_{HL} + Nt_G + (N-1)t_g, \tag{49}$$

because a cycle consists of the GATE transmission plus round trip propagation delay to the first ONU (the one with the shortest propagation delay) and all the traffic that is sent (which consists of the accumulated data traffic and the N REPORT messages). This gives

$$\mathbb{E}\Gamma_{HL} = \frac{(N+1)t_G + (N-1)t_g + 2\tau_{(1)}}{1 - \sum_{i=1}^{N}\rho_i}. \tag{50}$$

This scenario resembles the case of a single ONU treated in [61]. Inserting $\tau_{HL} := \frac{N+1}{2}t_G + \frac{N-1}{2}t_g + \tau_{(1)}$ and $\rho_t := \sum_{i=1}^{N}\rho_i$ in Eqn. (39) in [61], namely

$$D_{HL} = 2\tau_{HL}\frac{2 - \rho_t}{1 - \rho_t} + \frac{\rho_t}{2C(1 - \rho_t)}\left(\frac{\sigma_L^2}{\bar{L}} + \bar{L}\right) + \frac{\bar{L}}{C} \tag{51}$$

gives an approximation of the packet delay.

G. Stability of Limited Grant Sizing

For Limited grant sizing with SPD scheduling we evaluate the maximum throughput (stability limit) as follows. First, for arbitrary packet traffic, including self-similar traffic, we calculate the maximal cycle time by $t_0^{\text{start,max}} := 0$

$$t_i^{\text{start,max}} = \max(it_G + 2\tau_{(i)}, \ t_{i-1}^{\text{start,max}} + G_{i-1}^{\max} + t_g), \quad (52)$$

where $G_i^{\max}$ are the given maximal grant sizes ($G_0^{\max} = 0$). Then,

$$\Gamma_{\text{SPD}}^{\max} = t_N^{\text{start,max}} + G_N^{\max} \quad (53)$$

is the maximum cycle length (even in the heaviest traffic, no cycle will be longer than this). The stability condition for SPD is then

$$\rho_i \Gamma_{\text{SPD}}^{\max} < G_i^{\max} \quad (54)$$

for all i. (The stability limit for Largest Propagation Delay (LPD) first scheduling is obtained analogously by considering the τ_i in decreasing order in (52).) For homogeneous ONU loads $\rho_1 = \rho_2 = \cdots = \rho_N$ and maximum grant sizes $G_1^{\max} = \cdots = G_N^{\max}$, the total load is $\rho_t = N\rho_i$, resulting in the stability condition

$$\rho_{t,SPD}^{\max} < \frac{NG_i^{\max}}{\Gamma_{\text{SPD}}^{\max}}. \quad (55)$$

IV. EXPERIMENTAL PERFORMANCE ANALYSIS

We conducted a set of simulation experiments to: 1) validate our analytical models presented in Section III, and 2) quantify the improvements to cycle length and packet delay achieved by SPD grant scheduling under different operating conditions. We use an EPON simulator that we have developed using the CSIM discrete event simulation library [62]. We simulated an EPON with 32 ONUs and varied the maximum propagation delay to represent different EPON reaches. The following quad modal packet size distribution was used for all simulation experiments: 60% 64 bytes, 4% 300 bytes, 11% 580 bytes, and 25% 1518 bytes.

In the absence of any diversity in propagation delay, the impact of minimizing cycle length will be minor. In our experiments, we use a continuous (uniform) distribution for one-way (OLT-to-ONU) propagation delays with a minimum value of 6.68 μs and different maximum one-way propagation delay values. We modify this distribution slightly by forcing one ONU to have a minimum propagation delay value and another to have a maximum propagation delay value. The propagation delays for the other 30 ONUs are continuously distributed over the range.

We compare SPD scheduling to Largest Number of Frames (LNF) first scheduling which was previously demonstrated [25] to provide low average queueing delay compared to other scheduling policies [47]. We also present some results for Largest Propagation Delay (LPD) first scheduling which is the opposite of SPD to illustrate the range of possibilities for cycle length and packet delay. The average cycle length and average packet delay values presented in this section represent the mean of several independent runs that were constructed

a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

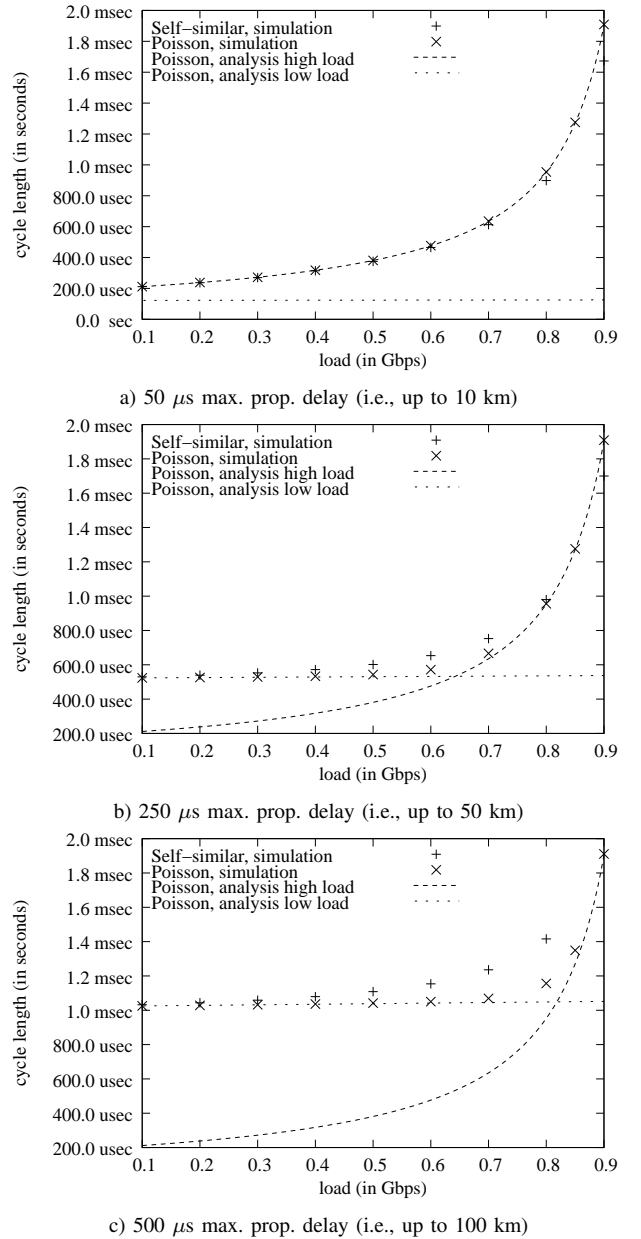c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 3. Average cycle length for SPD grant scheduling and Gated grant sizing.

using the batch means feature in the CSIM discrete event simulation library. The resulting statistical confidence intervals for Poisson traffic are smaller than the point marks in the plots.

We also compare with online IPACT scheduling [9], [24] where the sequence of grants to the ONUs is generally round-robin in the order in which the ONUs registered with the OLT. With the same set of random propagation delays used in the SPD simulations, we simulated and averaged many independent random permutations of the ONU granting sequence.

A. Gated Grant Sizing

In this section we present our results of the experiments we conducted using Gated grant sizing.

1) Cycle Length: Figure 3 shows the average cycle length as a function of the total load for SPD grant scheduling
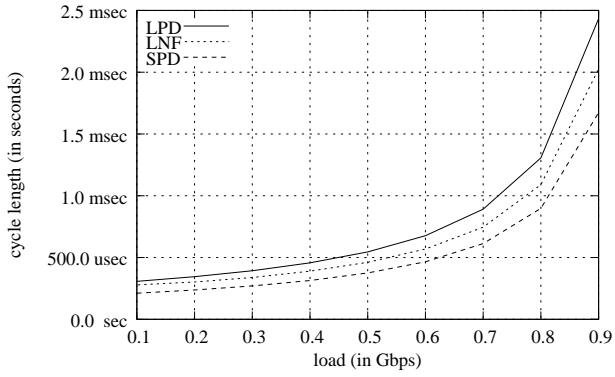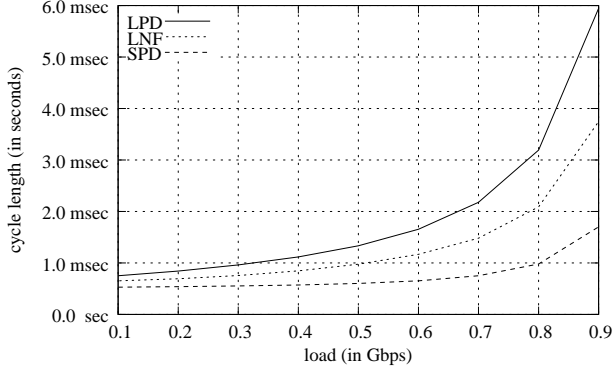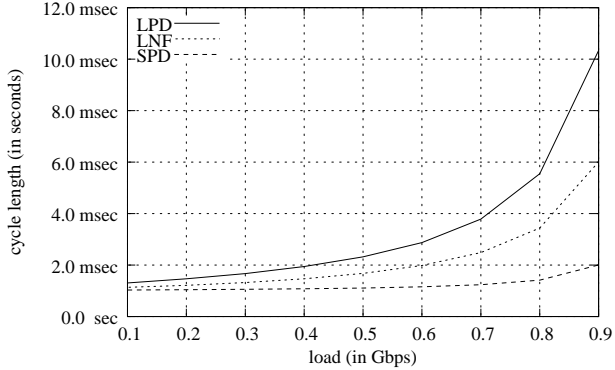
a) 50 μs max. prop. delay (i.e., up to 10 km)

a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 4.   Average cycle length with Gated grant sizing for self-similar traffic.

Fig. 5.   Average cycle length with Gated grant sizing for self-similar traffic as the load approaches the channel capacity.

by means of analysis using Eqs. (42) and (50), as well as simulation experiments using Poisson traffic sources and self-similar traffic sources. We observe from this figure that Eqs. (42) and (50) provide an excellent fit to the average cycle length measured in the simulation experiments.

Figure 4 shows the average cycle length for LPD, LNF, and SPD grant scheduling and varying propagation delay configurations. Figure 5 shows the same for load values approaching the channel limit. We observe from these figures that SPD always provides a lower average cycle length than LPD or LNF. The difference increases significantly with an increasing load and increasing maximum propagation delay. As an example, with a 500 μsec maximum propagation delay and load value of 0.9 the average cycle length was 2.0 milliseconds using SPD and 6.0 milliseconds using LNF.
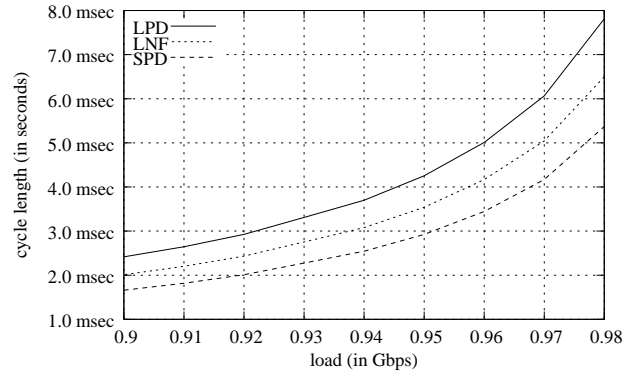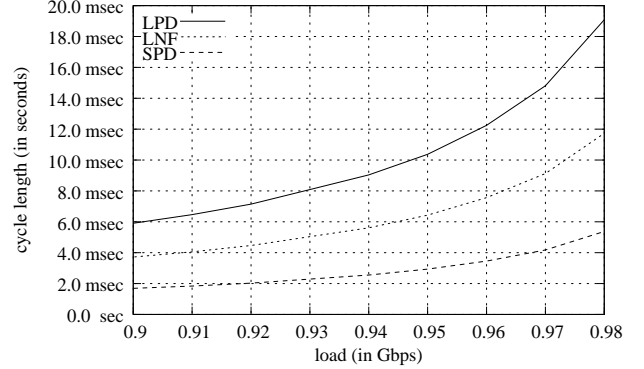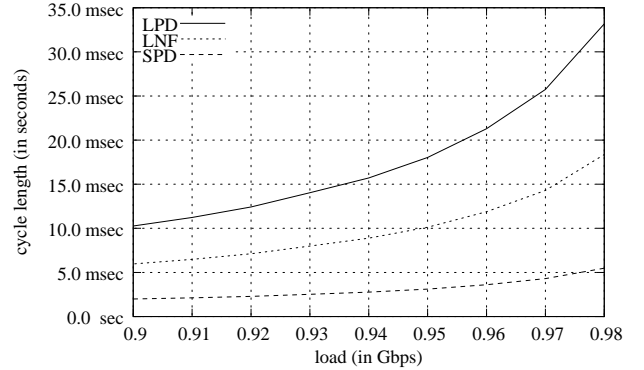
2) Packet Delay: Figure 6 shows the average packet delay for SPD grant scheduling by means of analysis using Eqs. (48) and (51), as well as simulation experiments using Poisson traffic sources and self-similar traffic sources. The measurements for average packet delay with the self-similar traffic sources are, as expected, much higher despite average cycle lengths similar to those observed with Poisson traffic sources. This is a result of a few very long cycles, whose lengths are observed once, that have many packets whose associated large delay is observed once for each of these packets.

Figure 7 shows the average packet delay for LPD, LNF, and SPD grant scheduling and varying propagation delay configurations for self-similar traffic. Figure 8 shows the same for load values approaching the channel limit. We observe from Figures 7 and 8 that SPD provides lower average packet
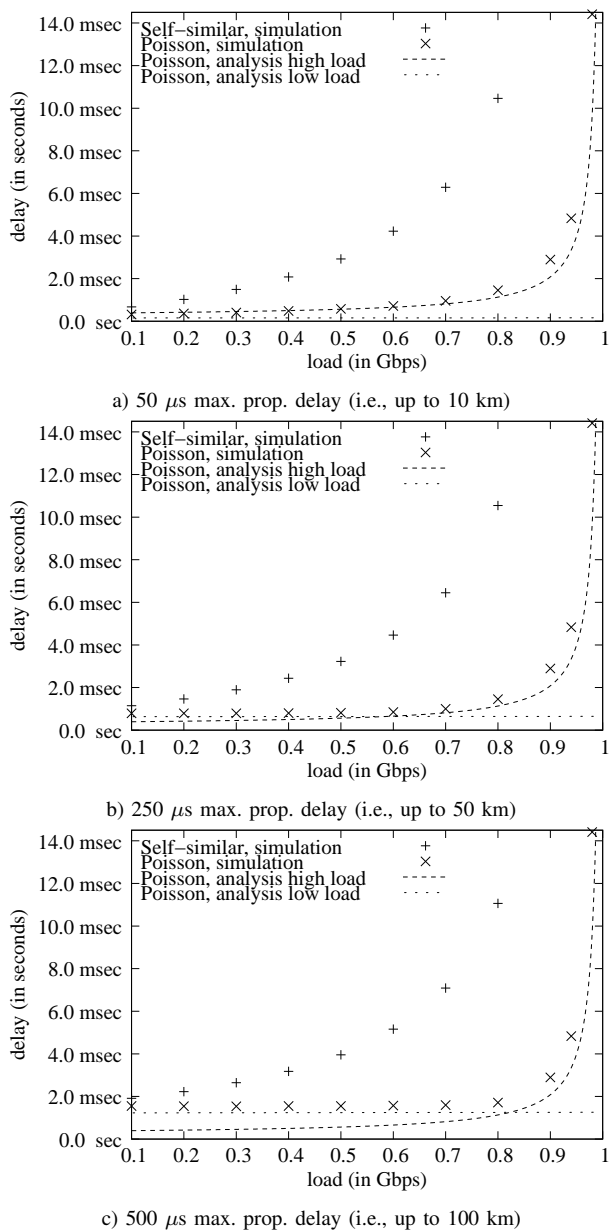
a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 6. Average packet delay for SPD grant scheduling and Gated grant sizing.

a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 7. Average packet delay with Gated grant sizing for self-similar traffic.

delay for the experiments with 250 μs and 500 μs maximum propagation delays. The difference with LNF becomes more pronounced at high load values. As an example, with a 500 μs maximum propagation delay and load value of 0.9 Gbps the average packet delay was 21.3 milliseconds using SPD and 26.5 milliseconds using LNF. At 50 μs, the difference between LNF and SPD is rather insignificant. However, it is worth noting that SPD clearly provides a lower average cycle length for all load values (see Figures 4 and 5) which leads to lower average packet delay. The mean packet delays with online IPACT scheduling (which are not included in the plots to avoid clutter) were only very slightly lower than the SPD packet delays; namely within 3% of the SPD delays.
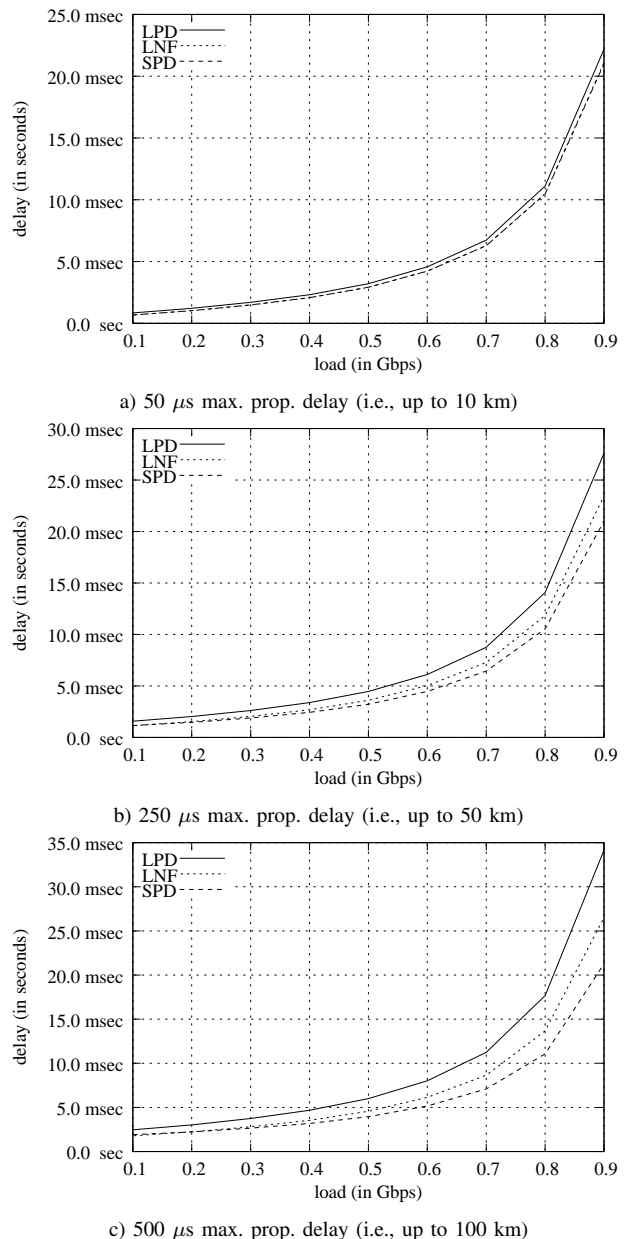
B. Limited Grant Sizing

In this section we present our results of the experiments we conducted using Limited grant sizing with $G_i^{\max} = 7188$ bytes, $\forall_i$.

1) Cycle Length: Figure 9 shows the average cycle length for LPD, LNF, and SPD grant scheduling for varying propagation delay configurations. We make two observations: 1) in all plots SPD provides lower average cycle length for all load values, and 2) as the maximum propagation delay is increased from 50 μs to 500 μs, SPD is able to maintain a maximum average cycle length around 2 milliseconds.

Exploring the second observation further, we see that the maximum average cycle length for LNF increases from around 2 milliseconds at 50 μs to close to 2.6 milliseconds at 500 μs. With a 500 μsec maximum propagation delay, the load at which the cycle length reaches its maximum is 0.9 Gbps for
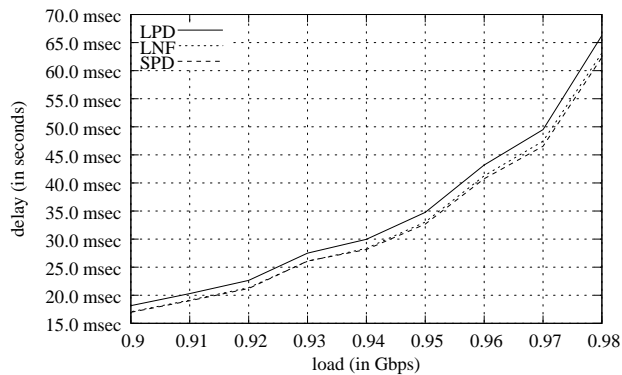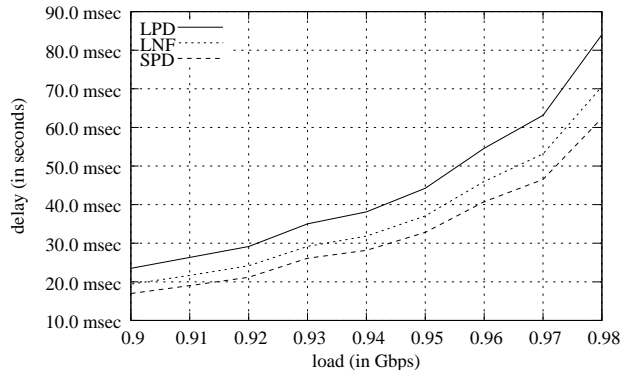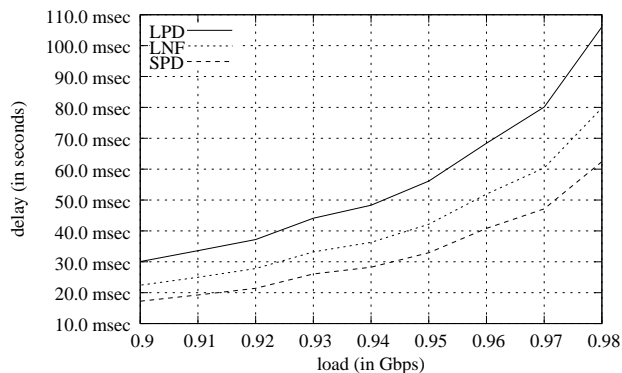
a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 8.   Average packet delay with Gated grant sizing for self-similar traffic as the load approaches the channel capacity.

a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 9.   Average cycle length for different grant scheduling policies with Limited grant sizing for self-similar traffic.

SPD and approximately only 0.62 Gbps for LNF. In Table I we compare the maximum average cycle length using Eq. (53) with the results from our simulation experiments. The data in this table indicates that the equation is within 0.7 % of the experimental data.

Our experimental data and Eq. (53) indicate that SPD is able to keep the maximum cycle length near 2 milliseconds as the maximum propagation delay is increased. Whereas, LNF is unable to do the same. Eq. (53) illustrates that the maximum cycle length is a function of the order of the ONUs as well as their propagation delays. LNF constantly changes the order of ONUs with respect to their propagation delays resulting in the variations that can be seen in its maximum average cycle length. Further, SPD is able to minimize the cycle length by ordering the ONUs in increasing order of propagation delay.
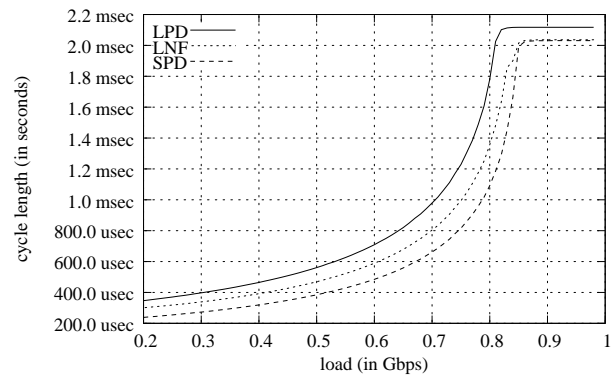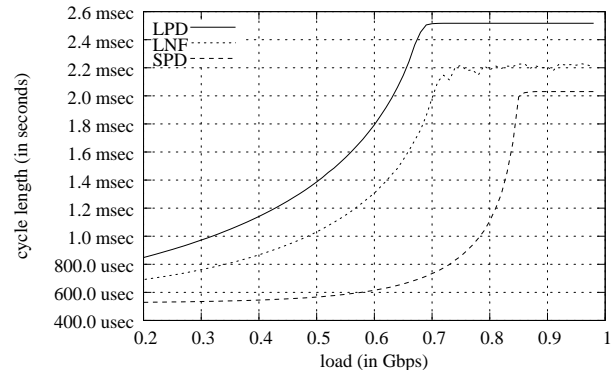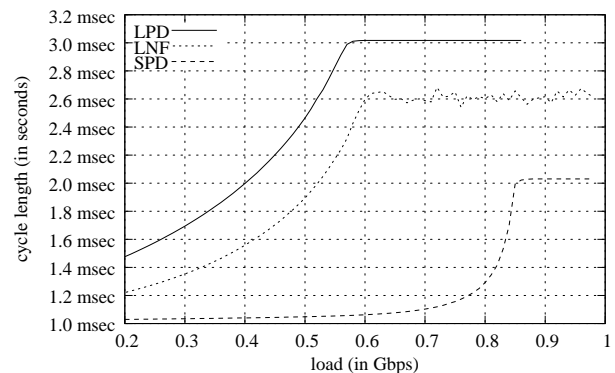
LNF does not order the ONUs according to their propagation delays and, as a result, does not minimize the cycle length.

2) Stability Limit: As the load increases and the grant sizes approach the prescribed maximum grant size $G_i^{\max}$, the cycle length approaches the maximum cycle length. That is, the average cycle length levels out at the maximum cycle length plotted in Fig. 9 and tabulated for SPD in Table I. Further increases in the load can not increase the cycle length, but result in infinite queue build-up in the ONUs, i.e., instability. The load value at which the average cycle length levels out to the maximum cycle length in Fig. 9 thus represents the stability limit, which is tabulated for SPD in Table II. We observe from Table II that Eqn. (55) very accurately characterizes the stability limit.

Turning to Fig. 9, we observe that the stability limit dif-

One-way prop. delay [μs]	$\Gamma_{SPD}^{\max}$ Eq. (53) [ms]	$\Gamma_{SPD}^{\max}$ sim. [ms]
50 (up to 10 km)	2.058	2.03
250 (up to 50 km)	2.058	2.03
500 (up to 100 km)	2.058	2.03

TABLE I

MAXIMUM AVERAGE CYCLE LENGTH (IN MILLISECONDS) FOR SPD FOR
LIMITED GRANT SIZING WITH $G_i^{\max} = 7188$ BYTES,$\forall_i$ FOR
SELF-SIMILAR TRAFFIC.

One-way prop. delay [μs]	$\rho_{t,SPD}^{\max}$ Eq. (55)	$\rho_{t,SPD}^{\max}$ sim.
50 (up to 10 km)	0.894	0.90
250 (up to 50 km)	0.894	0.90
500 (up to 100 km)	0.894	0.90

TABLE II

LOAD AT WHICH THE AVERAGE CYCLE LENGTH APPROACHES THE
MAXIMUM CYCLE LENGTH, WHICH IS EQUAL TO THE STABILITY LIMIT,
FOR SPD WITH LIMITED GRANT SIZING WITH $G_i^{\max} = 7188$ BYTES,$\forall_i$
FOR SELF-SIMILAR TRAFFIC.

ference between LNF and SPD increases significantly as the maximum propagation delay is increased. As an example, with a 500 μs maximum propagation delay, the stability limit is approximately 0.89 Gbps using SPD (the average cycle length converges to the maximum cycle length very slowly between a load of 0.85 and 0.89) and approximately 0.62 Gbps using LNF.

By scheduling the grants to the close-by ONUs first, SPD masks the long round-trip delays to the ONUs that are further away. This more efficient utilization of the upstream channel increases the stability limit substantially as the propagation delays become more diverse.

3) Packet Delay: Figure 10 shows the average packet delay for LPD, LNF, and SPD grant scheduling and varying propagation delay configurations. We observe that SPD provides lower average packet delay for all load values and reconfirm the higher stability limit for SPD. From additional simulations we found that online IPACT scheduling gives mean packet delays that are generally around 15–20% lower than for SPD offline scheduling. For instance, for 250 μs maximum propagation delay and a load of 0.8, SPD gives a mean packet delay of 0.153 s compared to 0.127 s with online IPACT. Within the offline scheduling framework, the SPD scheduling policy vastly reduces the packet delay compared to the LNF scheduling policy. Thus, the SPD scheduling policy makes it possible to reap the benefits of the offline scheduling framework with only relatively modest delay penalties compared to online scheduling.

It is instructive to compare the packet delay reductions with SPD compared to LNF grant scheduling for Gated grant sizing in Figs. 7 and 8 to the corresponding delay reductions for Limited grant sizing in Fig. 10. Clearly, for Limited grant sizing we observe substantially larger delay reductions. With Gated grant sizing, any reported queue size is served in one upstream transmission. In contrast, with Limited grant sizing, a large reported queue requires several maximum sized grants of size $G_i^{\max}$, and thus several cycles to transmit the traffic to the OLT. The cycle length minimizing SPD scheduling policy
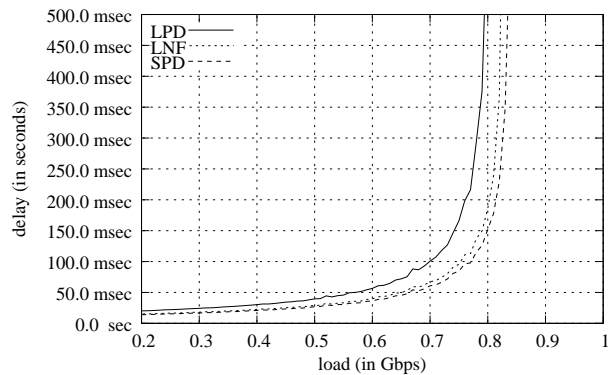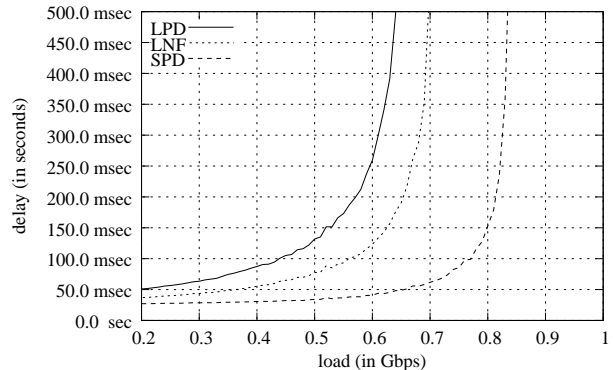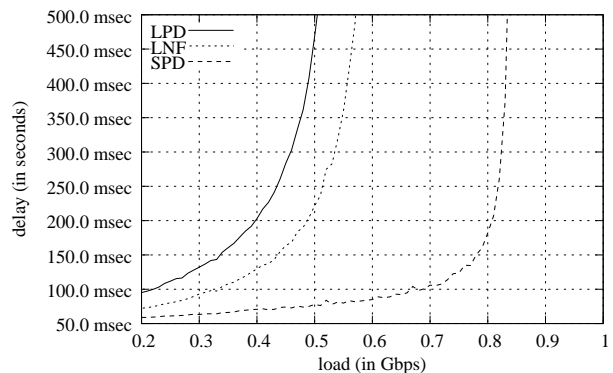

a) 50 μs max. prop. delay (i.e., up to 10 km)

b) 250 μs max. prop. delay (i.e., up to 50 km)

c) 500 μs max. prop. delay (i.e., up to 100 km)

Fig. 10. Average packet delay for different grant scheduling policies with Limited grant sizing for self-similar traffic.

leads hence to significantly more pronounced delay reductions for Limited grant sizing than for Gated grant sizing.

C. Engineering EPONs for Better Channel Utilization

In this final set of simulation experiments we wish to illustrate the utility of SPD grant scheduling in allowing close-by ONUs to be added to an EPON without taking bandwidth from existing ONUs. Essentially, the maximum channel utilization is significantly improved when adding close-by ONUs and utilizing SPD scheduling. We consider an EPON with Limited grant sizing with 4 ONUs with ONU-to-OLT propagation delays continuously distributed between 250 μs and 500 μs (i.e., ONU-to-OLT distances of 50 km to 100 km). We then added a varying number of close-by ONUs with ONU-to-OLT
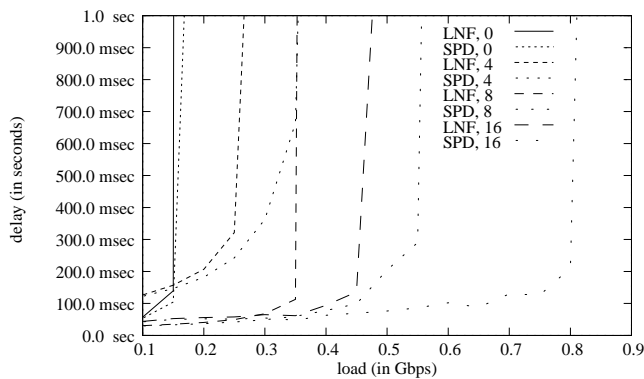
Fig. 11. Average packet delay for EPON with 4 ONUs between 50 km and 100 km from OLT with varying number of close-by ONUs (i.e., 0.5 km to 5 km from OLT) added.

propagation delays continuously distributed between 2.5 μs and 25 μs (i.e., ONU-to-OLT distances of 0.5 km to 5 km).

Figure 11 shows the average packet delay for different numbers of added close-by ONUs for both LNF and SPD. We observe that adding close-by ONUs increases the maximum achievable channel utilization (stability limit). The shorter propagation delays of the close-by ONUs allow these ONUs to be serviced while the GATE messages are propagating to the ONUs with the larger propagation delays and their upstream transmissions are propagating up to the OLT. Thus, the added close-by ONUs increase the channel utilization while almost not increasing the packet delay experienced by the far-away ONUs. With LNF scheduling the ONUs are ordered by their grant size, irrespective of their propagation delays leading to poor exploitation of this ability. SPD, on the other hand, always services the ONUs with shorter propagation delays first allowing them to mask the round-trip time to the ONUs with the larger propagation delays, and thus achieving substantially higher channel utilization than LNF.

V. Conclusion

In conclusion, we introduced a new EPON grant scheduling technique called Shortest Propagation Delay (SPD) first grant scheduling to exploit heterogenous propagation delays. We proved that SPD minimizes granting cycle length to within a small time period (number of ONUs times GATE message transmission time) and maximizes channel utilization. We analytically characterized the stability limit (maximum packet throughput or equivalently maximum channel utilization) for both Gated and Limited grant sizing for arbitrary traffic and characterized the cycle length and packet delay for Gated grant sizing for Poisson traffic. We have illustrated the utility of SPD through a set of simulation experiments. Specifically, we found that SPD can improve performance measures when using Gated grant sizing as well as Limited grant sizing. The most significant improvements came from its use with Limited grant sizing and long reach EPONs. In those circumstances, packet delay and channel utilization were significantly improved.

Another significant finding is the potential channel utilization improvement that is possible when using SPD grant

scheduling in conjunction with certain EPON design principles. Specifically, suppose there are numerous subscriber nodes that must be connected to a central office and a network engineer has some choices in how to layout several EPONs to connect all of the subscriber nodes to a central office. Our findings indicate that channel utilization can be significantly increased if network engineers construct EPONs such that each EPON contains some ONUs close to the OLT as well as ONUs that are further away from the OLT. ONUs that are within a short range of the OLT can fill the idle times in which the OLT waits for data from the ONUs that are further away.

An interesting avenue for future research arises from the convergence of fiber-based and wireless access networks, see e.g., [63], [64], which will potentially cover large geographic areas and thus have highly heterogeneous propagation delays. Further, the integration of medium access control on the fiber and wireless media may lump the propagation delay on the fiber and the wireless medium access delay together to lead to additional diversification of the round-trip delays experienced by the OLT.

References

[1] F. Effenberger, "An introduction to PON technologies," *IEEE Communications Magazine*, vol. 45, no. 3, pp. S17–S25, Mar. 2007.
[2] P. Green, *Fiber to the Home: The New Empowerment*. Wiley-Interscience, 2005.
[3] G. Keiser, *FTTX Concepts and Applications*. Wiley-IEEE Press, 2006.
[4] C. Lam, *Passive Optical Networks: Principles and Practice*. Academic Press, 2007.
[5] M. De Andrade, L. Gutierrez, and S. Sallent, "A distributed scheduling approach for Ethernet-based passive optical networks," in *Proceedings of IEEE Conference on Local Computer Networks*, Oct. 2007, pp. 205–206.
[6] C. Foh, L. Andrew, E. Wong, and M. Zukerman, "FULL-RCMA: a high utilization EPON," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 8, pp. 1514–1524, Oct. 2004.
[7] M. Hajduczenia, H. J. A. da Silva, and P. P. Monteiro, "EPON versus APON and GPON: a detailed performance comparison," *OSA Journal of Optical Networking*, vol. 5, no. 4, pp. 298–319, Apr. 2006.
[8] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON (ePON): Design and Analysis of an Optical Access Network," *Photonic Network Communications*, vol. 3, no. 3, pp. 307–319, July 2001.
[9] ——, "IPACT: A dynamic protocol for an Ethernet PON (EPON)," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 74–80, February 2002.
[10] M. Ma, Y. Zhu, and T. Cheng, "A systematic scheme for multiple access in Ethernet passive optical access networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 23, no. 11, pp. 3671–3682, November 2005.
[11] M. McGarry, M. Reisslein, and M. Maier, "Ethernet Passive Optical Network Architectures and Dynamic Bandwidth Allocation Algorithms," *IEEE Communication Surveys and Tutorials*, vol. 10, no. 3, pp. 46–60, October 2008.
[12] M. R. Radivojevic and P. S. Matavulj, "Implementation of intra-ONU scheduling for Quality of Service support in Ethernet passive optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 18, pp. 4055–4062, Sept. 2009.
[13] B. Skubic, J. Chen, J. Ahmed, L. Wosinska, and B. Mukherjee, "A comparison of dynamic bandwidth allocation for EPON, GPON, and next-generation TDM PON," *IEEE Communications Magazine*, vol. 47, no. 3, pp. S40–S48, Mar. 2009.
[14] J. Zheng and H. Mouftah, "Media access control for Ethernet passive optical networks: an overview," *IEEE Communications Magazine*, vol. 43, no. 2, pp. 145–150, February 2005.
[15] L. G. Kazovsky, W.-T. Shaw, D. Gutierrez, N. Cheng, and S.-W. Wong, "Next-generation optical access networks," *IEEE/OSA J. of Lightwave Tech.*, vol. 25, no. 11, pp. 3428–3442, Nov. 2007.
[16] J. Lazaro, J. Prat, P. Chanclou, G. T. Beleffi, A. Teixeira, I. Tomkos, R. Soila, and V. Koratzinos, "Scalable extended reach PON," in *Proc. OFC*, Feb. 2008.

[17] C. Michie, T. Kelly, I. Andonovic, and J. McGeough, "Reach extension of passive optical networks using semiconductor optical amplifiers," in *Proceedings of IEEE Int. Conf. on Transparent Opitcal Networks*, June 2008, pp. 194–197.

[18] H. Song, B. Kim, and B. Mukherjee, "Multi-thread polling: a dynamic bandwidth distribution scheme in long-reach PON," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 134–142, Feb. 2009.

[19] D. Shea and J. Mitchell, "Long-reach optical access technologies," *IEEE Network*, vol. 21, no. 5, pp. 5–11, Sep 2007.

[20] G. Talli and P. Townsend, "Hybrid DWDM-TDM long-reach PON for next-generation optical access," *IEEE/OSA Journal of Lightwave Technology*, vol. 24, no. 7, pp. 2827–2834, July 2006.

[21] R. Davey, J. Kani, F. Bourgart, and K. McCammon, "Options for future optical access networks," *IEEE Communications Magazine*, vol. 44, no. 10, pp. 50 – 56, Oct 2006.

[22] R. Davey, D. Grossman, M. Rasztovits-Wiech, D. Payne, D. Nesset, A. Kelly, A. Rafel, S. Appathurai, and S.-H. Yang, "Long-reach passive optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 3, pp. 273–291, Feb. 2009.

[23] P. Iannone and K. Reichmann, "Strategic and Tactical Uses for Extended PON," in *Extended Reach PON and Central Office By-Pass*. OFC 2008, Feb 2008, http://www.ofcnfoec.org/about_ofc/archive/2008/PDFs/08_Iannone.pdf.

[24] 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 Communications*, vol. 4, no. 1, pp. 89–107, Jan. 2002.

[25] M. McGarry, M. Reisslein, C. Colbourn, M. Maier, F. Aurzada, and M. Scheutzow, "Just-in-Time scheduling for multichannel EPONs," *IEEE/OSA Journal of Lightwave Technology*, vol. 26, no. 10, pp. 1204–1216, May 2008.

[26] J. Zheng and H. Mouftah, "A survey of dynamic bandwidth allocation algorithms for Ethernet Passive Optical Networks," *Optical Switching and Networking*, vol. 6, no. 3, pp. 151–162, July 2009.

[27] A. Banerjee, G. Kramer, and B. Mukherjee, "Fair sharing using dual service-level agreements to achieve open access in a passive optical network," *IEEE Journal on Selected Areas in Communications, Part Supplement*, vol. 24, no. 8, pp. 32–44, Aug. 2006.

[28] B. Chen, J. Chen, and S. He, "Efficient and fine scheduling algorithm for bandwidth allocation in ethernet passive optical networks," *IEEE Journal on Selected Topics in Quantum Electronics*, vol. 12, no. 4, pp. 653–660, July-August 2006.

[29] J. Chen, B. Chen, and L. Wosinska, "Joint bandwidth scheduling to support differentiated services and multiple service providers in 1G and 10G EPONs," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 1, no. 4, pp. 343–351, Sept. 2009.

[30] A. Dhaini, C. Assi, M. Maier, and A. Shami, "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, no. 1, pp. 277–286, Jan. 2007.

[31] M. Ma, L. Liu, and T. Cheng, "Adaptive scheduling for differentiated services in an Ethernet passive optical network," *OSA Journal of Optical Networking*, vol. 4, no. 10, pp. 661–67, 2005.

[32] F. Melo Pereira, N. L. S. Fonseca, and D. S. Arantes, "A fair scheduling discipline for Ethernet passive optical networks," *Computer Networks*, vol. 53, no. 11, pp. 1859–1878, July 2009.

[33] H. Naser and H. Mouftah, "A joint-ONU interval-based dynamic scheduling algorithm for Ethernet passive optical networks," *IEEE/ACM Transactions on Networking*, vol. 14, no. 4, pp. 889–899, Aug. 2006.

[34] A. Shami, X. Bai, N. Ghani, C. Assi, and H. Mouftah, "QoS control schemes for two-stage Ethernet passive optical access networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 8, pp. 1467–1478, Nov. 2005.

[35] J. Xie, S. Jiang, and Y. Jiang, "A dynamic bandwidth allocation scheme for differentiated services in EPONs," *IEEE Communications Magazine*, vol. 42, no. 8, pp. S32–S39, Aug. 2004.

[36] C. Assi, Y. Ye, S. Dixit, and M. Ali, "Dynamic bandwidth allocation for Quality-of-Service over Ethernet PONs," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1467–1477, November 2003.

[37] P. Choudhury and P. Saengudomlert, "Efficient queue based dynamic bandwidth allocation scheme for ethernet PONs," in *Proceedings of IEEE Globecom*, Nov. 2007, pp. 2183–2187.

[38] A. Shami, X. Bai, C. Assi, and N. Ghani, "Jitter performance in Ethernet passive optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 23, no. 4, pp. 1745–1753, Apr. 2005.

[39] J. Zheng, "Efficient bandwidth allocation algorithm for ethernet passive optical networks," *IEE Proceedings Communications*, vol. 153, no. 3, pp. 464–468, June 2006.

[40] K. Kanonakis and I. Tomkos, "Online upstream scheduling and wavelength assignment algorithms for WDM EPON networks," in *Proc. of European Conference on Optical Communications (ECOC)*, Sept. 2009, pp. 1.6.4.1–1.6.4.2.

[41] S. Bhatia and R. Bartos, "IPACT with smallest available report first: A new DBA algorithm for EPON," in *Proc. of IEEE International Conference on Communications*, June 2007, pp. 2168–2173.

[42] A. Kamal and B. F. Blietz, "A priority mechanism for the IEEE 802.3ah EPON," in *Proceedings of IEEE ICC*, May 2005, pp. 1879–1883.

[43] G. Kramer, A. Banerjee, N. Singhal, B. Mukherjee, S. Dixit, and Y. Ye, "Fair queueing with service envelopes (FQSE): a cousin-fair hierarchical scheduler for subscriber access networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, no. 8, pp. 1497–1513, Oct. 2004.

[44] Y. Luo and N. Ansari, "Bandwidth allocation for multiservice access on EPONs," *IEEE Communications Magazine*, vol. 43, no. 2, pp. S16–S21, Feb. 2005.

[45] ——, "Limited sharing with traffic prediction for dynamic bandwidth allocation and QoS provioning over EPONs," *OSA Journal of Optical Networking*, vol. 4, no. 9, pp. 561–572, 2005.

[46] Y. Zhu and M. Ma, "IPACT with grant estimation (IPACT-GE) scheme for Ethernet passive optical networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 26, no. 14, pp. 2055–2063, July 2008.

[47] J. Zheng and H. Mouftah, "Adaptive scheduling algorithms for ethernet passive optical networks," *IEE Proceedings Communications*, vol. 152, no. 6, pp. 643–647, October 2005.

[48] P. Sarigiannidis, G. I. Papadimitriou, and A. S. Pomportsis, "A high-throughput scheduling technique with idle timeslot elimination mechanism," *IEEE/OSA Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4811–4827, Dec. 2006.

[49] V. Sivaraman and G. N. Rouskas, "HiPeR-*l*: A high performance reservation protocol with *l*ook-ahead for broadcast WDM networks," in *Proceedings of IEEE Infocom*, 1997, pp. 1270–1277.

[50] D. Shea and J. Mitchell, "Architecture to integrate multiple PONs with long reach DWDM backhaul," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 2, pp. 126–133, Feb 2009.

[51] S. Photonics, "Luminent OLT Transceiver," http://www.sourcephotonics.com/datasheets/DS2201.pdf.

[52] S. Takahashi, K. Shiba, E. Mizuki, K. Makita, and A. Tajima, "Over 25-db dynamic range 10-/1-Gbps optical burst-mode receiver using high-power-tolerant APD," in *Proc. of OSA Conference on Optical Fiber Communication (OFC)*, Mar 2009, pp. 1–3.

[53] L. Wang, C. Schulz, S. Shang, and G. Tohmon, "Ultra-wide dynamic range receiver for noise loaded WDM transmission systems," *Optics Express*, vol. 16, no. 25, pp. 20 382–20 387, Dec 2008.

[54] S. Fukushima, A. Ohki, S. Kanazawa, and A. Okada, "Compact 10-Gb/s APD receiver module with a microsecond-order-response variable optical attenuator," in *Proc. of Annual Meeting of the IEEE Lasers and Electro-Optics Society (LEOS)*, Nov 2008, pp. 656–657.

[55] H. Lee, K. Reichmann, P. Iannone, X. Zhou, and B. Palsdottir, "A hybrid-amplified PON with 75-nm downstream bandwidth, 60 km reach, 1:64 split, and multiple video services," in *Proc. of OFC*, Feb 2007.

[56] F. Saliou, P. Chanclou, F. Laurent, N. Genay, J. Lazaro, F. Bonada, and J. Prat, "Reach Extension Strategies for Passive Optical Networks," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 1, no. 4, pp. C51–C60, Sep 2009.

[57] *ITU-T G.984.6: Gigabit-capable passive optical networks (GPON): Reach extension*.

[58] Alphion, "PON.ext Passive Optical Network Extension System," http://www.alphion.com/img/pdf/PB-PON.ext-1RU.pdf.

[59] M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, 3rd ed. Springer, 2008.

[60] D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research: Volume 1: Stochastic Processes and Operating Characteristics*. Courier Dover, 2003.

[61] F. Aurzada, M. Scheutzow, M. Herzog, M. Maier, and M. Reisslein, "Delay analysis of Ethernet passive optical networks with gated service," *OSA Journal of Optical Networking*, vol. 7, no. 1, pp. 25–41, Jan. 2008.

[62] "Csim (mesquite software)," http://www.mesquite.com.

[63] P. L. Tien, Y.-M. Lin, and M. C. Yuang, "A novel OFDMA-PON architecture toward seamless broadband and wireless integration," in *Proceedings of OFC*, Mar. 2009, pp. OMV2–1–OMV2–3.

[64] Z. Zheng and J. Wang, "A study of network throughput gain in optical-wireless (FiWi) networks subject to peer-to-peer communications," in *Proceedings of IEEE ICC*, June 2009.

Michael McGarry is an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Akron, OH. He received his B.S. in Computer Engineering from Polytechnic University, Brooklyn, NY in 1997. He received his M.S. and Ph.D. in Electrical Engineering from Arizona State University, Tempe, in 2004 and 2007, respectively. He is the recipient of the 2009 IEEE Communications Society Best Tutorial Paper award. His research interests include bandwidth forecasting, congestion control, and the optimization of MAC protocols.

Michael Scheutzow is a Professor in the Department of Mathematics at the Technical University Berlin, Germany. He received the Diploma in Mathematics from the Johann-Wolfgang-Goethe University Frankfurt/Main, Germany, in 1979. He received his Ph.D. in Mathematics from the University of Kaiserslautern, Germany, in 1983. He received his Habilitation in Mathematics from the University of Kaiserslautern in 1988. From 1988 through 1990 he was project leader with Tecmath GmbH, Kaiserslautern. Since 1990 he is Professor of Stochastics in the Department of Mathematics at the TU Berlin. From 1997 through 1999 he was Associate Chair of the department. He has visited the Univ. Carbondale, Ill., Rutgers Univ., Univ. of Rochester, Warwick Univ., and the MSRI at UC Berkeley.

Martin Reisslein (A96–S97–M98–SM03) is an Associate Professor in the School of Electrical, Computer, and Energy Engineering at Arizona State University (ASU), Tempe. He received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Germany, in 1994, and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1996. Both in electrical engineering. He received his Ph.D. in systems engineering from the University of Pennsylvania in 1998. During the academic year 1994–1995 he visited the University of Pennsylvania as a Fulbright scholar. From July 1998 through October 2000 he was a scientist with the German National Research Center for Information Technology (GMD FOKUS), Berlin and lecturer at the Technical University Berlin. From October 2000 through August 2005 he was an Assistant Professor at ASU. From January 2003 through February 2007 he served as Editor-in-Chief of the *IEEE Communications Surveys and Tutorials*. He currently serves as Associate Editor-in-Chief of the *IEEE Communications Surveys and Tutorials* and as Associate Editor for the *IEEE/ACM Transactions on Networking* and for *Computer Networks*. He has served on the Technical Program Committees of leading networking conferences, such as *IEEE Infocom*. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG–4 and H.264 encoded video, at `http://trace.eas.asu.edu`. His research interests are in the areas of video traffic characterization, wireless networking, optical networking, and engineering education.

Frank Aurzada studied mathematics at the Friedrich-Schiller University, Jena, Germany and the University of York, York, U.K. He received the Dipl.-Math. and Ph.D. degrees in mathematics from the Friedrich-Schiller University in 2003 and 2006, respectively. After completing his Ph.D., he joined the DFG Research Center Matheon at Technical University Berlin, where he is currently a Postdoctoral Researcher. His research interests lie in the queueing theoretic analysis of telecommunication networks, coding theory, and limit theorems in probability theory. He is currently the head of an Independent Junior Research Group at TU Berlin.