Just-in-Time Scheduling for Multichannel EPONs

Michael P. McGarry, *Member, IEEE*, Martin Reisslein, *Senior Member, IEEE*, Charles J. Colbourn, Martin Maier, *Member, IEEE*, Frank Aurzada, and Michael Scheutzow

***Abstract*—We investigate optical network unit (ONU) grant scheduling techniques for multichannel Ethernet passive optical networks (EPONs), such as wavelength division multiplexed (WDM) EPONs. We take a scheduling theoretic approach to solving the grant scheduling problem. We introduce a two-layer structure of the scheduling problem and investigate techniques to be used at both layers. We present an extensive ONU grant scheduling simulation study that provides: 1) insight into the nature of the ONU grant scheduling problem and 2) indication of which scheduling techniques are best for certain conditions. We find that the choice of scheduling framework has typically the largest impact on average queueing delay and achievable channel utilization. An offline scheduling framework is not work conserving and consequently wastes channel resources while waiting for all ONU REPORT messages before making access decisions. An online scheduling framework, although work conserving, does not provide the best performance since scheduling decisions are made with the information contained in a single ONU REPORT. We propose a novel online just-in-time (JIT) scheduling framework that is work conserving while increasing scheduling control by allowing the channel availability to drive the scheduling process. In online JIT, multiple ONU REPORTs can be considered together when making scheduling decisions, resulting in lower average queueing delay under certain conditions and a more effective service differentiation of ONUs.**

***Index Terms*—Dynamic bandwidth allocation (DBA), Ethernet passive optical network (EPON), media access control (MAC), scheduling, space division multiplexing (SDM), wavelength division multiplexing (WDM).**

I. Introduction

CURRENT Ethernet passive optical network (EPON) standards dictate a single channel used for downstream transmission and a single channel used for upstream transmission. The need for more passive optical network (PON) bandwidth capacity will drive up the utilization of multiple upstream and downstream channels. In an effort to provide more bandwidth capacity we can increase the bit-rate as well as utilize multiple upstream and downstream channels. The transition from increased bit-rate to utilizing multiple channels for an increase in bandwidth capacity will be a function of cost. At some point the transition to multiple transmission channels to increase bandwidth capacity will occur.

Besides an increase in bandwidth capacity there are additional benefits provided by utilizing multiple upstream and downstream channels. With multiple channels, several optical network units (ONUs) can transmit concurrently to the optical line terminal (OLT) thereby lowering the average queueing delay [1] experienced by the Ethernet frames queued at the ONUs. Multiple channels provide a method for dynamic bandwidth allocation (DBA) algorithms to reserve certain channels for certain traffic classes [2]. The discovery and registration process can be kept on a single channel which would allow transmissions on the other channels to be uninterrupted. Finally, selected channels can be used to provide all-optical OLT bypassing services [3].

Wavelength division multiplexed (WDM)-based multiple channel PONs were first proposed in the mid-1990s [4]. Recently, WDM PON architectures have regained interest as the enabling technologies have become mature [5]–[10]. Space division multiplexing (SDM) is another approach to channel separation, whereby each fiber strand carries a unique channel or channels. SDM can be combined with WDM for an even larger number of channels and service separation.

Dynamic bandwidth allocation (DBA) in multichannel EPONs can be viewed as consisting of grant sizing and grant scheduling. Grant sizing techniques have been examined in [11]–[13]. In this paper, we suppose that the grants are sized according to some existing technique and focus on the grant scheduling portion of the DBA problem for multichannel EPONs. We frame our investigation in the context of scheduling theory [14]. We model the grant scheduling problem using standard scheduling theory notation. We discover solution techniques for this model that result in a set of possible scheduling policies for producing a schedule given a set of ONU grants. More specifically, we partition the scheduling problem into: 1) a scheduling framework and 2) a scheduling policy operating within the adopted scheduling framework. Online scheduling, where the optical line terminal (OLT) schedules grants as soon as a REPORT message is received from an ONU, and offline scheduling where the OLT waits for REPORT messages from all ONUs before scheduling grants are the extreme ends of a continuum of possible scheduling frameworks. The choice of scheduling framework, as we will explain, depends on the level of control that is required by the scheduler. A hybrid between the two extremes provides generally the best performance.

This paper is organized as follows. In Section II we review related work on DBA for multichannel EPONs. In Section III we model the multichannel EPON ONU grant scheduling problem

Manuscript received June 26, 2007; revised December 14, 2007. This work was supported in part by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin. A preliminary overview of this work has appeared in *Proceedings of IEEE International Conference on Communications (ICC)*, Glasgow, U.K., pp. 2174–2179, June 2007.

M. P. McGarry and M. Reisslein are with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706 USA (e-mail: michael.mcgarry; reisslein@asu.edu).

C. J. Colbourn is with the Department of Computer Science and Engineering, Arizona State University, Tempe, AZ 85287-8809 USA (e-mail: charles.colbourn@asu.edu).

M. Maier is with the Institut National de la Recherche Scientifique, Montreal, QC H5A 1K6, Canada (e-mail: maier@emt.inrs.ca).

F. Aurzada and M. Scheutzow are with the Institute of Mathematics, Technical University Berlin, 10623 Berlin, Germany (e-mail: aurzada@math.tu-berlin.de; ms@math.tu-berlin.de).

Digital Object Identifier 10.1109/JLT.2008.919366

using standard scheduling theory notation. We then use this model to find good scheduling policies and discuss scheduling frameworks. In Section IV we discuss our online just-in-time (JIT) scheduling framework and how some scheduling policies can be adapted to be used within this framework. In Section V we present the results of our simulation study that analyzes the differences between the different scheduling techniques. We conclude in Section VI.

II. Related Work

We now discuss the related work on dynamic bandwidth allocation (DBA) for multichannel or WDM EPONs.

A. WDM Variants of IPACT

In [15], Kwong *et al.* propose a multiple upstream wavelength version of IPACT [12], called WDM IPACT-ST, whereby ST refers to single polling table. This algorithm keeps track of the wavelength available time of all upstream wavelengths. Upon receiving a REPORT from an ONU, the OLT schedules the ONU's next transmission grant on the next available wavelength. This algorithm assumes each ONU supports all wavelengths. This scheme is similar to the WDM IPACT we proposed in [16]. In [16] we supported differing ONU WDM architectures by selecting the next available supported wavelength for scheduling whereas WDM IPACT-ST does not allow for this flexibility. In [17], Clarke *et al.* propose SIPACT, another WDM extension of IPACT, that is similar to the WDM IPACT in [16].

B. Dynamic Bandwidth Allocation Schemes in Hybrid TDM/WDM Passive Optical Networks

Dhaini *et al.* [18] investigate WDM-PONs and dynamic wavelength and bandwidth allocation (DWBA) algorithms for two fundamental WDM-PON architectures they label A1 and A2. In A1, the ONUs are placed into wavelength sets and contain a fixed transmitter at the selected wavelength. In A2, the ONUs have tunable transmitters capable of transmission on several wavelengths. In both A1 and A2, the OLT has an array of fixed receivers for upstream reception. Dhaini *et al.* proceed to investigate the problem of dynamically allocating bandwidth in both time and wavelength. Their first approach, Static Wavelength Dynamic Time (SWDT), relies on architecture A1. SWDT statically assigns a wavelength channel to all grants from each ONU, this wavelength is the wavelength supported by the fixed transmitter at that ONU. Time is then managed using an existing single channel DBA.

Dhaini *et al.* then propose three variants of a dynamic wavelength and time bandwidth allocation. The first (DWBA-1) schedules ONUs after all REPORT messages have been received for a cycle. Further, DWBA-1 incorporates "fair" distribution of excess bandwidth. The second (DWBA-2) schedules underloaded ONUs upon receipt of their REPORT message and overloaded ONUs after receiving all ONU REPORT messages. When limiting grant sizes and distributing excess from ONUs not fully utilizing their guaranteed minimums, all REPORTs must be received in a cycle in order to know the excess from that cycle. Therefore, it makes sense to have overloaded ONUs wait until this information is available in order to properly size their grant. Finally, the third (DWBA-3) schedules all ONUs upon receipt of their REPORT message. Since the OLT needs to grant excess bandwidth, the authors create two grants in this approach. As soon as the ONU REPORT is received, one grants the guaranteed minimum; and after all REPORTs have been received, another grants the excess assigned to that ONU. There are two problems with that approach: 1) each overloaded ONU receives two grants which decreases efficiency due to more guard times, and 2) the split between the two grants will most likely not occur on frame boundaries causing one frame to be unnecessarily delayed till the next cycle. Therefore, DWBA-2 is more efficient.

C. Quality of Service in TDM/WDM Ethernet Passive Optical Networks

In [2], Dhaini *et al.* propose two approaches for providing QoS on a TDM/WDM EPON: 1) QoS-DBA-1, which uses the OLT scheduler (DWBA-2) of [19] with a Modified Deficient Weighted Round Robin (M-DWRR) intra-ONU scheduler, and 2) QoS-DBA-2 and QoS-DBA-3, which segregate constant bitrate expedited forwarding (EF) traffic [20] from assured forwarding (AF) [21] and/or best effort (BE) traffic by wavelength. Some wavelengths are allocated for EF traffic, and some are allocated for AF and BE traffic. In QoS-DBA-2, this segregation is strict where as in QoS-DBA-3 the unused capacity on the EF wavelengths can be utilized by AF and BE traffic.

D. Summary

The various WDM IPACT variants [15]–[17] do not consider the problem of grant scheduling. They simply apply a first-fit time and wavelength assignment. In [18], Dhaini *et al.* explore the problem of grant scheduling but only in the context of excess bandwidth distribution. In this paper we examine grant scheduling in a more general context and propose a new scheduling framework that is more efficient than either online or offline scheduling.

III. Grant Scheduling

The multichannel EPON scheduling problem can be formulated using the scheduling theory notation defined in [14]. Theoretical analysis of all the scheduling models discussed in this section can be found in [14]. Scheduling theory is concerned with scheduling a set of jobs with specific processing times to be executed on a set of machines as efficiently as possible with respect to an optimization criterion. We can view each ONU as representing a job, its grant size as defining its processing time, and the channels used for transmission on the EPON as representing the machines. In scheduling notation, a scheduling problem is defined by a triple $\alpha|\beta|\gamma$, where α describes the machine environment (e.g., single machine, parallel machines, etc.), β describes the processing characteristics and constraints, and γ describes the objective to be minimized.

For the formulation of the multichannel EPON grant scheduling problem in the scheduling theory notation, we let P denote identical parallel machines (channels) that define our machine environment. Our only processing characteristic or constraint is M_i which refers to machine (channel) eligibility constraints. Specifically, M_i is the set of machines (channels) that job (ONU) i can be executed (transmitted) on. Let C_i denote the time at which the transmission for ONU i is complete. With this notation we define the scheduling problem with

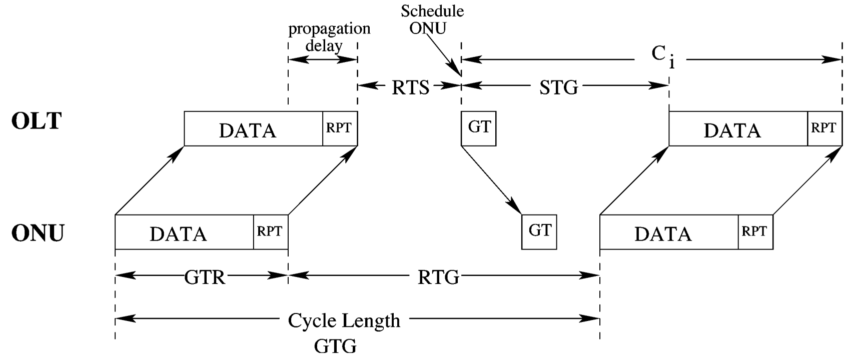


Fig. 1. Illustration of constituent delays of a scheduling cycle.

the objective to minimize the unweighted sum of the completion times as $P|M_i|\sum_i C_i$. The M_i processing constraint is required because each ONU has, in general, its own subset of supported channels. If all ONUs supported transmission on all wavelengths, we could remove the machine eligibility constraint to obtain $P||\sum_i C_i$.

Our performance objective in designing a scheduler for a WDM EPON is to lower the queueing delays experienced by frames in transit across the EPON and to increase the achievable resource (i.e., channel) utilization. To see how these performance objectives relate to the objectives from scheduling theory we first explore in detail the component delays in a scheduling cycle. We start by defining cycle length, which we also refer to as the GTG or GATE-to-GATE delay, as the time between back-to-back grants to an ONU. The component delays of a scheduling cycle are visualized in Fig. 1. GTR is the GATE-to-REPORT delay (since we append the REPORT at the end of the transmission window, GTR is equal to the transmission time of the grant), and RTG is the REPORT-to-GATE delay, that includes the propagation delay from ONU to OLT. Using GTR and RTG we express the cycle length as $\text{GTG} = \text{GTR} + \text{RTG}$.

The STG is the Schedule-to-GATE delay which is the time between the OLT scheduling an ONU's next grant to the time the grant starts being received at the OLT. The STG includes a propagation delay from OLT to ONU, a GATE message transmission time and propagation delay from ONU to OLT. The STG along with the grant time represents the completion time of an ONU's transmission from the point in time it is scheduled, i.e., $C_i = \text{STG} + \text{GTR}$. Since, the grant time (or size) is not determined by the scheduler, the scheduler can only work to minimize the variable portion of completion time, i.e., the STG. Minimizing C_i minimizes STG. The RTS is the REPORT-to-Schedule delay and is the delay from the OLT receiving a REPORT from an ONU to when the scheduling of the ONUs' REPORT is completed by the OLT. (The RTS includes the computation time for the schedule, which we neglect in our simulations in Section V.) Thus, REPORT-to-GATE (RTG) delay is composed of the RTS and STG delays, i.e., $\text{RTG} = \text{RTS} + \text{STG}$.

We introduce a layered approach to scheduling (see Fig. 2). We refer to the first layer as the **scheduling framework**, and the second layer as the **scheduling policy**. The scheduling framework is a logistical framework that determines when the OLT makes scheduling decisions, whereas the scheduling policy is a method for the OLT to produce the schedule. The OLT can produce a schedule with partial information about the ONU transmissions to be scheduled or after waiting to receive all of the information about the ONUs transmissions to be scheduled. Once the OLT has a set of ONU transmissions to schedule, the OLT uses a scheduling policy to create the schedule. The scheduling framework impacts the RTS delay by determining at what time after a REPORT message is received an ONU's next grant is scheduled. On the other hand, the scheduling policy impacts the STG delay by determining when an ONU grant is actually transmitted on a channel.

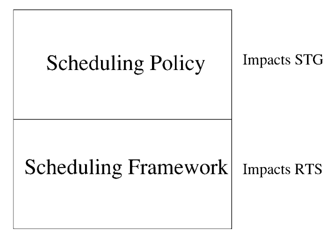


Fig. 2. Layered approach to scheduling.

In Section III-A, we discuss scheduling frameworks in more detail, and in Section III-B we discuss scheduling policies for multichannel EPONs.

A. *Scheduling Frameworks*

As mentioned above, the scheduling framework determines when the OLT produces a schedule. If the OLT produces a schedule as soon as any ONU REPORT is received without waiting for REPORT messages from other ONUs, this is referred to as an *online* scheduling framework. However, if the OLT were to wait for the REPORT messages from all the ONUs to be received before making scheduling decisions, this is referred to as an *offline* scheduling framework. The scheduling framework can be viewed as a continuum between the extremes of online and offline scheduling as illustrated in Fig. 3. On the online scheduling end of the continuum, the OLT only considers a single ONU REPORT in a scheduling decision. On the offline scheduling end of the continuum, the OLT considers all ONU REPORTs in a scheduling decision. Any scheduling framework that lies between online and offline is some form of an online scheduling framework because not all of the ONU REPORTs have been received. We will however reserve the term online scheduling framework to indicate the case where the OLT considers only one ONU REPORT at a time. In Section III-A1 we will explore offline scheduling in

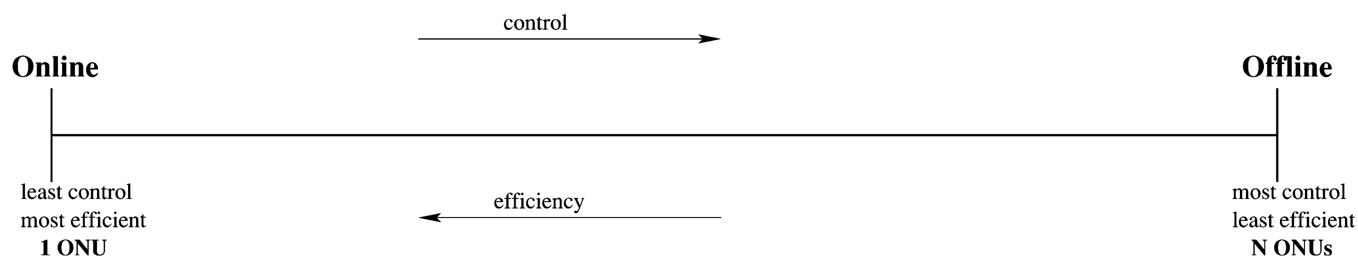


Fig. 3. Scheduling framework continuum (N: number of ONUs).

the context of multichannel EPONs, and in Section III-A2 we will explore online scheduling in the context of multichannel EPONs. We will see that both schemes have very different channel utilization characteristics.

1) Offline Scheduling: In an offline scheduling framework, scheduling decisions are made with full knowledge of all the jobs to be scheduled including their processing times for a particular scheduling cycle. Specifically for a multichannel EPON, an offline scheduling framework schedules ONU grants for transmission when the OLT has received the current MPCP REPORT messages from *all* ONUs. This allows the OLT to take the current bandwidth requirements of all ONUs into consideration in the grant sizing and scheduling, i.e., the scheduling pool contains all ONUs. The scheduling policy is executed after the OLT receives the end of the last ONU's gated transmission window. The RTS delay for the last ONU will be negligible; however the RTS may not be negligible for the other ONUs. This RTS delay introduces further queueing delays in the ONUs because it introduces additional delay in the cycle length (GTG) for an ONU. Waiting for all ONU REPORT messages to be received results in wasted channel capacity. The wasted capacity increases as the number of channels increases.

2) Online Scheduling: In an online scheduling framework, an ONU is scheduled for upstream transmission as soon as the OLT receives the REPORT message from the ONU, i.e., the scheduling pool contains one ONU. This is the scheduling framework that is at the far end of the online side of the scheduling framework continuum depicted in Fig. 3. This approach avoids wasted channel capacity by not keeping any channels idle while there is an ONU REPORT message for the OLT to act on.

B. Scheduling Policies

We now use our scheduling model developed in the beginning of Section III to find the best scheduling policies for a multichannel EPON that supports an evolutionary migration from single-channel EPONs to multichannel EPONs. The OLT uses these scheduling policies once a set of ONU grants to be scheduled has been determined by the scheduling framework.

1) Next Available Supported Channel: A simple scheduling policy for a multichannel EPON considers one ONU at a time and schedules the upstream transmission for that ONU on the wavelength channel that is available the earliest among the channels supported by that ONU. We refer to this scheduling policy as the *next available supported channel (NASC)* policy. NASC is our variation on an algorithm proposed by Graham [22] nearly 40 years ago called the *List* algorithm for identical parallel machines. This algorithm schedules jobs one by one and assigns them to the next available machine.

2) Parallel Machine Models and Solutions: We refer to [23] for a detailed exploration of scheduling policies for our parallel machine environment scheduling model. Least flexible job first (LFJ) with shortest processing time (SPT) first dispatching rule was shown to be a good heuristic for this model.

3) Unrelated Machine Models and Solutions: Another possible approach to modeling the multichannel EPON grant scheduling problem is to loosen our original parallel machine environment model by recognizing that $P|M_i|\sum_i C_i$ can be viewed as a special case of $R||\sum_i C_i$, where R refers to an unrelated machine environment where each machine executes a job at a different speed. For machines that are in M_i, we set the execution time to the processing time or grant length; for machines not in M_i, we set the execution time to infinity. In [23] we pointed out that a weighted bipartite matching formulation can optimally solve $R||\sum_i C_i$.

4) Summary: Using our parallel machine model, the results from scheduling theory indicate a few dispatching rules that can provide good scheduling policies. A dispatching rule (see [14, Secs. 14.1 and 14.2]) is a defined method of ordering jobs for dispatch in first fit fashion on available machines. Some examples of general dispatching rules are: Least Flexible Job (LFJ) first, Shortest Processing Time (SPT) first, and Largest Processing Time (LPT) first. The best scheduling policies for our multichannel EPON grant scheduling model are the dispatching rules discussed in Section III-B2; LFJ for machine eligibility restrictions, SPT for minimizing the sum of the completion times, and LPT for minimizing the maximum completion time. Other potential dispatching rules for the multichannel EPON grant scheduling problem are: Largest Number of Frames (LNF) first, Earliest Arriving Frame (EAF) first, and Earliest Average Arrival (EAA) first. LNF favors ONUs with more queued Ethernet frames, EAF favors ONUs that have the earliest arriving head-of-line (HOL) Ethernet frame, and EAA favors ONUs that have the earliest average Ethernet frame arrival time.

Dispatching rules can be used alone or grouped together to form composite dispatching rules. For multichannel EPON grant scheduling, LFJ can be combined with some of the other dispatching rules to create a composite dispatching rule that can provide better performance. The second dispatching rule is used to break ties from the first dispatching rule. Rather than using the second dispatching rule for tie breaking in the first dispatching rule, a weight can be set for each dispatching rule in the composite.

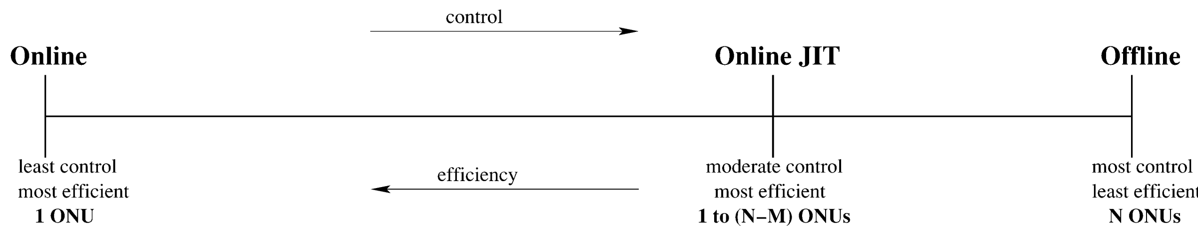


Fig. 4. Online JIT on the Scheduling Framework Continuum (N: Number of ONUs, M: Number of Channels).

The weighted bipartite matching (WBM) formulation that is proven optimal for minimizing the sum of the completion times for the unrelated machine environment scheduling model operates differently from the dispatching rules. Unlike the dispatch rules, the WBM scheduling policy results in a direct assignment of each ONU grant to a specific channel and time (position). There is no need to apply NASC for the scheduling. The output from the WBM scheduling policy fully characterizes the schedule. This is in contrast to the dispatching rules that specify the order in which the ONU grants are scheduled for first fit channel assignment according to NASC.

IV. Online JIT Scheduling

We now present a new scheduling framework that is a hybrid between offline and online scheduling discussed in Sections III-A1 and A2 respectively. We call this new scheduling framework online just-in-time (JIT) scheduling. The name indicates that scheduling is performed in just in time fashion. In our online JIT scheduling framework, ONUs are added to a scheduling pool as their MPCP REPORT messages are received by the OLT. When a wavelength becomes available, the ONUs in the pool are scheduled together according to the selected scheduling policy across all wavelengths. The ONUs that are scheduled so that their transmissions would start shortly (i.e., close to the one-way propagation delay from OLT to the ONU into the future) after the time they are scheduled are classified as "imminent"; the current schedule for these ONUs is considered firm. The OLT transmits GATE messages to these ONUs to inform them of their granted transmission window. The remaining ONUs are classified as "tentative" and can remain in the scheduling pool for the next scheduling round. Alternatively, all ONUs (i.e., both imminent and tentative) can always be firmly scheduled. We refer to the case where all ONUs are firmly scheduled as online JIT and the case where the tentative ONUs participate in future scheduling rounds as online JIT Tentative.

The online JIT scheduling framework gives the OLT more opportunity to make better scheduling decisions than standard online scheduling. ONUs are scheduled at the moment right before they potentially begin transmitting. To facilitate this on an EPON, we need to ensure the GATE message is transmitted by the OLT at least the one-way OLT-to-ONU propagation delay before we intend the ONU to begin transmission. In other words, the GATE message must be transmitted soon enough to accommodate the OLT-ONU-OLT round trip time (RTT) before we want to begin receiving the ONU's transmission at the OLT. Using the largest RTT in the EPON for this timing of the GATE message transmissions ensures that any ONU receives the GATE message in time. Since we desire the ONU to transmit as soon as the next wavelength becomes free, we need to schedule the ONUs in the pool at least an RTT before the next wavelength free time. Fig. 4 illustrates where the online JIT scheduling framework lies on the scheduling framework continuum. The online JIT scheduling framework can lie somewhere from the online scheduling framework up to a point just short of the offline scheduling framework. Let us consider the bounds of where the online JIT scheduling framework can lie with respect to the number of considered ONU REPORTs.

To obtain the lower bound, consider an EPON with very low traffic load, i.e., very few ONUs have traffic and those ONUs with traffic have small queue occupancies. In such a low traffic scenario, there are always free upstream channels. When the OLT receives an ONU REPORT, the OLT makes immediately a scheduling decision based on this one REPORT, i.e., the lower bound is 1 ONU REPORT.

To obtain the upper bound, consider a high traffic load scenario where all N ONUs report high queue occupancies resulting in grants that are larger than one RTT. In this high traffic scenario, the scheduling pool at the OLT contains all ONUs, except those that are currently transmitting (and will send their REPORTs at the end of their transmissions). That is, with N ONUs in an EPON with M upstream channels, there are $N-M$ ONUs in the scheduling pool, which is the upper bound on the number of REPORTs considered for online JIT scheduling. Thus, the online JIT scheduling framework can get very close to emulating an offline scheduling framework, especially with a small number of channels.

If the OLT has $N-M+k,\ k=1,2,\ldots,M$, ONUs in the scheduling pool, then k wavelengths are idle for one RTT to permit for the k REPORTs to reach the OLT and the corresponding grants to propagate to the ONUs. Hence, with the offline scheduling framework, for which $k=M$, the M upstream channels are not utilized for one RTT to allow for all N REPORTs to be received by the OLT and the GRANTs to propagate to the ONUs. Thus, the offline scheduling framework utilizes the channels less efficiently than the online and online JIT scheduling frameworks. On the other hand, both the online and online JIT scheduling frameworks are fully work-conserving, i.e., they do not let any upstream channel go unused while there is data to transmit. Importantly, note that the online JIT scheduling framework does not introduce any inefficiencies compared to the online scheduling framework, as further confirmed in Section IV-C. The only added complexity with the online JIT scheduling framework is keeping track of and scheduling the REPORTs in the scheduling pool, which can range from one to $N-M$ REPORTs, whereas only one REPORT is considered at a time in the online scheduling framework.

When using the online JIT Tentative scheduling framework, an ONU may participate in several scheduling rounds as "tentative" before it becomes firmly scheduled. It is possible that certain ONUs that are unfavorable to a particular scheduling policy can continuously be preempted by those that are more favorable. To prevent these ONUs from being starved of medium access, an aging mechanism is incorporated to keep these "less favorable" ONUs (or jobs) from being starved by the scheduler. A straightforward method to implement starvation prevention is to set a threshold at which an ONU is immediately scheduled on the next available wavelength regardless of the scheduling policy. This ensures that no ONU waits indefinitely for medium access. This threshold can be based on the number of participated scheduling rounds to adapt to changing cycle times.

A. Dispatching Rules for Online JIT

Dispatching rules or composite dispatching rules can be used without any modification within the online JIT scheduling framework. The dispatching rules result in an ordering of the ONUs in the scheduling pool. This ordering is used in conjunction with NASC to schedule an ONU grant at a specific time on a specific channel.

B. Weighted Bipartite Matching Adapted for Online JIT

The standard weighted bipartite matching (WBM) scheduling formulation [14] for minimizing the sum of the completion times needs to be modified to support an online scheduling framework. In any online scheduling framework, not all machines are immediately available for scheduling (i.e., they may still be processing jobs). We introduce this in the WBM formulation by setting an additive cost to a matching that is different for each machine. This additive cost is related to when the wavelength becomes available. We refer to this cost as a_{ji}, the availability cost of wavelength j for ONU i. Let t_i^{RTT} be the RTT delay for ONU i, t_i^{REPORT} be the time the REPORT message from ONU i is received at the OLT, Λ_j be the time when wavelength j is free, and t_i^{READY} be the time when ONU i is ready to transmit. Given $t_i^{\mathrm{READY}} = t_i^{\mathrm{REPORT}} + t_i^{\mathrm{RTT}}$, then $a_{ji} = \mid \Lambda_j - t_i^{\mathrm{READY}} \mid$. a_{ji} is the Euclidean distance between wavelength free time and ONU ready to transmit time. A weight can be used to control how much this availability cost affects the solution, we will use δ to represent this weight.

The following is the Integer Program that represents the WBM where k is the scheduling position, p_{ji} is the grant processing time for ONU i on channel j (either ONU grant time for supported channel j, or ∞ for nonsupported channel j), a_{ji} is the availability cost for channel j, x_{jki} are binary variables representing whether or not position k on machine (channel) j is selected for job (ONU) i, m is the number of machines (channels), and n is the number of jobs (ONUs):

$$\text{minimize} \sum_{j=1}^{m}\sum_{i=1}^{n}\sum_{k=1}^{n}(k \cdot p_{ji} + \delta \cdot a_{ji}) \cdot x_{jki} \quad (1)$$

subject to

$$\sum_{j=1}^{m}\sum_{k=1}^{n} x_{jki} = 1, \forall_i \quad (2)$$

$$\sum_{i=1}^{n} x_{jki} \le 1, \forall_j, \forall_k \quad (3).$$

The first constraint forces an ONU i to be assigned to only one scheduling position. The second constraint forces each scheduling position to be assigned to no more than one ONU. If a single ONU supplies traffic from multiple classes, each traffic class is treated as a separate job in the WBM formulation.

C. Stability Analysis

In this section we formally analyze the stability characteristics of the online JIT scheduling framework. Stability in the context of EPONs has so far been primarily examined for grant sizing techniques employing prediction of traffic newly arriving between sending a REPORT and the start of the corresponding upstream transmission, e.g., see [24]–[27]. These analyses consider the prediction control loop and examine controllability and stability of the grant sizing for rapid fluctuations in the traffic loads; whereby the grant size prediction is considered stable when ONUs receive a fair bandwidth share. The scheduling of the grants and the resulting utilization of the upstream transmission channels are not explicitly considered in the existing studies. Our stability analysis is fundamentally different from the existing analyses in that we examine whether the generated long-term traffic load can be accommodated on the upstream transmission channels. We consider grant sizing without prediction in our analysis. We remark that the online JIT scheduling framework could be used in conjunction with grant sizing techniques employing prediction; the analysis of such a combined system is left for future work.

Recall that we consider a multichannel EPON with N ONUs and M channels. Let $m_{ij} = 1$ if ONU i supports channel j, and $m_{ij} = 0$ if ONU i does not support channel j, for $i = 1, \ldots, N$ and $j = 1, \ldots, M$. Let λ_i denote the long-run average packet (Ethernet frame) generation rate at ONU i (in packets/second), and let $\bar{L}_i$ denote the average packet length (in bits). Further, let C denote the transmission bit-rate of an upstream channel (in bits per second). We define the relative loads $\rho_i = \lambda_i \bar{L}_i / C$ for $i = 1, \ldots, N$.

We begin the analysis by considering the upstream transmissions on a given upstream channel. We define the *upstream transmission* of an ONU to consist of the upstream transmission of Ethernet frames and the MPCP REPORT. Consider an arbitrary upstream transmission of a given ONU. Let D be a random variable denoting the transmission time of the payload data (Ethernet frames) in the upstream transmission. Let O_p be a random variable denoting the "proportional" overhead in the upstream transmission in terms of transmission time (i.e., number of overhead bits divided by channel transmission bit-rate C). The proportional overhead O_p accounts for the preamble of eight bytes for each Ethernet frame in the upstream transmission and the interpacket gap of 12 bytes between successive Ethernet frames in the upstream transmission. Note that O_p is given by a constant times the payload transmission time, i.e., $O_p = \omega \cdot D$ (whereby $\omega = o_p/\bar{L}$, with o_p denoting the fixed overhead per Ethernet frame and $\bar{L}$ denoting the mean of the considered Ethernet frame lengths). Let o_f be a constant denoting the "fixed" overhead in the upstream transmission, i.e., the MPCP Report, plus the guard time between successive upstream transmissions.

The *utilization* of the upstream channel during the considered upstream transmission is then

$$\eta = \frac{D}{D(1+\omega)+o_f}. \tag{1}$$

To examine the maximum achievable utilization, we initially consider two ONUs that transmit upstream on a given upstream channel. (We incorporate the impact due to the possibly restricted set of wavelengths supported by an ONU shortly.) With gated service and sufficiently high loads, the upstream transmission of an ONU is typically sufficiently long to mask the round-trip propagation REPORT-to-GATE delay for the other ONU. (For shorter upstream transmissions, we can consider more ONUs to mask the round-trip propagation delay.) Then, for any combination of scheduling framework and scheduling policy that ensures that there is exactly a guard time between successive upstream transmissions, the utilization on the considered upstream channel is equal to the η given in (1). The maximum achievable utilization arises when the transmission grants become very large such that $D(1+\omega) \gg o_f$ and is given by

$$\eta_{\max} = \frac{1}{1+\omega}. \tag{2}$$

We now turn to the constraints imposed by the transmission capabilities m_{ij} of the individual ONUs $i,\ i=1,\ldots,N$ on the upstream channels $j,\ j=1,\ldots,M$. We immediately observe that an ONU with a single transmitter can only transmit on one channel at a time. Hence, the ONU must not generate more traffic load than it can transmit, i.e., we have to require that $\lambda_i \bar{L}_i < C\eta_{\max}$, or equivalently that $\rho_i < \eta_{\max}$, for all $i=1,\ldots,N$.

We claim that if the EPON upstream transmission system is stable, then there are long-run average relative transmission rates $r_{ij} \geq 0$ such that

$$\sum_{j=1}^{M} r_{ij} m_{ij} = \rho_i, \quad \text{for all } i=1,\ldots,N \tag{3}$$

and

$$\sum_{i=1}^{N} r_{ij} m_{ij} < \eta_{\max}, \quad \text{for all } j=1,\ldots,M. \tag{4}$$

This can be seen as follows. Note that the relative transmission rates are obtained by normalizing long-run average transmission rates R_{ij} (in bits/second) of ONU i on channel j, by the channel bit-rate C, i.e., $r_{ij} = R_{ij}/C$. For any ONU i, the relative transmission rates over all channels that ONU i supports add up to ρ_i; thus, (3) follows. On the other hand, in the long run, one cannot send more over channel j than permitted by the maximum utilization $\eta_{\max}$ of the channel transmission bit-rate, which is (4).

If (3) and (4) hold, then one can construct a static periodic transmission strategy, similar to a time division multiplexing (TDM) transmission, that is indeed stable. To see this, note that (3) and (4) imply that we know how much traffic ONU i can send on channel j in the long run without causing stability problems. Namely, this is exactly r_{ij}. So, we allocate $r_{ij}/(\sum_{k=1}^{N} r_{kj})$ fraction of time of channel j to ONU i. This can be done in a periodic fashion resembling a TDM strategy. Then (3) ensures that this strategy accommodates all traffic load generated at each ONU and (4) guarantees that one does not send more on any channel in the long run than permitted by the maximum achievable utilization of the channels's transmission bit-rate, i.e., we see that this periodic transmission strategy is stable.

Any grant sizing mechanism that bases grants on actual queue occupancies (to avoid allocation of excess bandwidth that would then go unused), in combination with any work conserving scheduling framework and policy that ensures the minimum spacing between upstream transmissions (so that there is no unnecessary unused time on the upstream channels) achieves the same stability limit as the periodic transmission strategy. In particular, gated service grant sizing in conjunction with both the online and online JIT scheduling frameworks with any scheduling policy spacing upstream transmissions on a channel by no more than the guard time are stable if and only if (3) and (4) hold. We remark that when all ONUs support all channels, i.e., when $m_{ij}=1$ for all i,j, then (3) and (4) hold if and only if $\sum_{i=1}^{N} \rho_i < M\eta_{\max}$.

We briefly remark regarding a formal analysis of the delay performance of online JIT that online JIT scheduling always makes scheduling decisions no earlier than online scheduling. Hence, online JIT can make better scheduling decisions. However, it is relatively easy to find examples where online JIT and online scheduling behave in exactly the same way. In general, it can therefore only be proven that online JIT never makes worse scheduling decisions than online scheduling, which is straightforward to verify. (Note that the converse is not true, because online scheduling makes decisions with less information and hence it cannot mimic online JIT.)

V. Simulation Experiments

In this section we study different scheduling techniques that can be used for grant scheduling in multichannel EPONs by means of simulations. We developed a multichannel EPON simulation engine using the CSIM [28] simulation library. Each wavelength supports $C = 1$ Gbps transmission bit-rate, and the reported load corresponds to the payload data rate $\sum_{i=1}^{N} \lambda_i \bar{L}_i$. 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. The simulations were conducted using self-similar traffic sources [29] with a Hurst parameter of 0.75. The RTTs were uniformly distributed over $[13\mu\text{s},\ 100\mu\text{s}]$, which corresponds to distances of 2–15 km.

In our simulations we use the gated grant sizing technique which grants each ONU its full bandwidth request. The gated sizing technique has been demonstrated to achieve small delays in EPONs [30]. By fixing the grant sizing technique, we are comparing the scheduling aspects of the multichannel EPON.

A. Offline Scheduling Versus Online Scheduling

We refer the reader to our simulation study presented in [23] for a comparison of offline versus online scheduling. Our results showed that an online scheduling framework significantly outperformed the offline scheduling framework regardless of scheduling policy with respect to average queueing delay. The better scheduling decisions made by the scheduling policies used with the offline scheduling framework were not enough to overcome the RTS delay that dominated the cycle length. Therefore, the average queueing delay values were significantly higher with the offline scheduling framework. We provide further results for the comparison of offline and online scheduling in Section V-C.

B. Online Scheduling Versus Online JIT Scheduling

We want to study the impact of the type of scheduler and the ONU multichannel diversity. To vary the ONU multichannel diversity, we simulated an EPON with $N = 32$ ONUs and created three sets of ONU WDM configurations: WDM Mix 1, WDM Mix 2, and WDM Mix 3. WDM Mix 1 contains 16 ONUs that support all wavelengths, 8 that support half of the wavelengths, and 8 that support the other half. WDM Mix 1 provides ONU multichannel diversity without any single channel ONUs. By not having any single channel ONUs, the scheduling process has increased flexibility in wavelength assignment. WDM Mix 2 contains 16 ONUs that support all wavelengths, 6 that support half of the wavelengths, 6 that support the other half of the wavelengths, and 4 that only support one wavelength. WDM Mix 2 provides ONU multichannel diversity with some single channel ONUs. Finally, WDM Mix 3 provides less multichannel diversity with 8 ONUs that support all wavelengths and 24 ONUs that support only one wavelength. Of the three ONU WDM configurations, WDM Mix 3 is the most restrictive with respect to wavelength assignment.

In the experiments for WDM Mix 1 and WDM Mix 2, all ONUs generate the same traffic load. For the WDM Mix 3 experiments, each of the eight ONUs supporting all channels generates nine times the traffic load of each of the 24 ONUs supporting one channel; this load distribution gives a theoretical stability limit of 4 Gbps for the WDM Mix 3 scenario.

We varied the scheduler for each of the ONU WDM configurations. Two dispatching rule-based schedulers from our parallel machine scheduling model (see Section III-B2): LFJ-SPT and LFJ-LNF, and two weighted bipartite matching (WBM)-based schedulers with weight $\delta = 10$ from our unrelated machine model (see Section III-B3): WBM and WBM-LNF. These schedulers are all compared to an online scheduler: NASC that provides the baseline for performance.

Figs. 5, 7, and 8 show the average queueing delay plotted against load for WDM Mix 1 with $M = 8$ total wavelengths, WDM Mix 2 with $M = 8$ total wavelengths, and WDM Mix 3 with $M = 4$ total wavelengths, respectively.

For WDM Mix 1, plots a) and b) in Fig. 5 show up to a 10% decrease in average queueing delay at low and moderate loads provided by the WBM-based schedulers. The plotted confidence intervals indicate that the difference appears to be statistically significant. The confidence intervals were obtained through the CSIM batch means method with batches sized to minimize correlations, and individual batch means representing the average queueing delays of the Ethernet frames occurring in a given batch. The lower average queueing delay achieved by the WBM-based schedulers is due to more efficient wavelength scheduling decisions made by the scheduler. The WBM-based schedulers evaluate the costs of all possible wavelength assignments and select the lowest cost matching, i.e., the wavelength for which the availability time Λ_j most closely matches the time when the ONU is ready t_i^{READY}. From plots of the wavelength utilization of each of the eight wavelengths for each of the scheduling schemes, which are not included due to space constraints, we observed that all compared scheduling strategies achieved equally good load balancing. The WBM-based schedulers are simply making more efficient wavelength assignment decisions that are resulting in shorter cycle lengths. Fig. 6 shows the cycle length for the WBM-based schedulers compared to NASC.

At higher loads a different pattern emerges, we observe from plot b) in Fig. 5 that at high loads the schedulers that use LNF

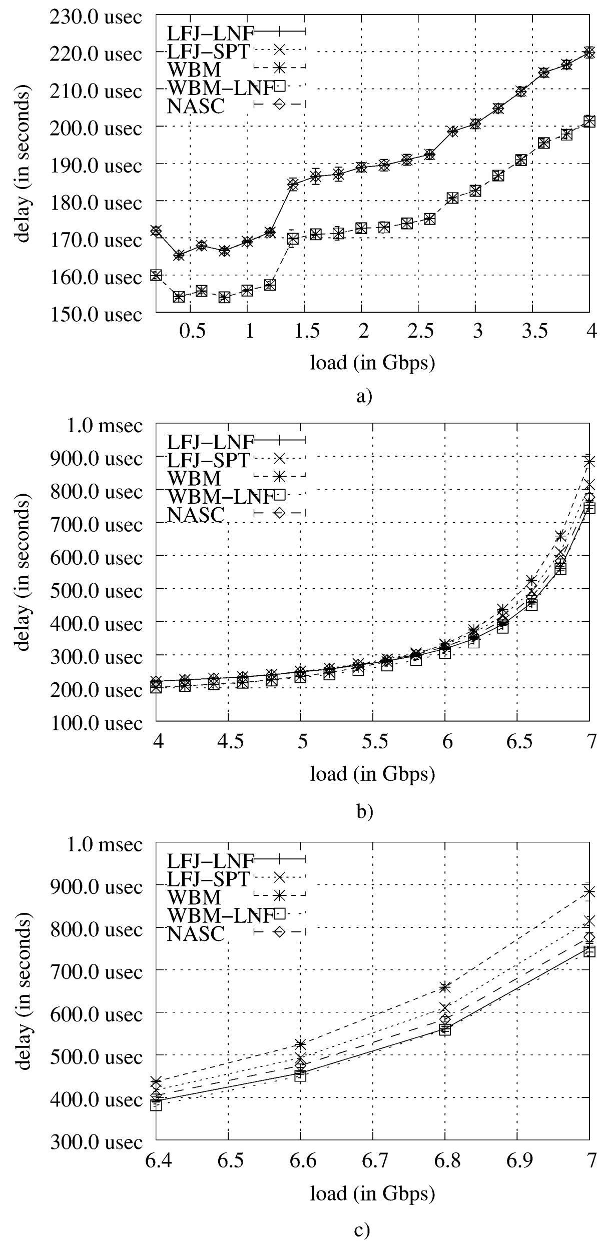


Fig. 5. Average queueing delay for WDM Mix 1, eight wavelength EPON. The WBM-based schedulers can provide lower queueing delays at lower loads. However, schedulers that incorporate the LNF dispatching rule provide lower queueing delays at higher loads. (a) Low load. (b) High load. (c) Very high load.

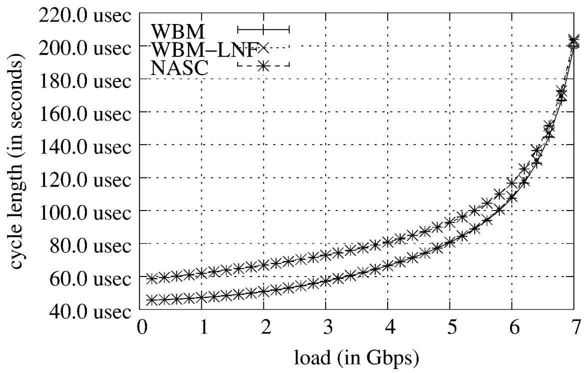


Fig. 6. Average cycle length for WDM Mix 1, eight wavelength EPON. The two WBM-based schedulers, whose curves fall on top of each other in the figure, provide lower cycle lengths due to better wavelength assignment.

(i.e., LFJ-LNF or WBM-LNF) provide a small improvement over the other schedulers. Plot c) in the same figure confirms this by zooming in on the higher loads. Again, the plotted confidence intervals appear to indicate a statistically significant difference. This indicates that at high loads the wavelength assignment has limited impact on average queueing delay. However, preferring ONUs with a larger number of queued frames can lower the average queueing delay. This is largely due to a frame sampling effect: ONUs with more frames have a larger impact on the average queueing delay measure than ONUs with fewer frames.

For WDM Mix 2, with less wavelength assignment flexibility due to the single wavelength ONUs, we observe from Fig. 7 a smaller decrease in average queueing delay provided by the WBM-based schedulers. The plots a) and b) in Fig. 7 indicate about a 3% decrease in average queueing delay. The plotted confidence intervals indicate that the difference is not statistically significant. At high loads, shown in plot b) and zoomed in plot c), we see the same pattern as seen for WDM Mix 1: at high loads the wavelength assignment has no impact on average queueing delay, but time ordering does have an impact due to a frame sampling effect.

For WDM Mix 3, with its limited ONU multichannel diversity and restrictive wavelength assignment capabilities, there is no measurable difference with respect to average queueing delays between the schedulers as observed from Fig. 8.

In summary, with increased ONU multichannel diversity and wavelength assignment flexibility, the WBM-based schedulers are able to provide lower average queueing delays. At high loads, the schedulers that favor ONUs with the largest number of frames provide lower average queueing delays, because of a frame sampling effect. Therefore, the online JIT scheduling framework has some utility in lowering the average queueing delay. However, its utility is much larger than this, as will be discovered in the next section.

C. Differentiated ONU Treatment Using Online JIT Scheduling

In this set of simulation experiments we explore how the online JIT scheduling framework can be used to provide differentiated treatment to ONUs without using an offline scheduling

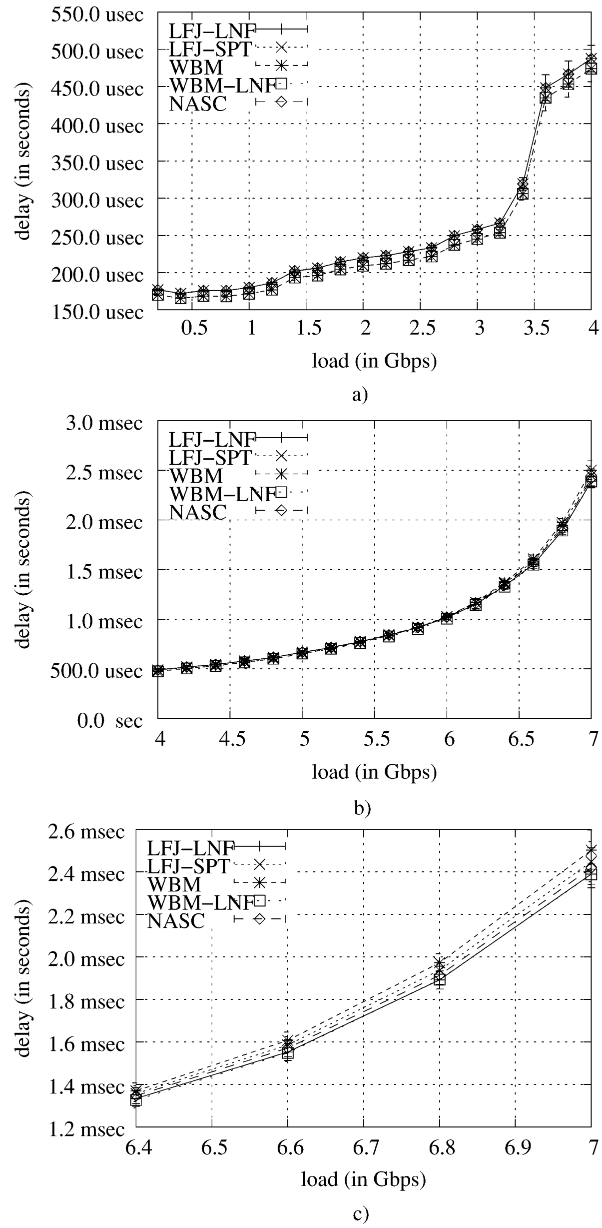


Fig. 7. Average queueing delay for WDM Mix 2, eight wavelength EPON. With less wavelength assignment flexibility as compared to WDM Mix 1, the reduction in average queueing delays achieved by the WBM-based schedulers is less pronounced. (a) Low Load. (b) High load. (c) Very high load.

mechanism. Avoiding use of an offline scheduling framework improves channel utilization and consequently lowers queueing delays. We have simulated the same EPON system described above for WDM Mix 1, WDM Mix 2, and WDM Mix 3. We now use a scheduler that always schedules two preferred ONUs, which support all wavelengths, ahead of any of the other 30 ONUs. The other 30 ONUs are scheduled with the LFJ-LNF dispatching rule.

Figs. 9–11 show plots of the average queueing delay experienced by all ONUs (labeled "Avg ONU") and the 2 preferred ONUs (labeled "Pref ONU") for the online JIT scheduling framework, the online JIT Tentative scheduling framework

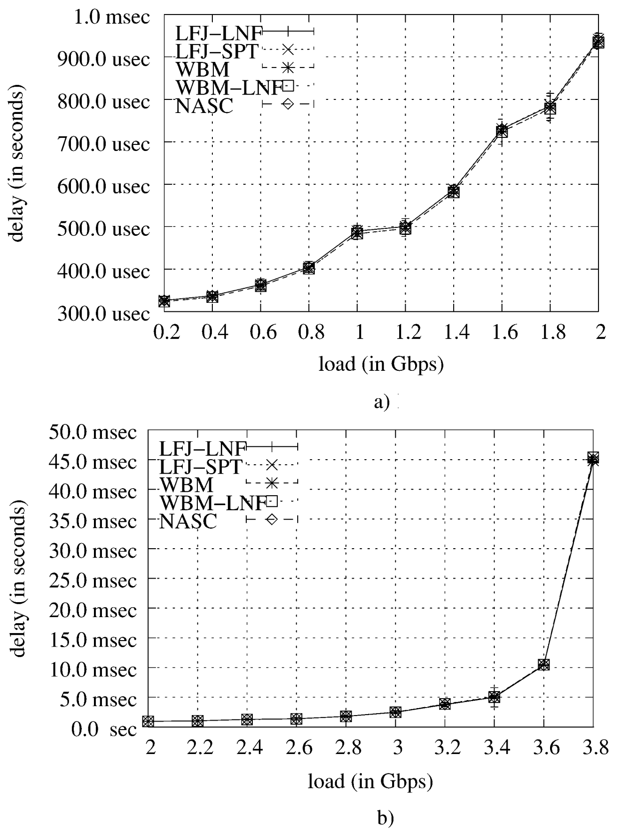


Fig. 8. Average queueing delay for WDM Mix 3, four wavelength EPON. With very limited wavelength assignment flexibility, the WBM-based schedulers do not provide a reduction in average queueing delay. (a) Low Load. (b) High load.

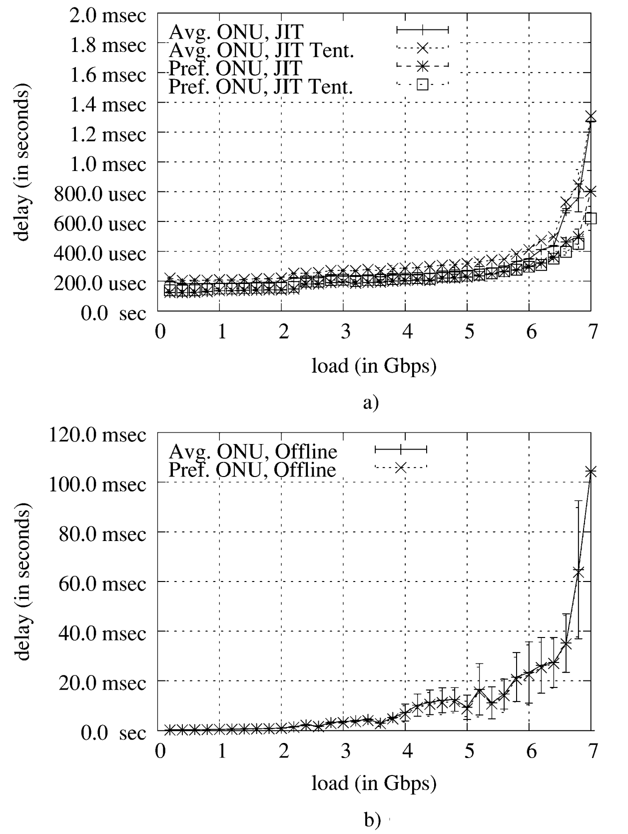


Fig. 9. Queueing delay for average ONUs versus the two ONUs given preferential scheduling (WDM Mix 1). Average ONUs are scheduled using the LFJ-LNF dispatching rule. The online JIT scheduling framework is able to provide significant differential treatment through scheduling. The oOffline scheduling framework does not provide differential treatment through scheduling. (a) Online JIT scheduling framework. (b) Offline scheduling framework.

with a starvation threshold set to 32 scheduling rounds, and the Offline scheduling framework.

Examining first the performance for the online JIT scheduling frameworks, we observe that they are able to provide differential treatment without using an offline scheduling framework. At high loads, the difference in average queueing delay becomes quite significant. For example, Fig. 10 shows that at a load of 6.8 Gbps, the preferred ONUs experience an average queueing delay of approximately 500 μs as opposed to approx. 1.6 ms for all ONUs. Comparing the figures for the different WDM Mixes, we see that as we move from WDM Mix 1 to WDM Mix 2, i.e., as we increase the number of single channel ONUs from zero to four, the average queueing delay increases but the queueing delay experienced by the two preferred ONUs stays the same. In Fig. 11 we see that when we increase the single channel ONUs to 75% of all ONUs and limit the EPON to four wavelengths, the average queueing delay increases significantly. However, the queueing delay experienced by the two preferred ONUs is significantly lower than the delay for the other ONUs. At a load of 3.8 Gbps, the delay for the preferred ONUs is nearly eight times smaller than the average queueing delay.

Comparing the online JIT scheduling framework with the online JIT Tentative scheduling framework, we observe a slight reduction in the average delay for the preferred ONUs with the online JIT Tentative scheduling framework as compared to the online JIT scheduling framework.

We summarize the main observations from comparing the online JIT scheduling framework with the offline scheduling framework as follows. First, we observe that the achievable maximum channel utilization (stability limit) is lower for offline scheduling as compared to online scheduling. Fig. 11, for instance, indicates that for offline scheduling of WDM Mix 3 the average ONU delays shoot up to very large values for loads around 3 Gbps, whereas we observe a similar jump in ONU delays for online scheduling at a load of 3.75 Gbps. The lower stability limit with offline scheduling is due to the nonwork conserving nature of offline scheduling, which forces the OLT to wait for all REPORT messages before making scheduling decisions. This waiting imposes idle times on the upstream channels which are avoided by the work conserving online scheduling frameworks, including online JIT and online JIT Tentative.

Second, we observe that the queueing delays are much larger for offline scheduling than online JIT scheduling. For a 6 Gbps load for WDM Mix 1, for instance, we observe from Fig. 9 average queueing delays around 20 ms for offline scheduling compared to less than 0.4 ms with online JIT. The larger delays with offline scheduling are mainly due to the increased RTS delays that increase the cycle lengths and subsequently increase the queueing delays.

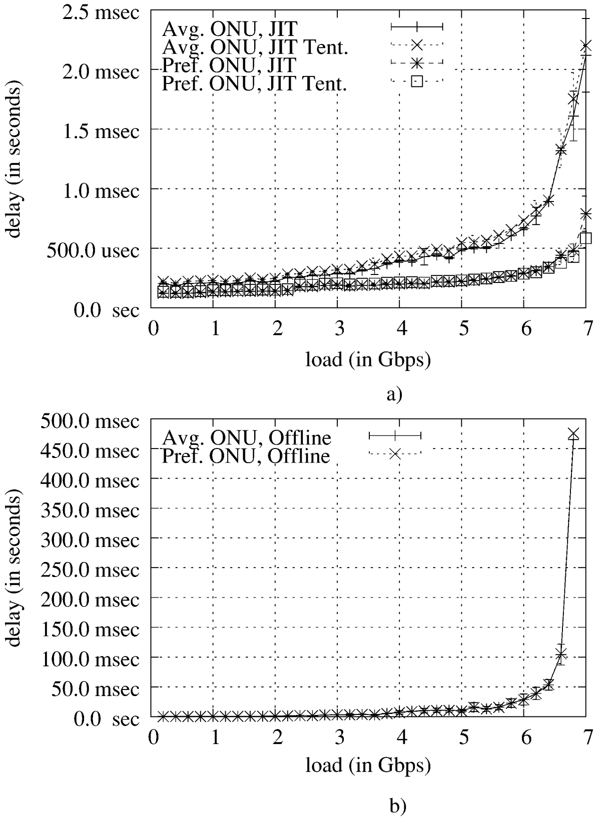


Fig. 10. Queueing delay for all ONUs versus the two ONUs given preferential scheduling (WDM Mix 2). Although the delay for the average ONU increases because of the single channel ONUs, the delay for the two preferred ONUs remains the same as with WDM Mix 1. (a) Online JIT scheduling framework. (b) Offline scheduling framework.

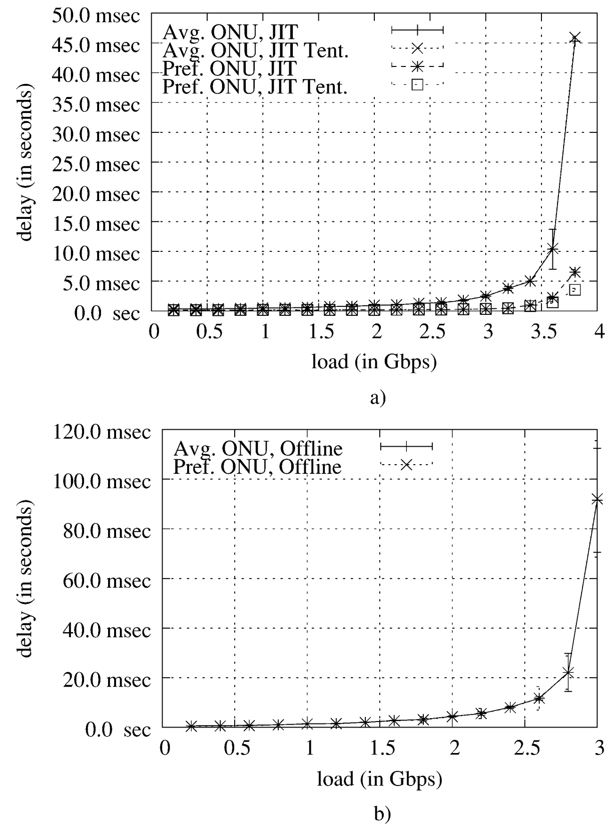


Fig. 11. Queueing delay for all ONUs versus the two ONUs given preferential scheduling (WDM Mix 3). The online JIT scheduling framework is able to provide very significant differential treatment. (a) Online JIT scheduling framework. (b) Offline scheduling framework.

Finally, turning to the differentiated ONU treatment, we observe from Figs. 9–11 that in contrast to online JIT, offline scheduling provides very little differentiation between the average ONUs and the preferred ONUs. This is mainly due to the fact that offline scheduling forces a relatively large RTS delay upon *all* ONUs, as illustrated by the delay components provided for WDM Mix 1 in Table I. In fact, for moderate to high loads, the RTS delay is the largest of the three delay components, which add up to the cycle length $GTG = RTS + STG + GTR$. In addition, the grant times, i.e., GTR delays, are relatively large compared to the STG delays, leaving little flexibility for differentiation due to reordering of the sequence of the upstream transmissions. In contrast, we observe from Table I that the RTS delays are relatively small for online JIT. At the same time, the STG delays are relatively large, compared to both RTS and GTR delays, providing significantly more flexibility in influencing the relative treatment of the ONUs through scheduling of the upstream transmissions.

In summary, the online JIT scheduling framework has the potential of reducing the average queueing delay experienced by all ONUs. However, its strongest utility appears to be for schedulers that provide differential treatment to ONUs. The OLT can benefit from an increased level of scheduling control without waiting for all ONU REPORT messages. Practical implementations may fine tune exactly when schedules are produced. It may be advantageous to purposefully leave a channel idle while waiting for more ONU REPORT messages to arrive at the OLT in an effort to gain a higher level of control.

TABLE I
AVERAGE REPORT-TO-SCHEDULE (RTS) DELAY, SCHEDULE-TO-GATE (STG) DELAY, AND GRANT TIME (GTR) IN MICROSECONDS FOR WDM MIX 1 WITH ONLINE JIT AND OFFLINE SCHEDULING FRAMEWORKS

	Online JIT			Offline		
Load	RTS	STG	GTR	RTS	STG	GTR
1.0	0.16	59.3	2.0	50.5	74.9	4.2
2.0	0.16	62.0	4.3	81.5	83.1	11.3
3.0	0.17	65.4	7.0	122.5	99.6	23.6
4.0	0.23	70.0	10.4	181.4	132.8	46.2
5.0	0.84	77.3	15.0	289.5	209.0	95.0
6.0	5.5	90.6	22.9	579.1	443.8	242.7
7.0	41.4	132.3	50.2	3416.1	3007.8	1850.8

VI. CONCLUSION

We have proposed: 1) a two-layer structure of scheduling in multichannel EPONs consisting of a scheduling framework layer and a scheduling policy layer, as well as 2) online just-in-time (JIT) scheduling, a novel work conserving scheduling framework. In the online JIT scheduling framework, channel availability, rather than ONU REPORT messages, drives the scheduling process. When a channel becomes available, the OLT makes an access decision with the information (i.e., REPORT messages) that has accumulated since the last channel became available. This gives the online JIT scheduling framework the ability to make better scheduling decisions as compared to an online scheduling framework that only considers one ONU REPORT message. Further, the online JIT scheduling framework is still work conserving and therefore is more efficient than the nonwork conserving offline scheduling framework where the OLT waits for all ONU REPORT messages to make access decisions.

In our simulation study, we found that with increased ONU multichannel diversity and wavelength assignment flexibility, the WBM-based scheduling policies used in the online JIT scheduling framework are able to provide lower average queueing delays. We also found that at high loads, the schedulers that favor ONUs with the largest number of frames provide lower average queueing delays, because of a frame sampling effect. Therefore, the online JIT scheduling framework has some utility in lowering the average queueing delay. However, it has stronger utility for use with schedulers that may provide differential treatment to ONUs. The OLT can benefit from an increased level of scheduling control, i.e., considering a larger scheduling pool, to differentiate ONU service. The only trade-off is the slightly increased complexity of online JIT, which requires the OLT to simultaneously consider and schedule up to close to as many ONU REPORTS as there are ONUs on the EPON; whereas, only one ONU REPORT at a time is considered and scheduled with online scheduling.

Future research could study how grant sizing techniques are affected by this new online JIT scheduling framework, and the impact of schedule generation time on performance.

REFERENCES

[1] M. Marsan and D. Roffinella, "Multichannel local area network protocols," *IEEE J. Sel. Areas Commun.*, vol. 1, no. 5, pp. 885–897, Nov. 1983.

[2] A. Dhaini, C. Assi, and A. Shami, "Quality of Service in TDM/WDM ethernet passive optical networks (EPONs)," in *Proc. IEEE ISCC*, Jun. 2006, pp. 616–621.

[3] M. Maier, M. Herzog, and M. Reisslein, "STARGATE: The next evolutionary step toward unleashing the potential of WDM EPONs," *IEEE Commun. Mag.*, vol. 45, no. 5, pp. 50–56, May 2007.

[4] N. Frigo, P. Iannone, P. Magill, T. Darcie, M. Downs, B. Desai, U. Koren, T. Koch, C. Dragone, H. Presby, and G. Bodeep, "A wavelength-division multiplexed passive optical network with cost-shared components," *IEEE Photon. Technol. Lett.*, vol. 6, no. 11, pp. 1365–1367, Nov. 1994.

[5] F. An, K. Kim, D. Gutierrez, S. Yam, E. Hu, K. Shrikhande, and L. Kazovsky, "SUCCESS: A next-generation hybrid WDM/TDM optical access network architecture," *J. Lightw. Technol.*, vol. 22, no. 11, pp. 2557–2569, Nov. 2004.

[6] Y. Hsueh, W. Shaw, L. Kazovsky, A. Agata, and S. Yamamoto, "Success PON demonstrator: Experimental exploration of next-generation optical access networks," *IEEE Commun. Mag.*, vol. 43, no. 8, pp. S26–S33, Aug. 2005.

[7] F. An, D. Gutierrez, K. Kim, J. Lee, and L. Kazovsky, "SUCCESS-HPON: A next generation optical access architecture for smooth migration from TDM-PON to WDM-PON," *IEEE Commun. Mag.*, vol. 43, no. 11, pp. S40–S47, Nov. 2005.

[8] Y. Hsueh, M. Rogge, S. Yamamoto, and L. Kazovsky, "A highly flexible and efficient passive optical network employing dynamic wavelength allocation," *J. Lightw. Technol.*, vol. 23, no. 1, pp. 277–286, Jan. 2005.

[9] C. Bock, J. Prat, and S. Walker, "Hybrid WDM/TDM PON using the AWG FSR and featuring centralized ligth generation and dynamic bandwidth allocation," *J. Lightw. Technol.*, vol. 23, no. 12, pp. 3981–3988, Dec. 2005.

[10] M. McGarry, M. Maier, and M. Reisslein, "WDM ethernet passive optical networks," *IEEE Commun. Mag.*, vol. 44, no. 2, pp. S18–S25, Feb. 2006.

[11] C. Assi, Y. Ye, S. Dixit, and M. Ali, "Dynamic bandwidth allocation for quality-of-service over ethernet PONs," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 9, pp. 1467–1477, Nov. 2003.

[12] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A dynamic protocol for an ethernet PON (EPON)," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 74–80, Feb. 2002.

[13] G. Kramer, B. Mukherjee, S. Dixt, Y. Y., and R. Hirth, "Supporting differentiated classes of service in Ethernet passive optical networks," *J. Opt. Networking*, vol. 1, no. 8, pp. 280–298, Aug. 2002.

[14] M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[15] K. Kwong, D. Harle, and I. Andonovic, "Dynamic bandwidth allocation algorithm for differentiated services over WDM EPONs," in *Proc. 9th Int. Conf. Communications Systems*, Sep. 2004, pp. 116–120.

[16] M. McGarry, "An evolutionary wavelength division multiplexing upgrade for Ethernet passive optical networks," Master's thesis, Arizona State Univ., Tempe, 2004.

[17] F. Clarke, S. Sarkar, and B. Mukherjee, "Simultaneous and interleaved polling: An upstream protocol for WDM-PON," in *Proc. Optical Fiber Communication Conf.*, Mar. 2006, p. 3.

[18] A. Dhaini, C. Assi, M. Maier, and A. Shami, "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks," *J. Lightw. Technol.*, vol. 25, no. 1, pp. 277–286, Jan. 2007.

[19] A. Dhaini, C. Assi, and A. Shami, "Dynamic bandwidth allocation schemes in hybrid TDM/WDM passive optical networks," in *Proc. IEEE CCNC*, Jan. 2006, vol. 1, pp. 30–34.

[20] V. Jacobson, K. Nichols, and K. Poduri, An expedited forwarding PHB Jun. 1999 [Online]. Available: http://www.ietf.org/rfc/rfc2598.txt, RFC 2598 (Proposed Standard) obsoleted by RFC 3246.

[21] J. Heinanen, F. Baker, W. Weiss, and J. Wroclawski, Assured forwarding PHB group June 1999 [Online]. Available: http://www.ietf.org/rfc/rfc2597.txt, RFC 2597 (Proposed Standard) updated by RFC 3260.

[22] R. Graham, "Bounds for certain multiprocessing anomalies," *Bell Syst. Tech. J.* , vol. 45, pp. 1563–1581, 1966.

[23] M. McGarry, M. Reisslein, M. Maier, and A. Keha, "Bandwidth management for WDM EPONs," *OSA J. Opt. Networking*, vol. 5, no. 9, pp. 637–654, Sep. 2006.

[24] H.-J. Byun, J.-M. Nho, and J.-T. Lim, "Dynamic bandwidth allocation algorithm in ethernet passive optical networks," *Electron. Lett.*, vol. 39, no. 13, pp. 1001–1002, Jun. 2003.

[25] J. M. Joo and Y. J. Ban, "Dynamic bandwidth allocation algorithm for next generation access network," in *Proc. OFC*, Mar. 2006.

[26] S. Yin, Y. Luo, N. Ansari, and T. Wang, "Bandwidth allocation over EPONs: A controllability perspective," in *Proc. IEEE Globecom*, Nov. 2006.

[27] S. Yin, Y. Luo, N. Ansari, and T. Wang, "Stability of predictor-based dynamic bandwidth allocation over EPONs," *IEEE Commun. Lett.*, vol. 11, no. 6, pp. 549–551, Jun. 2007.

[28] Csim (Mesquite Software) [Online]. Available: http://www.mesquite.com

[29] K. Park and W. Willinger, *Self-Similar Network Traffic and Performance Evaluation*. Hoboken, NJ: Wiley-Interscience, 2000.

[30] G. Kramer, *Ethernet Passive Optical Networks*. New York: McGraw-Hill, 2005.

Michael P. McGarry (M'98) received the B.S. degree in computer engineering from Polytechnic University, Brooklyn, NY, in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Arizona State University, Tempe, in 2004 and 2007, respectively.

He is currently a Senior Staff Scientist at Adtran and an Adjunct Professor at Arizona State University, Tempe. From 1997 through 2003 he was employed in industry by companies including PMC-Sierra and Yurie Systems (now Lucent Technologies). His research interests include congestion control and the optimization of MAC protocols for both optical access and mobile ad hoc networks.

Martin Reisslein (A'96–S'97–M'98–SM'03) received the Dipl.-Ing. (FH) degree from the Fachhochschule Dieburg, Dieburg, Germany, in 1994, and the M.S.E. degree from the University of Pennsylvania, Philadelphia, in 1996, both in electrical engineering, and the Ph.D. degree in systems engineering from the University of Pennsylvania in 1998.

He is currently an Associate Professor in the Department of Electrical Engineering at Arizona State University (ASU), Tempe. 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.

Dr. Reisslein served as Editor-in-Chief of the IEEE COMMUNICATIONS SURVEYS AND TUTORIALS from January 2003 through February 2007.

Charles J. Colbourn received the M.Math. degree in 1978 from the University of Waterloo, Waterloo, ON, Canada, and the Ph.D. degree in 1980 from the University of Toronto, Toronto, ON, both in computer science.

He has held academic positions at the University of Saskatchewan, the University of Waterloo, the University of Vermont, and is currently a Professor of Computer Science and Engineering at Arizona State University, Tempe. He is the author of *The Combinatorics of Network Reliability* (Oxford) and *Triple Systems* (Oxford). He is Editor-in-Chief of the *Journal of Combinatorial Designs* and serves on the editorial boards of *Networks*; *Discrete Mathematics*; *Journal of Combinatorial Theory (A)*; *Designs, Codes and Cryptography*; and others. He edited the standard reference work *The CRC Handbook of Combinatorial Designs*. He is the author of more than 250 refereed journal papers focussing on combinatorial designs and graphs with applications in networking, computing, and communications.

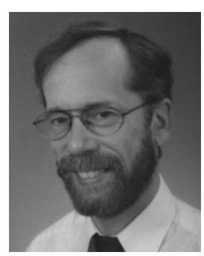
Dr. Colbourn was awarded the Euler Medal for Lifetime Research Achievement by the Institute for Combinatorics and its Applications in 2004.

Martin Maier (M'03) received the M.Sc. and Ph.D. degrees (both with distinctions) in electrical engineering from the Technical University Berlin, Berlin, Germany, in 1998 and 2003, respectively.

He is currently an Associate Professor with the Institut National de la Recherche Scientifique (INRS), Montreal, QC, Canada. In summer 2003, he was a Postdoctoral Fellow at the Massachusetts Institute of Technology (MIT), Cambridge. He was a Visiting Professor at Stanford University, Stanford, CA, from October 2006 through March 2007. His recent research activities aim at providing insights into technologies, protocols, and algorithms shaping the future of optical networks and their seamless integration with next-generation wireless networks. He is the author of the books *Metropolitan Area WDM Networks—An AWG Based Approach* (Kluwer, 2003) and *Optical Switching Network* (Cambridge University Press, 2008).

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 Friedrich-Schiller University in 2003 and 2006, respectively.

After completing his Ph.D., he joined the DFG Research Center Matheon at Technical University Berlin. His research interests lie in the queueing theoretic analysis of telecommunication networks, coding theory, and limit theorems in probability theory. He is currently a Postdoctoral Researcher at the Technical University Berlin, Berlin, Germany.



Michael Scheutzow received the Diploma in mathematics from the Johann-Wolfgang-Goethe University, Frankfurt/Main, Germany, in 1979, the Ph.D. degree in mathematics (magna cum laude) from the University of Kaiserslautern, Kaiserslautern, Germany, in 1983, and the Habilitation in mathematics from the University of Kaiserslautern in 1988.

He is currently a Professor in the Department of Mathematics, Technical University (TU) Berlin, Berlin, Germany. From 1988 through 1990 he was Project Leader with Tecmath GmbH, Kaiserslautern. Since 1990, he has been a Professor of stochastics in the Department of Mathematics, TU Berlin. From 1997 through 1999 he was Associate Chair of the department. He has visited the University of Carbondale, Rutgers University, the University of Rochester, Warwick University, and the MSRI at UC Berkeley.