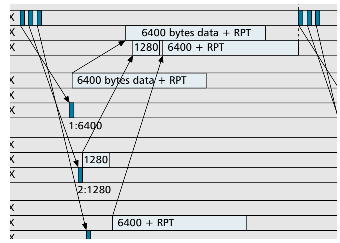

WDM ETHERNET PASSIVE OPTICAL NETWORKS

Michael P. McGarry and Martin Reisslein, Arizona State University
Martin Maier, Institut National de la Recherche Scientifique (INRS)

ABSTRACT

WDM EPONs not only allow for cautious pay-as-you-grow upgrades of single-channel TDM EPONs but also avoid linearly increasing polling cycle times for an increasing number of ONUs. In this article, we first provide a comprehensive overview of the state-of-the-art of TDM EPONs and recently reported dynamic bandwidth allocation algorithms, including decentralized scheduling schemes. After reviewing previous work on WDM EPONs, we address the requirements of WDM upgraded EPONs and make recommendations on an evolutionary WDM upgrade at the architecture, protocol, and dynamic bandwidth allocation algorithm levels, taking backward compatibility with MPCP and future-proofness against arbitrary WDM ONU structures into account. We describe and compare online and offline scheduling paradigms for WDM EPONs. Our simulation results indicate that online scheduling can achieve lower delays, especially at high loads. We outline areas of future research on WDM EPONs.

INTRODUCTION

Passive optical networks (PONs) have been considered attractive due to their longevity, low operational costs, and huge bandwidth. As a matter of fact, PONs are already widely deployed in the first/last mile of today's operational access networks [1]. PONs come in a number of flavors. The so-called asynchronous transfer mode (ATM) PON (APON) and broadband PON (BPON) are ATM-based systems. Gigabit PON (GPON), the successor of BPON, is able to support traffic other than ATM (e.g., telephony and Ethernet) in its native format by using time-division multiplexing (TDM) partitions and generic framing procedure (GFP) similar formats. Recently, Ethernet PONs (EPONs), standardized by the IEEE 802.3ah Ethernet in the First Mile (EFM) Task Force (http://www.ieee802.org/3/efm), have been attracting considerable attention from both industry and academia. EPONs aim at converging the low-cost equipment and simplicity of Ethernet and the low-cost fiber infrastructure of PONs. EPONs are a promising solution to provide sufficient bandwidth for emerging services such as videoconferencing, distributed gaming, IP telephony, and video on demand.

Current EPONs are single-channel systems; that is, the fiber infrastructure carries a single downstream wavelength channel and a single upstream wavelength channel, which are typically separated by means of coarse wavelength-division multiplexing (CWDM). In the upstream direction (from subscriber to network), the wavelength channel bandwidth is shared by the EPON nodes by means of TDM. In doing so, only one common type of single-channel transceiver is used network-wide, resulting in simplified network operation and maintenance. At present, single-channel TDM EPONs appear to be an attractive solution to provide more bandwidth in a cost-effective manner.

Given the steadily increasing number of users and bandwidth-hungry applications, current single-channel TDM EPONs are likely to be upgraded in order to satisfy the growing traffic demands in the future. Clearly, one approach is to increase the line rate of TDM EPONs. Note, however, that such an approach implies that *all* EPON nodes need to be upgraded by replacing the installed transceivers with higher-speed transceivers, resulting in a rather costly upgrade. Alternatively, single-channel TDM EPONs may be upgraded by deploying multiple wavelength channels in the installed fiber infrastructure in the upstream and/or downstream directions, resulting in WDM EPONs. As opposed to the higher-speed TDM approach, WDM EPONs provide a cautious upgrade path in that wavelength channels can be added one at a time, each possibly operating at a different line rate. More important, only EPON nodes with higher traffic demands may be WDM upgraded by deploying multiple fixed-tuned and/or tunable transceivers, while EPON nodes with lower traffic demands remain unaffected. Thus, using WDM enables network operators to upgrade single-channel TDM EPONs in a pay-as-you-grow manner where only a *subset* of EPON nodes may be upgraded gradually.

In this article we address the requirements of WDM upgrades of EPONs at the architecture, protocol, and bandwidth allocation algorithm levels. In particular, we outline an evolutionary WDM upgrade path of current single-channel TDM EPONs, taking backward compatibility and future-proofness against arbitrary EPON node structures into account.

 0163-6804/06/$20.00

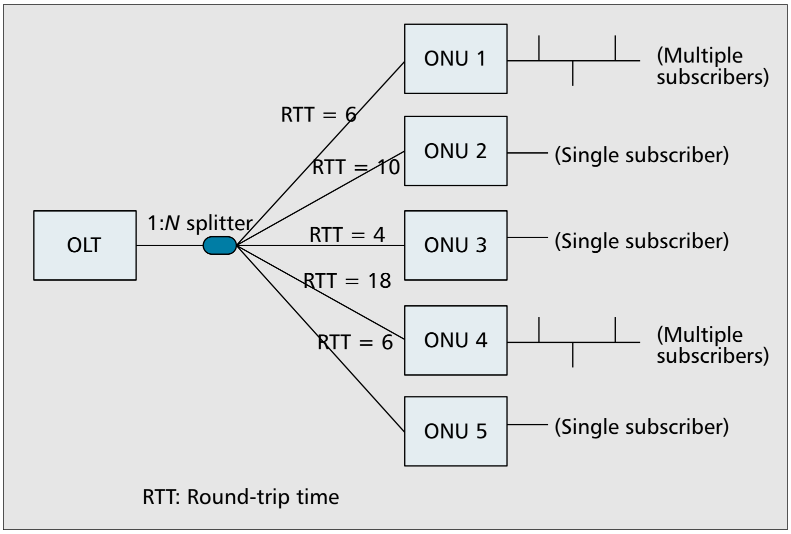

Figure 1. Network architecture of an EPON with one optical line terminal (OLT) and $N = 5$ optical network units (ONUs), each with a different round-trip time (RTT).

The remainder of the article is organized as follows. We provide a comprehensive overview of the state of the-art of single-channel TDM EPONs. We describe the proposed evolutionary WDM upgrade of single-channel EPONs. We then conclude the article.

General Overview of EPONs

In this section we describe the architecture and medium access control (MAC) protocol of single-channel TDM EPONs. Furthermore, we review the state of the art of dynamic bandwidth allocation (DBA) algorithms, which are used to assign upstream transmission time slots to EPON nodes in a dynamic and efficient way.

Architecture

In principle, EPONs (and PONs in general) may have any topology suitable for access networks. Among others, the topology may be a tree, tree-and-branch, ring, bus, or a combination of those in order to provide redundancy, e.g., double rings. Typically, EPONs have a physical tree topology with the central office located at the root and the subscribers connected to the leaf nodes of the tree, as illustrated in Fig. 1. At the root of the tree is an optical line terminal (OLT), which is the service provider equipment residing at the central office. The EPON connects the OLT to multiple optical network units (ONUs) (the customer premises equipment) through a 1:*N* optical splitter/combiner. An ONU can serve a single residential or business subscriber, referred to as fiber to the home/business (FTTH/B), or multiple subscribers, referred to as fiber to the curb (FTTC). Each ONU buffers data received from its attached subscriber(s). To support differentiated services each ONU may use priority queues, one for each traffic class. In general, the round-trip time (RTT) between the OLT and each ONU is different. Due to the directional properties of the optical splitter/combiner, the OLT is able to broadcast data to all ONUs in the downstream direction. In the upstream direction, however, ONUs cannot communicate directly with one another. Instead, each ONU is able to send data only to the OLT. Thus, in the downstream direction an EPON may be viewed as a point-to-multipoint network, and in the upstream direction an EPON may be viewed as a multipoint-to-point network. Due to this fact, the original Ethernet MAC protocol does not operate properly since it relies on a broadcast medium. Instead, the Multipoint Control Protocol (MPCP) arbitration mechanism is deployed, as discussed in the subsequent subsection. (We discuss architectural design alternatives for EPON that allow the use of the original Ethernet MAC protocol later.)

In the upstream direction all ONUs share the transmission medium. To avoid collisions, several approaches can be used. At present, TDM is the preferred solution. Given the aforementioned different connectivity in upstream and downstream directions of EPONs, the OLT appears to be best suited to arbitrate time sharing of the upstream channel, as discussed next.

Multipoint Control Protocol

To increase upstream bandwidth utilization, the OLT dynamically allocates a variable time slot to each ONU based on the instantaneous bandwidth demands of the ONUs, done best by means of *polling* [2]. To facilitate DBA and arbitrating the upstream transmissions of multiple ONUs, MPCP is deployed in EPON. Besides auto-discovery, registration, and ranging (RTT computation) operations for newly added ONUs, MPCP provides the signaling infrastructure (control plane) for coordinating the data transmissions from the ONUs to the OLT.

MPCP uses two types of messages to facilitate arbitration: REPORT and GATE. Each ONU has a set of queues, possibly prioritized, holding Ethernet frames ready for upstream transmission to the OLT. The REPORT message is used by an ONU to report bandwidth requirements (typically in the form of queue occupancies) to the OLT. A REPORT message can support reporting of up to eight queue occupancies of the corresponding ONU. Upon receiving a REPORT message, the OLT passes it to the DBA algorithm module. The DBA module calculates the upstream transmission schedule of all ONUs such that channel collisions are avoided. Scheduling can be done in two ways: inter-ONU scheduling and intra-ONU scheduling. Inter-ONU scheduling arbitrates the transmissions of different ONUs, while intra-ONU scheduling arbitrates the transmissions of different priority queues in each ONU. There are two possible implementations. Either inter-ONU scheduling is implemented at the OLT, and each ONU performs its own intra-ONU scheduling, or both inter-ONU and intra-ONU scheduling are implemented at the OLT. After executing the DBA algorithm, the OLT transmits GATE messages to issue transmission grants. Each GATE message can support up to four transmission grants. Each transmission grant contains the transmission start time and transmission length of the corresponding ONU. Each ONU updates its local clock using the timestamp contained in each received transmission grant. Thus, each ONU is able to acquire and maintain global synchronization. Each ONU sends backlogged Ethernet frames during its granted transmission window according to the corresponding intra-ONU scheduling. The transmission window may comprise multiple Ethernet frames; packet fragmentation is not allowed. As a consequence, if the next frame does not fit into the current transmission window, it has to be deferred to the next granted transmission window.

Note that MPCP does not specify any particular DBA algorithm. MPCP simply provides a framework for the implementation of various DBA algorithms, which are described in greater detail next.

Dynamic Bandwidth Allocation

To date, a plethora of DBA algorithms for single-channel TDM EPONs have been investigated. For a taxonomy and survey of DBA algorithms, we refer the interested reader to [3]. In the following we briefly review some of the most influential DBA algorithms. In addition, we describe recently proposed DBA algorithms not mentioned in [3].

Among others, the so-called interleaved polling with adaptive cycle time (IPACT) algorithm has attracted considerable attention [4]. In IPACT the OLT polls the ONUs individually and issues transmission grants to them in a round-robin fashion. Bandwidth is dynamically assigned to ONUs according to their reported queue occupancies. IPACT deploys in-band signaling of bandwidth requests by using escape characters within Ethernet frames instead of sacrificing an entire Ethernet frame for control (as done in MPCP), resulting in reduced signaling overhead. The OLT keeps track of the RTTs of all ONUs. As a result, the OLT can send out a grant to the next ONU before the current ONU has terminated its transmission, leading to *interleaved polling* and improved bandwidth utilization. In IPACT the polling cycle length is not static but adapts to the instantaneous bandwidth requirements of the ONUs (i.e., IPACT deploys an *adaptive cycle time*).

Different DBA algorithms to efficiently and fairly allocate bandwidth to ONUs and support differentiated services were examined in [5]. Differentiated services are supported by means of *priority queuing* and intra-ONU scheduling at each ONU. Specifically, only packets arriving before time $t_{request}$ are given high priority for transmission, where $t_{request}$ denotes the transmission time of the REPORT message. The order of transmission is based on their traffic classes. If packets arriving before $t_{request}$ are all scheduled, and the current transmission window can still accommodate more traffic, it will be allocated for packets arriving during the time interval $t_{grant} - t_{request}$ based on their priorities, where t_{grant} denotes the start time of the corresponding transmission window. Thus, all traffic classes of each ONU are scheduled in an efficient and fair manner by allowing all traffic classes to access the channel as requested and adhere to their respective priorities.

The aforementioned DBA algorithm with priority queuing guarantees service differentiation within the same ONU. However, due to the ordered transmission schedule of all ONUs, low-priority traffic of one ONU may be transmitted before high-priority traffic of another ONU. As a result, high-priority traffic may suffer from a larger queuing delay than some low-priority traffic. To mitigate this drawback, a two-layer DBA algorithm based on weighted fair queuing (WFQ) was proposed in [6]. Specifically, an ONU is allowed to report all its instantaneous traffic load for each traffic class separately. The OLT uses this information to proportionally allocate bandwidth according to the ratio of the demand of a single class to the total demand. The OLT first allocates the bandwidth to different traffic classes, then further distributes the bandwidth allocated to one class among all requesting ONUs. Accordingly, the DBA algorithm is called two-layer bandwidth allocation (TLBA). Within the same class, all ONUs fairly share the bandwidth following the max-min policy. More precisely, the shared upstream bandwidth is allocated evenly to all ONUs in order of increasing demand. No ONU is allocated more of the bandwidth than its demand. ONUs with unsatisfied demands evenly share the remaining bandwidth. TLBA is able to ensure each service class a minimum bandwidth based on its weight. Moreover, compared to priority-queuing based DBA schemes TLBA reduces the average queuing delay of high-priority and medium-priority traffic. Note that in TLBA the OLT needs to receive the REPORT messages from all ONUs to calculate the total bandwidth demand of each service class which results in a prolonged time interval between two adjacent frames.

To improve QoS metrics such as average delay, queue length, and frame loss, a multiservice DBA algorithm using bursty traffic prediction was examined in [7]. The so-called dynamic bandwidth allocation with multiple services (DBAM) algorithm employs class-based traffic prediction to take the packets arriving during the waiting time between REPORT and GATE messages at each ONU into account. Specifically, for each traffic class, a given ONU requests bandwidth for an estimation credit besides the already queued traffic. For a given ONU, the *estimation credit* is equal to the ratio of the ONU's waiting time encountered in the last polling cycle and the length of the last polling cycle.

Note that all the above mentioned DBA algorithms are centralized schemes. The OLT acts as the central control unit by performing inter-ONU scheduling or both inter-ONU and intra-ONU scheduling. Recently, research on *decentralized* DBA algorithms and distributed scheduling has begun [8, 9]. To enable distributed scheduling, however, the original EPON architecture has to be modified such that each ONU's upstream transmission is echoed at the splitter to all ONUs, each equipped with an additional receiver to receive the echoed transmissions. In doing so, all ONUs are able to monitor the transmission of every ONU and arbitrate upstream channel access in a distributed manner, similar to Ethernet LANs. Note that in such alternate EPON solutions both inter-ONU and intra-ONU scheduling take place at the ONUs without participation of the OLT. The reported performance results show that such decentralized EPONs and DBA algorithms are able to provide high bandwidth utilization. In the following we focus on EPONs as standardized in IEEE 802.3ah and refer the interested reader to [8, 9] for further details on decentralized DBA algorithms.

WDM EPONs

In this section we address the requirements of WDM EPONs and make our own recommendations on how to upgrade current single-channel TDM EPONs to multichannel WDM EPONs. After reviewing previous work on WDM EPONs, we describe our proposed architecture as well as WDM extensions to MPCP in detail. Finally, we outline and compare two broad paradigms for WDM-DBA in WDM EPONs.

Previous Work

The design and feasibility study of cost-effective WDM structures for OLT and ONU were addressed in [11]. The proposed WDM OLT structure consists of a multicarrier generator and supplies hundreds of optical carriers, thus greatly reducing the number of required laser diodes at the WDM OLT. Each ONU is assigned a separate pair of dedicated upstream and downstream wavelength channels. To decrease costs, ONUs deploy no light source, but simply modulate the optical carriers supplied by the OLT for upstream transmission. This remote modulation scheme realizes wavelength-independent ONUs, resulting in reduced costs and simplified operation and maintenance. Note that in the proposed architecture each pair of wavelength channels is dedicated to a different ONU. Thus, upstream wavelength channels are not shared among ONUs and no WDM-DBA is performed.

A WDM PON in which each upstream wavelength channel can be shared among multiple ONUs by means of TDM was examined in [11]. The work focuses on the design and feasibility study of cost-effective burst mode ONU transmitters that can operate on any wavelength channel without requiring wavelength tuning. Such so-called wavelength-selection-free transmitters require neither wavelength stability circuits nor network operators to stock spare transmitters for each wavelength channel, resulting in reduced costs. WDM-DBA algorithms were not discussed in greater detail.

The interconnection of multiple PONs of arbitrary topology was investigated in [12]. The proposed PON interconnection is highly scalable and provides an evolutionary upgrade path from single-channel TDM PONs to WDM PONs where each ONU is assigned a separate pair of dedicated upstream and downstream wavelength channels, and existing field-deployed PON infrastructures remain intact. The transmitter(s) at the OLT may be shared among all interconnected PONs. For upstream transmission, each ONU may have a different node structure. For instance, one ONU may deploy a single tunable transmitter while another ONU may deploy an array of fixed-tuned transmitters. In contrast, for downstream transmission ONU node structures are less flexible. More precisely, the receiver of each ONU has to operate on a different wavelength band. Thus, each ONU has to have a different wavelength-band-selective receiver, and each wavelength band must not be used by more than a single node at any given time. Different ONU node structures that provide a platform for deploying WDM-DBA were discussed and examined. However, no specific WDM-DBA algorithm was presented.

A WDM-DBA algorithm with online scheduling for WDM EPONs was described and investigated in [13]. In the considered WDM EPON all ONUs are equipped identically with an array of fixed-tuned transceivers, one for each upstream/downstream wavelength channel. The proposed WDM IPACT with a single polling table (WDM IPACT-ST) DWA algorithm is a multichannel extension of IPACT, discussed earlier. In WDM IPACT-ST, transmission windows are assigned to ONUs in a round-robin fashion, allowing them to transmit on the first available upstream wavelength channel (first fit). It was shown that the resultant WDM IPACT-ST EPON outperforms a single-channel TDM IPACT EPON in terms of delay. This is due to the fact that in TDM EPONs the polling cycle time increases linearly with the number of attached ONUs, as opposed to WDM EPONs, which use multiple wavelength channels simultaneously to accommodate an increasing number of ONUs while maintaining a short polling cycle.

The WDM EPON presented in [14] aims at integrating both APON and EPON. The proposed so-called byte size clock (BSC) protocol is scalable in bandwidth assignment since heavy users may be assigned a single wavelength, whereas light users may share a single wavelength. In BSC time is divided into periodically recurring time frames. Each frame consists of dedicated reservation minislots, one for each ONU, and data slots, which are assigned on demand. Users send request packets in their assigned minislots at start of frame i. ONUs then transmit their respective data packets in accordance with the OLT grants received in frame $(i - 1)$. Once the request packets of frame i are received by the OLT, it computes the grants and broadcasts them back to the ONUs in frame $(i + 1)$. Note that a data packet has to go through a delay of at least one frame due to the reservation. It was shown that this delay can be reduced by *pre-allocating* a minimum number of dedicated data slots to certain ONUs which can increase in subsequent frames if the ONUs make reservations in their minislots. As a result, part of the data packets can be sent without reservation, leading to improved throughput-delay performance and decreased queue length of the corresponding ONUs. The performance of the preallocation BSC protocol can be further improved by means of *delta compression*. By using delta compression to compute the delta (difference) between packets and transmitting only the delta instead of the original packets, the need to reserve a large number of data slots is avoided as well as the number of preallocated time slots is kept to a minimum. On the downside, in BSC all nodes need to be synchronized, and the resultant TDM frame time structure does not comply with IEEE 802.3ah.

Architecture

The WDM upgrade of single-channel TDM EPONs will very likely occur over long periods of time in a pay-as-you-grow manner. Hence, the type of WDM ONU node structures in a given EPON can differ as current technology and economic constraints as well as service provider preferences dictate.

Given that the main driver for supporting WDM on an EPON is expansion of the bandwidth available on the EPON, it is not advantageous for this goal to have a tunable transceiver in the OLT. By having a tunable transceiver, only a single wavelength channel can be used at any given time. This would not expand the bandwidth currently available with a single fixed-tuned transceiver. In fact, it would actually provide less bandwidth because of the dead time imposed every time there is a wavelength switch due to the non-zero transceiver tuning time. Therefore, it appears reasonable that any WDM upgrade of an EPON has an array of fixed-tuned transceivers in the OLT, one for each operating wavelength channel.

Our envisioned goal for managing WDM in the EPON in an evolutionary manner is to manage the different wavelength channels to increase the available bandwidth of the EPON without imposing a particular WDM architecture on the ONU, thereby allowing the ONUs to take on whatever architecture is preferred at the time they are upgraded, possibly using transceivers with different tuning times and tuning ranges. ONUs should also be able to be upgraded incrementally as needed (e.g., adding new fixed-tuned and/or tunable transceivers incrementally). The evolutionary WDM upgrade of EPONs should not impose any particular WDM ONU architecture, thus allowing these decisions to be dictated by economics, state-of-the-art transceiver manufacturing technology, and service provider preferences. Note that the proposed evolutionary and flexible WDM upgrade path not only increases capacity but also meets key requirements of PONs [15]. Specifically, evolutionary WDM upgrade allows for cautious pay-as-you-grow upgrades and thus helps operators realize their survival strategy for highly cost-sensitive access networks.

We note that network operators are expected to deploy only a small number of different WDM ONU structures and incremental upgrades. Regardless of this, our proposed evolutionary WDM upgrade of EPONs does not impose any particular WDM ONU architecture, thus allowing these decisions to be made by network operators. We also note that although we can allow for arbitrary ONU WDM node structures, it is reasonable to maintain a common channel that all ONUs support for transmission and reception, whereby the legacy EPON channel appears to be the best candidate for the common channel. Such a common channel for reception allows the OLT to forward broadcast frames to all ONUs with a single transmission on one wavelength channel.

WDM Extensions to MPCP

To guarantee compliance of the aforementioned evolutionary WDM upgrade of EPON with IEEE 802.3ah, we have to extend MPCP accordingly. The following recommended WDM extensions to MPCP enable the OLT to schedule transmissions to and receptions from ONUs on any wavelength channel(s) supported by the OLT and the respective ONU.

Discovery and Registration — For backward compatibility, the discovery and registration of ONUs take place on the original wavelength channel of TDM EPONs. During the registration process a discovered ONU conveys the following information about its WDM architecture to the OLT:

- *TX_type* and *RX_type*: 2 bits each indicating the transmitter and receiver type, respectively, by using the following assigned values: 0 = no WDM, 1 = fixed-tuned, 2 = tunable, and 3 = reserved.
- *TX_tuning_time* and *RX_tuning_time*: 16 bits each indicating the tuning time of the transmitter and receiver, respectively, as an integer multiple of unit time, such as microsecond (if TX_type/RX_type = 2).
- *Wavelength_id_type*: 1 bit indicating the encoding scheme of the supported wavelengths by using the following assigned values: 0 = two-level hierarchical encoding scheme (waveband identifier/bitmap of supported wavelengths within waveband), 1 = flat encoding scheme (bitmap of supported wavelengths).

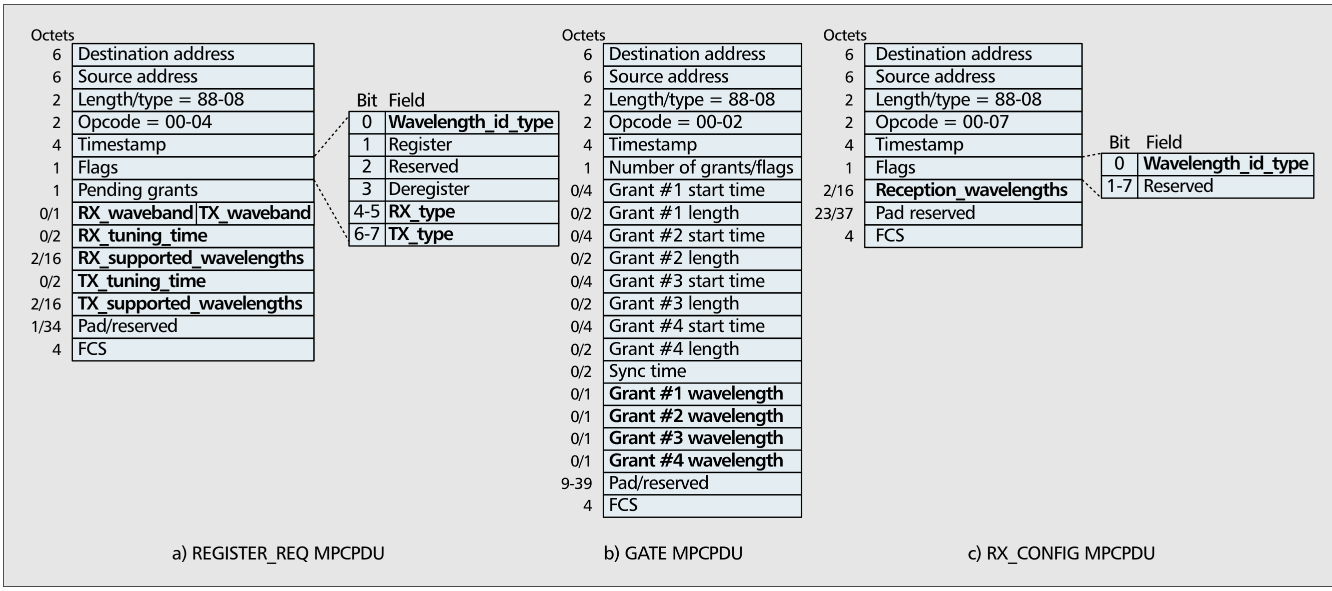

FIGURE 2. WDM extensions to MPCP protocol data units (PDUs): a) REGISTER_REQ; b) GATE; c) the proposed RX_CONFIG (extensions are shown bold).

- *TX_waveband* and *RX_waveband*: 4 bits, each indicating the identifier of the supported waveband of the transmitter and receiver, respectively. Wavebands are defined in compliance with the WDM channel spacings specified in ITU-T G.694.1 (if Wavelength_id_type = 0).
- *TX_supported_wavelengths* and *RX_supported_wavelengths*: 16 bits each (Wavelength_id_type = 0) or 128 bits each (Wavelength_id_type = 1) indicating the bitmap of the supported wavelengths of the transmitter and receiver, respectively.

This information is mapped into the reserved fields of the REGISTER_REQ MPCP protocol data unit (PDU), as shown in Fig. 2a.

Upstream Coordination — To facilitate the need for the OLT to assign a specific wavelength channel for the upstream transmission from a given ONU to the OLT, an 8-bit wavelength identifier (allowing for support of up to 256 unique wavelengths) is issued along with every transmission grant by the OLT by using the reserved fields of the GATE MPCPDU, as depicted in Fig. 2b.

Downstream Coordination — To let the OLT (re)configure the receiving wavelength(s) of a given ONU, currently no appropriate MPCPDU exists. We propose two new MPCPDUs: RX_CONFIG (Opcode = 00-07) and RX_CONFIG_ACK (Opcode = 00-08). The OLT sends the RX_CONFIG MPCPDU to (re)configure a given ONU's receiver(s). The ONU acknowledges the (re)configuration by sending the RX_CONFIG_ACK MPCPDU to the OLT. As illustrated in Fig. 2c, the Reception_wavelengths field of the RX_CONFIG MPCPDU consists of 16 or 128 bits, depending on the applied encoding scheme given in the respective Flags field. The RX_CONFIG_ACK MPCPDU (not shown in the figure) consists of the echoed Flags and Reception_wavelengths fields.

We note in closing on the protocol aspects that the WDM EPON maintains compliance with the IEEE 802.1d bridging by assigning one logical link ID (LLID) for each ONU, irrespective of the number of physical wavelength channels supported by the ONU. The IEEE 802.1d bridging uses the LLID (which is associated with a logical port number) to keep an entry in a filtering database that specifies to which logical port a frame with a particular destination address is forwarded. Assigning one LLID for each ONU allows for a unique entry in the filtering database; the multiple physical wavelength channels to an ONU are simply used as another dimension (in addition to the time dimension) for DBA as explored in the next section.

WDM Dynamic Bandwidth Allocation

Whereas in conventional TDM single-channel EPONs the DBA problem is limited to scheduling the upstream transmissions on the single wavelength channel, in WDM EPONs the DBA problem is expanded to scheduling the upstream transmissions on the different upstream wavelengths supported by the ONUs.

In other words, in WDM EPONs not only decisions on when and for how long to grant an ONU upstream transmission, but also on which wavelength channel to grant the upstream transmission are required. In this section we outline and compare two broad paradigms for dynamically allocating grants for upstream transmissions on the different upstream wavelengths in a WDM EPON: *online* and *offline scheduling*.

Online Scheduling — In an online scheduler a given ONU is scheduled for upstream transmission as soon as the OLT receives the REPORT message from the ONU. In other words, the OLT makes scheduling decisions based on individual requests and without global knowledge of the current bandwidth requirements of the other ONUs. A basic online scheduling policy for the WDM EPON is to schedule the upstream transmission for an ONU on the wavelength channel available earliest among the channels supported by the ONU, which we refer to as the next available supported channel (NASC) policy. The amount of the bandwidth (i.e., the length of the granted transmission) allocated to an ONU can be determined according to any of the existing DBA mechanisms for single-channel EPONs [3].

Figure 3 illustrates online scheduling for an EPON with three ONUs. Notice that the 2560-byte upstream transmission from ONU 2 is scheduled on the earliest available supported wavelength, wavelength 1, and is timed by the OLT such that it is separated from the preceding transmission on wavelength 1 by ONU 1 by a guard interval.

Offline Scheduling — In an offline scheduler the ONUs are scheduled for transmission once the OLT has received current MPCP REPORT messages from *all* ONUs, allowing the OLT to take into consideration in the scheduling the current band-

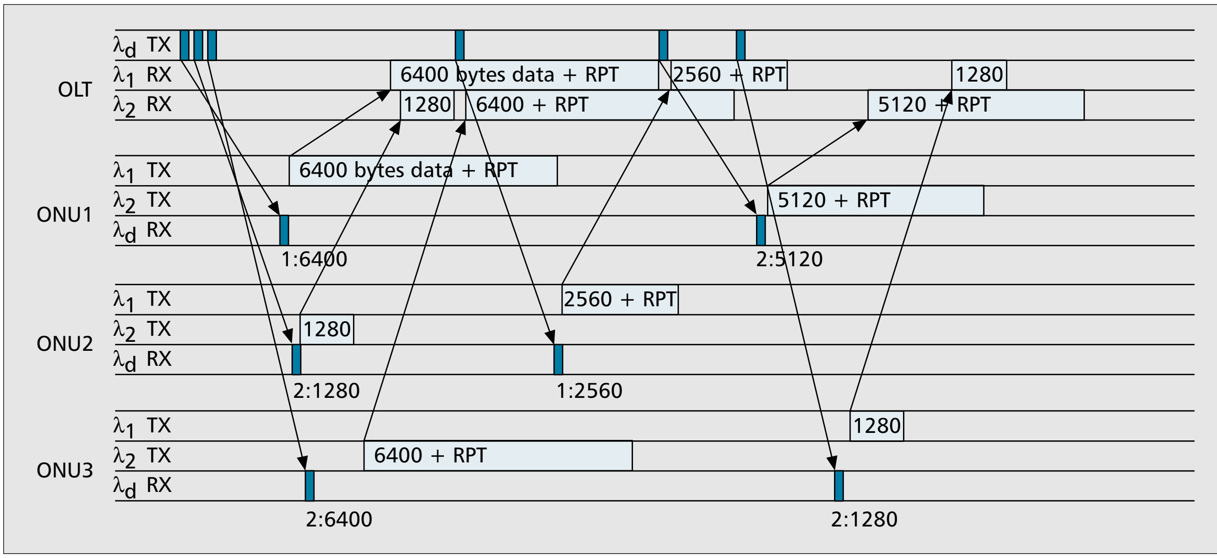

FIGURE 3. Illustration of online scheduling with NASC policy. The illustration includes one downstream wavelength λ_d and two upstream wavelengths, λ_1 and λ_2, which are supported by all three ONUs. Each ONU reports its queue occupancy in the REPORT message, which is appended to the current upstream transmission. Upon receipt of a REPORT message the OLT immediately schedules the next upstream transmission for the corresponding ONU and sends a GATE message (illustrated by the dashed message) indicating the wavelength and length (in bytes in the illustration) of the granted transmission to the ONU.

width requirements from all ONUs. Since an offline scheduler makes scheduling decisions for all ONUs at once, all of the REPORTs, which are usually appended to the end of the data stream of a gated transmission window, from the previous cycle must be received, as illustrated in Fig. 4. This requires that the scheduling algorithm be executed after the OLT receives the end of the last ONU's gated transmission window. Due to this, a gap between scheduling cycles is introduced, which we refer to as the inter-scheduling cycle gap (ISCG). The length of the ISCG on a wavelength channel is equal to:

- The computation time of the schedule
- The transmission time for the grant (64 bytes)
- The RTT to the first ONU scheduled on the wavelength in the next round

The offline scheduling of WDM EPONs can be viewed in terms of classical scheduling theory [16]. In particular, we may view ONUs as jobs, the bandwidth requests of the ONUs as the processing times of the jobs, and the upstream wavelength channels as machines. From the wide variety of scheduling algorithms that can be employed for the WDM EPON, we consider the least flexible job (LFJ) first policy in our simulation comparison of the online and offline scheduling paradigms. The LFJ first policy is optimal in that it minimizes the length of the schedule when:

- The wavelengths supported by the less flexible ONUs are a subset of the wavelength channels supported by the more flexible ONUs (as is the case for the example in our simulations).
- Equally sized bandwidth units are scheduled.

The LFJ first policy first schedules transmissions by the ONUs that support the fewest wavelength channels (i.e., least flexible ONU) at the earliest available time on the supported channels.

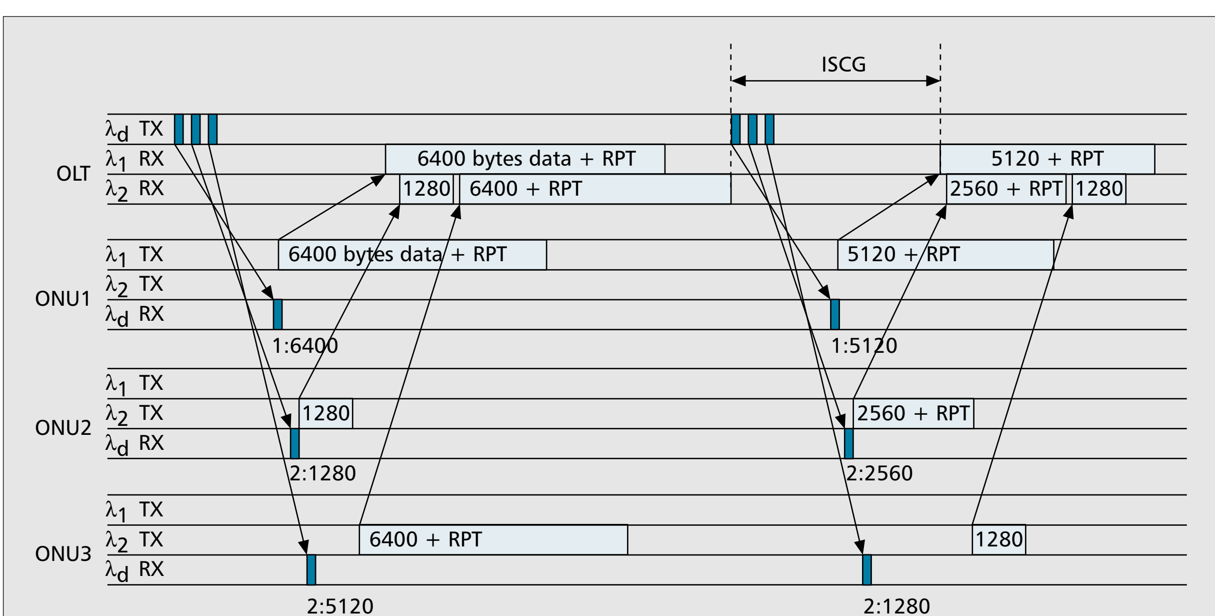

FIGURE 4. Illustration of the offline scheduler, which introduces the ISCG between successive cycles.

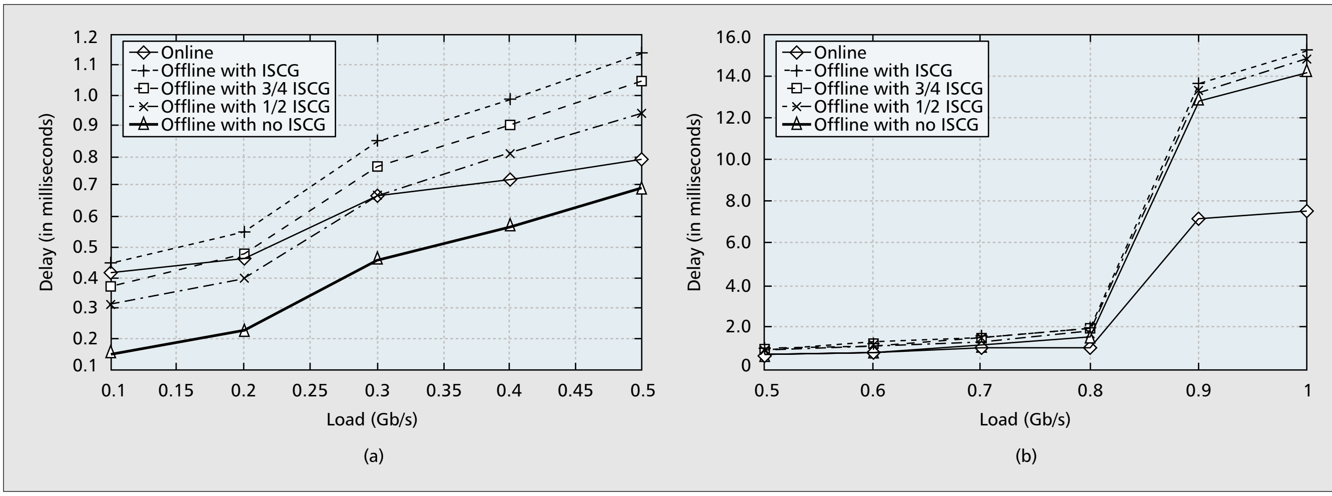

FIGURE 5. Comparison of packet delay with online and offline scheduling on a five-wavelength EPON as a function of load: a) lower load; b) higher load.

SIMULATION RESULTS

In our simulations we consider an EPON consisting of 10 ONUs (five lightly loaded non-WDM ONUs and five heavily loaded WDM ONUs that generate twice as much self-similar traffic with Hurst parameter 0.75 as the lightly loaded ONUs). We investigate the WDM upgrade of the heavily loaded ONUs, which we assume support transmission on all wavelengths supported on the EPON using a single tunable transmitter with negligible tuning time. Each wavelength supports 1 Gb/s link speed. The loads are measured with respect to this link speed. A load of 0.8 represents a total traffic load of 0.8 Gb/s in the network. The load is calculated based on the Ethernet data frames (header + payload + frame check sequence trailer), and does not include the MPCP control frames or the preamble and inter-frame gap (IFG).

For our simulations the RTT was randomly generated according to a uniform distribution U[100 μs, 200 μs], which corresponds to ONU distances of 15–30 km from the OLT. This makes the minimum ISCG = [100 μs + (64 · 8 bit)/(10^9 b/s)] = 100.5 μs, and the maximum ISCG = 200.5 μs. We assumed negligible schedule generation time.

We compare the queuing delay of packets when an offline and online scheduler is used for scheduling the ONUs' upstream transmissions. For online scheduling we consider the NASC scheduler in conjunction with the gated DBA mechanism [4], which grants each ONU the full bandwidth request; for offline scheduling we consider the LFJ first policy.

Our results in Fig. 5 indicate that at low loads the offline scheduler with the ISCG achieves about the same queuing delay as the online scheduler; with masking of a part of or the entire ISCG, the offline scheduler could achieve smaller delays. As the load increases, the delay of the offline scheduler grows faster, and the online scheduler achieves generally lower queuing delays at high loads. To explain these findings we first take another look at Fig. 4. Notice that the offline scheduling introduces for some ONUs a delay from the receipt of their request (RPT) at the OLT to the commencement of the scheduling as the OLT is waiting until the requests from all ONUs are received, which we refer to as the report-to-scheduling (RS) delay. This RS delay is visible in Fig. 4 for ONU1 and ONU2, whose requests can only be scheduled after the 6400-byte upstream transmission from ONU3 is received at the OLT. In addition, all ONUs experience the ISCG (unless it is masked by some technique).

To explain the delay behaviors at the different load levels first consider the extreme case of a very low load such that there is typically only one (or a few) ONU(s) with upstream traffic to be served at any time. In this situation the online and offline scheduling are almost identical. To see this, note that the RS delay in the offline scheduler is negligible in this situation since the upstream schedule consists only of the upstream data plus report transmission of the one ONU with traffic (plus the upstream report transmissions of the other ONUs). As long as only one ONU at a time transmits or reports upstream traffic, there is essentially no delay due to waiting for reports from the other ONUs. At the same time, with only one ONU having upstream traffic, the ISCG in the offline scheduler does not add any significant additional delay compared to online scheduling. To see this, note that with only one ONU having upstream data at a time in online scheduling, there are no upstream transmissions from other ONUs ongoing as the report from the one ONU with data propagates up to the OLT and the grant down to the ONU. In other words there are no other ONUs to take advantage of interleaving of upstream transmission cycles while the request/grant of a given ONU propagates up to/down from the OLT.

On the other hand, with increasing load the offline schedule becomes increasingly longer, consisting of scheduled transmissions from several (if not all) ONUs, leading to significant RS delays for the ONUs contained at the beginning of the schedule. At the same time, the transmission cycles from several ONUs with upstream traffic can be interleaved in online scheduling, whereas with offline scheduling all ONUs are jointly served in successive cycles, and no upstream transmissions can take place during the ISCG.

CONCLUSION

We have surveyed the existing mechanisms and protocols for single-wavelength-channel EPONs and examined their extension to multi-wavelength-channel WDM EPONs. In particular, we have reviewed the recent research on WDM extensions of single-wavelength-channel EPONs. We have provided architectural guidelines for WDM EPONs and a WDM extension of the Multipoint Control Protocol. The extensions are backward compatible with single-wavelength EPONs and do not impose a particular ONU WDM architecture, allowing for a flexible and evolutionary WDM upgrade path for EPONs. We have also outlined and compared two broad paradigms for dynamically allocating the WDM upstream bandwidth, online and offline scheduling. We have found that online scheduling, which makes bandwidth allocations based on individual ONU requests, tends to result in lower packet delays at medium and high traffic loads. Offline scheduling, on the other hand, which makes bandwidth allocations based on a collection of requests from all ONUs, introduces an inter-scheduling cycle gap and report-to-scheduling delay not experienced by an online scheduler.

There are several broad areas for future research on WDM EPONs. As indicated by our comparison results, online scheduling appears to be a promising paradigm for WDM EPONs. Research challenges for online scheduling in WDM EPONs include mechanisms for ensuring fairness and QoS. Fairness and QoS are often more easily achieved with an offline scheduler that has a global view of the ONUs' bandwidth requirements. For offline scheduling, the research challenges include mechanisms for efficiently masking the ISCG. An interesting direction may be the development of hybrid schedulers that incorporate elements from online and offline scheduling to achieve low delays, and at the same time ensure fairness and prescribed levels of QoS.

ACKNOWLEDGMENT

We are grateful to Prof. Ahmet Keha from the Scheduling Group in the Department of Industrial Engineering at Arizona State University for insightful discussions on scheduling in WDM EPONs.

REFERENCES

[1] M. Abrams *et al.*, "FTTP Deployments in the United States and Japan — Equipment Choices and Service Provider Imperatives," *IEEE/OSA J. Lightwave Tech.*, vol. 23, no. 1, Jan. 2005, pp. 236–46.

[2] J. Zheng and H. T. Mouftah, "Media Access Control for Ethernet Passive Optical Networks: An Overview," *IEEE Commun. Mag.*, vol. 43, no. 2, Feb. 2005, pp. 145–50.

[3] M. P. McGarry, M. Maier, and M. Reisslein, "Ethernet PONs: A Survey of Dynamic Bandwidth Allocation (DBA) Algorithms," *IEEE Commun. Mag.*, vol. 42, no. 8, Aug, 2004, pp. S8–S15.

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

[5] C. M. Assi *et al.*, "Dynamic Bandwidth Allocation for Quality-of-Service Over Ethernet PONs," *IEEE JSAC*, vol. 21, no. 9, Nov. 2003, pp. 1467–77.

[6] J. Xie, S. Jiang, and Y. Jiang, "A Dynamic Bandwidth Allocation Scheme for Differentiated Services in EPONs," *IEEE Commun. Mag.*, vol. 42, no. 8, Aug. 2004, pp. S32–S39.

[7] Y. Luo and N. Ansari, "Bandwidth Allocation for Multiservice Access on EPONs," *IEEE Commun. Mag.*, vol. 43, no. 2, Feb. 2005, pp. S16–S21.

[8] C. H. Foh *et al.*, "FULL-RCMA: A High Utilization EPON," *IEEE JSAC*, vol. 22, no. 8, Oct. 2004, pp. 1514–24.

[9] S. R. Sherif *et al.*, "A Novel Decentralized Ethernet-Based PON Access Architecture for Provisioning Differentiated QoS," *IEEE/OSA J. Lightwave Tech.*, vol. 22, no. 11, Nov. 2004, pp. 2483–97.

[10] J.-I. Kani *et al.*, "A WDM-Based Optical Access Network for Wide-Area Gigabit Access Services," *IEEE Commun. Mag.*, vol. 41, no. 2, Feb. 2003, pp. S43–S48.

[11] D. J. Shin *et al.*, "Hybrid WDM/TDM-PON With Wavelength-Selection-Free Transmitters," *IEEE/OSA J. Lightwave Tech.*, vol. 23, no. 1, Jan. 2005, pp. 187–95.

[12] Y.-L. Hsueh *et al.*, "A Highly Flexible and Efficient Passive Optical Network Employing Dynamic Wavelength Allocation," *IEEE/OSA J. Lightwave Tech.*, vol. 23, no. 1, Jan. 2005, pp. 277–86.

[13] K. H. Kwong, D. Harle, and I. Andonovic, "Dynamic Bandwidth Allocation Algorithm for Differentiated Services over WDM EPONs," *Proc. IEEE Int'l. Conf. Commun. Sys.*, Singapore, Sept. 2004, pp. 116–20.

[14] C. Xiao, B. Bing, and G. K. Chang, "An Efficient Reservation MAC Protocol with Preallocation for High-Speed WDM Passive Optical Networks," *Proc. IEEE INFOCOM*, Mar. 2005. Miami, FL.

[15] N. J. Frigo, "A Survey of Fiber Optics in Local Access Architectures," *Opt. Fiber Telecommun.*, IIIA, 1997, pp. 461–522.

[16] M. Pinedo, *Scheduling: Theory, Algorithms, and Systems*, Prentice Hall, 2nd ed., 2002.

BIOGRAPHIES

MICHAEL MCGARRY (michael.mcgarry@asu.edu) received a B.S. degree in computer engineering from Polytechnic University in 1997 and an M.S. degree in electrical engineering from Arizona State University in 2004. He is currently pursuing a Ph.D. degree in electrical engineering at Arizona State University. His research interests are in the areas of MAC protocol design for optical and mobile ad hoc networks.

MARTIN MAIER (maier@ieee.org) is an associate professor at Institut National de la Recherche Scientifique (INRS), Montréal, Canada. He was educated at the Technical University of Berlin, and received Dipl.-Ing. and Dr.-Ing. degrees (both with distinctions) in 1998 and 2003, respectively. Currently, his research activities focus on evolutionary WDM upgrades of optical access and metro networks. He is the author of the book *Metropolitan Area WDM Networks — An AWG Based Approach* (Springer, 2003).

MARTIN REISSLEIN (reisslein@asu.edu) is an associate professor in the Department of Electrical Engineering at Arizona State University, Tempe. He received his Ph.D. in systems engineering from the University of Pennsylvania in 1998. 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 of Berlin. He maintains an extensive library of video traces for network performance evaluation, including frame size traces of MPEG-4 and H.263 encoded video, at http://trace.eas.asu.edu

EDITORIAL/*from page S2*

This year and next month, a special event is taking place, a conference that promises to dazzle (the technologist) and puzzle (the conservatist). It is OFC/NFOEC 2006, the product of two consolidated photonic conferences, entangled in a relationship of strength and technological grandeur. OFC/NFOEC 2006 will take place in… where else… Anaheim, California, from 5 to 10 March. It will have more than 750 presentations focused on the hottest topics such as FTTx, reconfigurable optical add-drop multiplexers (ROADM) and many other networking topics as well as hot topics of the most comprehensive optical and photonic technology, and the latest breakthroughs and innovations. In addition, it will provide a thorough view of a spectrum of commercial products and services. It will also have hundreds of exhibitors where the visitor will once again be amazed with the technology innovation and business sense of these companies.

Anaheim, the host of this anchor event in optical networks and photonic technology, once again will demonstrate its hospitality to an international audience in a festive atmosphere, which in our experience has been most rewarding in terms of papers and exhibitors. See you in Anaheim.

For more information visit http://www/ofcnfoec.org

CONFERENCE PREVIEW/*from page S4*

an excellent opportunity for attendees to learn about new products, cutting-edge technology and vital information at the forefront of communications. New selections include Alternate Broadband Access: Wired and Wireless Technologies for the Last Mile, Quantum Cryptography and Quantum Information, and Passive Optical Components and Filtering Technologies. New offering are also available in hands-on workshops including both Outside Plant Hands-on Testing and Troubleshooting and Outside Plant Splicing, Testing and Troubleshooting for FTTX as well as Silicon Microphotonics: Technology Elements and the Roadmap to Implementation. The entire 2006 offering is available at www.ofcnfoec.org/conference_program/Short_Course.

The OFC/NFOEC Exhibit Floor has also been expanded for 2006. The exhibition is the ideal place to build and maintain professional contacts while broadening your knowledge of the companies that lead the optical communications industry in product development and technological advances. An impressive line up of more than 650 exhibitors will share space with the new FTTH Center, an Emergency Restoration Demonstration Area, the Video Theater and its series of new generation topics in digital DVD formats, the Plastic Optical Fiber Trade Organization Market Opportunities Presentation, as well as the International Theater and Exhibitor Showcases.