Jitter Performance in Ethernet Passive Optical Networks

Abdallah Shami, *Member, IEEE*, Xiaofeng Bai, *Student Member, IEEE*, Chadi M. Assi, *Member, IEEE*, and Nasir Ghani, *Senior Member, IEEE*

***Abstract*—Ethernet passive optical networks (EPONs) have emerged as one of the most promising access network technologies. Propelled by rapid price declines in fiber optics and Ethernet components, these architectures combine the latest in optical and electronic advances and are poised to become the dominant means of delivering gigabit broadband connectivity to homes over a unified single platform. As this technology matures, related quality of service (QoS) issues are becoming a key concern. This paper proposes a novel dynamic scheduling algorithm, termed hybrid granting protocol (HGP), to support different QoS in EPON. Specifically, the proposed dynamic scheduling algorithm minimizes packet delay and jitter for delay and delay-variation sensitive traffic (e.g., voice transmissions) by allocating bandwidth in a grant-before-report (GBR) fashion. This considerably improves their performance without degrading QoS guarantees for other service types. Detailed simulation experiments are presented to validate the effectiveness of the proposed algorithm.**

***Index Terms*—Dynamic bandwidth allocation (DBA), ethernet passive optical network (EPON), packet delay variation, quality of service.**

I. Introduction

RECENTLY, there has been a dramatic increase in the amount of network data traffic, primarily driven by the rising number of Internet users demanding increased data rates. Moreover, a wide range of increasingly bandwidth-intensive services are continuing to emerge, e.g., storage extension/virtualization, grid computing, packet video teleconferencing, etc., [1]. As traffic demands continue to grow, network carriers and operators have steadily scaled their backbone capacities by deploying high-speed dense wavelength-division-multiplexing (DWDM) technology within their backbones and even metro-regional cores. Concurrently, end-user local-area networks (LANs) have also seen their tributary speeds progressively increase from 100 Mb/s upwards to 1.0 Gb/s and beyond. In all, the net result is a growing gap between the capacity of backbone networks and critical "last-mile" access infrastructures connecting to end-user networks, i.e., termed "access bottleneck."

Since the cost for access technologies is usually prohibitive for household implementations, there is actually not a prevalent and dedicated infrastructure deployed that promises to carry broadband and data-dominated services in the end users' neighborhood. As a matter of fact, the most widely deployed "broadband" solutions today are Digital Subscriber Line (DSL) and Cable Modem (CM) networks.

Although DSL and CM provide much more bandwidth than 56 Kb/s dial-up lines, they cannot offer enough bandwidth for the emerging bandwidth-intensive services such as Video-On-Demand (VoD) and two-way video conferencing. Clearly, more radical revolutions in access networks are required to propel an entire optical architecture that is optimized for Internet Protocol (IP) data traffic. The remaining voice and video services will converge into a digital format to share the integrated full-service platform. Considering the inherent cost-effectiveness of access connections, the next-generation access network should minimize the necessary on-site engineering overhead as well as the per-subscriber cost for the best scalability. Passive Optical Network (PON) is a technology viewed by many as an attractive infrastructure for the next generation access network. A PON is a point-to-multipoint optical network with no active elements in the signals' path from source to destination. The only interior elements used in PON are passive optical components, such as optical fiber, splices, and splitters. PON minimizes the number of optical transceivers, central office terminations, and fiber deployment. In addition, since there is no mandatory power supply and maintenance in the fiber plant, PON also offers scalability in the most economical way. Given the widely deployed Ethernet ports and ubiquitous Ethernet-originated end-user traffic, Ethernet-based Passive Optical Networks (EPONs) appears to be a natural candidate for the next-generation broadband access networks. In particular, an EPON basically comprises of an optical line terminal (OLT) residing in the central office connected to multiple optical network units (ONUs) near subscribers' locations. Meanwhile, there is only a single optical landline deployed between the OLT and a passive optical splitter/combiner, which is located near ONUs to dispatch/collect optical signals. Various ONU deployments exist, as per different architectures such as fiber-to-the-home (FTTH), fiber-to-the-building (FTTB), and fiber-to-the-curb (FTTC).

Currently, numerous research and standardization efforts are underway in order to evolve more capable EPON solutions at highly cost-effective price points. It is the purpose of this work to address some of these issues through devising and demonstrating several novel dynamic bandwidth allocation algorithms for EPON that will help increase the viability of optical access over a broader range of subscriber access scenarios. Specifically, this paper proposes a novel dynamic scheduling algorithm,

Manuscript received October 15, 2004; revised January 4, 2005. This work was supported in part by the National Sciences and Engineering Research Council of Canada (NSERC).

A. Shami and X. Bai are with the Department of Electrical and Computer Engineering, the University of Western Ontario, London, ON N6A 5B9, Canada (e-mail: ashami@eng.uwo.ca; xbai6@uwo.ca).

C. M. Assi is with Concordia Institute for Information Systems Engineering, Concordia University, Montreal, QC H3G 1M8, Canada.

N. Ghani is with the Department of Electrical and Computer Engineering, Tennessee Tech University, Cookeville, TN 38505 USA.

Digital Object Identifier 10.1109/JLT.2005.844510

termed hybrid granting protocol (HGP), to support guaranteed QoS for different types of applications in access networks. The proposed dynamic scheduling algorithm minimizes the packet delay and jitter for delay-sensitive applications in EPON, such as narrowband voice. This scheduling algorithm allocates capacity in a grant-before-report (GBR) fashion and considerably improves related delay performance without degrading QoS support for other service types. We conduct detailed simulation experiments to study the performance of the proposed algorithm and validate its effectiveness. The rest of this paper is organized as follows. Section II presents a background motivation for our work and subsequently, Section III presents our new protocol. Detailed performance analysis is presented in Section IV and final conclusions are covered in Section V.

II. Overview and EPON Architecture

An EPON setup is basically a point-to-multipoint (1:N) optical access network. In the downstream direction of an EPON, Ethernet frames are broadcast by the OLT and are selectively received by each ONU. While in upstream direction, all of the ONUs must contend for shared capacity link, and this requires an appropriate access protocol. Now herein, the direct application of existing carrier sense multiple access with collision detection (CSMA/CD) schemes is very difficult owing to large OLT-ONU propagation delays, i.e., distances typically 10–25 km. Hence many researchers have tabled time-division multiple-access (TDMA) schemes to achieve reliable, robust channel sharing in optical access networks [2]–[6]. This approach allows the ONUs to share a single upstream wavelength in which the OLT allocates time slots to each ONU to transmit its backlogged traffic.

Since the optical splitter is designed for one-way upstream communication, EPON prohibits direct information exchange amongst ONU nodes. As result, the OLT is the only network element that can arbitrate capacity allocation within the EPON. One benefit of centralizing intelligence at the OLT is that it allows for a very cost-effective design with lower ONU costs. Now to implement TDMA schemes, upstream data transmission can be scheduled using static bandwidth allocation (SBA), where each ONU is preassigned a fixed timeslot to send its backlogged packets at the full capacity of the link. Clearly, SBA is a simple scheme and is suitable for supporting predictable constant bit rate traffic. However, owing to end-user service diversity, next-generation access networks are expected to provide bundled support for a wider range of services, including video and data services. It is well-known that such aggregated traffic profiles exhibit nonuniform behaviors, thereby mandating advanced bandwidth allocation schemes [7].

In order to achieve statistical multiplexing in EPON architectures, the IEEE 802.3ah Ethernet in the First Mile Task Force has developed a multipoint control protocol (MPCP) [8]. MPCP defines a message-based mechanism to facilitate real-time information exchange between the OLT and each ONU. The control messages defined in MPCP are REPORT, GATE, REGISTER_REQ, REGISTER, and REGISTER_ACK. Instead of specifying a particular scheduling approach, MPCP only provides a basic mechanism for developing a wide range of bandwidth allocation schemes. The exact choice of such mechanisms is left to vendor discretion. Hence, by using MPCP an OLT is capable of scheduling transmission using various dynamic bandwidth allocation (DBA) strategies amongst ONUs [4].

To date, various DBA algorithms have been proposed for EPONs. Most notably, the interleaved polling scheme with adaptive cycle time (IPACT) [6] requires the OLT to poll every ONU and dynamically assign it bandwidth before transmission. This bandwidth is allocated according to the buffer occupancy status of the ONU, which is reported by the ONU through its REPORT message. Any unrequested bandwidth will not be granted and the scheduling frame size is therefore not fixed. In order to better utilize the leftover bandwidth from ONUs with smaller traffic backlogs, the authors in [5] proposed a DBA scheme in which ONU nodes were partitioned into two groups—underloaded and overloaded—according to their minimum guaranteed transmission window sizes. Here, total bandwidth saved from underloaded group is reallocated to overloaded ONUs to improve efficiency. This approach assumes that the saved bandwidth will always be fully occupied by overloaded ONUs, which is not necessarily true. In [10], the authors proposed an advanced dynamic bandwidth algorithm and introduced a measure for total excess bandwidth requested by overloaded group to combat the aforementioned deficiency. Also, the authors in [11] proposed a new concept of DBA, where the ONUs are partitioned into bandwidth guaranteed (BG) group and nonbandwidth guaranteed (non-BG) group. Each ONU is allocated one or multiple evenly divided and distributed granules of the scheduling frame.

Since EPON carries different types of service over a shared platform, QoS is a very important concern for all DBA scheduling algorithms. Recently, various studies have looked at multiservice QoS support in EPON settings. For example, the authors in [2] employed priority queuing and intra-ONU queue scheduling to provide differentiated service for each service class and to eliminate the light load penalty. While in [12] the authors proposed a per-queue based logical link identifier (LLID) to offer QoS and fairness according to a global perspective. Namely, each priority queue in the same ONU node will get its individual grant information in the GATE message. However, the disadvantage of this scheme is that per-queue granting can potentially increase idle times due to packet fragmentation at each queue.

Consider the fact that high priority services are very delay sensitive but tend to be of a narrow-band nature. Conversely, medium and low priority services are more delay tolerant but generally have a wide-band nature. Hence it is reasonable to schedule high priority service frequently with smaller bandwidth grants and to schedule medium and low priority class services less frequently but with larger bandwidth grants. Along these lines, the authors in [13] have tabled a *stratified round robin* (SRR) scheme basis. Furthermore, [14] decouples the generation of grant window sizes from the decision of transmission start times. Specifically, grants for different purposes (e.g., SBA, DBA, auto-discovery, etc) are generated independently and then scheduled and converged into the downstream flow by the OLT. This approach, however, requires much more bandwidth to be dedicated for downstream control messaging. In [17], the authors introduced the *Hybrid Slot-Size/Rate* (HSSR) protocol to support QoS in EPON networks. This is achieved by separating

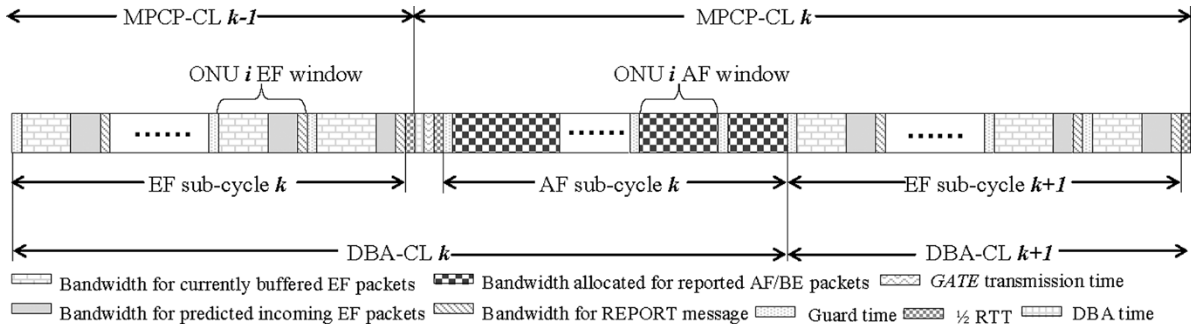

Fig. 1. Illustration of DBA-CL and MPCP-CL.

the transmission of high priority and low priority packets into steady part and dynamic part respectively within one scheduling frame. Here the frame size is assumed to be fixed (1 ms) and SBA is applied in the steady part. This fixed frame size limits the upstream efficiency in the context of highly bursty traffic, which typically occupies the dynamic part. The fairness of sharing the dynamic part in HSSR is provided by a counter that is weighted by the amount of backlogged low priority data at each ONU. However, this mechanism does not yet explicitly guarantee a minimum bandwidth for inputting low priority streams, especially when one or more ONUs pump a large amount of low priority data into the network. In real network scenario, both the high burstiness of aggregated traffic and per-ONU bandwidth guarantee should be taken into consideration.

Most of the above proposals are aware of the fact that delay sensitive traffic should be treated in a specialized manner within the OLT DBA stage. Nevertheless, these schemes do not address the inherent limitation of using a grant-after-report (GAR) allocation strategy in a uniform manner for all service types. In this paper we propose a novel, improved hybrid granting (HG) protocol that uses a different granting scheme for GATE messages to improve QoS support in EPON. The proposed protocol is adapted to variable scheduling frame size and guarantees a minimum bandwidth for each ONU in every frame.

III. Proposed Scheduling Algorithm

The proposed algorithm is motivated by the need to further improve the earlier scheduling algorithms, i.e., as tabled in [2], [5], [6], [10], [11], [14], [17]. Briefly, let us consider the operation of REPORT/GATE messaging in EPON as detailed in [4], [5]. Here, the OLT sends a GATE message to each ONU (according to its last REPORT message) at the beginning of every scheduling frame. In turn, every ONU sends its REPORT message to report its buffer occupancy to the OLT at the end of its assigned transmission window. Upon receiving a REPORT message from every ONU, the OLT performs DBA and the next frame is then launched. The DBA algorithm proposed in [5] describes the basic operations of MPCP in an EPON. Hereafter, we refer to this algorithm as regular EPON scheduler.

In order to support different classes of service with different packet delay and delay jitter requirements, we introduce three prioritized service classes to represent different types of service—expedite forwarding (EF) with the highest priority for strictly delay sensitive services that is typically constant bit rate (CBR) voice transmission, assured forwarding (AF) with medium priority for services of nondelay sensitive variable bit rate (VBR) services such as video stream, and best effort (BE) with the lowest priority for delay tolerable services which include web browsing, background file transfer and e-mail applications.

A. GAR- and GBR-Based Allocation

In a regular EPON scheduler, MPCP is always implemented in a grant-after-report (GAR) fashion, i.e., the grant always follows the reported information that it is based upon. In this scheme only packets that are reported by the last REPORT message of their parent ONU are eligible for transmission in the current scheduling frame. This "report-grant-transmit" operation defines a minimum packet queuing delay of one scheduling frame for each service type. Nevertheless, this might not be the optimum solution for delay-sensitive applications.

A key motivation for DBA is the inherently nondeterministic nature of bit-rates for AF and BE service types. Conversely, the amount of traffic offered by EF end-users is fully-deterministic and this permits the OLT to use a grant-before-report (GBR) strategy for such services during DBA. Specifically, the GATE message may include grant information for the expected EF traffic that will arrive before the next transmission start time of its parent ONU. This preallocation scheme guarantees a maximum packet queuing delay of one scheduling frame for EF services.

It is helpful to clearly define the term transmission cycle before any further discussion. Inside each scheduling frame, we refer to the time interval between the moment when the foremost GATE message is completely received and the moment when the lattermost REPORT message is completely transmitted, as a transmission cycle. Precisely, a scheduling frame sequentially consists of OLT DBA time, transmission and propagation delay of the foremost GATE message, transmission cycle duration and propagation delay of the lattermost REPORT message. Note the transmission delay of the lattermost REPORT message is scheduled into the lattermost transmission window, this is shown in Fig. 1. Since a guard time is scheduled before each transmission window to perform ranging for different OLT-ONU distances and to provide transmitter switching time, we premise a uniform OLT-ONU distance for simplicity. Thus, both the propagation delays of the foremost GATE message and the lattermost REPORT message are one half of round-trip time (RTT).

B. EF Subcycle and AF Subcycle

We now present a two-cycle allocation scheme for EF and AF/BE services in EPON. First of all, we assume that the OLT

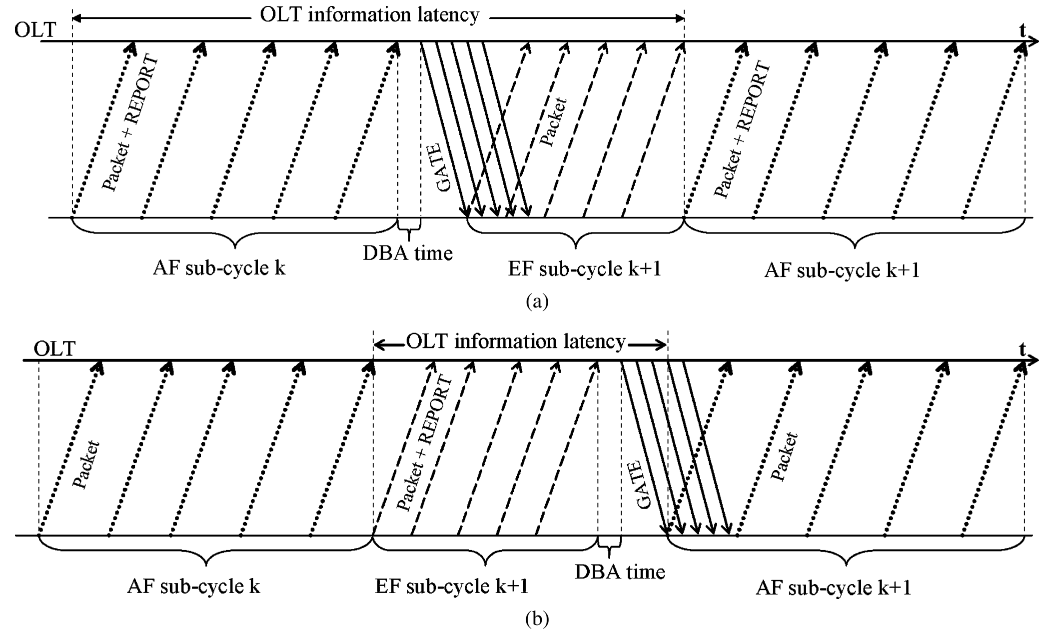

Fig. 2. (a) Illustration of OLT information latency in nonshift situation. (b) Illustration of OLT information latency in shifted situation.

applies GBR allocation for EF services and retains GAR allocation for AF and BE services, i.e., via GATE message. Hence EF bandwidth allocation is not based upon the last REPORT message of every ONU. Instead, when the OLT is performing DBA, it needs to determine the prospective bandwidth demand for the EF services at each ONU before deciding the allocations for other service types. Now in order to determine the bandwidth demand for EF services, it is necessary for the OLT to accurately predict the prospective transmission start time of each ONU in the next transmission cycle, thereby the expected arriving EF traffic at that moment could be estimated accordingly. Nevertheless, the prospective transmission start time of an ONU can not be decided before bandwidth is allocated to all ONUs with earlier transmission times. This contradiction may prohibit the simultaneous implementation of GBR and GAR allocations.

To achieve feasible operation, we extract the EF data transmission of each ONU out of its transmission for AF and BE packets. Namely, we build an EF subcycle and an AF subcycle in the above defined transmission cycle to transmit EF data and AF/BE aggregated data respectively. Therefore, EF bandwidth allocation for one ONU and transmission start time for the next ONU can be decided sequentially in the EF subcycle. Since EF service is narrowband and prioritized over AF and BE services, EF subcycles will always be granted bandwidth on "as-needed" basis and ahead of AF subcycles. Therefore, the available bandwidth for AF subcycle is contingent upon the bandwidth allocated to the EF subcycle along with the total available bandwidth defined by the maximum transmission cycle time (MTCT) parameter.

C. DBA Cycle and MPCP Cycle

In order to properly illustrate our proposed approach we introduce two concepts—DBA cycle (DBA-CL) and MPCP cycle (MPCP-CL). A DBA-CL is defined by the above mentioned MTCT and determines the minimum guaranteed bandwidth for each ONU in a transmission cycle, i.e., sum of EF subcycle and AF subcycle durations. Meanwhile, a MPCP-CL refers to a physical operational round of the MPCP messages between the OLT and every ONU. In other words, DBA-CL is a computation-based concept and MPCP-CL is operation based. At the beginning of a MPCP-CL, the OLT sends GATE message to each ONU to inform it of its granted window size and transmission start time for the next EF and AF subcycles, respectively. Here, the window size for EF subcycle is allocated in GBR fashion and AF subcycle in GAR fashion. At the end of its AF transmission window, each ONU sends a REPORT message to the OLT to request bandwidth for its buffered AF and BE packets. Note that now it is not necessary for the ONU to request bandwidth for EF packets.

As mentioned earlier, the EF subcycle in DBA-CL should be scheduled ahead of AF subcycle. However, this is not a ubiquitous rule for both DBA-CL and MPCP-CL. In the previous scheduling schemes, both DBA-CL and MPCP-CL were completely overlapped within a scheduling frame, albeit these two concepts are not necessarily always "co-phased". Now we shift them in Fig. 1 for further refinement. As shown in Fig. 1, each DBA-CL begins with an EF subcycle and ends with an AF subcycle, while a MPCP-CL begins with an AF subcycle and ends with an EF subcycle (some related overheads are shown in the figure). Hence, the REPORT message is sent at the end of each EF transmission window by the ONU. The reason for this shifting is to purposely delay the REPORT message sending time at each ONU. Namely, this delay enables the OLT to obtain more up-to-date buffer occupancy information from each ONU. This point is further illustrated in Fig. 2(a) and (b). Hence we can see that in the nonshifted case, the OLT only takes the buffer occupancy information of an ONU updated to the moment that is two subcycles (one EF subcycle and one AF subcycle) ahead of the next AF transmission start time of this ONU. However, in the shifted case the OLT can obtain buffer occupancy information of an ONU updated to a point in time that is about one subcycle (one EF subcycle or one AF subcycle, or a mixture of these two) ahead of the next AF transmission start time of the ONU. With this improved, more up-to-date information, the DBA module of the OLT can allocate more bandwidth to the next AF subcycle in case that the MTCT is not met.

D. Dynamic Bandwidth Allocation (DBA)

Now a GATE message includes two grant sections—one for the AF subcycle of the current DBA-CL and the other for the EF subcycle of the next DBA-CL. Hereafter, we refer to this granting approach as Hybrid Granting (HG) and the following discussion is premised upon the DBA-CL. To achieve per-ONU bandwidth guarantee, the transmission window size allocated to the first section should be ensured with the residual amount of the minimum guaranteed bandwidth for every ONU. Note that the minimum guaranteed bandwidth is already partially occupied by the EF subcycle of the current DBA-CL, which is already expired. Besides this guarantee, any unused portion of an underloaded ONU's minimum guaranteed bandwidth should be reasonably shared by other overloaded ONUs, in case of necessity. In order to formularize this DBA policy, it is helpful to consider some necessary parameter definitions:

C	OLT link capacity (b/s).
n	Total number of ONU nodes.
$B_i^{\min}$	Minimum guaranteed bandwidth for ith ONU (sum of EF and AF guarantees).
B^{tl}	Total available upstream bandwidth in a DBA-CL defined by MTCT.
$T_{cl}^{\max}$	Maximum transmission cycle time (MTCT).
L_{rep}	Bit length of REPORT message.
g	Guard time.
w_i	Preassigned weight for i-th ONU in sharing the B^{tl}, where $\sum_{i=1}^{n} w_i = 1$.
R_i	Reserved EF rate for ith ONU (b/s).
$t_{i,k}^{\text{star\_c}}$	*EF subcycle* transmission start time of ith ONU in kth DBA-CL.
$c_{i,k}$	Granted *EF subcycle* window size for ith ONU in kth DBA-CL, where $c_{i,k} = R_i \times (t_{i,k}^{star\_c} - t_{i,k-1}^{star\_c}) + L_{rep}$.
$r_{i,k}$	Requested *AF subcycle* window size by ith ONU in kth DBA-CL.
$v_{i,k}$	Granted *AF subcycle* window size for ith ONU in kth DBA-CL, where $v_{i,k} \le r_{i,k}$.
M	Set of ONU nodes that require more bandwidth than their minimum guarantee (overloaded ONUs), i.e., $c_{i,k} + r_{i,k} > B_i^{\min}$, in kth DBA-CL.
N	Set of ONU nodes that require less bandwidth than their minimum guarantee (underloaded ONUs), i.e., $c_{i,k} + r_{i,k} < B_i^{\min}$, in kth DBA-CL.

Because two independent *subcycles* (EF and AF) require two sets of guard times to delimit packets transmitted by different ONUs in one DBA-CL, the total available upstream bandwidth in a DBA-CL, i.e., B^{tl}, should be determined as:

$$B^{tl} = C \times (T_{cl}^{\max} - 2 \times g) \tag{1}$$

whereas the minimum guaranteed bandwidth for ith ONU, i.e., $B_i^{\min}$ can be decided by its weight w_i and B^{tl} as:

$$B_i^{\min} = w_i \times B^{tl} \tag{2}$$

From above, the total excessive bandwidth v_k^{ex} saved by N (underloaded ONUs) and total extra bandwidth v_k^{dem} demanded by M (overloaded ONUs), in kth DBA-CL, are given as:

$$v_k^{ex} = \sum_{l \in N} \left(B_l^{\min} - c_{l,k} - r_{l,k}\right) \qquad \left(Note\ for\ N: c_{l,k} + r_{l,k} < B_l^{\min}\right) \tag{3}$$

and

$$v_k^{\text{dem}} = \sum_{j \in M} \left(c_{j,k} + r_{j,k} - B_j^{\min}\right) \qquad \left(Note\ for\ M: c_{j,k} + r_{j,k} > B_j^{\min}\right) \tag{4}$$

Using these definitions, the aforementioned two sections in the *GATE* message can be computed as follows:

$$v_{i,k} = \begin{cases} r_{i,k} & v_k^{\text{dem}} \le v_k^{\text{ex}}\ or \\ & r_{i,k} \le B_i^{\min} - c_{i,k} \\ B_i^{\min} - c_{i,k} + v_{i,k}^{\text{ex}} & Otherwise \end{cases} \tag{5}$$

$$c_{i,k+1} = R_i \times \left(t_{i,k+1}^{star\_c} - t_{i,k}^{\text{star\_c}}\right) + L_{\text{rep}} \tag{6}$$

where $v_{i,k}^{\text{ex}}$ is the portion of v_k^{ex} eventually shared by ith $(i \in M)$ ONU. The background principle for deciding $v_{i,k}$ is as follows: if the total unused bandwidth saved by underloaded ONUs can satisfy the total extra demanded bandwidth by overloaded ONUs, every $r_{i,k}$ will be fully satisfied. Otherwise, underloaded ONUs are granted $v_{i,k}$ as their request, i.e., $r_{i,k}$, while overloaded ONUs are granted their guaranteed bandwidth first and then share a portion of v_k^{ex}, which is $v_{i,k}^{\text{ex}}$.

In order to reach reasonable sharing, $v_{i,k}^{\text{ex}}$ should be evaluated according to the proportion of individual extra bandwidth demanded by every overloaded ONU, termed as $v_{i,k}^{\text{dem}}$, respecting to the total extra bandwidth demanded by all overloaded ONUs, i.e., v_k^{dem} as:

$$v_{i,k}^{\text{ex}} = v_k^{\text{ex}} \times \frac{v_{i,k}^{\text{dem}}}{v_k^{\text{dem}}}. \tag{7}$$

It is important to emphasize that when the OLT performs DBA in k-th DBA-CL, every $t_{i,k+1}^{\text{star\_c}}$ is determinable and hence $c_{i,k+1}$ can be carried by the *GATE* message, along with $v_{i,k}$.

The granted EF window $c_{i,k+1}$ here might not be fully utilized by the ONU since Ethernet packets are not allowed to be fragmented. In this situation some fractional bandwidth will be wasted. Unlike AF and BE classes, however, EF packets are generally equal-sized. Therefore, the OLT can only grant EF window suitable for the transmission of integer number of EF packets, and register any extra fraction for the next EF granting. Explicitly, we can rewrite the granted EF window size as:

$$c_{i,k+1} = floor\left\{\frac{1}{L_{\text{EF}}} \times \left[\left(t_{i,k+1}^{\text{star\_c}} - t_{i,k}^{\text{rep\_c}}\right) R_i + f_{i,k}\right]\right\} \times (\text{IPG} + P + L_{\text{EF}}) + (\text{IPG} + P + L_{\text{rep}}) \tag{8}$$

and

$$f_{i,k} = \text{mod}\left[\left(t_{i,k}^{\text{star\_c}} - t_{i,k-1}^{\text{rep\_c}}\right) R_i + f_{i,k-1}, L_{\text{EF}}\right] \tag{9}$$

TABLE I
SIMULATION PARAMETERS

Number of ONUs	32
Upstream/downstream link capacity	1 Gbps
OLT—ONU distance (uniform)	20 Km
Maximum transmission cycle time (MTCT)	2 ms
Guard time	1 μs
OLT DBA time	neglected

where L_{EF} denotes the bit length of an EF packet, and $f_{i,k}$ is the registered fraction for ith ONU in kth DBA-CL. Here we also considered the 96 bits interpacket gap IPG and 64 bits preamble P in front of each packet when computing $c_{i,k+1}$. The IPG and P should also be included in every $r_{i,k}$.

E. Inner ONU Queue Scheduling

As mentioned before, the AF transmission window granted to an ONU is an aggregated value for both AF and BE services. In order to prevent AF services from inadvertently monopolizing this granted window allocation, it is desirable to further employ intra-ONU scheduling schemes [2], [5], i.e., only reported packets are eligible to access the link in accordance with their service priorities.

This new HG approach specifically protects the service latency of delay-sensitive EF service and permits the provider to reserve an explicit bandwidth for it. Nevertheless, the upstream network overhead is slightly higher as we have to introduce an extra set of guard times, see Fig. 1. However, these new guard times do not necessarily equal the number of ONU nodes in the network. The reason here is that the granted window size for each ONU in every AF subcycle can fluctuate widely. Hence in the case where an ONU is not granted any bandwidth in its AF transmission window, it will essentially waste a guard time. At this moment, this ONU will not be scheduled in the next AF subcycle and hence its associated guard time will be cancelled. Numerical results of the overall network throughput are presented in the next Section.

IV. PERFORMANCE EVALUATION

Detailed simulation studies are conducted to validate the performance of the proposed HG protocol. The scheme is tested using a discrete event simulator developed in C++, and the key simulation parameters are summarized in Table I. To closely emulate the self-similar property of AF and BE traffic (as generated by multiple LAN subscribers), we generate detailed self-similar traffic models for all ONUs. Furthermore, in order to combat the extreme uncertainty of self-similar traces and deliver conclusive results, the outcomes of multiple repeated simulation runs are averaged for each result. Since EF service is narrowband, it is allocated 20% of the traffic load at each entry point to the network. The remaining 80% of the load is evenly distributed between AF and BE services, i.e., 40% each. To simplify the simulations, we also assume that the total network load is evenly distributed amongst all ONUs and the ONUs are equally weighted.

A. Average Packet Delay

Fig. 3 compares the average packet delay of EF, AF, and BE services for both the regular EPON scheduler and our proposed

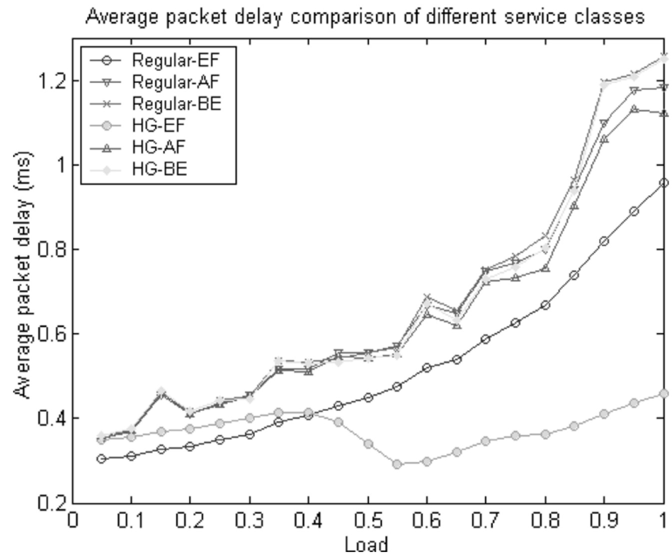

Fig. 3. Comparison of average packet delay.

HG scheduler. We can explicitly identify the specific case of HG scheduler for EF service in the graph. These results show that the HG scheme reduces average EF packet delay by over 50% as compared to the regular EPON scheduling scheme, i.e., for both medium and heavy load scenarios. This is a notable improvement.

Note that Fig. 3 shows a slight increase in EF packet delay at lighter loads for the HG scheme as compared to the regular EPON scheduling scheme. This occurs due to the aforementioned guard-time issue. Namely, in HG scheduler, when the load is very light the DBA-CL is so small that some ONUs may not have any EF packets buffered in their queues on the next EF transmission start time. Hence the OLT only allows such ONUs to send a REPORT message in the next EF subcycle. However, this REPORT message still needs a guard time for scheduling, and hence leads to the slight increase in delay. Conversely, with the regular scheduling scheme the REPORT message does not require a separate guard time in such situations. Moreover, in HG scheduler the availability of more up-to-date ONU buffer occupancy information at the OLT also tends to keep the DBA-CL larger than the one in regular EPON scheduler. This also contributes to the slight increase in delay performance at lighter load points.

Nevertheless, by having more accurate ONU buffer occupancy information at the OLT, the HG scheme can always maintain lower average packet delays for AF and BE services, Fig. 3.

B. Network Throughput

The proposed HG scheduler tends to keep a larger DBA-CL than the regular EPON scheduler. Therefore, the DBA-CL duration in HG scheduler approaches its maximum value determined by MTCT faster when the network is heavily loaded. Fig. 4 shows that the HG scheduler offers the same level of throughput as the regular EPON scheduler if the offered load is less than 0.8. This is achieved by reducing the messaging frequency between the OLT and ONUs with larger DBA-CL durations. For loads higher than 0.8 Fig. 4 shows that the proposed HG scheduler offers 83.5% throughput, compared with 85% in the regular

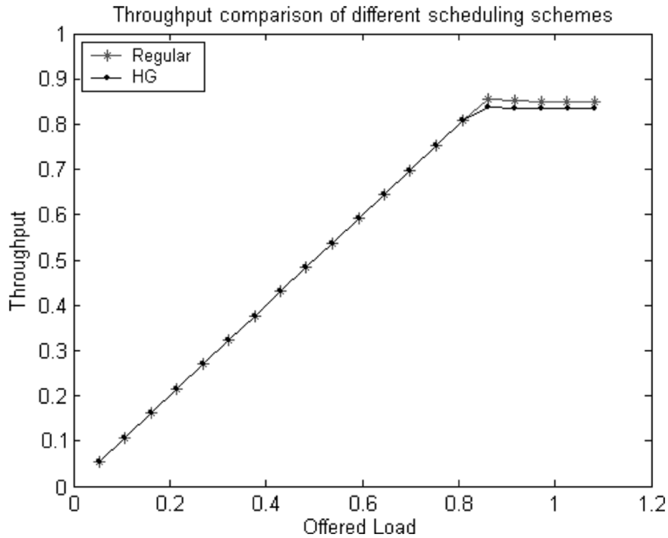

Fig. 4. Comparison of network throughput.

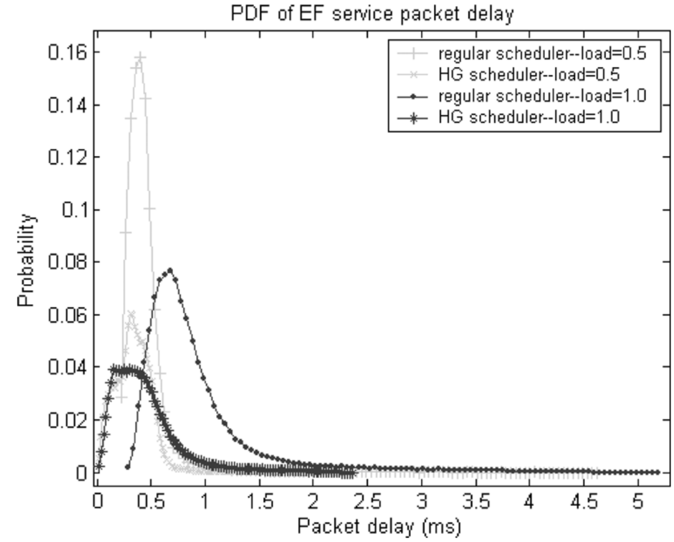

Fig. 5. Simulation estimated pdf.

EPON scheduler. Considering the improved delay performance (nearly 50% in medium and heavy load scenarios) of delay sensitive services, therefore 1.5% throughput degradation can be tolerated.

C. EF Service Packet Delay Variation

Another positive contribution of the HG protocol is the improved packet delay variation for EF services. Before developing into deeper discussion, it is worthwhile to acknowledge the following fact. Since EF traffic is nonbursty, it is reasonable to assume that the inter-arrival time of two successively arrived EF packets is greater than the transmission time of the first EF packet by the ONU. We assume that all EF packets are equal-sized as 70 bytes for easier discussion. The above inter-arrival time and packet transmission time are therefore both constant. Hence the packet delay variation of two consecutively departed EF packets from the same ONU in the same transmission window, that is referred to as intrawindow jitter (A-WJ), can be formularized as

$$J_i = D_i - D_{i-1} = T_{\text{trans}} - T_{\text{int}} \equiv C \tag{10}$$

where J_i is the ith delay jitter within the window, D_i is the ith packet delay within the window, T_{int} and T_{trans} are the interarrival time and packet transmission time respectively. Note that J_1 is not defined here. It is clear that the A-WJ is always a minus constant value C, because $T_{\text{int}} > T_{\text{trans}}$. Namely, in the same transmission window, EF packet delay decreases by a constant rate, while its actual value depends on the delay value of the first departed EF packet in the transmission window, i.e., D_1. Therefore, a greater D_1 leads to a series of larger delay, and vice versa. The variation of D_1 between two consecutive transmission windows (consecutive EF windows for HG scheduler), that is referred to as interwindow jitter (E-WJ), consequently maps the distribution property of the total EF delay sequence for the parent ONU. In other words, with more fluctuated E-WJ, EF packets tend to be continuously overdelayed or underdelayed, respecting to their mean delay value. Also, the total EF delay sequence appears more dispersed. Whereas less fluctuated E-WJ tends to keep EF packet delay distributed evenly around their mean delay value and the total EF delay sequence more centralized. To validate these theoretical deductions, we tracked the EF packet delay of a randomly selected ONU at different loading for both regular EPON scheduler and HG scheduler in one simulation process. Fig. 5 shows the probability density function (pdf) of EF service packet delay at half and full loading scenarios. It is shown that in both scenarios the EF delay sequence presents centralization with all data points condensed before 2.5 ms for HG scheduler and dispersion with enough number of data points in a long tail until 5.5 ms for regular EPON scheduler. Comparing the variance of the sampled delay sequence (σ^2) reveals a numerical explanation of this discrepancy, where $\sigma^2 = 0.0325\ \text{ms}^2$ at half loading and $\sigma^2 = 0.1058\ \text{ms}^2$ at full loading for HG scheduler, versus $\sigma^2 = 0.0554\ \text{ms}^2$ at half loading and $\sigma^2 = 0.3095\ \text{ms}^2$ at full loading for regular EPON scheduler. Since similar observations were obtained at half loading scenario, the following discussion is assisted only with the simulation results at full loading scenario. ha

As we mentioned earlier, since EF packet delay within one transmission window follows a stationary decreasing process, the larger variance (σ^2) for regular EPON scheduler implies more serious fluctuation of E-WJ, and hence introduces distribution continuity in the delay sequence. The autocorrelation function (ACF) of the error vector $\overline{V} = \overline{D} - mean(\overline{D})$ visionably explores this inherent continuity, where $\overline{D}$ is the original EF delay sequence. Due to the existence of different σ^2 for the two scheduler (regular and HG), it is reasonable to normalize one error vector to the other, in order to reach comparability. Here we revise the error vector of regular EPON scheduler by the following equation:

$$\overline{V}_{\text{reg\_nom}} = \overline{V}_{\text{reg}} \times \sqrt{\frac{\sigma_{HG}^2}{\sigma_{\text{reg}}^2}} \tag{11}$$

where $\overline{V}_{\text{reg}}$ and $\overline{V}_{\text{reg\_nom}}$ are the error vectors of regular EPON scheduler before and after the normalization, while σ_{reg}^2 and σ_{HG}^2 are the σ^2 for regular EPON scheduler and HG scheduler,

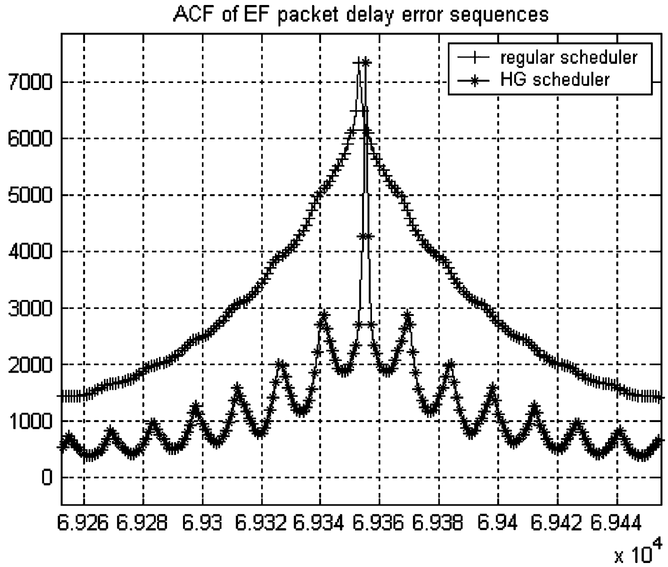

Fig. 6. EF packet delay error vector ACF.

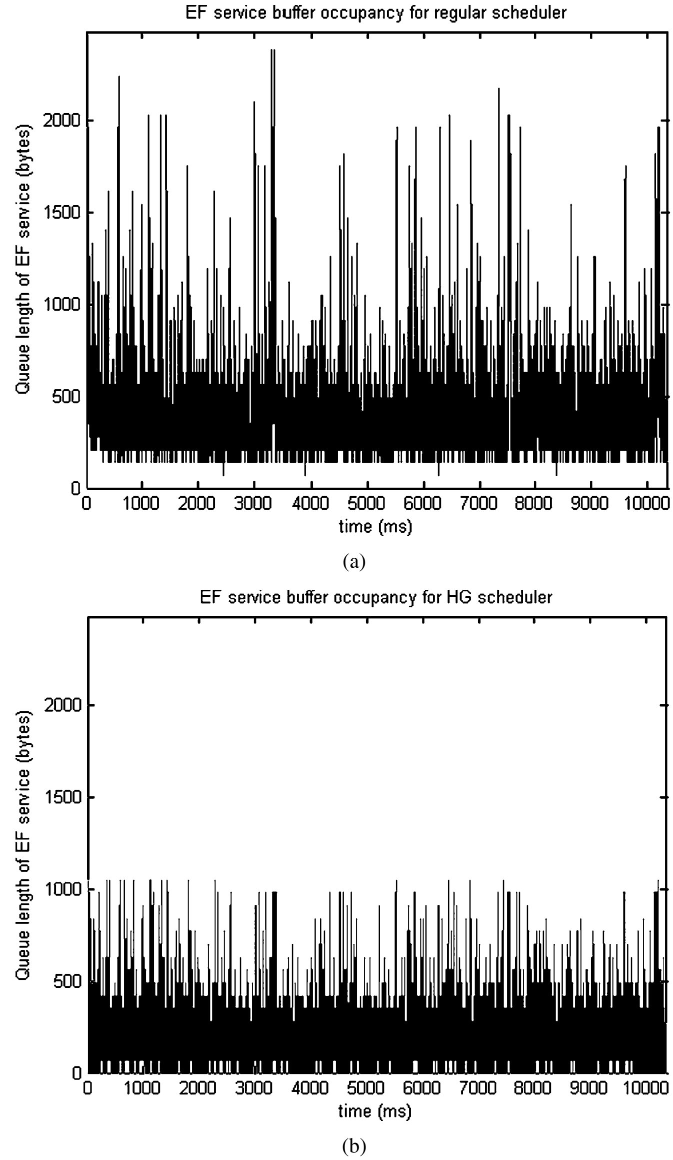

Fig. 7. (a) EF queue length (regular scheduler). (b) EF queue length (HG scheduler).

respectively. Fig. 6 expands the central fraction in the ACF of $\overline{V}$ for regular EPON scheduler and HG scheduler. The stronger positive values in the ACF for regular EPON scheduler explicitly verified the aforementioned continuity in the corresponding EF delay sequence. While the EF delay sequence for HG scheduler has more opportunities to cross its mean, i.e., more oscillation. Note this stronger oscillation does not constitute more severe A-WJ.

The improved E-WJ for EF service in HG scheduler can be attributed to its proper separation between EF and AF/BE transmissions. Specifically, in regular EPON scheduler, the EF packet transmission start time of each ONU can fluctuate widely within the DBA-CL due to random behaviors of interfering AF and BE streams at other ONUs. However, in HG scheduler this transmission start time is only affected by the total length of the DBA-CL and not by individual variations of other specific ONUs. Overall, this separation notably mitigates the above "floating" start time effect and leads to reduced E-WJ for EF services.

D. ONU Buffer Occupancy

The primary beneficiary of the above E-WJ improvement in HG scheduler may be the ONU buffer occupancy, or packet loss ratio. It is understandable that a larger E-WJ incurs a larger number of EF packets buffered in the queue, which requires more buffering space. Thus with a limited physical buffering space at each ONU, severely fluctuated high priority (EF) queue length, which is enforced by severely fluctuated E-WJ, will certainly result in a greater packet loss ratio for low priority services. Because packets may be dropped by serious transient buffer overflow or constant bandwidth starvation, the buffering space was assumed infinite in our simulation in order to examine the performance of different schedulers regarding to buffer occupancy. Fig. 7(a) and (b) sketch the queue-length fluctuation of EF service for regular EPON scheduler and HG scheduler. More fluctuated EF queue length for regular EPON scheduler in the figure matches its severe E-WJ performance very well. Also, the static strip at the bottom of Fig. 7(a) reflects the improvement of GBR fashion over GAR fashion respecting to buffer occupying process.

V. Conclusion

The issue of QoS support in EPON is of growing importance to network designers. In this paper, we propose a novel hybrid granting (HG) protocol to specifically improve the performance of EF services in EPON. The scheme properly separates EF and AF/BE bandwidth allocation and uses a grant-before-report (GBR) strategy to enforce EF guarantees. Furthermore, we also table a pair of "phase-shifted" DBA and MPCP cycle schemes to reduce updating latencies for ONU buffer occupancy information at the OLT. Besides average packet delay, we develop comprehensive analysis on the refined performance regarding to different system parameters. Overall, simulation results confirm that the HG scheduler yields notable improvements in average packet delay, delay variation, buffering space utilization performance of EF services without degrading QoS support for other

AF, and BE services. Nonetheless, the performances of AF and BE services are yet to be further improved and QoS to be better delivered. Since these nonuniform traffics with varying packet size constitute the dominant bit flow in EPON, research effort in this aspect may be rewarded with potential breakthrough of system performance, such as service delay, delay jitter, traffic rate control, packet loss ratio, etc. These are also the central engagements of our future work.

REFERENCES

[1] K. G. Koffman and A. Odlyzko, "Internet growth: is there a "Moore's Law" for data traffic?," in *Handbook of Massive Data Sets*. Norwell, MA: Kluwer, 2001.

[2] N. Ghani, A. Shami, C. Assi, and M. Y. A. Raja, "Intra-ONU bandwidth scheduling in ethernet passive optical networks," *IEEE Commun. Lett.*, vol. 8, no. 11, pp. 683–686, Nov. 2004.

[3] International Engineering Consortium (IEC). [Online]Available: http://www.iec.org

[4] G. Kramer and G. Pesavento, "Ethernet passive optical network (EPON): Building a next-generation optical access network," *IEEE Commun. Mag.*, vol. 40, pp. 66–73, Feb. 2002.

[5] C. Assi *et al.*, "Dynamic bandwidth allocation for quality-of-service over ethernet PONs," *IEEE J. Sel. Areas Commun.*, vol. 21, pp. 1467–1477, Nov. 2003.

[6] G. Kramer, B. Mukherjee, and G. Pesavento, "Interleaved polling with adaptive cycle time (IPACT): a dynamic bandwidth distribution scheme in an optical access network," *IEEE Commun. Mag.*, vol. 40, pp. 74–80, Feb. 2002.

[7] H. Jeong, D. McNickle, and K. Pawlikowski, "Generation of self-similar time series for simulation studies of telecommunication networks," in *Proc. 1st Western Pacific/3rd Australia-Japan Workshop on Stochastic Models*, Christchurch, New Zealand, September 1999.

[8] IEEE 802.3ah Task Force Home Page. [Online]Available: http://www.ieee802.org/3/efm

[9] NSF Workshop Report. (2003, Oct.) Residential Broad-band Revisited: Research Challenges in Residential Networks, Broad-band Access, and Applications. [Online]Available: http://cairo.cs.uiuc.edu/nsfbroadband

[10] A. Shami, X. Bai, C. Assi, and N. Ghani, "Quality of service in two-stage ethernet passive optical access networks," in *Proc. IEEE ICCCN*, Chicago, IL, Oct. 2004.

[11] M. Ma, Y. Zhu, and T. Cheng, "A bandwidth guaranteed polling MAC protocol for ethernet passive optical networks," in *Proc. IEEE INFOCOM*, San Francisco, CA, Apr. 2003.

[12] G. Kramer. On Configuring Logical Links in EPON. [Online]Available: http://www.ieeecommunities.org/epon

[13] S. Ramabhadran and J. Pasquale, "Stratified round robin: a low complexity packet scheduler with bandwidth fairness and bounded delay," in *Proc. IEEE SIGCOMM*, Karlsruhe, Germany, Aug. 2003.

[14] H. Lee, T. Yoo, J. Moon, and H. Lee, "A two-step scheduling algorithm to support dual bandwidth allocation policies in an ethernet passive optical network," *ETRI Journal*, vol. 26, pp. 185–188, Apr. 2004.

[15] (2003, Oct.) NSF Workshop on Research Challenges in Broad-band Residential. [Online]Available: http://cairo.cs.uiuc.edu/nsfbroadband/

[16] M. Subramanian *et al.*, "QoS and bandwidth management in broad-band cable access network," *Comput. Networks*, vol. 43, pp. 59–73, 2003.

[17] F. An *et al.*, "A new dynamic bandwidth allocation protocol with quality of service in ethernet-based passive optical networks," in *Proc. International Conference on Wireless and Optical Communication (WOC 2003)*, Banff, Canada, Jul. 2003.

Abdallah Shami (M'03) received the B.E. degree in electrical and computer engineering from the Lebanese University, Beirut, Lebanon, in 1997, the M.S. degree in computer and electrical engineering from the Saint Joseph/Lebanese University, Beirut, Lebanon, in 1998, and the Ph.D. degree in electrical engineering from the graduate school and university center of the City University of New York, New York, in September 2002.

At the City University of New York he held the Irving Hochberg Dissertation Fellowship and a GTF teaching fellowship. In September 2002, he joined the Department of Electrical Engineering at Lakehead University, ON, Canada, as an Assistant Professor. Since July 2004, he has been with the University of Western Ontario, London, ON, Canada, where he is currently an Assistant Professor in the Department of Electrical and Computer Engineering. His current research interests are in the area of data/optical networking, EPON, optical packet switching, Internet protocol over wavelength-division multiplexing, and software tools.

Xiaofeng Bai (S'04) received the B.Eng. degree in communication engineering from Shandong University, Shandong, China, in 1997. He is currently working toward the M.E.Sc. degree in Electrical and Computer Engineering at the University of Western Ontario, London, ON, Canada.

His research interests are in the area of Quality of Service in broad-band access networks, medium access control protocol design, and network traffic modeling.

Chadi M. Assi (M'03) received the Bachelor's degree in engineering from the Lebanese University, Beirut, Lebanon, in 1997 and the Ph.D. degree from the graduate center of the City University of New York, New York, in April 2003.

He was a Visiting Researcher at Nokia Research Center, Boston, MA, from September 2002 until August 2003, working on quality of service in optical access networks. He joined the Concordia Institue for Information Systems Engineering (CIISE) of Concordia University, Montreal, QC, Canada, in August 2003 as an Assistant Professor. His research interests are in the areas of optical networks control, provisioning and restoration, and ad hoc networks.

Dr. Assi received the Mina Rees Dissertation Award from the City University of New York in August 2002 for his research on wavelength-division-multiplexing optical networks.

Nasir Ghani (S'96–M'97–SM'01) has a wide range of experience in a broad range of networking-related fields and has held senior technical positions at Sorrento Networks, Nokia Research, IBM, and Motorola Codex. Currently, he is with the Electrical and Computer Engineering Department at Tennessee Tech University, Cookeville, TN, where his research focus includes metro/edge networks, access networks, optical networks, Internet protocol routing, and network survivability. He has published more than 40 refereed journal and conference papers, several book chapters, and has two patents granted.

Dr. Ghani was full conference Co-Chair for The International Society for Optical Engineers (SPIE) OPTICOMM 2003 and has served as a program committee member for numerous other IEEE, SPIE, Association for Computing Machinery (ACM), and International Electrotechnical Commission (IEC) conferences. Currently, he is also a Co-Chair for the Optical Networking Symposium for IEEE ICC 2006 and 2007 and also IEEE GLOBECOM 2006. Dr. Ghani is an Associate Editor for the IEEE COMMUNICATIONS LETTERS and has also served as a Guest Editor for IEEE NETWORK.