Performance evaluation of GPON vs EPON for multi-service access

T. Orphanoudakis\textsuperscript{1,*,†}, H.-C. Leligou\textsuperscript{2}, E. Kosmatos\textsuperscript{2} and J. D. Angelopoulos\textsuperscript{2}

\textsuperscript{1}University of Peloponnese, Karaiskaki str., 22100 Tripoli, Greece
\textsuperscript{2}National Technical University of Athens, Heroon Polytexneiou 9, Zographou Campus, Athens, Greece

SUMMARY

Recently both ITU and IEEE have standardized solutions for passive optical networks (PONs) operating at gigabit per second line rates and optimized for the transport of packet-based traffic to improve the efficiency of previously standardized broadband PONs, which used the ATM cell as the data transport unit. The efficiency and performance of PON systems depend on the transmission convergence layer and mainly on the implemented medium access protocol. Although the latter is not part of the standards and left to the implementer, the standards describe a set of control fields that constitute the tool-set for the media access control (MAC) operation. Though starting from a common and quite obvious basis, the two standards present significant differences with the legacy of Ethernet marking the IEEE approach, while the emphasis of ITU is on demanding services. In this paper we compare the efficiency and performance of the two systems assuming the implementation of as close as possible MAC protocols. The target is twofold: assess and compare the traffic handling potential of each of the two standards and identify the range of applications they can support. Useful insight can also be gained to the MAC tools that could be designed into the next generation extra large WDM PONs. Copyright \textcopyright 2008 John Wiley \& Sons, Ltd.

Received 4 December 2007; Revised 13 May 2008; Accepted 13 August 2008

KEY WORDS: passive optical networks; EPON; GPON; multi-service access networks; performance evaluation; TDMA multiplexing; quality of service

1. INTRODUCTION

In response to the steadily increasing demand for bandwidth and networking services for residential users as well as enterprise customers, passive optical networks (PONs) have emerged as a promising access technology that offers flexibility, broad area coverage and cost-effective sharing of the
expensive optical links compared with the conventional point-to-point (P2P) transport solutions. In addition, they inherently concentrate traffic and greatly reduce the number of input ports in the access multiplexer, both important for the cost-sensitive residential access market. Owing to these advantages, PONs have generated during the last decade substantial commercial activity also reflected in the study of several standardization bodies. Since the initial standardization of ATM-based PONs (APONs or alternatively named in ITU-T G.983.1 standard [1] Broadband PONs—BPONs) newer standards support multi-gigabit rates and better adapt to the packet-based Internet applications. In January 2003, the GPON (Gigabit PON) standards were ratified by ITU-T and were included in the G.984.x series of ITU-T Recommendations [2, 3]. Driven by a closed group of worldwide system vendors and national telecom operators, they are designed to support a mix of TDM, ATM and packet-based services, reaching symmetrical transmission rates of up to 1.244 or 2.488 Gb/s. At the same time IEEE, through the activities of Ethernet in the first mile (EFM) 802.3ah group, has standardized a Gigabit Ethernet-friendly technology [4] called Ethernet PON (EPON) with the objective to leverage the great success of Ethernet as a LAN technology and exploit the economies of scale that the dominance of Ethernet has generated. The objective of EFM efforts has been to combine a minimal set of extensions to the IEEE 802.3 media access control (MAC) and MAC control sub-layers with a family of physical (PHY) layers. In addition, a mechanism for network operations, administration and maintenance (OAM) is included to facilitate network operation and troubleshooting. Since the development of the two standards was progressing in approximately the same time there was also a liaison between the 802.3ah EFM Task Force and the ITU-T SG 15 group with the goal to align specifications to the extent possible especially for the PON transceiver equipment, which is considered to have the highest impact on the overall system cost. Despite these efforts the two standards diverged significantly due to the different requirements and the interests of the participants in each group. The similarities between the two are limited to a common wavelength plan and minimum and maximum power setting for the transceivers. The two standards differ substantially in the properties of the physical medium-dependent layer, transmission convergence (TC) layer, OAM capabilities and the MAC layer, which is the focus of this study.

Although PONs can achieve economical deployment and operation, which are a major concern to operators and service providers, high and fair resource utilization is equally important. To this end the bandwidth allocation mechanism, guided by the MAC protocol, should be designed so as to optimally trade off efficiency with specific performance guarantees enabling applications with different requirements. Support of different quality of service (QoS) levels is also crucial for the success of this technology, since it is tightly associated with the offer of new services like triple-play (real-time multimedia content delivery, telephony and data) to residential users. It is also needed to provide low-cost alternatives to businesses, obviating resort to lower rate (622 Mbps, 1.24 Gbps) legacy SDH networks for backhauling between fixed locations (e.g. between buildings, large-scale campuses or even cellular network base stations). Owing to the multiple access nature of PONs in the upstream direction, the performance of a PON in terms of delay, delay variation and throughput strongly depends on the upstream bandwidth allocation function of the MAC residing at the optical line termination (OLT). The evaluation of access mechanisms in the shared upstream channel of PONs has attracted special attention lately, due to the increased interest for new services, like peer-to-peer streaming, data-center and transparent LAN interconnection, mobile backhauling, etc, which present significantly higher requirements than asymmetric interactive Internet services. However, no direct
performance comparisons of the upstream traffic multiplexing (the downstream direction is similar and presents no challenge) between EPON and GPON systems have been published, although it is of significant interest for operators as interactive multimedia services gain momentum. Therefore, in this paper we proceed to compare these two technologies focusing mostly on traffic performance and how this is affected by the MAC layer design choices. We first provide in the following section an overview of the technologies and operational parameters adopted by each standard. Then in Section 3 we describe the overall service policies in order to define QoS metrics that should be taken into account at the MAC layer while in Section 4 we discuss the details of the implementation of suitable MAC protocols based on time division multiple access (TDMA) for each case. On the basis of computer simulation of the MAC implementations, we quantify and compare the performance results in Section 5 providing concluding remarks in Section 6.

2. STANDARDIZED PON OPERATIONAL PARAMETERS UNDER GPON AND EPON

Before attacking the MAC functionality, it is expedient to give a brief comparative outline of the physical medium parts of the standards and how they affect PON operation as well as present the available messages and fields related to the MAC functionality. The tree-shaped topology of PONs is the result of passive splitting of fibers extending from the OLT that represents the ‘root’ and reaching multiple optical network termination units (ONUs) that represent the ‘leafs’ of the tree. This architecture also results in the broadcast operation in the downstream direction, while in the upstream channel an aggregate data flow is generated by means of burst transmissions from the active ONUs in a TDMA fashion. The activation of each ONU’s transmitter and window of operation is controlled by the OLT.

2.1. Physical layer

At the physical layer both standards adopted the same wavelength plan (1310/1490 nm wavelength band for the upstream/downstream direction, respectively) and transceivers with similar characteristics. GPON, however, adopted an additional optical distribution network class (ODN class C) resulting in 5 dB higher maximum signal losses. This results in the increased largest split/maximum range capabilities of GPON. Additionally, the upstream rate and power settings in GPON are configurable providing higher flexibility. The higher line rates impose more stringent requirements on the burst-mode receiver, which has to perform amplitude and phase recovery by means of the burst preamble preceding each upstream burst. EPON adopted a fixed line rate and configurable preamble, while in GPON a fixed-sized preamble is used for each upstream rate, which is minimized to improve upstream bandwidth utilization. The bandwidth efficiency of GPON is also significantly increased by the adoption of non-return-to-zero (NRZ) coding whereas EPON can only provide a net 1000 Gb/s data rate with respect to the 1250 Gb/s line rate due to the adoption of 8b/10b encoding, chosen to simplify the receiver design. A comparative reference of the basic features of the two standards is provided in Table I.

2.2. Transmission convergence (TC) layer

The purpose of the TC layer is to reconcile the point-to-multi-point (P2MP) shared physical link of the PON tree with the higher layer protocols, which have been designed for use in P2P links.
### Table I. Summary of standardized GPON and EPON operational parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>EPON</th>
<th>GPON</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optical distribution network (ODN) classes</td>
<td>A, B</td>
<td>A, B, C</td>
</tr>
<tr>
<td>Max./differential reach</td>
<td>20 km/20 km</td>
<td>60 km/20 km</td>
</tr>
<tr>
<td>Min./Max. split (transmission convergence layer)</td>
<td>16/N/A</td>
<td>64/128</td>
</tr>
<tr>
<td>Line rate (up/down)</td>
<td>1250/1250 Mb/s</td>
<td>155.52-622.08-1244.16-2488.32/Mb/s</td>
</tr>
<tr>
<td>Coding</td>
<td>8b/10b</td>
<td>NRZ (+scrambling)</td>
</tr>
<tr>
<td>Upstream data rate</td>
<td>1000 Mb/s</td>
<td>Equal to up. line rate</td>
</tr>
<tr>
<td>Opt. loss</td>
<td>15/20 dB</td>
<td>15/20/25 dB</td>
</tr>
<tr>
<td>Upstream burst timing</td>
<td>Laser turn on/off: 512 ns (max), AGC setting &amp; CDR lock 400 ns (max)</td>
<td>Guard: 25.6 ns, Preamble: 35.2 ns (typical), Delimiter: 16.0 ns (typical)</td>
</tr>
</tbody>
</table>

It provides for (de)multiplexing frames (from) to the MAC layer based on a unique identifier per ONU indicated inside the frame header. Frames are filtered based on this identifier and classified into separate queues emulating discrete logical P2P links between the OLT and the multiple ONUs of the different customer interfaces. This allows reaping the economic advantages of sharing the expensive physical resources, while keeping the appearance of a bunch of dedicated links with ordinary protocols above the TC protocol.

In EPON the TC is carried out very simply by a single addition to the common Ethernet framing: the replacement of the 8-byte Ethernet preamble by the so-called logical link identifier (LLID), which is used in EPON to emulate a P2P operation. By means of the LLID, EPON ONUs filter the received packets, forwarding to the upper layers only packets carrying the LLID value assigned to each one of them by the OLT during the registration phase. IEEE 802.3ah specified a P2P emulation (convergence) sub-layer as part of the Ethernet stack reconciliation sub-layer resulting in the equivalent operation (at the MAC and higher layers) of a collection of P2P links over the actual P2MP network. It is also worth stressing that only integral Ethernet frames are exchanged between the OLT and each ONU.

The GPON TC specification (G.984.3) is more ambitious and can accommodate any kind of higher-layer frames, packets, ATM cells as well as native TDM services. GPON adopts a fixed duration frame structure (of 125μs), which can carry TDM traffic, ATM cells and variable length frames/packets in any proportion, as shown in Figure 1. In the downstream, each frame starts with the physical layer control block (PCBd in the figure). Any client traffic other than TDM or ATM (e.g. Ethernet frames or IP packets) is encapsulated in GEM frames using the so-called GPON encapsulation method (GEM). This protocol allows for segmentation and reassembly of Ethernet frames (something not permitted in EPON) allowing for higher-frame utilization. The only required overhead field in each upstream burst transmission is the so-called upstream physical layer overhead (PLOu), which together with a small guard-band provides the means for synchronization, error-free reception and frame delineation among different ONU bursts. Optionally, additional header fields may be included (dictated by the OLT in the upstream bandwidth map) that facilitate physical layer OAM functions (PLOAmu 13-byte messages as defined in G.983.1), power level sequencing (120-byte PLSu field) and dynamic bandwidth reporting (2–5 byte DBRu field).
2.3. MAC layer

Both standards allow for the implementation of a system-specific dynamic bandwidth allocation (DBA) mechanism: ONUs may dynamically request transport capacity by sending their queue length reports and the OLT schedules the time slots for ONUs based on their requests; thus, adapting to traffic fluctuations. In both cases only the type of messages that should be exchanged during operation are defined, in order to guarantee interoperability, without specifying exact algorithms that can be employed especially for bandwidth allocation. This is left open to the vendors and network providers to handle according to their specific requirements.

In the EPON case, the multi-point control protocol (MPCP) is used to arbitrate the upstream access. MPCP uses two types of messages (encapsulated in Ethernet frames) during normal operation for arbitration of packet transmissions: the REPORT message used by an ONU to report the status of its queues to the OLT (up to eight reported in a single message) and the GATE messages issued by the OLT and indicating to the ONUs when and for how long they are allowed to transmit in the upstream channel. Each GATE message can support up to four transmission grants targeting individual service entities within the same ONU (i.e. data queues). In the upstream, the granted ONU transmits (possibly) multiple Ethernet frames—as many integral packets as can fit into the allocated transmission slot, since fragmentation is not allowed—from one or more queues preceded by the indispensable physical layer overhead. It may also transmit REPORT messages in order to request additional grants. In EPONs, the traffic streams arriving at the ONUs from the customer premises are kept in queues. In compliance to the 802.1p prioritization scheme, it is possible to inject the traffic in up to eight (8) logically separate, possibly prioritized, queues holding Ethernet frames, with different QoS requirements, to allow for the enforcement of different service mechanisms.

In GPON, the MAC overhead is limited to a small encapsulation header attached in front of each upstream and downstream frame. In the upstream, the DBA report (DBRu) is used (Figure 1) to signal a request to the MAC controller for an allocation of upstream transmission with as many bytes. Based on this information, the OLT defines the upstream time allocations in the upstream bandwidth map field of the downstream frame. The bandwidth map contains a variable number of access structures defining which ONU and for how long it is allowed to transmit.


Figure 1. The downstream and upstream frame structure in GPON.
The transmissions are assigned to each queue, uniquely identified by the so-called Alloc-ID field. Each queue (Alloc-ID) can aggregate streams per traffic class or be used for finer flow levels depending on the implementation. Further multiplexing of traffic is possible using the Port-ID field of GEM frames (just as the virtual path/virtual circuit, VP/VC fields are used in the ATM part). A GPON can support almost 4000 Alloc-IDs in total with the 12-bit-long relevant field, but note that the first 254 Alloc-IDs are reserved as ONU identifiers, which are also used during the setup/activation of a (new) ONU. Thus, multiple queues can be implemented in each ONU, limited in theory only by the address space of the 12-bit Alloc-ID field, but in practice by cost, complexity and performance, which suffer as the number of managed queues rises.

3. DESIGN CONSIDERATIONS FOR EFFICIENT SUPPORT OF DYNAMICALLY EVOLVING SERVICE BUNDLES

The objective of this paper is to prepare the grounds for a fair comparison of the two protocols under common scenarios as well as to derive relevant metrics for comparison and quantitative performance results. In this section, we classify common requirements and considerations for resource allocation and specify the operational parameters selected for the implementation and simulation of the MAC protocol in each enabling a direct evaluation and comparison.

Both EPON and GPON have been designed to fit all FTTx solutions and all service categories. From the traffic engineering point of view, the difference among the various cases is the number of input streams sharing the upstream bandwidth and the QoS requirements of the users in combination to the tariffs they are willing to pay. The most demanding real-time interactive applications require a one-way delay of 1.5 ms while many studies have verified that a maximum of 2 ms transmission cycle time is an acceptable value [4, 5]. To handle QoS with lower complexity and cost, services are grouped into behavior aggregates (classes) with a similar set of requirements providing scalability and flexibility (also following the 802.1P approach).

In a previous study, the authors in [3, 6] have presented a GPON-aligned MAC protocol adopting the concept of four discrete queues for serving PON traffic aggregates corresponding to the T-CONTs, which remained as a legacy from the BPON DBA specification G.983.4 [7], while in [8] they presented a 802.3ah compliant DBA scheme for EPONs, which also used four discrete service policies based on the same traffic aggregation requirements with four queues per ONU. It was shown that it can efficiently support any QoS requirement using the MAC control tools provided by the EPON standard. Building on the ground laid by the above study, we also use the four priority class queues in this comparison study. The QoS requirements assumed for each of them as well as example applications are shown in Table II. This table also indicates the service guarantees that the MAC protocol offers.

All PON MAC mechanisms in their quest for fairness have to visit all active ONUs periodically giving rise to the concept of the scheduling period $D_m$, which is the time interval until the scheduling goes back to the same ONU. This period is central to the way PONs allocate upstream transmissions even when this is not a fixed value but varies slowly in response to the changing ONU activity. The scheduler will decide the allocations after collecting the newer information about packet arrivals, which is also completed in one such period but also a new round of allocations is already decided for a complete round and not on a packet-by-packet basis as in centralized multiplexers (because it would not cover fairly all ONUs). This $D_m$ parameter that must be kept low for both GPON and EPON is the starting point for understanding the DBA operation and
more details can be found in an earlier study [3, 8–11]. While DBA has been well understood from studies in BPONs [12–14], EPONs triggered new research efforts and approaches tailored to their specific needs. A review of the most well-known mechanisms is provided in [15]. Studying the relevant literature [9–11, 16–18], it has been recognized that strict isolation between elastic and real time traffic is required to provide performance guarantees. In addition, bandwidth pre-allocation (unsolicited allocations based on static reservations) for high-priority traffic is the only means to provide acceptable access delay and combat the barrier arising from the large round-trip delays of request-allocation schemes.

In line with this strategy, in the performance comparison presented in Section 5, the MAC protocol serves CoS1 queues periodically, allocating an adequate number of unsolicited grants in every scheduling cycle $D_m$. This way the operator actually guarantees service to a contracted peak rate $R_{p1}$ and a strict delay bound $D_{\text{max}}$, which can be derived as a function of $D_m$. The scheduling period $D_m$ is used to calculate the bytes to be allocated to each queue to achieve the desired service rate. Hence, considering the case where unsolicited grants cover the sustainable rate $R_{s2}$ of the second class, the total number of unsolicited grants for the $i$th ONU (UG$_i$) in bytes is expressed as follows: $\text{UG}_i = (R_{p1i} + R_{s2i}) \times D_m$ (rates expressed in bytes/s). The remaining unallocated part of each scheduling period $D_m$ is distributed dynamically in a weighted manner as described in [8] and a service weight $w_i$ can be used to enforce proportional sharing of the upstream transmission window among ONUs to guarantee the portion reserved for CoS3 queues. Since this represents an initial allocation of bandwidth based on the assumption that all ONUs appear backlogged during an interval $D_m$, the algorithm dynamically redistributes transmission intervals based on the actual still unserved requests from CoS3 and CoS4 queues following the max–min fair sharing algorithm. Finally, CoS4 is served as best effort, i.e. whenever unallocated slots exist. Obviously, an appropriate service configuration phase should be implemented during which the negotiation of the above parameters, acceptance control and resource reservation will be performed.

### 4. THE INTRICACIES OF TDMA OPERATION IN PACKET-BASED PONS

As a basis for our comparison we take the EPON compliant MAC protocol proposed by the authors in [8], which extends the approach investigated in [17] to collectively handle four allocation strategies with enhanced bandwidth efficiency and optimized scheduling of granted upstream
transmission windows. This is totally aligned with the GPON MAC protocol described in [6] with respect to the service class definition and the differences in the time scheduling of upstream allocations are only the ones arising from the different MAC mechanisms implied in each standard. The GPON protocol allows scheduling of upstream transmissions per ONU queue (Alloc-ID) at arbitrarily small scheduling periods (integral multiples of the 125μs enabling even the support of TDM services). Therefore, in order to bring the two concepts close enough to allow for a fair comparison of their MAC (GPON as described in [6] and EPON in [8]), the $D_{m}$ scheduling period will be chosen small enough to bring the source granularity of EPON close to that of GPON.

The impact of the scheduling period $D_{m}$ stems from the two fundamental differences between the two technologies as also identified in [19]. The first is the time consumed for the transmission of the preamble, the delimiter plus the required guard time (denoted as $T_{pre}$ in [8]), which introduces larger overhead in EPONs when multiple burst transmissions are attempted by one ONU interleaved with transmissions from other ONUs within the same scheduling period. The second is the segmentation functionality provided by the GEM encapsulation in GPON only, which allows for complete utilization of each upstream allocation independently of its length, whereas in EPONs small upstream slot allocations increase the probability for waste of bandwidth. Therefore, EPON has been designed assuming a large enough scheduling period $D_{m}$, otherwise bandwidth efficiency suffers. However, we will show that this can only be achieved by penalizing delay as well as jitter performance.

The overhead introduced by different scheduling alternatives is graphically explained in Figure 2 where upstream burst allocations and actual data transmissions (covering a time window $D_{m}$) from different ONUs and CoS queues are shown. In Figure 2(a), a possible scheduling of grants accommodating the requests collected in earlier polling cycles is provided. The order of the allocations (which target individual queues) affects the achieved efficiency. In Figure 2(a), higher-priority allocations precede lower-priority allocations. Taking into account that the duration of the CoS 1 allocations is fixed, the position of the two higher-priority allocations in every scheduling cycle $D_{m}$ are fixed, forming sub-frames per CoS (a scheduling discipline also selected in [17, 18]). In EPONs, part of the allocations may be wasted since the exact ONU queue occupancy and packet delineation are not known by the grant scheduler at the OLT. It is then very likely that the leftover time at the end of the allocated window does not match the length of the next packet in the first-in–first-out queue. This phenomenon called unused slot remainder (USR) is also shown in Figure 2 only for the second ONU as an example. Figure 2(b) shows the alternative of serving all CoS queues from the same ONU before allocating slots to other ONUs creating sub-frames.
per ONU. Obviously, this schedule introduces less physical layer overheads and the efficiency improvement depends on the scheduling period $D_m$ as well as on the number of supported queues and ONUs. The longer the $D_m$, the higher the efficiency achieved (minimizing physical layer overheads). However, assuming that this will also be the service period for all services (including delay sensitive ones), the scheduling period also directly affects (and is close to) the delay observed by CBR-like services. To achieve a delay guarantee of 1.5 ms for voice services \([4, 5]\), a scheduling period of about equal duration should be selected.

5. PERFORMANCE EVALUATION

We now come to the main objective of this study, which is to compare the efficiency of EPONs and GPONs as multi-service broadband access systems when operating under the full extent of their dynamic grant scheduling capabilities with which the relevant standards have endowed them. The assessment was based on the following performance metrics: (i) average delay per class of service; (ii) delay variation for real-time services and (iii) bandwidth utilization of the shared upstream channel. A simulation model was developed using the OPNET simulator including 16 ONUs, each equipped with four CoS queues for both systems.

The offered load is shared uniformly among all ONUs. Several scenario sets were carried out under a varied set of parameters to cover many operational conditions. The traffic mix included on average 10% high-priority traffic. This service class is expected to serve mostly narrowband traffic also associated with a higher tariff limiting its contribution to the overall network load as argued in \([17]\). The second, third and fourth priority were injecting at 15, 20 and 55% of the total load, respectively \([20]\). Using the same split allows for direct comparison with the results presented in Reference \([8]\). High-priority sources were of constant bit rate generating short fixed-size packets periodically (a model compatible with voice traffic), whereas the sources for the other three types of traffic were of the ON–OFF type (modeling self-similar Internet traffic), with different burstiness factors, namely 2, 5 and 5 for second, third and fourth priority, respectively. Given the choice of Ethernet as the dominant client protocol, the widely used tri-modal distribution was adopted for the packet size. This is based on extensive actual measurements, which showed it to be a quite good approximation of IP applications originating in Ethernet networks. It consists of packet sizes of 64, 500, 1500 bytes appearing with probability of 0.6, 0.2 and 0.2, respectively, according to \([21]\). The 1500 byte size is the Ethernet limit, whereas 60% of packets are short ACK packets with the rest approximated by their 500 byte average.

Typical overhead values were selected according to the supported MAC operational parameters by each standard. The guard-band and physical layer overhead (i.e. $T_{\text{pre}}$) were assumed equivalent to 1 $\mu$s for EPON, while for GPON (see also Table I) this is an order of magnitude lower (GPON operates with only 15 bytes overhead for burst ONU transmission). As discussed in the previous section, a final important parameter is the scheduling period $D_m$. In order to proceed to the comparison of these two technologies, we exploit the fact that the GPON MAC protocol presented in \([6]\) is functionally equivalent to the EPON protocol presented in \([8]\) assuming that the $D_m$ window adopted in EPON can be selected arbitrarily small (down to the 125 $\mu$s granularity of GPON). The EPON protocol we evaluate employs the most commonly proposed in the literature \([8–11, 16–18]\) MAC mechanisms but, as we will show through our simulation results, requires rather large $D_m$ values (1 and 2 ms are the values most frequently assumed in the literature) in order to achieve adequate bandwidth efficiency trading off packet delay and jitter performance.
The GPON MAC though can operate for smaller $D_m$ values improving delay performance without the high cost on bandwidth loss of EPON. We will demonstrate these performance trade-offs by varying the $D_m$ period from sub-millisecond values (750 µs) up to 2 ms.

5.1. Delay performance

The first conclusion regarding the efficiency of the EPON protocol and the limitations in the achievable throughput can be drawn from the study of the average access delay as a function of the offered load for each of the four classes (cumulative for all ONUs) shown in Figure 3. The implemented service differentiation and DBA algorithm achieves perfect isolation and guards high-priority classes as expected. First and second priority remain stable even at full load, since for the simulated scenario they represent the 10 and 15% of the offered load, respectively, and their service has been guaranteed through pre-programmed grants. The average delay though of the three highest-priority classes with QoS guarantees is inversely proportional to the scheduling period $D_m$, as expected. However, when improving delay by decreasing $D_m$, channel utilization decreases as shown by the results for CoS4. Note that this class serves best-effort traffic allocating the remaining bandwidth up to the available capacity. Congestion for this class appears for traffic loads above 90% of the channel bandwidth, when the 2 ms period $D_m$ is used, but drops inversely proportionally to the value of $D_m$, to values as low as 70% for sub-millisecond period $D_m$ values. Consequently, improving average delay for CoS1 and CoS2 traffic (real-time applications) so that the EPON MAC could offer similar to the GPON MAC delay bounds [6] comes at the cost of overall channel utilization for reasons that will also be clarified in Section 5.3.

5.2. Jitter performance

Apart from the average delay, it is equally important to observe the deviation from an ideal CBR service for CoS1, i.e. the probability density function (PDF) of the delay, which quantifies packet delay variations (i.e. jitter). These curves (cumulative among all ONUs) are depicted in Figure 4 for CoS1, since as shown in Table II this class is targeting real-time applications and jitter performance in this case is of interest (the relative jitter performance for CoS2–CoS4 is discussed in [8]). Results for two indicative total load values (a light loading of 40% and a higher loading...
near system saturation of 70%) are shown in Figure 4. The implemented MAC protocol limits the maximum delay, which rarely exceeds the $D_m$ bound. A nearly uniform distribution of delay values across the whole spectrum of values up to $D_m$ is observed in the PDF, reflecting the fact that packet arrivals can of course occur at any time of the scheduling window designated by $D_m$ in the EPON MAC. This effect is discussed in detail in [8, 17]. To illustrate the impact of the MAC algorithm on jitter performance, we show in Figure 5 the PDF of the cumulative (across all ONU) packet delay variation again only for the real-time class CoS1, where jitter minimization is desirable, expressed as the one-way-ipv (interpacket delay variation metric as defined in [22]). The optimal performance would lead to a jitter PDF curve sharply centered on the axis origin (i.e. approaching zero delay variation). Deviations from the optimal performance can be observed in Figure 5 (smaller peaks of the curve around specific values), which are the result of unavoidable inaccuracies introduced in practical implementations by pointer computation round-offs even at relatively light loading (see detailed discussions in [8, 17]). These deviations in EPONs can be brought down to negligible values by reducing the scheduling interval $D_m$, albeit at the penalty of badly reduced bandwidth utilization as shown in the previous section. To achieve this lower protocol overhead and more flexible mechanisms allowing frequent polling of requests (i.e. make the OLT aware of new packet arrivals at ONU queues earlier) should be implemented. This led to the adoption in GPON of the fine granularity with minimum cost (protocol overhead as will also be discussed in the following and final section), so that each ONU can be granted as frequently as required in multiples of 125μs.

5.3. Bandwidth efficiency

As regards system efficiency, we analyze in this section the GPON features that lead to superior performance. Apart from the inefficiency introduced by the 8/10 bit coding adopted in EPONs, which limits the available upstream rate to 1 Gb/s instead of 1.24 Gb/s in GPON, their efficiency when transporting Ethernet traffic is also better in terms of allocated slot utilization due to GEM encapsulation, which combats the USR effect as explained earlier. Although GEM encapsulation introduces a header of 5 bytes on each upstream frame or segment of frame, it allows segmentation and reassembly of Ethernet frames achieving a nearly perfect fit. Hence, while in GPON the
inefficiency stems from the GEM overhead, in EPON inefficiency is introduced by the USR, the longer physical layer overhead and the longer reporting message (3 bytes per queue in GPON vs 64 bytes for all eight queues in EPON).

The gain of GPON is quantified in Figure 6, which shows the extra payload bytes that a GPON ONU would transmit compared with an EPON ONU for the same sequence of upstream time allocations as a percentage of the total transmitted payload. This gain is due to the complete suppression of the USR by GEM, which turns out to save more bandwidth than lost to the 5-byte overhead per packet, as well as the lower protocol overheads of GPON. This gain slightly decreases as system load rises from 35 toward 65% because longer upstream windows are granted to the ONUs; hence, the USR in EPON (which is limited by the maximum packet size) decreases as a percentage of the allocated time. The total GPON gain due to the adoption of segmentation and re-assembly ranges from 2 to 20%, depending on the selected $D_m$, and must be added on top of the 25% higher transport efficiency of GPON due to the 8/10bit coding in EPON. The overall bandwidth efficiency (maximum throughput achieved in each case) taking into account all the above effects is summarized in Figure 7. Obviously not as good a trade-off between bandwidth
efficiency and delay variation as can be reached in GPON is possible in EPON, since the reduction of \( D_m \) in order to make polling more frequent (and bring down average access delay and delay variation for delay sensitive applications) results in high losses due to the lack of segmentation mechanisms and the USR effect.

6. CONCLUSIONS

The scope of this paper was to compare the EPON and GPON technologies mainly in terms of MAC efficiency relying on computer simulation to assess traffic behavior and identify the performance trade-offs. It is shown that GPON enjoys improved performance by the more elaborate MAC features introduced into the standard and specifically: segmentation and reassembly, reduced physical layer (including burst preambles and line coding) and reduced MAC protocol overheads. It is important to note that the above GPON features do not only improve average performance
(delay, utilization) but also delay variation, which is equally important for the support of real-time applications, an essential requirement of modern multi-service broadband access networks. GPON can efficiently perform frequent polling of requests (i.e. detect and serve newly arrived packets at ONU queues) resulting in reduced packet delay and jitter without the waste of bandwidth that EPON would have to suffer if the same service frequency was attempted. These features stem from the requirements that have been adopted during the design phase of each standard. EPONs aimed at exploiting the widespread and mature Ethernet technology for reducing component development effort, design cycles and overall cost. GPONs on the other hand aimed at higher line rates with higher efficiency accepting higher receiver circuit costs while targeting a set of mechanisms for flexible traffic multiplexing, detailed traffic management specifications and QoS guarantees with better control of network resource allocation. Of course the overall merit of the two competing solutions cannot be judged on the performance alone. Many previous examples show that winning technologies possess many features. Cost (capital and operational), ease of deployment, personnel expertise and support of user needs play an important role while best performance is not necessarily a decisive factor.

REFERENCES


AUTHORS’ BIOGRAPHIES

**Theofanis G. Orphanoudakis** (S’93–M’98) received his Dipl Ing degree in Electrical and Computer Engineering in 1995 and the PhD in Telecommunications in 1998, both from the National Technical University of Athens (NTUA). He is currently with the University of Peloponnese as a visiting assistant professor. Prior to joining the university he was with Ellemedia Technologies as a Technical Manager. His research interests include design and implementation of high-speed networking systems, performance evaluation of MAC protocols, traffic engineering and network processor architectures. He has led several research and development projects in the above areas and has published over 40 papers in international journals and conferences.

**Helen C. Leligou** received the Dipl Ing and PhD degrees, both in electrical and computer engineering, from the National Technical University of Athens (NTUA), Athens, Greece, in 1995 and 2002, respectively. Her research interests lie in the area of access control mechanisms in broadband networks including HFC, PON, WDM metro and core networks. She has participated in several EU-funded research projects in the above areas. Since 2007, she has been a Lecturer at the Technological Educational Institute of Halkida.

**Evangelos Kosmatos** received the Dipl Ing and PhD degrees, both in electrical and computer engineering, from the National Technical University of Athens (NTUA), Athens, Greece, in 2002 and 2007, respectively. His research interests lie in the area of network control plane design, resource allocation in broadband networks and intelligent multimedia content delivery. He has participated in several EU and national research projects in the above areas.
John D. Angelopoulos (S’92–M’95) received the electrical engineering degree from the National Technical University of Athens (NTUA) in 1973, the MSc degree in automatic control from Nottingham University in 1977 and the PhD degree in telecommunications from NTUA in 1992. He is currently a Professor in the Automation Engineering Department of the Technological Institute of Piraeus. His research interests cover most aspects of high-speed communications with special focus on shared medium access protocols and implementations.