Efficient medium arbitration of FSAN-compliant GPONs

H. C. Leligou*,†, Ch. Linardakis, K. Kanonakis, J. D. Angelopoulos and Th. Orphanoudakis

Institute of Communication and Computer systems (ICCS), Department of Electrical and Computer Engineering, National Technical University of Athens, 9 Heroon Polytechniou Street, Zographou, 15773, Athens, Greece

SUMMARY

The steadily rising demand for multimedia and data services, the falling cost and omnipresence of Ethernet and the maturity of passive optical networks (PON) technology, promise to radically change the landscape in the local loop. The heart of a gigabit PON system (recently standardized by FSAN/ITU) is the medium access controller (MAC), which arbitrates access to the upstream link among users with fluctuating traffic demands and effects the multiplexing and concentration policy. At the same time, it has to safeguard the service quality and enforce the parameters agreed in the service level agreements (SLAs) between the users and the service provider. In this paper, a MAC protocol designed to serve any mix of services according to their quality of service (QoS) needs, employing four priority levels along with a high number of logically separate data queues is presented. The architecture and implementation in hardware of a MAC algorithm capable of allocating bandwidth down to a resolution of a byte with QoS differentiation is the focus of this paper. It employs the bandwidth arbitration tools of the FSAN/ITU G.984.3 standard and maps SLA parameters to GPON service parameters to create an efficient, fair and flexible residential access system.

KEY WORDS: passive optical networks (PON); gigabit-PON (GPON); FSAN; access control; MAC protocol; dynamic bandwidth allocation (DBA); quality of service (QoS)

1. INTRODUCTION

The steadily increasing demand for bandwidth [1], fuelled by streaming and multimedia applications, generates high interests in deploying and operating cost-efficient broadband access networks. Cost efficiency is determined by the simplicity of the technology (affecting installation, operation and maintenance costs) and scalability. The ideal solution should also target delivering bundled voice, data, and video services (also called triple-play services) to an end-user subscriber over a single network. Today, many alternatives exist for upgrading the

*Correspondence to: H. C. Leligou, Institute of Communication and Computer Systems (ICCS), Department of Electrical and Computer Engineering, National Technical University of Athens, 9 Heroon Polytechniou Street, Zographou, 15773, Athens, Greece.
†E-mail: nelly@telecom.ntua.gr

Contract/grant sponsor: GIANT; contract/grant number: EU IST 2001-34523
existing access network to broadband such as digital subscriber line (DSL) or cable TV networks. For example, VDSL (very high bit rate DSL) offers up to 50 Mb/s to the customers and upgraded hybrid fibre coax (HFC) networks (i.e. upgraded CATV networks having fibre running between a video head-end or hub to a curb-side optical node, with the final drop to the subscriber being coaxial cable) provide a shared bandwidth of about 36 Mb/s. However, VDSL can meet the above high capacities over short copper loops whereas the shared nature of the HFC networks (between hundreds or thousands of homes) results in frustrating slow speed during peak hours. Fibre to a certain extent is therefore needed in the access network.

Although optical fibre has conquered the long and medium haul transmission plant, it has stumbled in the access part. The reason is that optical systems are not cost effective at low traffic volumes per link, as is the case in the local loop. Some form of resource sharing is necessary and the passive optical network (PON) tree topology can achieve this by letting many customers share most of the feeder fibre and the optical transceiver on the network side. At the same time, they allow traffic concentration to reach levels where optical technology becomes cost effective while they create one high-speed input port into the access multiplexer instead of several hundreds inefficient low-rate ones, further reducing the cost of the access node.

The increased interest in PON technology as an attracting solution for the deployment of efficient broadband access networks is reflected by recent standards introduced by the full service access network (FSAN) group of ITU-T and the Ethernet in the first mile (EFM) group of IEEE which cover PON solutions operating at gigabit rates (i.e. GPONs). Both can efficiently transport packet-based traffic avoiding the cell tax, paid by ATM-based PON systems [2]. APONs present the following disadvantages compared to Ethernet PONs (EPONs): higher protocol overhead, lower bandwidth, increased costs and not straightforward integration of LANs into future optical Ethernet-based WANs [3, 4]. The protocol overhead introduced by ATM for segmenting and transporting large variable-length IP packets into fixed-size 53-byte cells is considered an increased waste of resources (also called ‘cell tax’) and can reach levels above 10% [3]. Moreover, a dropped or corrupted ATM cell will invalidate an entire IP datagram, while the remaining cells carrying the portions of the same IP datagram will propagate further, thus consuming network resources unnecessarily. The overall efficiency of EPON is also considered higher compared to APON not only because of the higher bit rate (1 Gb/s versus 155/622 Mb/s) but most important because ATM did not live up to its promise of becoming an inexpensive technology, whereas the large numbers of Ethernet component and system vendors and manufacturing volumes make economics more favourable and integration of LANs transparent for Ethernet-based WANs [5].

FSAN placed emphasis on the ability of a gigabit PON (GPON) system to support all service needs with any quality of service (QoS) demand and high efficiency [6]. It adopts fixed periodic framing accommodating TDM and ATM needs, so that services with very strict requirements can be serviced at the right moment, temporarily interrupting data packets, hence the need for fragmentation [7]. The PON architecture, as shown in Figure 1, allows traffic from different customer terminations, which are grouped in optical network units (ONUs), to be multiplexed using TDMA (time division multiple access) under the arbitration of the MAC (medium access control) protocol in the upstream direction. The user–network interfaces are implemented in the ONUs, which may lie at the customer curb, building or even home, serving one or more users. Access control is exercised by the OLT (optical line termination) at the root of the tree, which lies in the central office.
The MAC has to apply the service level agreements (SLAs) taking into account QoS requirements ensuring that the contracted peak rates and maximum burst sizes will not be violated. Thus, it must include in its algorithms policing functions, which are part of the arbitration mechanism. To this end, the MAC protocol as executor of the TDMA multiplexing in the upstream of the PON is of prime importance for cost effectiveness, fairness, traffic profile control and QoS guarantees [7, 8].

The upstream link is a shared medium in which several users send bursts arbitrated by the MAC controller in a collision-free manner as depicted in Figure 1. Two snapshots are shown: at time $t_1$, several bursts have been emitted from the ONUs, which later (at $t_2$) have been successfully multiplexed under the guidance of the MAC controller in the feeder fibre with only a guard-band between them. This TDMA technique requires that the ONUs are synchronized on a frame basis (i.e. to align transmissions within the upstream frame boundaries). The OLT MAC controller arranges ONUs burst transmissions within each upstream frame assuming they are equidistant. Distance differences though between the OLT and each ONU cause different propagation delays, which in turn dictates a phase alignment of these burst transmissions so as to avoid collisions at aggregation points. To create a common timing for the upstream frame, a ranging procedure during activation and registration measures the distance differences between the OLT and each ONU [6], which is then instructed to add a suitable delay that makes them all appear equidistant. When granted access, each ONU transmits a burst of variable size consisting of the physical layer overhead upstream (PLOu) field, other optional overheads—such as the dynamic bandwidth report upstream (DBRu) to report its queue status and the power levelling sequence upstream (PLSu)—and of course the payload [9]. The MAC controller decisions travel in the ‘bandwidth map’ (BWmap) field of every fixed-size downstream frame, which contains a variable number of access structures, each indicating the ONU queue identifier (called AllocID) field, two pointers, which indicate the starting and ending point of the upstream transmission, and flags controlling the transmission of overhead blocks. (Details can be found in the FSAN/ITU-T standard [6], as well as in Reference [7].) Thus, each pair of start and stop pointers indicates to the addressed ONU the time and, implicitly, the number of bytes (allocation bytes, AB) it can transmit from a specific queue. Each of the supported 4 K queue buffers the traffic from different user interfaces/applications. To avoid the waste from reporting packet

![Figure 1. Concept of TDMA operation in upstream.](image-url)
boundaries, the reports give total queue length and as the MAC controller remains unaware of packet limits, segmentation and reassembly is implemented using the GPON encapsulation method (GEM) [6]. GEM allows the extraction of variable length packets from fixed length frames and reconstruction of those spanning frame boundaries and is used to transport both Ethernet and TDM traffic over GPONs.

The FSAN standard for GPONs [6] does not specify a MAC protocol, since this is not necessary for openness and interoperability, but just the framing and the fields, which contain the tools for announcing queue sizes (i.e. the accumulated data burst sizes that request access to the shared upstream medium) and granting ONUs the access rights in a TDMA fashion. The focus of this paper is on the design and implementation of an efficient MAC protocol, which arbitrates access to the GPON and its relation to the agreed service level parameters, which it has to enforce. Its main features include the capability to dynamically adapt the upstream bandwidth allocation to the temporal traffic properties and the efficient support of all kinds of services, while it can be implemented with low-cost microelectronics technology. The algorithm used by the MAC controller presented in this paper has been initially outlined in Reference [10] and is presented in Section 2 in a rigorous manner, probing further on system level design and hardware/software partitioning issues. The MAC algorithm needs parameters which express the QoS requirements from the SLAs. A method to calculate these parameters on the basis of the SLAs is presented in Section 3. The MAC controller as implemented in hardware for the demonstrator of the IST-GIANT project [7, 11, 12] is described in Section 4 while the performance of the MAC protocol is assessed in Section 5.

2. THE MAC ALGORITHM

In a bursty traffic environment, peak-rate bandwidth allocation for all kinds of services leads to low utilization levels. The answer is the adoption of a dynamic bandwidth allocation (DBA) scheme, which allows for the exploitation of multiplexing gain, still providing performance guarantees, as will be shown in Section 4. The presented MAC controller bases its allocation decisions on queue status reports (reflecting the traffic fluctuations) as well as on the service parameters negotiated during the activation phase by means of management tools expressing the SLA. Depending on QoS requirements, the AllocIDs (representing data queues at the ONUs) are associated during service activation with one of four T-CONT (traffic container) types, which are summarized in Table I (there is also a T-CONT 5, which is not really a different class but a way to allow a mix of the other four, see Reference [6]). The term and concept are a legacy from APONs [13]. The type of applications that better suit each T-CONT are indicated in the second column of the table while the service guarantees that each T-CONT provides are included in columns 3 and 4.

Our MAC design is based on the requirement that all the above combinations of throughput and delay guarantees must be supported and implements the following mechanisms: T-CONT 1 traffic does not employ DBA, but uses fixed periodic grants related to the peak rate of the SLA to offer strict delay guarantees, while the T-CONTs 2–4 are based on a reservation method; thus, they are serviced up to the contracted rates and only when non-empty queues are reported (as also indicated in column 5 in the table). This way their portion of bandwidth is exploited when these sources are silent. The detection of arrivals is based on polling, since the bandwidth delay product of the GPON precludes any collision resolution protocols, as was also the case for
Table I. Mapping of GPON MAC service parameters and service level traffic profile descriptors to the four traffic CONTainers defined by FSAN.

<table>
<thead>
<tr>
<th>TC</th>
<th>Applications</th>
<th>Throughput</th>
<th>Delay</th>
<th>GPON service</th>
<th>MAC parameters</th>
<th>Traffic descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant bit rate (leased line)</td>
<td>Fixed</td>
<td>Strict guarantee</td>
<td>Periodic upon activation</td>
<td>SImax, ABmin</td>
<td>PIR, PBS, Dm</td>
</tr>
<tr>
<td>2</td>
<td>Variable bit rate (voice, video)</td>
<td>Assured</td>
<td>Bounded</td>
<td>Periodic, validated by requests</td>
<td>SImax, ABmin</td>
<td>PIR, PBS, Dm</td>
</tr>
<tr>
<td>3</td>
<td>Better than best effort</td>
<td>Assured</td>
<td>No guarantee</td>
<td>Periodic, validated by requests</td>
<td>SImax, ABmin</td>
<td>GIR, GBS</td>
</tr>
<tr>
<td>4</td>
<td>Best effort</td>
<td></td>
<td>No guarantee</td>
<td>Dynamic based on request and availability</td>
<td>SImax, ABmin</td>
<td>PIR, PBS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dynamic</td>
<td></td>
<td>Dynamic assignment up to PIR</td>
<td>SImax, ABmin</td>
<td>PIR, PBS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>assignment</td>
<td></td>
<td>(used for polling)</td>
<td>SImax, ABmin</td>
<td>PIR, PBS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>up to PIR</td>
<td></td>
<td>Dynamic based on request and availability</td>
<td>SImax, ABmin</td>
<td>PIR, PBS</td>
</tr>
</tbody>
</table>

APONs [8]. T-CONT-3-type queues service to a provisioned rate is guaranteed (assured) while they may be assigned surplus bandwidth upon request and availability. Their service rate cannot, however, exceed an agreed peak information rate (PIR) as is also the case for T-CONT-4-type queues where no throughput guarantees apply.

When a new flow (AllocID) is activated, a resource management module running in software in the OLT checks the availability of the resources needed to satisfy the requirements of the new stream, ensuring that the total guaranteed service rate of all the active AllocIDs will not exceed system capacity. Depending on the associated T-CONT type, the software calculates up to four service parameters: the maximum and minimum service interval (denoted hereafter as $S_{\text{Imax}}$ and $S_{\text{Imin}}$ and expressed in multiples of the frame duration, i.e. 125 μs) and the minimum and surplus allocation bytes (denoted hereafter $A_{\text{Bmin}}$ and $A_{\text{Bsur}}$ and expressed in bytes), as shown in the ‘MAC parameters’ column in Table I. The guaranteed service rate is expressed as: $A_{\text{Bmin}}/S_{\text{Imax}}$, while the surplus (non-guaranteed) service rate is expressed as: $A_{\text{Bsur}}/S_{\text{Imin}}$. The guaranteed rate and the surplus rate sum up to the allowed peak rate for T-CONT 3 while for T-CONT 4, the $S_{\text{Imax}}$ is expressing the polling period, since $A_{\text{Bmin}}$ is equal to the DBRu report message length. (The last column will be discussed in the following section).

In every frame, the MAC algorithm allocates first, the guaranteed bandwidth (as described in Figure 2) and then the surplus, if there is still available space in the frame (as shown in Figure 3). In both figures, Assign$(i, \text{Allocation_bytes})$ denotes the preparation of the standard-compliant

```plaintext
Frame_bytes = 19440 /* for 1.24Gb/s GPON upstream rate */
/* Guaranteed Bandwidth Allocation phase */
/* inspect all S_{Imax} timers */
for (each AllocID i)
  if (S_{Imax}_timer(i) = 1)
    Allocate_guaranteed(i)
    S_{Imax}_timer(i) = S_{Imax} /*negotiated parameter */
  else
    S_{Imax}_timer(i) = S_{Imax}_timer(i) - 1;
/* process Allocate_guaranteed (i) */
if (T_CONT(i) = 1 or T_CONT (i) = 4)
  Allocation_bytes = A_{Bmin}
else
  Allocation_bytes = min\{A_{Bmin}, Request(i)\}
Frame_bytes = Frame_bytes - Allocation_bytes
Request(i) = Request(i) - Allocation_bytes
Assign(i, Allocation_bytes)
```

Figure 2. MAC algorithm pseudo-code: allocation of guaranteed part of bandwidth.

/* Surplus Bandwidth Allocation phase */
/* update of all T_CONT 3 and 4 SImin timers */
for (each AllocID i)
  if ((T_CONT(i) = 3 or T_CONT (i) = 4) and SImin_timer(i) > 1)
    SImin_timer(i) = SImin_timer(i) - 1
/* inspect T_CONT 3 requests and SImin timers */
N = TC3_start_pointer
for (each AllocID i)
  if (Frame_bytes = 0)
    TC3_start_pointer = N + i - 1
    return
  else
    if (T_CONT(N+i) = 3 and Request(N+i ) > 0 and SImin_timer(N+i) = 1)
      allocate_surplus(N+i)
/*inspect T_CONT 4 requests and minSI */
L = TC4_start_pointer
for (each AllocID i)
  if (Frame_bytes = 0)
    TC4_start_pointer = L + i - 1
    return
  else
    if (T_CONT(L+i) = 4 and Request(L+i ) > 0 and SImin_timer(L+i) = 0)
      allocate_surplus(L+i)
/* allocate_surplus(i) */
Allocation_bytes = min{ABsur, Request(i), Frame_bytes}
Assign(i, Allocation_bytes)
SImin_timer(i) = SImin(i); /*negotiated parameter */
Frame_bytes = Frame_bytes - Allocation_bytes;
Request(i) = Request(i) - Allocation_bytes;

Figure 3. MAC algorithm pseudo-code: allocation of surplus part of bandwidth.

access structure granting Allocation_bytes to AllocID i. SImax_timer is a count down counter decreased by one every frame duration while SImax is the MAC service parameter, fixed for the lifetime of each AllocID. Also, it is assumed that every time a queue report for AllocID j arrives, the reported length is stored in the Request[j] element of the request matrix.
For the surplus bandwidth allocation to T-CONT 3, setting all SImin parameters equal and varying the ABsur parameters, a weighted round robin service is enforced. If all the ABsur parameters are also set equal, the service becomes equivalent to a simple round robin. The same applies for T-CONT 4 AllocID service. However, the algorithm as well as its implementation that will be presented in Section 4, support different SImin value for every AllocID, since the service parameters affect both the performance and the efficiency of the system, as will be discussed in Section 3.

A GPON according to FSAN should support both status reporting (SR) and non-status reporting ONUs, i.e. ONUs that do not have the capability to report the queued traffic. Still some implicit application of DBA is possible for NSR ONUs albeit with a certain inevitable inefficiency. Obviously, for T-CONT 1 traffic no difference exists since no reporting is applicable. However, for T-CONT types 2–4 the adopted MAC mechanism is modified to consider that whenever an NSR ONU sends upstream in response to an allocation a burst that does not occupy all the allocated bytes, its queue has become empty. No further allocation will be provided to such an ONU until the next polling time. The ‘polling’ in this case involves a grant for a pre-determined payload size. Should a full burst be returned, the allocation follows the above-described algorithm (as if the queue was full); otherwise, the queue is considered empty. Thus, in the NSR case, the MAC controller will stop allocating bandwidth only when an empty queue is surmised, which is always accomplished at the expense of an underutilized last allocation.

3. SELECTING SI, AB VALUES ON THE BASIS OF SLA-BASED TRAFFIC DESCRIPTORS

In our GPON system, a service rate $R$ is effected by the MAC by allocating AB every SI (service interval), i.e. $R = AB/SI$. Obviously many AB, SI pairs can provide the required rate, but high SIs increase delay while low ABs cause inefficiency. To achieve the best trade-off between transport efficiency and delay, the higher SI that satisfies the maximum tolerated delay (with some safety margin) has to be used and then the AB value can be calculated. If no delay guarantee exists, the AB is first selected in a way that would not violate the peak burst size (PBS).

When a new connection is accepted, depending on the service class and hence the applicable T-CONT, suitable traffic descriptors (listed in Table I) are negotiated including a peak info rate (PIR), a maximum access delay $D_m$ and a PBS. In the case of T-CONT 3, a guaranteed info rate (GIR) and guaranteed burst size (GBS) are used as well. To calculate the $S_{\text{max}}$ as a function of the maximum delay ($D_m$) the application tolerates, a relation between them can be expressed as follows:

$$D_m = 2 \times S_{\text{max}} + 2 \times T_p + 2 \times T_{\text{pr}}$$

The above formula is derived based on the following observation: due to the reservation approach in T-CONTs 2–4, the worst-case latency is at most $S_{\text{max}}$ for a polling opportunity to send a DBRu field. On top of this the fixed propagation time required for the DBRu field to reach the OLT (denoted $T_p$ hereafter) should be added, plus the scheduling time required for the MAC controller to schedule a grant (which is again $S_{\text{max}}$ in the worst case), plus the fixed
propagation ($T_p$) of the grant to reach the ONU. Finally, the time required for the transmission and decoding of the messages ($T_{pr}$) in each node should also be taken into account.

Having defined the $S_{\text{max}}$ value from the above relationship, the $A_{\text{min}}$ is calculated on the basis of peak information rate (PIR) for T-CONTs 1 and 2 AllocIDs, and guaranteed information rate for T-CONT 3 AllocID. The $A_{\text{min}}$ parameter should be augmented to account for the GPON encapsulation headers and the queue report field. For T-CONT 3 AllocIDs, $A_{\text{min}}$ is checked versus Guaranteed Burst Size parameter, while for T-CONT 4 AllocIDs, $S_{\text{max}}$ represents its polling period.

The ($A_{\text{sur}}, S_{\text{min}}$) pair of parameters are defined only for T-CONTs 3 and 4 AllocIDs and are calculated based on the PIR. First, the $A_{\text{sur}}$ is selected so that it is lower than the PBS as well as the buffer space allocated to the respective flow at the ONU side. Having selected $A_{\text{sur}}$, the $S_{\text{min}}$ is calculated as a function of the peak rate.

Finally, as regards connection acceptance policy, a simple conservative solution is to stop accepting new connections when the sum of PIRs of T-CONT 1 and T-CONT 2 plus the GIRs of T-CONT 3 reach system capacity. Depending on operator tariff policy, it may be desirable to offer better opportunities to non-guaranteed traffic, in which case the blocking of new connections could stay well below 100% of capacity (e.g. 80–90%) to leave some more room for best-effort traffic. Otherwise, the non-guaranteed traffic will only gain access whenever the variability of the guaranteed parts leaves the total instantaneous load below system capacity (which is not of course rare when bursty services are multiplexed, but can guarantee that T-CONT 4 will not be blocked out for prolonged intervals whenever T-CONTs 2 and 3 are near their peak rates).

4. THE MAC CONTROLLER IMPLEMENTATION ARCHITECTURE

Static bandwidth allocation for constant peak rate service of high-priority T-CONT 1 traffic in our system is scheduled during connection (AllocID) activation. These pre-arbitrated grants as well as minimum bandwidth allocation for polling of requests together with all service configuration parameters related to each AllocID are prepared in the management plane and provided to the MAC controller by the system host microprocessor (on board controller, OBC) executing all control and management plane software. The computationally intensive bandwidth allocation per frame and BW map construction evidently though cannot be executed in software by any available instruction set architecture, within the 125 μs frame period. Therefore, in the following we describe the custom hardware implementation of the MAC controller and we discuss the implications of the MAC algorithm with regard to an efficient system design.

The above MAC design has been implemented for the IST-GIANT system demonstrator placing the controller’s hardware on a Xilinx Virtex II FPGA device. The selection of reconfigurable logic is justified by the fact that this technology can provide devices with very high densities and fast enough, enabling future upgrades and reconfiguration at competitive cost. Although the relatively high cost of high-end FPGAs has made them known widely as a best match for prototyping, the economies of scale from the production of GPON OLT systems due to the shared cost and low production quantities may prove custom ASIC development more expensive. As we demonstrate in this section, an efficient implementation of the above MAC design can fit into a middle-sized FPGA achieving a good cost–performance trade-off.
The interfaces as well as the internal organization of the MAC controller are shown in Figure 4. The OBC configuration parameters, except from Slmin and Slmax, which are stored in individual memory blocks, are stored in the ‘BW allocation parameters’ (BAP) memory. The queue status information, which is necessary to enforce the DBA MAC, is reported by the ONUs in the DBRu fields and received from the upstream deframer. The MAC controller schedules next the appropriate grants (which is the most challenging task implemented by its core functional blocks) and provides the downstream framer with the calculated BW map. Two blocks are responsible for the FSAN standard-compliant frame construction. The access structure construction block organizes the information in standard-compliant access structures calculating the start and stop pointers, defining the flags indicating the overheads transmissions (based on information provided by the ‘ONU service’ block) and appending the CRC. The BWmap forwarder entity is responsible for the implementation of the ranging procedure and the forwarding of the prepared access structures to the downstream framer. The ‘GLTP2 serial interface’ block, appearing in the right-hand side announces requests for PLSu and physical layer OAM blocks.

The task of dynamically scheduling grants to fairly allocate surplus bandwidth upon request as well as enforcing the maximum configured transmission rate is a challenging one even in a custom hardware implementation. Since we expressed the bandwidth allocation parameters in terms of bytes and time intervals to achieve the desired granularity, several counters and timers need to be maintained both for the periodic and dynamic service. To allocate the guaranteed part of bandwidth, executing the logic described in Figure 2, up to 4096 timers (each corresponding to a different AllocID) have to be inspected and updated. A typical implementation of a timers management unit would employ a common memory to store timer values and sequential memory accesses to decrement and inspect for expiry. This solution though suffers in terms of throughput, since for the target operation frequency of 77.5 MHz the sequential service
would consume 85% of the frame duration, squeezing the time available for the allocation of the surplus part of bandwidth. The fastest possible solution would require a fully parallel implementation where all timers would be individually triggered by the same clock (the frame clock), each of them being associated with an expiration flag fed to a priority enforcer. However, this latter option consumes significant hardware resources, especially in our target FPGA device. In the case of the IST-GIANT demonstrator, we opted for the first option, i.e. to keep the timers in memory and update them sequentially. Additionally to overcome the drawback the bandwidth allocation phase was organized in a pipelined way: the expired timers identified during a certain frame were kept in the ‘timers FIFO’ and were serviced in the following frame time by the ‘timers service’ entity, which defines the number of bytes to be allocated to each AllocID based on the ABmin and request values.

Once the guaranteed part of bandwidth has been allocated (i.e. the modules appearing in the upper block in Figure 4 have completed their operation), and if there are still available bytes in the frame, the blocks allocating the surplus bandwidth (included in the lower block) are activated. The surplus bandwidth allocation is triggered dynamically by the non-zero requests. To accelerate the inspection of requests, two memory regions (one for T-CONT 3 and another for T-CONT 4) organized in 16-bit words are implemented, where a one-bit flag per AllocID is stored, indicating non-zero requests for the relevant AllocID. T-CONT 3 flags are inspected prior to T-CONT 4 requests in a round robin manner. Each time a non-empty queue has been detected, the relevant Smin timer value is inspected as well, to enforce peak rate policing. If this has expired, indicating that the respective AllocId has been allocated surplus bandwidth at least Smin frames earlier, the AllocId is allocated the min(ABsur, Request(i), Frame bytes). Thus, from a traffic management point of view the overall design effectively implements a 1.24 Gb/s hierarchical scheduler-policer.

Regarding the memory requirements of the MAC design in the BAP memory, 39 bit are devoted to each AllocID:

- 13 bit are required to code each AB parameter (min and sur), allowing for the assignment of whole frame times (19 440 byte),
- 8 bit are required for coding the ONU ID,
- 3 bit for the T-CONT type and
- 2 bit for control information.

The ‘Request_matrix’ memory is 21-bit wide, to accommodate the decoded queue length value in bytes. (Non-linear coding is employed for the queue length, which is measured in blocks of 48 byte, as is mandated in Reference [6].) Considering 10-bit wide SI parameters, 100 bit of RAM memory per supported AllocId are required plus two memories of first-in first-out (FIFO) structure: the BWmap FIFO where up to 1024 access structures (9 byte each) are stored and the timers FIFO (512 × 12 bit). Hence, to support 4096 AllocIds, the overall memory requirements reach 425 kbit, which can easily be accommodated on-chip.

A MAC controller following this architecture has been implemented for the demonstrator of the IST-GIANT in a Xilinx FPGA device. The design has a gate equivalent of 2 400 000 gates (accounting also the memory requirements) and occupies less than 30% of the Virtex II XC2V2000 resources for an internal clock frequency of 77.5 MHz. To this end, this prototype implementation proves that the OLT MAC controller can meet its performance targets even with FPGA technology. Replacing the FPGA with ASIC technology in a commercial product...
would easily increase operation (clock) frequency by a factor of two or three as is necessary to support 2.5 Gb/s in the upstream channel.

5. PERFORMANCE EVALUATION

To complete the assessment of the proposed MAC design, we present in this section typical performance results that demonstrate its efficiency in multiplexing traffic with different QoS requirements. Given the lack of analytical tools, a model of the GPON consisting of 32 ONUs was developed, using the OPNET simulation software. Each ONU is equipped with 4 queues, which can be assumed to store traffic of T-CONT 2, 3 or 4 type. It is worth noting that for T-CONT 1 the delay has a deterministic behaviour with well-defined limits, so no simulation is needed. The upstream rate is 1.24416 Gb/s. Once the overhead is subtracted this leaves a net capacity of about 1 Gb/s. Here, the performance is given as a function of offered load. (Further simulation results investigating a large number of operational conditions are presented in Reference [10].)

For the set of simulations presented here, exponential inter-arrival times were used. Regarding the packet length frequencies the following 3 lengths were generated with respective probability: 60% for 64-byte long packets (representing TCP Ack packets), 20% for 500-byte long and 20% for 1500 byte (about the maximum Ethernet length). That is, the length distribution followed the widely used tri-modal length distribution, which was found to accurately reflect IP data traffic from LAN-located end-user terminals. The load distribution between ONUs and T-CONTs was uniform. The polling period (maximum time between queue reports) was 1.25 ms, i.e. 10 frames. To test the system behaviour under different offered loads the mean packet generation interval was varied from run to run and the AB parameter used by the MAC controller was modified accordingly. Regarding T-CONT 3 AllocIDs, a GIR = 1/3PIR was used.

The results are shown in Figure 5. When the total offered load is below 0.9 Gb/s, all the injected traffic experiences low access delay (below 2 ms) with the service rate being equal to the source rate. Since the employed MAC algorithm is not a simple prioritized round robin, the

![Figure 5. Delay versus offered load for T-CONTs 2, 3 and 4.](image-url)

Copyright © 2005 John Wiley & Sons, Ltd.

delay experienced by T-CONT 2 sources strongly depends on the service parameter choice, hence can be higher than that of T-CONT 4. (This can happen, for example, if the SImax values are chosen for 3 ms tolerated delay and the PIR of T-CONT 4 is higher than the source rate.)

When the total offered load is above 0.9 Gb/s, the system cannot satisfy the entire offered load. The traffic classes that need quality guarantees do not feel any overload and only T-CONT 4 traffic suffers the congestion. The relevant queues are growing causing packet losses. In real-life, these will be detected by the TCP closed-loop control mechanisms and the source rate will be accordingly regulated. T-CONT 2 sources are still observing low delay values due to the employed service strategy. It is worth stressing that still T-CONT 2 is serviced based on DBA, i.e. it is not allocated the peak-rate bandwidth unless non-empty queue is reported, allowing for efficiency improvements. Also T-CONT 3 traffic enjoys good performance. However, when the total offered load goes above 1.6 Gb/s (i.e. T-CONTs 2 and 3 sum up to 1 Gb/s, which is the net capacity), then T-CONT 3 service also becomes unstable (its service rate is equal to the GIR which is lower than its source rate) and only T-CONT 2 enjoys good performance thanks to its prioritized service.

6. CONCLUSIONS

The MAC protocol in an FSAN-compliant G-PON is the executor of the multiplexing policy and as such it controls the traffic behaviour on which the cost effectiveness of the system is based. The dynamic bandwidth allocation principle together with the four priority levels and the weighted round robin policy can satisfy a dynamic mix of services covering all residential needs.

The MAC algorithm is implemented in hardware because of the high speed required, while the service provisioning is handled by the embedded controller, which provides the required service parameters that govern the MAC allocation extremes. Although the allocations match the dynamic arrival fluctuations, the enforcement of these extremes turns the MAC into a policing tool as well. The overall effect is a system that manages to concentrate the traffic as it passively travels towards the single feeder fibre and the single input port to the access multiplexer, thus bringing it to the gigabit per second levels where optics can become cost effective.

ACKNOWLEDGEMENTS

The work presented in this paper has been partially funded by EU IST 2001-34523 GIANT project.

REFERENCES


AUTHORS’ BIOGRAPHIES

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 PhD dissertation was in the area of access control mechanisms in broad-band networks. She is currently investigating access mechanisms for high-speed networks for tree-topology hybrid fibre/coax (HFC) and GPON systems and wavelength-division-multiplexed metro-ring networks.

Charalambos Linardakis (S’02) received the Dipl Ing and PhD degrees, both in electrical and computer engineering, from the National Technical University of Athens (NTUA), Athens, Greece, in 2000 and 2004, respectively. He is a Hardware Designer at the Telecommunications Laboratory in NTUA. His interests focus on wavelength-division-multiplexed access networks, passive optical network access networks, and convergence of TCP/IP and broad-band telecommunications.

Konstantinos Kanonakis was born in Heraklion Crete, Greece, in 1982. He received his Diploma in 2004 and is currently pursuing his PhD from the Faculty of Electrical and Computer Engineering of the National Technical University of Athens (NTUA), Greece. His main research interests are in the area of traffic engineering, architectures and control protocols for optical networks and their implementation in hardware.
John D. Angelopoulos (S’92–M’95) received the electrical engineering degree from the National Technical University of Athens (NTUA), Athens, Greece in 1973; the MSc degree in automatic control from Nottingham University, Nottingham, U.K., in 1977; and the PhD degree in telecommunications from NTUA in 1992. He worked for 11 years in the research and development section of the aerospace and telecommunication industries before joining the Technological Institute of Piraeus in 1989, where he is Professor in the Automation Engineering Department teaching data-network-related courses. He is also participating in NTUA research activities in the areas of high-speed networks. He has been involved in several RACE, ESPRIT, ACTS, and IST projects with emphasis on broadband access systems (passive optical networks, hybrid fibre/coax (HFC), and wireless). His research interests cover most aspects of high-speed communications with special focus on shared medium access protocols and implementations. Dr Angelopoulos is a Member of the Technical Chamber of Greece.

Theofanis G. Orphanoudakis 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. He is currently at Ellemedia Technologies leading R&D activities in the area of broadband networking components. Since 2005 he is also with the University of Peloponnese as an assistant professor. From 1996 to 1999, he was a research associate at the Telecommunications Laboratory of the National Technical University of Athens. He has participated in several EU research programs and has led projects in the areas of broadband networking components and broadband access networks. His research interests include design & implementation of broadband access networks, performance evaluation of MAC protocols for shared broadband systems, traffic engineering in high speed IP and ATM networks, scheduling and network processor architectures. He has several publications in the above-mentioned areas.