Prioritized Multiplexing of Traffic Accessing an FSAN-Compliant GPON

John D. Angelopoulos, Helen-C. Leligou, Theodore Argyriou, and Stelios Zontos

National Technical University of Athens,
School of Electrical and Computer Engineering
Telecommunications Systems Laboratory
9 Heroon Polytechniou Street, Zographou, 15773, Athens, Greece
{jangel, nelly, teo, zontos}@telecom.ntua.gr

Abstract. The Gigabit per second PON, now under standardization by the FSAN group in ITU-T, constitutes the most promising approach for the photonisation of the local loop, an elusive target of telecom operators for more than a decade. The traffic multiplexing / concentration on which the system efficiency (and hence low cost) is based, relies upon the dynamic MAC protocol to allocate bandwidth according to momentary needs, following traffic fluctuations. To serve each traffic class according to its needs, four priority levels are employed along with a high number of logically separate queues. This paper presents and evaluates the FSAN access control algorithms and discusses the choice of suitable traffic service parameters that optimize system performance.

Keywords. FSAN, GPON, PON, EFM, Shared access, reservation MAC.

1   Introduction

The current initiative of FSAN (Full Services Access Network) consortium to standardize a Gigabit per second Passive Optical Network, constitutes the most promising approach to the photonisation of the local loop. The incentive for such technology lies in cost benefits stemming from the fact that they need less fiber and less costly optical interfaces at the central office (one optical interface serves the entire network) but also achieve high traffic concentration as appropriate for low cost residential access systems [1]. After a series of developments starting in the late 80s in BT, the need for a standard led in the mid-1990s a group of major network operators to establish the FSAN (Full Service Access Network) consortium with a mandate to define a common standard for PON equipment, so that vendors and operators could create a competitive market for the last mile. Already a 155 Mb/s PON system has been specified and later adopted as the ITU-T G.983 series of standards known as APONs/ BPONs [2], [3]. In APON, data is transported via fixed-size packets of 53 octets, implementing a segmentation and reassembly process, for variable length packets. The small, fixed-size packets are deemed essential for carrying delay-sensitive traffic, making APON appropriate for services like voice, video and data.

On the other hand the provisioning of ATM virtual paths and virtual channels appears to add important complexity in the PON system, while the five-byte header added in each 48-byte payload consumes a large amount of upstream bandwidth. Trying to eliminate these drawbacks of APONs and following the major drive of the increasing packet traffic due to the omnipresence of Ethernet LANs and IP, the IEEE through the Ethernet in the First Mile (EFM) [5] group are also working on standardizing a new form of PON called EPON (Ethernet PON). In EPON data is transmitted in variable length packets of up to 1,518 bytes avoiding the frequent and time-consuming segmentation and reassembly procedure of APON.

In a parallel activity, but with emphasis in QoS and efficiency as well as IP support, the FSAN group, initiated in 2001 a new effort for standardizing PON networks operating at bit rates above 1 Gb/s and supporting multiple services, and operation, administration, maintenance and provisioning (OAM&P) functionality and scalability. The result is the Gigabit PON (GPON), intended to support all service needs of access networks by either packet or TDM transport.

The work presented in this paper was carried out in the framework of IST project GIANT (GIgaPON Access NeTwork) [4], which targets the design, implementation and demonstration of such an Ethernet-based GPON system. It will support all kind of services from very strict QoS down to plain best-effort. The studies of the GIANT project have heavily influenced FSAN through contributions of consortium members and particularly Alcatel Bell, the consortium leader, but in turn the GIANT system has been aligned to FSAN decisions as they become available resulting in an FSAN-compliant system.

2   GPON Operation

The PON access technology is based on a passive star fibre network, which connects a number of ONUs (Optical Network Units) at the subscriber side to one OLT (Optical Line Termination) in the local exchange, as shown in Figure 1. The traffic streams arriving at the ONUs from the customer premises are kept in logically separate queues identified by their Allocation Identifier (AllocID) in FSAN terminology. Depending on QoS requirements, the AllocIDs belong to one of five T-CONT (Traffic Container) types. In the downstream direction, the OLT broadcasts the packets and their destination ONU extracts them based on the knowledge of which AllocIDs are locally assigned. In upstream, an aggregate data flow is generated by the active ONU transmitters via a TDMA mechanism controlled by the MAC controller in the OLT, which allocates the channel to each ONU in turn, according to a MAC algorithm. As with all PONs, a ranging procedure is relied upon for eliminating the problems arising by the difference in distances due to the tree topology. The AllocID assignments take place during the activation and registration process of each ONU. The choice in FSAN was to handle each AllocID separately without multiplexing those with the same QoS in one queue, due to the will of operators to keep detailed control of the customer traffic so as to check compliance with SLAs.
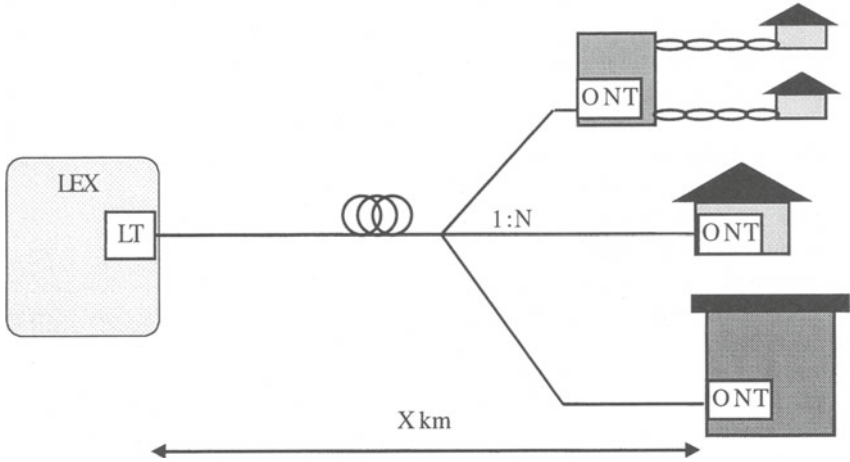
Fig. 1. PON structure

In the downstream direction frames of fixed size are transmitted every 125us consisting of a Physical Control Block downstream (PCBd) header of variable length and a payload field. The PCBd contains information about synchronization, bit error rate estimation, PLOAM messages [2] and a Bandwidth Map field, which is created by the GIANT MAC algorithm and contains the grants that enable upstream transmissions from the referenced ONU and AllocID. More specifically, the bandwidth map (shown in Figure 2) contains a number of allocation structures, which distribute the upstream bandwidth to the AllocIDs. Each allocation structure contains apart from the AllocID field two pointers, which indicate the starting and ending point of the upstream transmission in the channel (all with reference to the downstream frames, which play the role of time reference for the system). Thus, by the difference of the start and stop pointer each ONU learns the time that is granted to it for upstream transmission which can be compiled in the number of bytes it can transmit (Transmission Bytes, TB) avoiding, thus, collisions and at the same time exploiting in an efficient way the upstream bandwidth.
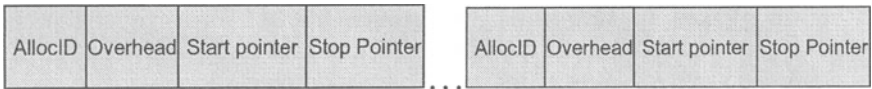

Fig. 2. BW Map structure

The upstream frame is created by a number of transmissions (shown in Figure 3) from one or more ONUs as dictated by the successive structures of the bandwidth map. Moreover, apart from user data, four PON overhead types can be inserted in the upstream frame. These types are:

- The physical layer overhead (PLO) which is always present at the start of an ONU upstream burst and contains the indispensable preamble, allowing proper physical layer operation (e.g. bit / byte alignment) of the burst mode upstream link.
- The physical layer operation, administration and management field (PLOAM), which is responsible for management functionalities, such as ranging, activation of ONU and alarm transfer.
- The power leveling sequence field (PLSu) which transfers information about the level of power at lasers of ONUs as been seen by the OLT.
- Finally, the Dynamic Bandwidth Report (DBR) field which reports to the OLT the queue length of each AllocID, allowing for the operation of the reservation based MAC algorithm according to the DBA (Dynamic Bandwidth Allocation) principle.

The inclusion of any of the above fields, except the first one, is controlled by appropriate flags sent by the OLT in the downstream direction.
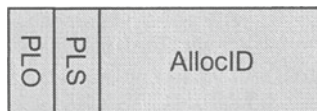

Fig. 3. Upstream frame format

In contrast to EPON, the GPON implements segmentation and reassembly for reasons of efficiency avoiding the overhead of reporting the boundaries of packets. So frames must be reconstructed in the egress of the system and the adopted method is called GEM (GPON Encapsulation Mode).

As mentioned, FSAN places emphasis on QoS support and seeks to control by means of the MAC protocol each traffic stream so as to be able to effect the SLA (service level agreement) and provide the required quality per user and stream. The quality class, and hence the service received, are determined by assigning each queue to one of the five Traffic Containers (T-CONTs) which follow different service policies.

The five traffic classes of FSAN are a legacy from the APON DBA specification G.983.4 [3] keeping the same term: T-CONT. Trying to offer different levels of QoS GPON determines four different T-CONTs corresponding to ATM's Class of Services:

- T-CONT-1 based on unsolicited periodic permits granting fixed payload allocations and intended for the emulation of leased line services,
- T-CONT-2 for VBR traffic and applications with both delay and throughput requirements, such as video and voice,
- T-CONT-3 for better than best effort services offering service at a guaranteed minimum rate while any surplus bandwidth is assigned only upon request and availability,
- T-CONT4 for purely best-effort services (browsing, FTP, SMTP, e.t.c.), and as such is serviced only upon bandwidth availability up to a provisioned maximum rate,

- •   T-CONT5 is a combined class of two or more of the other four T-CONTs. It is now left to the ONU to choose which queue to service.

The MAC controller services the traffic belonging to each T-CONT starting from T-CONT 2, 3 and 4 with absolute priority i.e. no traffic of a lower priority is serviced while any requests for higher priority exist. The operation of the algorithm is presented in detailed next.

3   The MAC Algorithm

The MAC controller is responsible for the assignment of both the guaranteed and the surplus part of the bandwidth to the active queues. The service principle is a prioritised weighted round robin. Service of T-CONT1 is based on unsolicited periodic grants, which create a virtual leased line allocating a fraction of the link bandwidth inflexibly to each customer, (i.e. no variation according to the offered load, any excess is wasted as in a leased line). The priority order of the rest is: T-CONT2 first, then 3 and 4 while the weights follow SLA (Service Level Agreement) parameters. Unlike T-CONT 1, T-CONTs 2,3 and 4, use the reservation approach which regulates allocation according to fluctuating needs which in FSAN is called DBA (Dynamic Bandwidth Allocation). The information required for the proper operation of the MAC controller is based on the queue length reports sent for every AllocID, as well as the service level parameters negotiated during the activation phase by means of management tools.

Each flow is allocated at activation time two parameters: SDI (Successive Data Interval, i.e the time distance between two successive grants) and TB (Transmit Bytes, i.e how many bytes are to be sent with each allocation). Upper and lower bounds of these parameters are defined in the service agreement and are related to the specified peak and average rates as well as the allowed variation. This provides the tool to specify a guaranteed part (based on minTB, MaxSDI) allowing the surplus bandwidth to be assigned dynamically up to the peak rate (defined by MaxTB, MinSDI) by varying the actual values of TB and SDI in each allocation.

In more detail, for T-CONT1 the maximum and minimum TB and SDI values are equal (to keep delay variation zero). For T-CONT2, although the same relation between maximum and minimum TB and SDI values exists, the respective allocations are issued only on condition of outstanding requests i.e. unsatisfied reports. For T-CONT3, maximum and minimum values are different, resulting in the differentiation of guaranteed and surplus bandwidth assignments, while for T-CONT4, the maximum grant interval is theoretically infinite, providing no guarantees. The MAC controller in GIANT uses timers to enforce the extreme values while relying on DBA to decide each round of allocations by inspecting the "request" table where past unserviced requests are stored, reflecting the queue fill level. The examination of this table follows the round robin discipline within each T-CONT while absolute priority is observed among T-CONTs.

As expected, the polling frequency is a critical parameter for DBA performance, since nothing can be guaranteed a service faster than the polling cycle as the round trip for reservation and the processing times must also be taken into account. For example, in

order to satisfy the maximum of 3ms of round trip delay budget for real-time services, a maximum polling interval of 500μs has to be adopted to guarantee an access delay below 1.5ms.

4 Performance versus Offered Load

To study the performance of the MAC algorithm, given the lack of analytical tools due to the high system complexity, a series of computer simulations have been carried out. In this section, the overall performance of the MAC is evaluated versus the total offered load under uniform loading among all sources.

It is worth noting that for T-CONT 1 the delay has a deterministic behaviour with well-defined limits so no simulation is needed. In contrast, the evaluation of T-CONT 2 access delay is very important since it is through this T-CONT that delay sensitive applications will be serviced based on a dynamic mechanism seeking greatly improved efficiency when compared with the rigid and wasteful T-CONT1 approach. For T-CONT 3 and 4 the metric of greater interest is the throughput rather than the delay, provided that the latter stays within reasonable limits (hundreds of ms).

The model consisted of 32 ONUs each supporting T-CONT 2, 3 and 4 with only one AllocID per T-CONT, i.e. totally 96 queues. The sources generating the traffic load were following the widely used for end-user data systems tri-modal length distribution model, which reflects IP data traffic length distribution from LANs. Exponential interarrival times were used. So, the packet length frequencies were about 60% of 64 byte long packets, 20% of 500 byte long and 20% of 1500 bytes while the load distribution among ONUs and T-CONTs was uniform. The polling period (maximum time between queue reports) was 1.25ms, i.e.10 frames.

The parameters that change from run to run were, first, the mean of time intervals between packet generation (decreasing for increasing load), in order to vary the total load, and second the negotiated parameter TB on which the MAC controller bases the allocations beside the AllocID queue report (implicit request).

Regarding T-CONT3 two options are investigated. As T-CONT 3 has in its specification a guarantied part and a surplus part, in the first scenario group the assured bandwidth for T CONT 3 is about 2/3 and in the other it is 1/3 of the provisioned bandwidth on the basis of its SLA.

As shown in Figure 4, as the total offered load increases, queues of T-CONT 4 suffer first the congestion at about 0.9Gbps due to the prioritized service. Hence, for total offered load below 0.9Gbps, all traffic is serviced and the observed delay remains in the order of ms for all types of traffic. When the total offered load is above 0.9Gbps, all traffic is protected except best-effort no-guarantee T-CONT 4 traffic which suffers all the congestion. It is worth noting that the sources did not contain any closed loop congestion control (i.e. TCP-like), which would in real life come into action to reduce the offered load. It was chosen to focus on the MAC mechanism and to exclude from the model interference from other network elements that are encountered by a flow in an end-to-end travel through a network since such an approach

would involve many other assumptions about the rest of the network that do not play a role in the MAC evaluation.
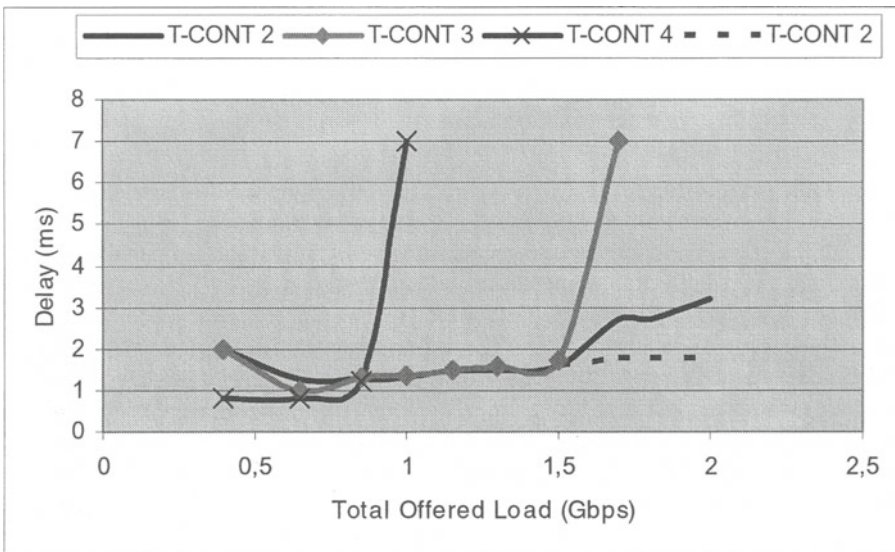

Fig. 4. Impact of increasing offered load

Queues of T-CONT 3 type are serviced to the demanded rate even when the total offered load is up to 1.6Gbps, i.e. above the nominal link rate. This of course is not expected to happen thanks to combined action of admission control and policing. So it can be considered a simulation of misbehaving T-CONT4 sources, which, as a result, have caused overflowing queues, (protecting within the designed limits the service received by T-CONT2 and 3 traffic). Of course, when the offered load reaches beyond 1.6Gbps, (at which time the offered load of T-CONT 2 and T-CONT 3 exceeds 1.06Gbps), the surplus bandwidth is not enough for the full service of T-CONT 3, which also gradually enters unstable conditions.

Focusing on the performance of T-CONT 2, this is as expected better than that of the other two, in both scenarios where the guaranteed part of T-CONT 3 is 2/3 or just 1/3 of the T-CONT 3 bandwidth. However in the scenario of high proportion of guarantied BW to T-CONT3, the access delay of T-CONT 2 start increasing earlier (solid line) than when only 1/3 is guaranteed (dotted line), though not as much as the delay of T-CONT3 which has a lower priority. In the second case, the delay for T-CONT 2 is kept lower than 2 ms up to a total load of 2Gbps as shown by the dotted line. Despite the fact that such unrealistically high loads are to be prevented by SLAs, they have a value in evaluating the MAC under extreme conditions and checking the effect of instantaneous overloads, which cannot be excluded while they stay within the tolerance provided by the specified leaky bucket buffering of the policing unit.

5   The Impact of TB and SDI on Performance

To study the impact of the choice of TB and SDI on performance, five sets of simulation scenarios were carried out, changing in each scenario the values of TB and SDI while keeping constant their ratio TB/SDI, which is the service rate and must equal the provisioned source rate. The polling period (used for ONUs to inform the OLT about their queue lengths) also coincides with the SDI. Again only T-CONT2 type traffic was used in all scenarios.
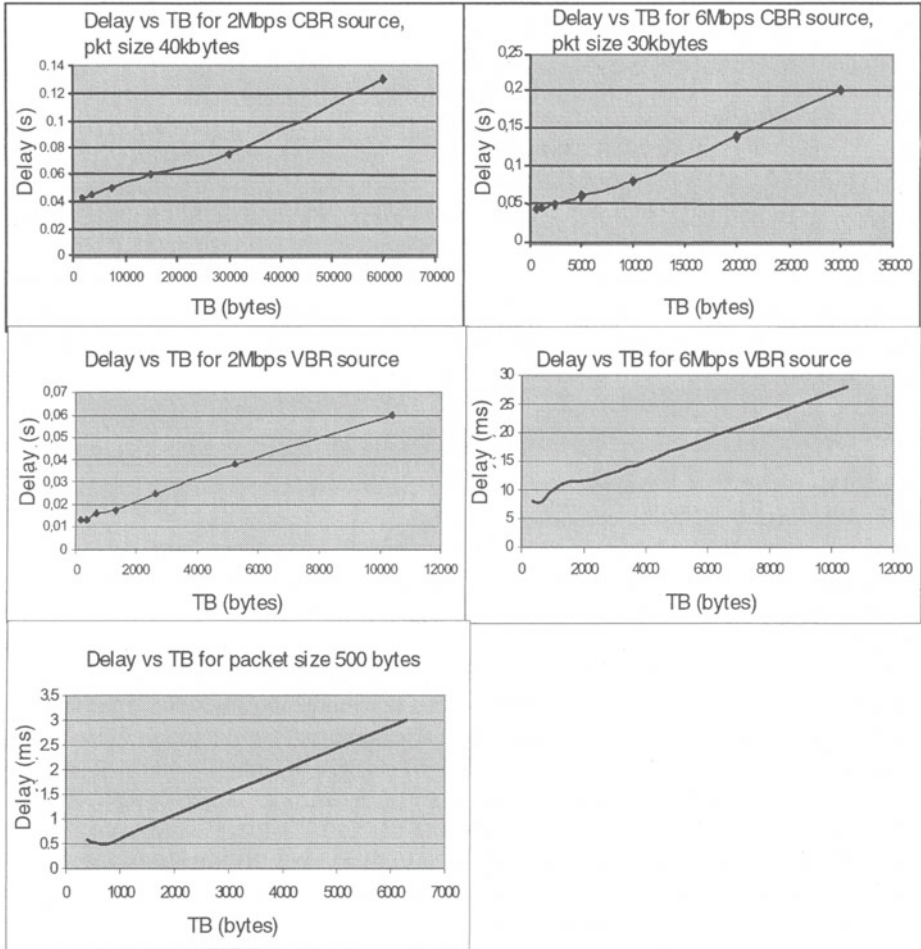

Fig. 5. Simulation results for testing TB and SDI

The sources employed in these five sets of scenarios were:
CBR: 25Mbps, packet size=500 bytes, interval =160µs,
CBR: 2Mbps, 10Kbytes every 40ms
CBR: 6Mbps, 30kbytes every 40ms

VBR: (mean rate 2Mbps) peak-rate=6Mbps, Ton=13ms, Toff=27ms, pkt size = 1300bytes

VBR: (mean rate 6Mbps) peak-rate=10Mbps, Ton=24ms, Toff=16ms, pkt size = 1300bytes

The results of these five sets of scenarios are shown in the figure 5.

It is obvious from the results that under all circumstances the system achieves better performance when the MAC serves each ONU with smaller TB, but more frequently (smaller SDI also). Even if longer packets take longer to have the last fragment reach the OLT, still the higher frequency of service ends in better performance. With smaller packets this is obviously more prominent. So we can conclude that the choice of TB and SDI does not depend upon complex decision but it can be based on a simple trade-off between delay performance and efficiency. Obviously for a high number of AllocIDs the overheads that accompany each upstream transmission make the choice of frequent service inefficient and thus indirectly affect the performance through the reduced net bandwidth available to payload. The other performance limit is of course dictated by the polling frequency, which affects the worst reaction time in a case of new arrivals (new arrivals can only be first reported with a polling allocation). Thus the choice of SDI should be the one that is just lower by the necessary margin of the worst case polling period required by the service and then the TB will be the one that meets the maximum rate. Any more frequent service will come at the cost of efficiency at no extra benefit while any less frequent will at times exceed delay limits.

6   The Impact of Polling

Dynamic Bandwidth Allocation is based on reporting to achieve high performance and utilisation at the same time. The reporting frequency affects the experienced delay since each packet that arrives at the ONU's data buffers needs a transmission permit to be sent in the upstream. The ONU waits for a permit to report its buffer status, which depends on the reporting frequency, sends upstream the report and then the OLT schedules the data permit. Hence, the reporting frequency has a strong impact on the worst-case delay, which is a critical parameter for the support of real-time applications.

The evident solution is to increase the reporting frequency regulated by the OLT, to minimise delay. However, such an action leads to efficiency degradation, since more often reports would travel announcing empty queues particularly with bursty sources.

To improve efficiency, the MAC also uses reports that travel piggy-backed in upstream frames carrying payload, in which case the SDI coincides with the reporting period. However, mainly because the value of the SDI affects the experienced delay, in the set of simulations run to demonstrate the impact of the reporting frequency in performance, the successive data interval and the reporting frequency are not assumed equal.

Table 2 presents the results from scenarios where both T-CONT 2 and 3 queues were activated and different traffic profiles were tested. The frequency that queues were

Table 1. The impact of the reporting frequency

Al-locID	Source type	T-CON T	Poll-ing Inter-val (frames)	Av. Delay (ms)	Number of Empty queue announce-ments	Max re-ported value
1	CBR	2	4	0.8	0.51%(20 of 3915)	48
2	CBR	2	8	1.5	0 of 2102	95
4	CBR	2	16	2.5	0 of 1135	161
6	Tri-modal/expo	2	4	0.85	1.19% (48 of 4022)	177
7	Tri-modal/expo	2	8	1.53	0.19% (4 of 2125)	192
9	Tri-modal/expo	2	16	2.6	1.18% (14 of 1190)	301
11	CBR	3	4	0.75	2.12% (86 of 4049)	48
12	CBR	3	8	1.13	0.19% (4 of 2122)	85
14	CBR	3	16	2	0 of 1208	170
16	Tri-modal/expo	3	4	0.83	3.74% (149 of 3980)	99
17	Tri-modal/expo	3	8	1.4	1.64% (35 of 2131)	99
19	Tri-modal/expo	3	16	2.3	0.27% (3 of 1092)	252

polled to report their lengths were either 4, 8 or 16 frames (remember frame duration is 125), and it was in general different to the service period (Max SDI).

In these runs, 20 queues with different service parameters were tested. CBR sources had a rate equal to 16 Mbps while sources with tri-modal traffic pattern a rate equal to 16.66 Mbps. So the system total offered load was about 600 Mbps. The traffic profile of the sources, the T-CONT type and the polling interval are given in columns 2, 3 and 4 respectively. The results include the average delay, the number of queue reports announcing empty queues and the maximum queue reported value, listed in columns 5, 6 and 7. The number of empty queue announcements allows for the estimation of efficiency degradation, which should be traded off with the performance ameliora-tion.

A first important result is that the delay observed by sources of the same profile and serviced at the same rate (successive data interval) is increasing as the polling period increases. For example, compare sources corresponding to AllocIDs 1, 2 and 4. Should the polling frequency continue increasing, the delay would become greater than 3ms, which is the delay budget for the access domain for real-time applications (e.g. telephony).

Comparing T-CONT 2 to T-CONT 3 queues of the same profile and service parame-ters, (for example source 1 and 11), they seem to have identical performance, demon-strating the impact of the reporting period (polling). Even in the case that the sources have different traffic profiles, still the average delay stays at the same levels. For example, sources 1 and 6 which are of CBR and tri-modal traffic pattern respectively, generate packets at the same mean rate and their average delays differ only by 0.05 ms. This small difference is due to the fact that T-CONT 3 reports empty queue more frequently than T-CONT 2, leading to some inefficiency. The fact that the reporting frequency is the main factor that affects system performance, can be also deduced from the similar difference in average delay between the cases of 4, 8 and 16 frames DBA reporting period, regardless of traffic generation by tri-modal or CBR traffic pattern and T-CONT 2 or T-CONT 3 traffic classes.

The reporting period for all the sources included in the table is chosen so that per-formance remains acceptable for real-time services. Any increase of the reporting period to more than 20 frame times leads to considerably higher delay values.

As regards efficiency, the 6th column shows that the higher the polling frequency and the burstier the source, the higher the percentage of reports announcing empty queues, hence wasting resources. This leads to the recommendation that for services that do not require strict delay guaranties (e.g. for T-CONT4 or some T-CONT3 queues), the polling period should be much greater than 16 frames, particularly since such source types are expected to be much more bursty.

7   Conclusions

The cost-effective multiplexing of a variety of traffic in a GPON relies upon a dynamic MAC protocol that allows support of many services with response matching the fluctuating demand. The delay performance is dominated by the polling period so for services with strict delay requirements frequent polling below 16 frames should be chosen. For the non-real-time services, efficiency dictates larger polling values. The performance evaluation based on computer simulations shows that the FSAN GPON can satisfy any mix of service classes thanks to its prioritized MAC service policy with quite satisfactory efficiency.

References

1. J.D. Angelopoulos, I.S. Venieris, G.I. Stassinopoulos, "A TDMA based Access Control Scheme for APON's," IEEE/OSA Journal of Lightwave Technology, Special Issue: Broadband Optical Networks, Vol. 11, No. 5/6, May/June 1993, pp. 1095-1103.
2. ITU G.983.1: "Broadband optical access systems based on passive optical networks (PON)".
3. ITU-T, G.983.4, Study Goup 15, "A Broadband Optical Access System with increased service capability using Dynamic Bandwidth Assignment".
4. A. Houghton, "Supporting the Rollout of Broadband in Europe: Optical Network Research in the IST Program", IEEE Communications Magazine, vol. 41, no 9, pp. 58-64, September 2003
5. G. Kramer, G. Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access Network", IEEE Communications Magazine, vol. 40, issue 9, pp. 66-73, February 2002