Analytical model for the IPACT dynamic bandwidth allocation algorithm for EPONs

Bart Lannoo,[1],* Lieven Verslegers,[1] Didier Colle,[1] Mario Pickavet,[1] Maurice Gagnaire,[2] and Piet Demeester[1]

[1]*Department of Information Technology (INTEC), Interdisciplinary Institute for Broadband Technology (IBBT)-Ghent University, G. Crommenlaan 8, 9050 Gent, Belgium*
[2]*Département Informatique et Réseaux, École Nationale Supérieure des Télécommunications (ENST), rue Barrault 46, 75634 Paris, France*
*Corresponding author: bart.lannoo@intec.ugent.be

We describe the research that has been performed in the field of the dynamic bandwidth allocation algorithm, interleaved polling with adaptive cycle time (IPACT), for Ethernet passive optical networks (EPON). The main focus has been on modeling packet delay analytically. To derive the packet delay, an important part of the analysis will also focus on the cycle time. It is assumed that the traffic load is symmetric, that packet arrivals are Poisson distributed, and that packets have fixed size. Simulations were performed to prove the accuracy of the analytical model. Some extensions and limitations of the model are treated, including asymmetric traffic load, packet size distribution, self-similar traffic, and differentiated services. © 2007 Optical Society of America

OCIS codes: 000.4430, 060.4250.

1. Introduction

Passive optical networks (PONs) are optical networks that do not contain active elements between source and destination but only passive optical components such as optical fiber, couplers, splices, and splitters. A PON can be deployed in several topologies, e.g., tree, ring, or bus. The most common, tree topology, will be the one discussed in the analysis. A PON is made up of an optical line terminal (OLT) connected to the optical networks units (ONUs) by optical fiber that splits up at a passive optical splitter. The OLT is located at the local exchange and connects the access to the metro backbone. The ONU can reside at the curb [fiber to the curb (FTTC)] or at the end-user location [fiber to the home or building and (FTTH or FTTB, respectively)]. Due to the use of passive components, a PON is multipoint to point (MP2P) in the upstream direction and point to multipoint (P2MP) in the downstream direction. Currently, there are three important PON standards: Ethernet PON (EPON), broadband PON [(BPON), based on the ATM PON (APON) standard], and gigabit PON (GPON). More detailed descriptions can be found throughout the literature; for example, in [1,2]. The focus of this paper is on the EPON.

Since in the upstream direction multiple ONUs share a common channel, an arbitration mechanism is necessary for upstream data transmission. In the case of EPON, such a mechanism is delivered by the multipoint control protocol (MPCP). It supports time slot allocation to the ONUs by the OLT. The MPCP is not concerned with a specific bandwidth allocation scheme; it provides a framework that is intended to facilitate the implementation of bandwidth allocation algorithms. Interleaved polling with adaptive cycle time (IPACT) is a possible dynamic bandwidth allocation (DBA) scheme, proposed by Kramer *et al.* [3].

In this paper, we present a thorough analysis of IPACT, which focuses on modeling cycle times and packet delay analytically. To our knowledge, this analysis is one of the first in this direction. In the past, Park *et al.* [4] and Bhatia *et al.* [5] already obtained interesting analytical results for the IPACT algorithm. However, their approaches and assumptions were different from the model presented in this paper. In Section 2, more details will be given about EPON, MPCP, and the IPACT bandwidth allocation scheme. The elaborate Section 3 deals with the developed analytical model and veri-

fies the obtained results by extensive simulations. In Section 4, some extensions and limitations of the model will be discussed. Section 5 concludes this paper.

2. EPONs and MPCP—IPACT

2.A. Ethernet Passive Optical Networks

As the name suggests, EPONs use Ethernet frames to encapsulate data. As with gigabit Ethernet, EPON has a nominal bit rate of 1.25 Gbits/s (physical layer). Due to $8B/10B$ encoding, it is effectively 1 Gbits/s (data layer). The Ethernet standard has been defined for both shared media and full-duplex point to point (P2P) links. Ethernet PON is in some respects a combination of both [6]: in the downstream direction, the EPON behaves as a shared medium (frames transmitted by the OLT reach all ONUs). In the upstream direction, due to the directional properties of the passive coupler, frames sent by an ONU only reach the OLT, not the other ONUs. This is comparable with P2P (or MP2P, if one considers all ONUs together). However, there is a difference with real P2P links: collisions can occur between traffic from different ONUs. Therefore, a protocol is required to manage the access to the network upstream.

2.B. Multipoint Control Protocol

The MPCP (developed by the IEEE 802.3ah task force and part of the EPON standard) is the arbitration mechanism, supporting time slot allocation to the ONUs by the OLT. The MPCP provides a framework, intended to facilitate the implementation of bandwidth allocation algorithms for the EPON, by providing signaling infrastructure (64 byte Ethernet control messages: GATE and REPORT) for coordinating upstream data transmission.

   The bandwidth allocation approach is time-division multiplexing (TDM): every ONU will be allowed to send its data in a specific time slot, according to the adopted algorithm. The REPORT message, sent upstream, informs the OLT about the state of the queues (containing packets ready for upstream transmission) at the ONU and can report up to 13 queue occupancies for one ONU (queues can be prioritized). Upon receiving a REPORT message, the OLT relies on the bandwidth allocation algorithm for determining the upstream transmission schedule. After the execution of the algorithm the GATE message, sent downstream, is used to issue transmission grants. A grant contains the transmission start time and the transmission length. To accomplish the scheduling of ONU transmission slots, a mechanism to synchronize distributed events to a central master counter (the OLT) is required. For this purpose, the control messages contain time stamps. Thanks to this mechanism of synchronization, the OLT can order ONUs to send their data so that transmission windows will not overlap.

2.C. Interleaved Polling with Adaptive Cycle Time

Several DBA algorithms exist and an overview can be found in [7]. They are not part of the EPON standard. IPACT [3] is the scheme that is treated further on. Interleaved polling with adaptive cycle time is a DBA scheme, in which the OLT polls the ONUs individually and issues grants to them in a round-robin fashion. The OLT keeps a polling table containing the number of bytes waiting in each ONU's buffer and the round-trip time (RTT) to each ONU. The OLT then sends a GATE message to an ONU to grant a transmission window allowing it to immediately send a certain amount of bytes. The transmission time of a GATE message is determined by taking the RTT to the concerned ONU and the transmission window of the previous ONU into account, so that packets from different ONUs do not overlap in time. In fact, transmission windows are only separated by a guard time, which provides protection for RTT fluctuations. At the end of a transmission window, an ONU reports its queue size(s) to the OLT by transmitting a REPORT message. The OLT uses this information to update its polling table and to determine the next granted transmission window.

   There exist several different services disciplines, i.e., ways for the OLT to determine the granted window size W_i (bits) for ONU i, depending on the requested window V_i (bits). The cycle times and the packet delay for the following three disciplines will be analyzed in detail.

- **Fixed service**. This service discipline ignores the window, requested by the ONU; instead, the OLT will always grant the maximum window to all ONUs. This causes the cycle time to be constant and maximal: $W_i = W_{MAX}$.
- **Gated service**. This approach imposes no limit on the size of the granted transmission window; the ONU is always authorized to send the amount of bytes it requested. This means that the cycle time will grow with the traffic load: $W_i = V_i$.
- **Limited service**. With this service discipline, an ONU is granted its requested transmission window, but not more than W_{MAX}. With this approach, the cycle time is variable, but it will not surpass a certain limit: $W_i = \min(V_i, W_{MAX})$.

Other possible service disciplines consist of trying to predict how many bytes an ONU will hold at the moment its transmission window begins. If the OLT manages to do so, all packets arriving in a cycle will be sent in the first transmission window (counting from their arrival). This way, one can decrease packet delay. The simplest approaches here are to add a constant credit to the requested window or to multiply the requested window size by a constant (linear credit). It is possible to think of much more complicated prediction mechanisms. However, it should be noted that overestimating will cause bandwidth to be lost.

3. Analytical Model

An EPON, consisting of an OLT and N ONUs, and using the DBA scheme IPACT to regulate the upstream bandwidth, is studied. In the analysis, packets are assumed to follow a Poisson arrival process with bit rate λ (Mbits/s) and to have a fixed size (B bits). Further on, traffic is assumed symmetric, i.e., traffic charge is the same for all ONUs (which are at the same distance from the OLT). The system to be modeled looks fairly simple at first sight. However, trying to capture all functionality and interdependencies into formulas will prove to be rather complex, and approximations will have to be made to allow numerical results to be obtained. Results will be verified with various simulations (in NS-2). Table 1 shows the parameters used throughout numerical calculations and simulations.

It is not sufficient to just concentrate on the end results (i.e., packet delays); examining intermediate processes often turns out to be helpful to understand the mechanisms that play a part in it. An important part will focus on the cycle times, which turn out to be much more complex than might have been expected. This cycle time is defined as the time between the start of two successive transmission windows for a fixed ONU. The cycle time analysis will then prove to be useful in predicting packet delays.

Due to the physical properties of optical fiber and lasers, transmission links are limited in bandwidth. For the EPON, this bandwidth is standardized at 1 Gbits/s (at the data link layer). Pushing the traffic load beyond this limit will cause the system to become instable. At that point, no meaningful results can be obtained anymore in simulations or numerical and analytical calculations. For gated service, cycle times will keep growing, causing time delays to increase as well. For limited and fixed service, buffers will not run empty an infinite number of times, also causing time delays

Table 1. Simulation Parameters

Symbol	Explanation	Value
N	Number of ONUs	16
λ	ONU arrival rate (Poisson traffic)	from 5 to 57.5 Mbits/s
T_{fiber}	Two-way delay on the EPON	200 μs
T_{proc}	Processing time	35 μs
T_{guard}	Guard time	1.5 μs
B	Packet size (network layer)	12,000 bits (=1500 bytes)
B_{eth}	Ethernet overhead	304 bits (=38 bytes)
B_{req}	REPORT or request message size	576 bits (=64 bytes+8 bytes preamble)
R_U	Upstream bandwidth on the EPON	1 Gbits/s
P_{max}	Maximum transmission window (fixed or limited)	10 packets

to increase continuously. The 1 Gbits/s bandwidth has to be shared amongst the N ONUs. This would mean that, if all ONUs have the same service level agreement, in a first approximation the bandwidth per ONU is equal to $1/N$ Gbits/s. However, one must take several sources of overhead into consideration, that cause the available bandwidth to be lower: the guard time T_{guard}, the time consumed by REPORT messages, and the Ethernet overhead.

An important remark to prevent confusion is the point of view that is taken throughout the analysis. Sometimes this is the system as a whole (the EPON with N ONUs), at other times this is the ONU. Generally speaking, for gated and limited service, the analysis of cycle times, queue sizes, and packet delays will take the ONU's viewpoint for lower traffic loads and the system's viewpoint for higher traffic loads. For fixed service, the ONU's viewpoint is taken. This distinction in approach should be clear from the use of λ_i for the ONU i's traffic load (Mbits/s) and Λ for the aggregate traffic load, logically calculated as: $\Lambda = \sum_{i=1}^{N} \lambda_i$ where N designates the number of ONUs in the EPON. In the majority of cases, all ONUs will be assumed to have the same traffic load λ, so that $\Lambda = N\lambda$.

3.A. Fixed Service
Remember that this service discipline ignores the window requested by the ONU. It always grants the maximum window to all ONUs and the cycle time will be constant and equal for all traffic loads,

$$T_{\text{cycle}} = \frac{NB_{req} + N(B + B_{eth})P_{\max}}{R_U} + NT_{\text{guard}}. \tag{1}$$

With the parameter values from Table 1, this results in a 2.0 ms cycle time. Since the cycle time is constant, the system can be considered at discrete moments that are T_{cycle} apart and located immediately after an ONU has sent packets in its granted transmission window. When deriving the packet delay, the sending of packets will be considered not to consume time, in a first approximation. An extra term will be ultimately added to compensate for this.

If $Q(n)$ is the queue size (packets) of an ONU at $t = nT_{\text{cycle}}$, then $Q(n)$ is a discrete homogeneous Markov chain, which means the next queue size [at $t = (n+1)T_{\text{cycle}}$] only depends on the past through the present (instant nT_{cycle}) and is also independent of n itself. This makes it possible to define transition probabilities and a transition matrix describing the system evolution in a probabilistic way,

$$p_{i,j} = \Pr[Q(n+1) = j | Q(n) = i], \tag{2}$$

$$P = \begin{pmatrix} p_{0,0} & p_{0,1} & & \cdots & & & p_{0,M} \\ p_{1,0} & p_{1,1} & p_{1,2} & & \cdots & & \\ p_{2,0} & p_{2,1} & p_{2,2} & & & \cdots & \\ \cdots & p_{3,1} & \cdots & \cdots & & & \\ 0 & & \cdots & p_{M-1,M-2} & p_{M-1,M-1} & p_{M-1,M} \\ 0 & 0 & & p_{M,M-2} & p_{M,M-1} & p_{M,M} \end{pmatrix}. \tag{3}$$

The transition probabilities are formulated below, and Poisson properties allow concrete formulas for these probabilities to be obtained

$$p_{i,0} = \sum_{k=0}^{P_{\max}-i} \Pr[N_{T_{\text{cycle}}} = k] = \sum_{k=0}^{P_{\max}-i} \exp\left(-\frac{\lambda}{B}T_{\text{cycle}}\right) \frac{\left(\frac{\lambda}{B}T_{\text{cycle}}\right)^k}{k!} \quad \text{for } i \leqslant P_{\max}, \tag{4a}$$

$$p_{i,j} = \Pr[N_{T_{\text{cycle}}} = P_{\max} + j - i] = \exp\left(-\frac{\lambda}{B}T_{\text{cycle}}\right) \frac{\left(\frac{\lambda}{B}T_{\text{cycle}}\right)^{P_{\max}+j-i}}{(P_{\max}+j-i)!} \quad \text{for } i \geqslant 0,$$

$$j > 0 \text{ and } j - i \geqslant -P_{\max}, \tag{4b}$$

$$p_{i,j} = 0 \quad \text{for } i,j \geqslant 0 \text{ and } j - i < - P_{\max}. \tag{4c}$$

It is only possible for a queue to send all its packets in one transmission window if the number of packets present is equal to or smaller than the maximum transmission window [Eq. (4a)]. The probability for such a transition is then given by the sum of the probabilities of having no more packet arrivals than the maximum transmission window minus the number of packets present [Eq. (4b)]. The probability of having a transition of i packets in the queue at an instant n to j packets at an instant $n+1$, is equal to having $P_{\max} + j - i$ arrivals in a period T_{cycle}, where $P_{\max}$ is the number of packets in a maximum transmission window [Eq. (4c)]. The probability of having a transition from i to j is zero, if j is smaller than $i - P_{\max}$, because the transmission window is restricted.

For an infinite buffer, there is no limit for the number of packets in the queue. For the analysis, this means that the transition matrix would have to be of infinite dimension. Because no results could be obtained in that way, the matrix's dimension has to be limited to some sufficiently large value M (set to 200 in the calculations). However, limiting the dimension of the matrix causes the sum of the elements of a line in the transition matrix no longer to be one, a necessary property for a Markovian matrix. This problem is solved by defining

$$p_{i,M} = 1 - \sum_{j=0}^{M-1} p_{i,j}. \tag{5}$$

To find the stationary distribution of queue sizes, a linear system of equations must be solved

$$\pi P = \pi, \tag{6}$$

$$\sum_{i=0}^{M} \pi_i = 1, \tag{7}$$

where π is the vector giving the probabilities of the queue occupancies 0 to M, in steady state. The average queue size $\hat{Q}$ at the discrete moments in time considered is given by

$$\hat{Q} = \sum_{i=0}^{M} \pi_i i. \tag{8}$$

To convert $\hat{Q}$ to the average queue size in continuous time two additive terms Q_1 and Q_2 must be adopted. Since Poisson arrivals are uniformly distributed over time (i.e., they are not bursty), the average queue size in continuous time will be equal to the average queue size exactly in between the discrete moments considered earlier (assuming the time it takes to send the packets is negligible). Therefore the first additive term is given by the average number of arrivals in $T_{\text{cycle}}/2$ s. A time T_{cycle} divided by two is considered on the assumption that the mean value of the average number of arrivals in one cycle time occurs in the middle of this cycle time.

$$Q_1 = \sum_{i=1}^{\infty} i \, \Pr(N_{T_{\text{cycle}}/2} = i) = \frac{T_{\text{cycle}}}{2} \frac{\lambda}{B}. \tag{9}$$

To be even more correct, one can add a second term, Q_2, that accounts for the time consumed by sending the packets:

$$Q_2 = \frac{\dfrac{\lambda}{B}}{\dfrac{R_U}{B + B_{eth}}} \times \frac{\dfrac{\lambda}{B} T_{\text{cycle}}}{2}. \tag{10}$$

This last formula can be interpreted as the percentage of time the ONU is sending packets multiplied by the mean value of the average number of packets that are to be sent in the transmission window. The division by two now indicates that this mean value is found in the middle of the transmission window. The term Q_2 is, generally
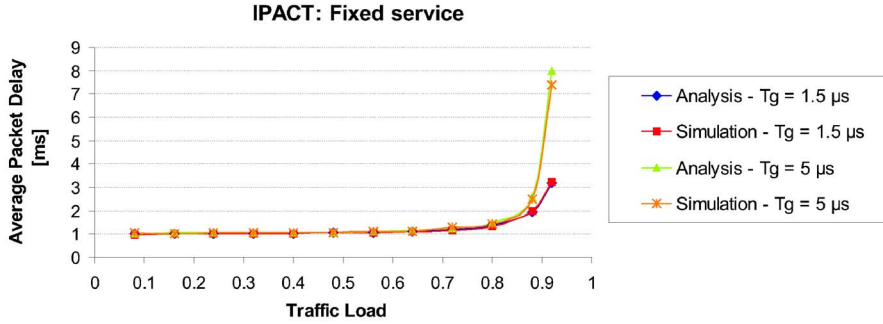
Fig. 1.   Average packet delay for fixed service: simulation–analysis comparison.

speaking, much smaller than the sum of $\hat{Q}$ and Q_1. The average queue size $\bar{Q}$ in continuous time is given by the following sum: $\bar{Q} = \hat{Q} + Q_1 + Q_2$.

The average time delay (s) follows from Little's law, which says that the average number of packets in a stable system (i.e., queue size $\bar{Q}$) is equal to their average arrival rate, multiplied by their average time in the system (i.e., the waiting time or time delay $\bar{W}$). This gives

$$\bar{W} = \frac{B\bar{Q}}{\lambda}. \tag{11}$$

Figure 1 shows that the results from simulation and analysis practically coincide for an EPON with 16 ONUs at 20 km from the OLT, a maximum transmission window of ten 1500 byte IP packets and a guard time of 1.5 and 5 μs. The load of every ONU is varied from 5 to 57.5 Mbits/s, which corresponds to a load ranging from 0.08 to 0.92, taking a maximum load per ONU of 62.5 Mbit/s (at the data link layer) into account. Due to several sources of overhead, the available bandwidth at the network layer will be somewhat lower. The only small discrepancies appear at very high traffic load, but even this can be explained by the fact that near instability, the system becomes more susceptible to small variations, due to the probabilistic nature of the Poisson traffic.

If the traffic load is sufficiently low, packets will most likely be sent in the next transmission window, which starts on average half a cycle (1.0 ms) later. This causes the first part of the graph to be nearly constant. For higher traffic loads, packets are more likely to wait for several cycles, so the average packet delay can surpass the cycle time. Comparing the difference in results for the two guard times leads to the conclusion that this parameter can improve the system's performance, in terms of packet delay, especially at high traffic loads. This is because lowering the guard time increases the stability limit. This, combined with the asymptotic behavior near instability explains the large improvement in packet delay.

3.B. Gated Service

For the gated service discipline, analysis becomes even more complex, due to the varying cycle times. Figure 2 gives an idea of this strong variation for a (high) 57.5 Mbits/s traffic load per ONU. The figure also clearly shows how there exist periods of shorter and longer cycle times. This leads to the conclusion that successive cycle times influence each other.

The analysis for gated service takes the cycle time as its starting point. A first approximation supposes the cycle time tends to a length in which as many packets arrive as can be sent

$$T_{\text{cycle}} = N\left(T_{\text{guard}} + \frac{B_{req}}{R_U}\right) + E[N_{T_{\text{cycle}}}]\frac{1B + B_{eth}}{R_U}. \tag{12}$$
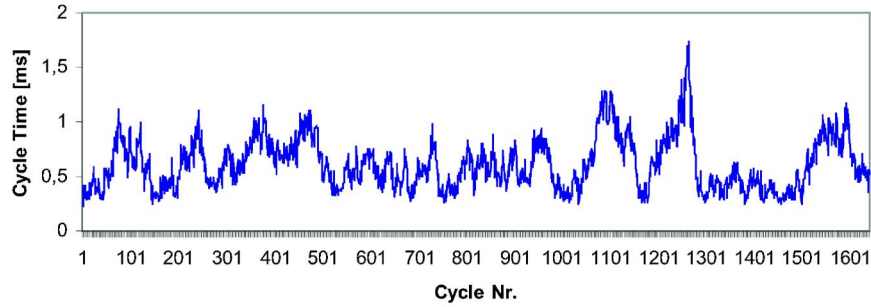
57.5 Mbits/s

Fig. 2.   Evolution of cycle times for gated service.

With the expected number of packet arrivals for Poisson traffic,

$$E[N_{T_{\text{cycle}}}] = T_{\text{cycle}} \frac{\Lambda}{B}.$$

Substituting this value in Eq. (12), and solving for T_{cycle} gives (note that the cycle time is proportional to T_{guard}):

$$T_{\text{cycle}} = \frac{N\left(T_{\text{guard}} + \dfrac{B_{req}}{R_U}\right)}{1 - \dfrac{\Lambda(B + B_{eth})}{R_U B}}. \tag{13}$$

By verifying Eq. (13) with simulations, it proves to be only partially correct. For high traffic loads (as from 45 Mbits/s), it predicts the average cycle time almost perfectly, whereas for lower traffic loads there is a big discrepancy. The declaration is that the cycle time cannot become lower than a minimum value, and ignoring this can lead to an underestimation of the cycle times, especially in the case of a low traffic load. It is inherent to IPACT that a GATE message cannot be sent before the previous REPORT message from the same ONU is received. This is because the grant needs information (requested window size) contained in the previous request. The minimum cycle time is thus determined by the distance between OLT and ONUs. As optical fiber has a refractive index $n = 1.5$, a 20 km fiber between ONU and OLT causes the cycle time to be at least $T_{\text{fiber}} = 200$ μs. To this value must still be added the time it takes to generate and interpret the REPORT and GATE message, called T_{proc}, in simulations chosen to be 35 μs. The absolute minimum for the cycle time is thus given by

$$T_{\text{cycle}}^{\min} = T_{\text{fiber}} + T_{proc}. \tag{14}$$

A more correct analysis distinguishes between low and high traffic load (such a distinction was also made during the analysis from Bhatia *et al.* [5]). A rule of thumb can be used to make this distinction. The probability of having more packet arrivals than can be sent in a minimum cycle time is given by

$$1 - \sum_{k=0}^{P} \exp\left(-\frac{\Lambda}{B} T_{\text{cycle}}^{\min}\right) \frac{\left(\dfrac{\Lambda}{B} T_{\text{cycle}}^{\min}\right)^k}{k!}. \tag{15}$$

Here P is the maximum number of packets that can be sent in a minimum cycle time:

$$P = \left\lfloor \left(T_{\text{cycle}}^{\min} - N\left(T_{\text{guard}} + \frac{B_{req}}{R_U}\right)\right) \frac{R_U}{B + B_{eth}} \right\rfloor. \tag{16}$$

If the probability given by Eq. (15) is lower than 5%, the analysis for low traffic is the most appropriate; for higher probabilities, the more complex analysis for high

traffic loads is more suitable. An approximation is proposed, which will also lead to results very close to the ones obtained through simulation.

3.B.1. Low Traffic Load

For low traffic load, the minimum cycle time (under traffic load) differs slightly from the physical minimum and can have different values in different cycles because it is also determined by P, the number of packets that are sent. Note that REPORT messages are sent at the end of the transmission window. If this were not the case, then the minimum cycle time under traffic load would be equal to the absolute minimum cycle time. Now, the minimum cycle time under traffic load becomes in a first approximation:

$$T_{\text{cycle}}^{\text{min}'} = T_{\text{cycle}}^{\text{min}} + \frac{P(B + B_{eth})}{R_U}. \tag{17}$$

One could have expected Eq. (17) to be correct, but it turns out that the ONUs that are polled right before a specific ONU also have their influence on the cycle time experienced by that ONU. What complicates analysis is the fact that ONUs can cluster, which causes ONUs to influence cycle times of ONUs that are polled successively. The clustering effect is caused by several (two or more) successive ONUs, sending out REPORTs to the OLT in a certain cycle that are only separated by a guard time. This phenomenon is especially important at low traffic loads, since such a succession of REPORT messages is a result of ONUs that have received a minimum transmission window during the considered cycle. It is then clear that the data sent upstream in the next cycle by the first ONU will have its effect on the start of the transmission window for the next ONUs. For instance, if the first ONU sends one packet, the transmission window for the second ONU will be delayed by the duration of sending one packet. Another consequence is that the different transmission windows normally begin a whole number of packets apart.

It is mainly the ONU and (in the case of clustering) the ONUs that are served before that ONU, that will determine the cycle time. The average cycle time can then be estimated as

$$\overline{T_{\text{cycle}}} = T_{\text{cycle}}^{\text{min}} + (1 + C)\frac{\lambda}{B}T_{\text{cycle}}^{\text{min}}\frac{B + B_{eth}}{R_U}. \tag{18}$$

This is the minimum cycle time plus the time necessary for sending the packets. The factor of C indicates the effects of clustering. Based on the explanation of clustering, C is defined as the average number of REPORT messages right before the considered ONU. An approximation for C can be derived by taking into account that the beginnings of the different transmission windows are normally a whole number of packets apart. The $N-1$ REPORTs from the other ONUs can only fall at $P+1$ places before the considered ONU. The clustering factor C is then the average of a binomial distribution with $N-1$ trials and a chance $1/(P+1)$ of success, and so it can be defined as $C=(N-1)/(P+1)$.

To calculate the packet delay, it is important to mention that a packet that arrives will not be sent in the first transmission window (counted from its arrival). Indeed, the ONU has to send a REPORT for the packets that have arrived during the cycle and wait for the GATE to arrive. Given the uniform distribution of Poisson arrivals (by uniform is meant not bursty), a packet arrives at the queue on average half-way between two transmission windows. Therefore on average it stays in the queue for one-and-a-half cycles. For low traffic loads, the cycle time shows only small fluctuations above its minimum value, which means that it suffices to know the average cycle time to derive the packet delay. A good approximation of the mean packet delay is given by

$$\bar{W} = \frac{3}{2}\overline{T_{\text{cycle}}}. \tag{19}$$

3.B.2. High Traffic Load

For high traffic load, the cycle time is most often determined by the aggregate traffic load, i.e., the traffic load of all ONUs together, with successive cycle times being cor-

related. The logical explanation is that, as the cycle becomes longer, the expected number of packet arrivals $(\Lambda/B)T_{\text{cycle}}$ increases as well. In this case, statistical properties of Poisson traffic allow an approximate distribution of cycle times to be derived. A first assumption, necessary for rendering the analysis feasible, is that the cycle time only takes the following discrete values:

$$T^m_{\text{cycle}} = \frac{m(B + B_{eth}) + NB_{req}}{R_U} + NT_{\text{guard}} \quad \text{for } m \geqslant 0.$$ (20)

Further, the physical limitation and the effect of the ONU's traffic load still have their influence and will be taken into account by considering the minimum cycle time $T^{\text{min}'}_{\text{cycle}}$ to be [with $\overline{T_{\text{cycle}}}$ the value obtained in Eq. (18), for low traffic load]:

$$T^{\text{min}'}_{\text{cycle}} = T^k_{\text{cycle}},$$

with

$$k = \min\left\{ m : \frac{m(B + B_{eth}) + NB_{req}}{R_U} + NT_{\text{guard}} > \overline{T_{\text{cycle}}} \right\}.$$ (21)

By consequence, the value of m in Eq. (20) has to be restricted to $m \geqslant k$. The basic idea for deriving the distribution is to abstract the evolution of cycle times as a discrete Markov chain. This makes it possible to define transition probabilities from one specific cycle time to another, and to make up a transition matrix P. For transitions to cycle times that are bigger than the minimum cycle time (under traffic load), the transition probability is given by

$$p_{i,j} = \Pr[T_{\text{cycle}}(n+1) = T^{j+k}_{\text{cycle}} | T_{\text{cycle}}(n) = T^{i+k}_{\text{cycle}}]$$

$$= \exp\left(-\frac{\Lambda}{B} T^{i+k}_{\text{cycle}} \right) \frac{\left(\frac{\Lambda}{B} T^{i+k}_{\text{cycle}} \right)^j}{j!} \quad \text{for } i \geqslant 0 \text{ and } j > 0.$$ (22a)

The probability of transition to the minimum cycle time (under traffic load) is logically given by

$$p_{i,0} = \Pr[T_{\text{cycle}}(n+1) = T^k_{\text{cycle}} | T_{\text{cycle}}(n) = T^{i+k}_{\text{cycle}}]$$

$$= \sum_{j: T_{\text{cycle}} \leqslant T^{\text{min}'}_{\text{cycle}}} \exp\left(-\frac{\Lambda}{B} T^{i+k}_{\text{cycle}} \right) \frac{\left(\frac{\Lambda}{B} T^{i+k}_{\text{cycle}} \right)^j}{j!} \quad \text{for } i \geqslant 0.$$ (22b)

Just as in the case of the queue size distribution for fixed services, a vector π giving the probabilities for the different cycle times in steady state, from 0 (i.e., minimum cycle time) to M can be calculated. The average cycle time for high traffic load is then easily obtained as

$$\overline{T_{\text{cycle}}} = \sum_{i=0}^{M} \pi_i T^{i+k}_{\text{cycle}}.$$ (23)

The distribution π can then be used to derive the average packet delay. In a first approximation, one could take the same approach as for low traffic load. However, this leads to serious underestimations, since for high traffic load, big fluctuations in cycle time take place. Knowledge of only the mean cycle time is then no longer sufficient. Supposing all cycle times have equal probability, a packet is more likely to arrive in a longer than in a shorter cycle. For this, one must consider that the probability that a packet falls in a cycle of a certain length is also proportional to this length. The probability for a packet to arrive in a cycle of duration T^{i+k}_{cycle} (again using the discrete approximation), is then

$$\widetilde{\pi}_i = \frac{\pi_i T^{i+k}_{\text{cycle}}}{\sum_{j=0}^{M} \pi_j T^{j+k}_{\text{cycle}}}.$$ (24)
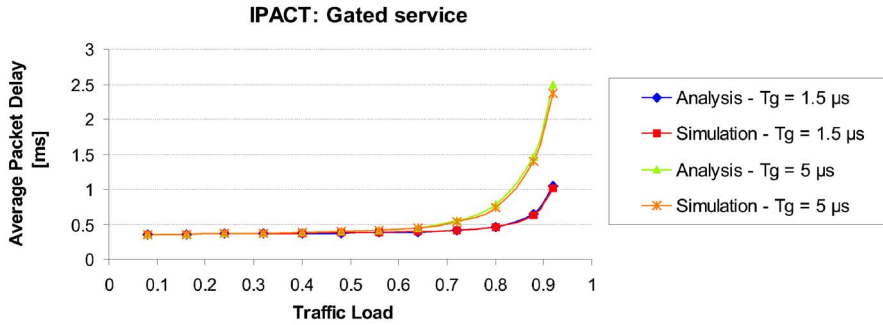
Fig. 3.   Average packet delay for gated service: simulation–analysis comparison.

The denominator normalizes the probability. With the same reasoning as the derivation of Eq. (19), and by using $\widetilde{\pi}_i$ instead of π_i in Eq. (23), one can estimate the average packet delay for high traffic load to be

$$\bar{W} = \frac{3}{2}\sum_{i=0}^{M} \widetilde{\pi}_i T_{\text{cycle}}^{i+k}. \tag{25}$$

Figure 3 shows how the analysis suits the simulated results, again for an EPON with 16 ONUs at 20 km, a guard time of 1.5 and 5 μs and a load varying from 5 to 57.5 Mbits/s for every ONU. The analysis predicts the packet delay well, but slightly overestimates for higher traffic loads, which is explained by the fact that the analysis is approximate. As for fixed service, one can see that for the first part of the graph, the packet delay increases very slowly; this is the domain determined by the ONU's traffic and by the traffic of the ONUs that are polled right before that ONU. In this domain, the average cycle time is still very close to its minimum value. For higher traffic loads, the aggregate traffic load becomes the determining factor and the packet delay increases quickly.

Comparing between different guard times leads to the conclusion that this parameter strongly influences the packet delay. The explanation follows naturally from the simplified model for the cycle time [Eq. (13)] and the fact that a packet stays on average one-and-a-half cycles in the queue.

3.C. Limited Service
Limited service shows some properties similar to fixed service and some similar to gated service. The cycle time for limited service can, in a first approximation, be analyzed in a way that resembles the gated service analysis. Since for limited service, ONU i's transmission window is limited to $P_{\max}$ packets, the cycle time, contrary to gated service, will not become larger than the following value:

$$T_{\text{cycle}}^{\max} = \frac{NB_{req} + NP_{\max}(B + B_{eth})}{R_U} + NT_{\text{guard}}. \tag{26}$$

Note that this is the cycle time for fixed service. In the analysis for gated service, the cycle time was only constrained to a sufficiently large value in order to be able to obtain numerical results. For limited service, it makes sense to limit the dimension of the transmission matrix to a value that corresponds to $T_{\text{cycle}}^{\max}$. A similar set of equations will have to be solved in order to calculate the approximate cycle time distribution, from which the average cycle time will follow.

However, for small maximum transmission windows (e.g., $P_{\max}=3$), there is a significant difference between the simulation and analytical result for the highest traffic load (Fig. 4). For this smaller transmission window in combination with a high traffic load, the probability of having more packet arrivals than can be sent in the transmission window is no longer negligible. If a packet cannot be sent in its first requested transmission window, it will have an influence over multiple cycle times. The analysis does not take this into account and there does not seem to be a feasible way of extending the analysis to include this feature. At lower loads, a model analogue to gated service gives fairly good results (as can be noticed in Fig. 4).

The cycle time analysis is no longer valid for traffic loads near instability. Consequently, an analysis of the packet delay, similar to the one for gated service, will fail
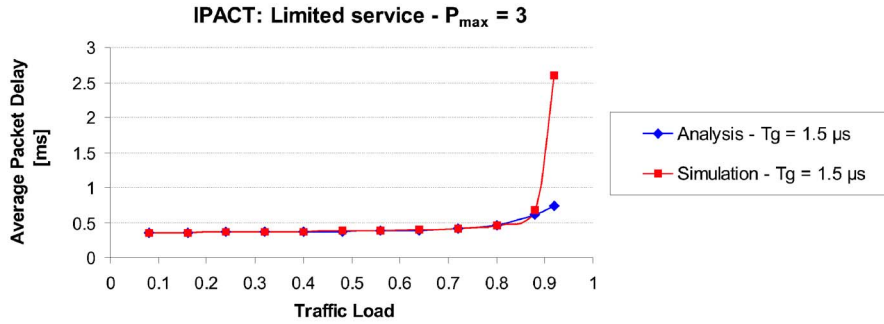
Fig. 4. Average packet delay for limited service: simulation–analysis comparison.

in predicting packet delays for high traffic loads. A reasonable idea seems to model the process of packets waiting for multiple cycles in the same way as for fixed service. However, in combination with the varying cycle times, this is no longer feasible.

Nonetheless, the fixed and gated service analysis provide the insight to understand limited service qualitatively. All processes together cause the system to become too complex for a complete numerical analysis similar to the fixed or gated service.

4. Extensions and Limitations of the Model

The above analytical model only studies simple traffic sources: symmetric load, Poison arrival rates, and fixed packet sizes. Possible extensions of this analytical model have been investigated, including asymmetric traffic load, a packet size distribution, and self-similar traffic. Finally, differentiated services are also briefly mentioned.

For asymmetric traffic load, the fixed service analysis still applies since the transmission window does not depend upon its own traffic load or on the traffic load of the other ONUs. Consequently, all formulas and results still apply.

For gated service, the situation is more complicated and the analysis cannot merely be repeated. Consider, for instance, the case of all ONUs having the same packet arrival rate of 55 Mbits/s, except for one tagged ONU, for which the traffic charge is varied. Based on Eq. (15), the analysis for high traffic analysis has to be applied in this example. Further, the aggregate traffic load has to be adapted to $\Lambda = \lambda_1 + (N-1) \times 55 \times 10^6$. To estimate the minimum cycle time in the high traffic load analysis, the average cycle time can reasonably be assumed to be

$$\overline{T_{\text{cycle}}} = T_{\text{cycle}}^{\min} + (1+C)\frac{\Lambda}{NB}T_{\text{cycle}}^{\min}\frac{B+B_{eth}}{R_U}. \qquad (27)$$

Even though serious approximations are made throughout the analysis, the results still show a fairly good prediction, as depicted in Fig. 5. So the gated service analysis still has its value for asymmetric traffic. The arrival rate for the tagged ONU was varied from 10 to 100 Mbits/s. For the whole EPON, this corresponds to a traffic load only varying from 0.835 to 0.925.

An analysis that includes packet size distribution is difficult for fixed service, due to an unused remainder of the transmission window. For gated service, this problem
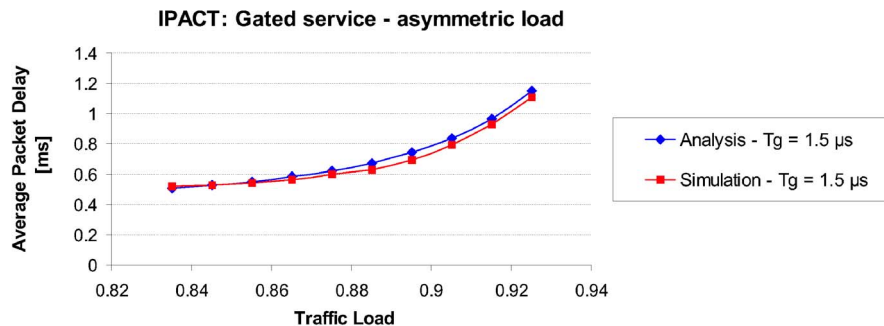

Fig. 5. Gated service and asymmetric traffic: simulation–analysis comparison.

does not exist, and fairly good results can be obtained by repeating the analysis from Subsection 3.B with an average packet size.

Other types of traffic, such as self-similar traffic, are too complex for analytical methods to be used. Because of the much more unpredictable behavior for, e.g., the queue sizes and cycle times, an analysis similar to the one performed for Poisson traffic is no longer possible. Although self-similar traffic is much more common in access networks, our analytical model based on Poisson traffic still provides us with a lot of useful insights for the general performance of IPACT.

To end this section, note that EPONs are essentially conceived to support differentiated services: voice communications, standard and high-definition video, video conferencing, and data traffic. Because MPCP is part of the IEEE 802 family of standards, intra-ONU scheduling is by default strict priority scheduling. This means that packets corresponding to a certain class will only be sent if no packets are present in queues corresponding to higher priority classes.

Differentiated services are only of importance together with the fixed and limited service disciplines. For gated service, packets are sure to be sent in their first requested transmission window, and, therefore, priority scheduling would only put higher priority packets up front in the window. For fixed service, an analysis that logically follows from the model performed in Subsection 3.A can be formulated. It will consider, e.g., separated packet arrival rates per service class and take into account that the highest priority packets are the first ones to be served in the transmission window. Finally, for limited service a similar analysis will again be infeasible.

5. Conclusions

The EPON with MPCP-IPACT proves to be a very complex system to analyze, yet it still allows the packet delay to be derived in the case of fixed and gated service discipline and Poisson traffic with constant size packets. This analysis also turns out to be valuable in more general cases.

The importance of a good understanding of the cycle time is shown. For fixed service, an approach is suggested in which queue sizes are considered at discrete moments in time, resulting in an analytical packet delay that corresponds to simulation. For gated service, a distribution of cycle times is derived. This proves useful when determining the packet delay. Again simulation and analysis match well. For limited service, it is shown how a similar analysis is not feasible. However, knowledge of fixed and gated service provides additional insight.

The model is also investigated to see if and how it can be extended to include asymmetric traffic, packet size distribution, self-similar traffic, and differentiated services. This raises the key question of to what extent mathematics can capture reality, and when the quantitative analysis should be replaced by just qualitative analysis: at a certain point, a mathematical analysis becomes infeasible and needs to be replaced by simulation.

Acknowledgment

References and Links

1. M. Gagnaire, *Broadband Local Loops for High-Speed Internet Access* (Artech House, 2003).
2. D. Gutierrez, K. S. Kim, S. Rotolo, F. An, and L. G. Kazovsky, "FTTH standards, deployments and research issues," presented at the 8th Joint Conference on Information Sciences, Salt Lake City, Utah, 21–26 July 2005.
3. G. Kramer, B. Mukherjee, and G. Pesavento, "Interleaved polling with adaptive cycle time (IPACT): a dynamic bandwidth distribution scheme in an optical access network," Photonic Network Commun. **4**, 89–107 (2002).
4. C. G. Park, D. H. Han, and B. Kim, "Packet delay analysis of dynamic bandwidth allocation scheme in an Ethernet PON," in *Proceedings of the 4th International Conference on Networking* (Springer, 2005), pp. 161–168.
5. S. Bhatia, D. Garbuzov, and R. Bartos, "Analysis of the gated IPACT scheme for EPONs," in *Proceedings of the IEEE International Conference on Communications* (IEEE, 2006), pp. 2693–2698.
6. G. Pesavento, "Ethernet passive optical network (EPON) architecture for broadband access," Opt. Networks Mag. **4**, 107–113 (2003).
7. M. P. McGarry, M. Maier, and M. Reisslein, "Ethernet PONs: a survey of dynamic bandwidth allocation (DBA) algorithms," IEEE Commun. Mag. **42**, 58–515 (2004).