Energy Efficient PONs with Service Delay Guarantees

L. Valcarenghi*, M. Chincoli†, P. Monti‡, L. Wosinska‡, and P. Castoldi*
*Scuola Superiore Sant'Anna, Pisa, Italy
†CNIT, Pisa, Italy
‡KTH, Kista, Sweden

***Abstract*—Passive Optical access Networks (PONs) are currently the major contributor to the energy consumption budget of fixed optical networks. In PON, the largest part of the energy consumption is due to the equipments at the customer premises.**

This paper proposes a method for maximizing energy savings while providing services with delay guarantees (i.e., frame delivery time and frame delay variation). The method combines service-based variable sleep period and a queueing theory model to compute the optimal sleep time. Simulation results prove the effectiveness of the method for a Poisson frame arrival process.

***Index Terms*—Energy saving, PON, service, delay, guarantees.**

I. Introduction

The increasing number of broadband network subscribers results in massive deployment of fiber access networks, where Passive Optical Networks (PON) is a dominating architecture due to some attractive features, such as relatively low power consumption, cost and resource efficiency. Customer premises equipments (i.e., Optical Network Units — ONUs) of PONs are contributing with over 65% of the overall PON power consumption [1]. Therefore, they are currently considered as the main research target for improving PON energy efficiency. Standardization authorities, industries and researchers are proposing several methods for decreasing ONU energy consumption [2]–[7], including a method based on cycles of sleep and waking periods, called *cyclic sleep* or *fast sleep*. However, cyclic sleep, as well as the other proposed methods, might cause an increase in the delay experienced by the transmitted data, which can be critical for the Quality of Service (QoS) requested by the users. On the other hand the QoS degradation can be minimized by an optimal scheduling of sleeping periods according to the supported services.

With this aim, the concept of service-based variable sleep period was introduced in [8] to save energy in ONUs while avoiding service degradation in the presence of traffic with demanding requirements (e.g., regarding delay). According to this approach, each class of service (CoS) is assigned a specific sleep period. If multiple CoS are received by one ONU, the sleep period of the most demanding class is selected. In [9], instead of varying the sleep period based on service information, the sleep length is computed as a function of the estimated frame interarrival time. In [9] two methods of estimating frame inter-arrival time are combined with two strategies for choosing the length of the sleep period as a function of the estimated inter-arrival time. The experimental evaluation of the four proposed methods showed good energy savings but the methods in which sleep period is not bounded suffered from some latency drawbacks. In [10] an analytical model to estimate the energy consumption of an Energy Efficient Ethernet link (i.e., a link in which IEEE 802.3az is implemented) is presented. The model is based on M/G/1 queue with server timeout and activation time. In [11] a method is proposed to compute the maximum sleep time when cyclic sleep mode is implemented in the ONU while guaranteeing service delay requirements. The method is based on modeling the upstream data transmission by means of a M/G/1 queue with vacation. Results show that the method is capable of achieving a limited average delay while maximizing the sleep time and, consequently, energy saving.

This paper focuses on service-based variable sleep period because, in general, users connected to ONUs may dynamically modify their service subscriptions (e.g., request a video streaming to watch a movie for some hours), possibly changing the overall delay requirements. However such changes have a period of hours, therefore sleep time does not need to be often updated. The method to compute the sleep time value is based on the predicted delay statistics that frames will experience. Such prediction is made according to a queueing model of the implementation of cyclic sleep in the PON. The objective of the method is to maximize the energy savings while providing delay-related guarantees to the services subscribed by the ONU. Thus, differently from [9], the sleep time is a function of services to which the ONU subscribed rather than of the estimated packet inter-arrival time. Moreover, stemming from [8], the proposed method more strictly relates the sleep time dependence to the service class delay requirements and to the predicted delay statistics rather than to the downstream frame statistics. Furthermore, differently from [10], the sleep time is not fixed but computed based on the proposed model and function of the service subscriptions. Finally, differently from [11] it considers not only the average delay but also the delay variation (i.e., the jitter) as a Quality

978-3-901882-46-3

of Service (QoS) constraint.

Our results show that the proposed cyclic sleep with service-based variable sleep period is effective in guaranteeing average frame delivery time. Moreover, by utilizing a simplified model to compute the fame delay variation, the method succeeds also in providing services with delay variation guarantees.

II. Cyclic sleep with service-based variable sleep period

In this section the system description and its model are outlined. In particular the behavior of the proposed cyclic sleep with service-based variable sleep period is detailed.

A. System description

A Time Division Multiple Access (TDMA) PON is considered. The choice of switching an ONU to sleep is based on downstream traffic only. Upstream traffic scheduling is not taken into account but it can be locked to downstream scheduling. The ONU implements a slightly modified version of the Sleep and Periodic Wake-up (SPW) (i.e., fast sleep) [8] whose behavior is depicted in Fig. 1 and Fig. 2. Sleeping cycles are triggered by the Optical Line Terminal (OLT) by means of `Request` messages. `Request` messages carry a field, called sleeping time T_{sl}, whose value is the time for which the ONU is set to sleep. Every `Request` message is always acknowledged by the ONU by means of an `ACK` message. When the T_{sl} period expires (i.e., $T_{sl} = 0$), the ONU wakes up and sends a `Confirmation` message to the OLT to check whether it can sleep for one more cycle. When the ONU is asleep, downstream frames destined to that ONU are buffered by the OLT. Upon `Confirmation` reception, the frames that arrived to the OLT before the `Confirmation` message are sent to the ONU. Then, the OLT sends to the ONU a new `Request` message.

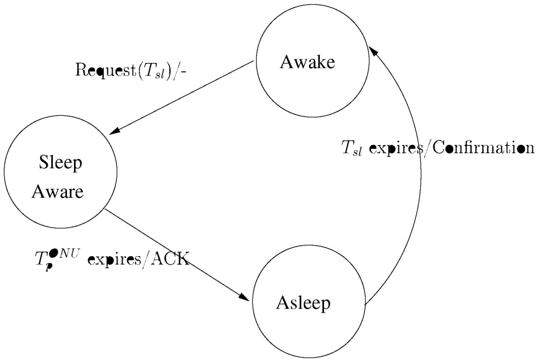

Figure 1. Modified Sleep and Periodic Wake-up ONU Finite State Machine (FSM) state transition diagram (condition/action).

The main differences between the modified SPW proposed in this paper and the SPW proposed in [8] are the decision of when the ONU is switched to sleep and the duration of the sleeping time T_{sl}. The decision to switch the ONU to sleep is not based on a downstream frame inter-arrival time threshold as in [8]. The ONU always alternates a period of activity T_a and a sleep period T_{sl}. The period of activity T_a of the ONU is used for exchanging `Confirmation`, `Request`, and `ACK` messages. Furthermore, during T_a, all the frames arrived to the OLT before the reception of the ONU `Confirmation` message (e.g., frames $n-2$, $n-1$, and n in Fig. 2) are sent to the ONU. This latter operation requires a variable time T_{DS}.

The value assigned to T_{sl} is not constant but it is based on the services the ONU subscribed. It is assumed that the OLT is constantly aware of the services each ONU subscribed. Every service is categorized into a specific CoS that is characterized by specific absolute bandwidths (i.e., data rates) and delay constraints. The CoS can be derived, for example, from ITU-T standards such as [12] and [13]. By leveraging this information, the OLT dynamically builds a table containing the bandwidth and delay requirements per each registered ONU. The bandwidth and the delay requirements depend on the services the ONU subscribed. If the set of services subscribed by the ONU changes, the T_{sl} is also changed by the OLT.

Although this paper assumes that T_{sl} is carried by `Request` messages, the value of T_{sl} could be notified to the ONU through management and control interface layer (OMCI) messages (defined in ITU-T Rec. 984.4 for Gigabit-capable PON — GPON — and ITU-T Rec. G.988 for 10 Gigabit-capable PON — XG-PON) because service subscription periods are typically long. In this way, `Request` messages need not carry the value of T_{sl}.

B. System Model

If frame interarrival times are exponentially distributed and the queue at the OLT is infinite, the system OLT-ONU with the modified SPW can be modeled as a polling system with gated service, exponentially distributed frame interarrival time, and generic service time [14].

In a PON with the modified SPW, it is the ONU that polls the OLT for downstream frames. Polling times correspond to the instants at which `Confirmation` messages reach the OLT. At these instants the ONU is up and, if queued frames exist at the OLT, they are transmitted to the ONU. The frame transmission is gated because only frames that arrived to the OLT before the `Confirmation` message can be transmitted to the ONU. Then, upon exchange of `Request` and `ACK` messages the ONU goes back to sleep for a period T_{sl} after which the OLT is polled again. In the polling system with gated service [14] the following quantities are defined: the average client service time $\overline{S}$, the second moment of the service time $\overline{S^2}$, the average reply interval (i.e., the time needed to switch service from one station to the other) $\overline{V}$, the second moment of the reply interval between stations $\overline{V^2}$, the client average arrival rate λ, the number of stations to be served in the system N. The average time spent by a client in the queue $E[W_q]$ is expressed by:

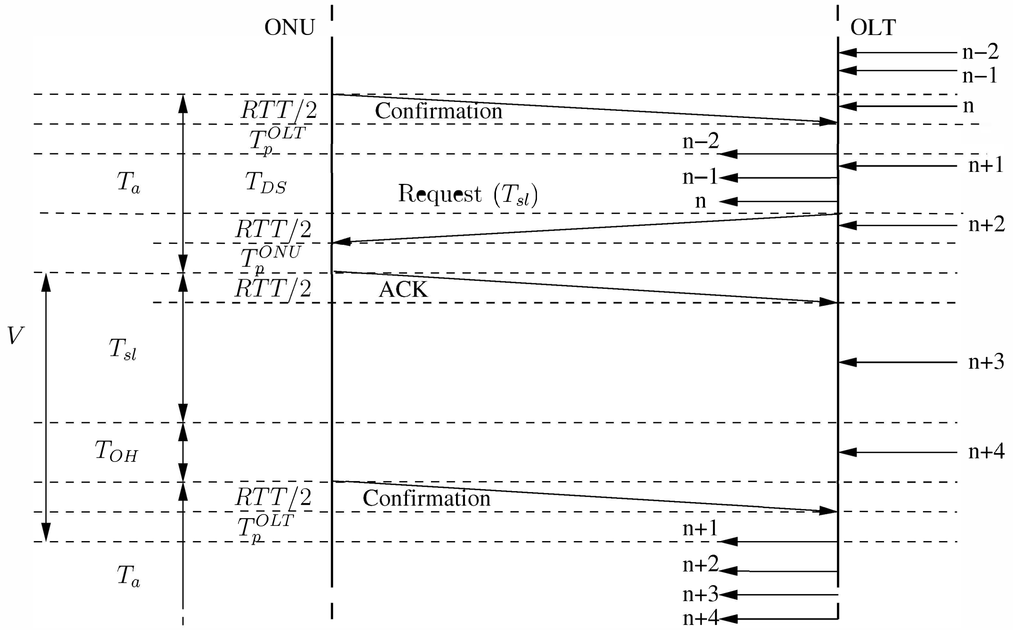

Figure 2. Modified Sleep and Periodic Wake-up.

$$E[W_q] = \frac{\overline{V^2} - \overline{V}^2}{2\overline{V}} + \frac{N\overline{V}(1+\lambda\overline{S})}{2(1-N\lambda\overline{S})} + \frac{\lambda\overline{S^2}}{2(1-N\lambda\overline{S})} \qquad (1)$$

In this paper a simplified polling system with one station and one server is considered because only one ONU and one OLT are assumed. The following mapping between the single OLT-ONU system and a polling system is performed: the ONU is the server, the OLT is the station, the downstream frame is the client, the service time is the time required to transmit a frame and its statistics depend on the frame size distribution, the reply interval is a function of T_{sl}. The reply interval corresponds to the time between when the OLT terminates sending downstream frames to the ONU (i.e., when the ONU receives and processes the `Request` message that terminates downstream frame transmission and sets the ONU to sleep) and when the OLT is polled again by the ONU to request whether it can sleep longer (i.e., when the OLT receives and processes the `Confirmation` message sent by the ONU). In the OLT-ONU system the reply interval V is defined as the ONU idle time, that is the time for which the ONU does not receive downstream frames. As shown in Fig. 2, the ONU idle time is the sum of the ONU sleep time T_{sl}, the PON round trip time RTT, the overhead time T_{OH} (i.e., the time required to the ONU to wake up from the sleep state), and the control frame processing time at the OLT T_p^{OLT} and at the ONU T_p^{ONU}. If, for simplicity, the processing times are assumed to be negligible, the average idle time is constant (because all of its components are constant) and it can be expressed as $\overline{V} = T_{sl} + T_{OH} + RTT$. Therefore Eq.(1) can be rewritten as:

$$E[W_q] = \frac{\overline{V}(1+\lambda\overline{S})}{2(1-\lambda\overline{S})} + \frac{\lambda\overline{S^2}}{2(1-\lambda\overline{S})}, \qquad (2)$$

that represents the average time spent by a frame in the OLT queue.

For scenarios in which the system load is small (i.e., $\lambda\overline{S} << 1$) and $\lambda\overline{S^2} << 1$, Eq.(2) can be approximated as:

$$E[W_q] \approx \frac{\overline{V}}{2} \qquad (3)$$

For example, in a scenario with constant size frames (e.g., 1250 bytes), 10Gb/s transmission rate, $\lambda = 62.5$ frame/s (i.e., 625 kb/s) arrival rate, $T_{sl} = 1$ ms, $T_{OH} = 2$ ms, and $RTT = 0.4$ ms the two contributions to Eq.(2) are about 1.7x10^{-3} and 3.13x10^{-11} respectively. Therefore $E[W_q]$ is approximately 1.7x10^{-3}.

In such cases, it is possible to write:

$$E[W_q] \approx \frac{\overline{V}}{2} = \frac{T_{sl} + T_{OH} + RTT}{2}. \qquad (4)$$

To compute the energy efficiency η provided by the modified SPW with service-based variable sleep period the following procedure is followed. Based on Little's theorem, the average number of frames in the OLT queue $E[N_q]$ can be expressed as:

$$E[N_q] = \lambda E[W_q]. \qquad (5)$$

Because of the gated service behavior of the OLT, $E[N_q]$ is also the average number of frames sent during T_{DS}. Therefore, the average time spent to send the queued frames $E[T_{DS}]$ is:

$$E[T_{DS}] = \lambda\overline{S}\left(\frac{\overline{V}(1+\lambda\overline{S})}{2(1-\lambda\overline{S})} + \frac{\lambda\overline{S^2}}{2(1-\lambda\overline{S})}\right). \tag{6}$$

Under the same assumption utilized for simplifying Eq.(2), it is possible to write:

$$E[T_{DS}] \approx \lambda\overline{S}\frac{\overline{V}}{2} = \lambda\overline{S}\frac{T_{sl}+T_{OH}+RTT}{2}. \tag{7}$$

As depicted in Fig. 2 the time for which the ONU is “on” T_{ON} can be written as:

$$T_{ON} = T_a + T_{OH}, \tag{8}$$

disregarding the processing times at the OLT and at the ONU. Because $T_a = E[T_{DS}] + RTT$, Eq.(8) can written as:

$$\begin{aligned} T_{ON} &= \lambda\overline{S}\frac{T_{sl}+T_{OH}+RTT}{2} + RTT + T_{OH} = \\ &= \frac{\lambda\overline{S}T_{sl} + (\lambda\overline{S}+2)(RTT+T_{OH})}{2} \end{aligned} \tag{9}$$

The energy efficiency η achieved by implementing the modified SPW with service-based variable sleep period can be written as:

$$\begin{aligned} \eta = 1 - \frac{E_{sl}}{E_{nosl}} &= \frac{P_{ON}T_{ON} + P_{sl}T_{sl}}{P_{ON}(T_{ON}+T_{sl})} \\ &= \frac{(\lambda\overline{S}+2\alpha)T_{sl} + (\lambda\overline{S}+2)(RTT+T_{OH})}{(\lambda\overline{S}+2)(T_{sl}+RTT+T_{OH})}, \end{aligned} \tag{10}$$

where E_{nosl} is the energy consumed by the ONU when sleep mode is not implemented, E_{sl} is the energy consumed by the ONU when the modified SPW with service-based variable sleep period is implemented, and $\alpha = P_{sl}/P_{ON}$ is the power ratio, that is the ratio between the power consumed when the ONU is alseep and when the ONU is “on”.

If subscribed services present also delay variation constraints, the IP packet Delay Variation (IPDV) must be considered as well. Although this paper deals with Ethernet frames and Frame Transfer Delay (FTD) and Frame Delay Variation (FDV) are defined in [15], constraints on IP packet Transfer Delay (IPTD) and IP packet Delay Variation (IPDV) are specified in [13] for IPTD and IPDV but not for FDV and FTD in [15]. Therefore, in this paper IPTD and IPDV are considered equivalent to FTD and FDV.

The definition of the IPDV utilized in this paper is the one reported in [13]:

$$IPDV = IPTD_{upper} - IPTD_{min}, \tag{11}$$

where $IPTD_{upper}$ is the $1-10^{-3}$ quantile of the IPTD and $IPTD_{min}$ is the minimum IPTD. Based on the system description as depicted in Fig. 2, the values of $IPTD_{min}$ and $IPTD_{upper}$ can be approximated as:

$$IPTD_{min} = \overline{S} + RTT/2 \tag{12}$$

$$\begin{aligned} &IPTD_{upper} \approx IPTD_{max} = \\ &= T_{DS,max} + RTT + T_{sl} + T_{OH} + \overline{S} + RTT/2. \end{aligned} \tag{13}$$

Eq.(12) is obtained by considering that a frame arriving just before the arrival of the `Confirmation` message at the OLT, will be transmitted immediately, if no frames are awaiting in the OLT queue. In Eq.(13) $IPTD_{max}$ is the delay experienced by a frame arriving just after the arrival of the `Confirmation` message at the OLT. Such frame will be the first frame to be transmitted during the next ONU period of activity but it will need to wait for the ONU to wake up. Based on Eq.(12) and Eq.(13) $IPDV$ can be written as:

$$\begin{aligned} IPDV &= IPTD_{upper} - IPTD_{min} = \\ &= T_{DS,max} + RTT + T_{sl} + T_{OH} \end{aligned} \tag{14}$$

In case the system load is small (i.e., $\lambda\overline{S} << 1$), $T_{DS,max}$ contribution can be neglected and therefore Eq.(14) becomes:

$$IPDV \approx RTT + T_{sl} + T_{OH} \tag{15}$$

C. Sleep Time Calculation

The proposed cyclic sleep with service-based variable sleep period utilizes the model presented in section II-B to compute the sleep time. The calculation is also based on the specific delay constraints requested by the services subscribed by the ONU.

Table I
Delay requirement and data rate of typical services.

Class ID	Service Type	QoS Class	IPTD Delay [ms]	IPDV [ms]	Data Rate B_i [b/s]
1	Web Browsing	5	U	U	30.4k
2	Internet Relay Chat	3	400	U	1k
3	Multimedia on Web	4	1000	U	28.8k-500k
4	Voice over IP	0	100	50	5.3k-64k

Tab. I reports delay requirements specified in [13] for different QoS classes. The reported values are upper bounds for the aforementioned parameters considering end-to-end IP packet transfer [1]. Although such values are specified for IP packets, similar performance parameters are specified for Ethernet frames, such as Frame Transfer Delay (FTD)

---

[1] The letter U means unspecified. In this situation a value of 1000 ms is considered in the paper.

and Frame Delay Variation (FDV) [15] but bounds are not provided. Tab. I also maps some of the services defined in [12] for IP-based networks into the classes defined in [13] and specifies typical service data rates.

The sleep time T_{sl} computation for an ONU is based on the equations relating it to the expected frame transfer delay, proportional to $E[W_q]$, or the FDV. Its value is impacted by the number and the type of services subscribed by the ONU and on the service delay constraints. Moreover, in the T_{sl} computation, some standardized PON timeouts must be considered. For example, in EPON, the T_{sl} value must be bounded by the time after which the ONU is deregistered: 1s as specified in IEEE-802.3-2008 page 288 (`mpcp_timer`, `mpcp_timeout`). In XG-PON, T_{sl} must be bounded by the value of $Isleep$ to which the timer $Tsleep$ (i.e., the time for which the ONU remains in sleep mode) is initialized. $Isleep$ is 4 bytes long and it counts units of 125μs. Therefore the $Isleep$ maximum value is 149 hrs (see page 98 of ITU-T G.987.3, page 76 of ITU-T G.988, and [16]). Additionally the T_{sl} value must consider the value of the OLT timer Ter_i (see ITU-T G.987.3 page 91) that is the latest instant at which an upstream burst is expected from sleeping or dozing ONU i before the OLT attempts to force the ONU i to wake up. Finally, another timer to be considered is the Loss of Burst counter $Clob$ that is the counter of missing upstream bursts at the OLT (see page 98 of ITU-T G.987.3). If $Clob$ expires the OLT declares Loss of Burst (LoB) against the ONU.

If the IPTD and the minimum guaranteed bandwidth are the only considered constraints, Fig. 3 shows the algorithm utilized to compute the sleep time for an ONU. For each service j subscribed by the ONU, the OLT computes the value of the average frame arrival rate λ_j as:

$$\lambda_j = \frac{B_j}{8 f_s}, \tag{16}$$

where B_j is the minimum bandwidth requested by service j and f_s is the average frame size. Because of the assumption of exponential frame interarrival times the overall frame interarrival rate is $\lambda = \sum_{j=1}^{S} \lambda_j$, where S is the set of services subscribed by the ONU. Moreover the maximum tolerable IPTD $IPTD^{max}$ is computed as the minimum among the IPTD constraints of the services subscribed by the ONU:

$$IPTD^{max} = min_{j \in S} IPTD_j. \tag{17}$$

Then the maximum T_{sl} based on IPTD constraints T_{sl}^{IPTD} is computed by substituting $E[W_q] = IPTD^{max} - \overline{S} - RTT/2$ in Eq.(4) as per [17] and by solving it for T_{sl}:

$$T_{sl}^{IPTD} = 2(IPTD^{max} - \overline{S} - RTT) - T_{OH} \tag{18}$$

The computed T_{sl}^{IPTD} value is transmitted by the OLT to the ONU. Such procedure is recalled every time an ONU (un)subscribe a service.

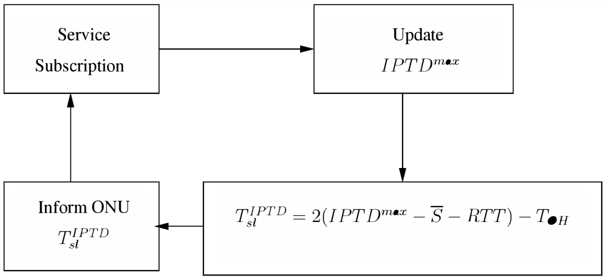

Figure 3. Service-based variable sleep period calculation at OLT

If subscribed services present also IPDV constraints, the computation of T_{sl} based on IPDV constraints T_{sl}^{IPDV} is as follows. As in the case of $IPTD^{max}$, $IPDV^{max}$ is computed as:

$$IPDV^{max} = min_{j \in S} IPDV_j. \tag{19}$$

Therefore, based on Eq.(15), the sleep time duration based on the $IPDV$ constraint T_{sl}^{IPDV} is:

$$T_{sl}^{IPDV} = IPDV^{max} - RTT - T_{OH}. \tag{20}$$

If services subscribed by the ONU have both IPTD and IPDV constraints the sleep time value is:

$$T_{sl} = min\{T_{sl}^{IPTD}, T_{sl}^{IPDV}\}. \tag{21}$$

III. Power and Performance Assessment

A. Simulation Scenario

A single OLT and ONU system is simulated using the event driven simulator OPNET Modeler®. The end-to-end communication is assumed to consist of a sequence of spans, and the OLT-ONU connection is part of this sequence. The values of some of the parameters utilized in the performance evaluation are summarized in Tab. II.

Table II
Simulation parameters

Parameter	Symbol	Value
Overhead time	T_{OH}	2 ms
Round trip time	RTT	0.4 ms
Power consumption in sleep mode	P_{sl}	1W
Power consumption in active mode	P_{ON}	10W
Number of spans end-to-end	n	8

Data frames arrive at the OLT with inter arrival times exponentially distributed, thus generating a Poisson frame arrival process. Service data rates are generated by combining the frame arrival rate and the frame size based on Eq.(16). The respective characteristics and constraints for each service are shown in Tab. III. In the table it is assumed that an end-to-end connection consists of eight spans that contribute equally and additively to the considered end-to-end parameters. Therefore, because the access span is one-eighth of the end-to-end connection, the constraints of the access segment are assumed to be simply one-eighth of the end-to-end constraints. This assumption is a pessimistic assumption for the delay variation constraint because it assumes that the delay variation values for each span positively cumulate.

Table III
SERVICE CHARACTERISTICS

Class ID	Frame payload Size [Byte]	Data Rate B_i [b/s]	IPTD Delay [ms]	IPDV [ms]
1	1500	30.4k	125	125
2	560	1k	50	125
3	1067	256k	125	125
4	200	64k	12.5	6.25

The considered Ethernet payload size of each frame carrying Web Browsing services is the maximum Ethernet frame size because web browsing data packet can be as large as 1500 bytes [18]. For Internet Relay Chat the case of Yahoo Messenger Protocol v9 is used as a reference [19]. Typically Yahoo Messenger protocol data section does not exceed 1 kByte. The data rate and the payload size for Multimedia on Web are specific for the case of a single-layer MPEG-4 encoded video with rate control at a bit rate of 256kb/s [20]. In this case the average frame size is 1.067 kBytes. For the VoIP service, G.711 encoding is considered with a payload size of 160 byte (plus overhead) per frame [21].

In addition to synthetic traffic traces, performance evaluation is performed with real traffic traces. Statistics obtained from the traces reported in [22] are also used to generate traffic with a Poisson distribution with the same characteristics (i.e., frame size and bit rate) for sake of comparison. In particular, the traces under consideration refer five specific week days, i.e., from 06/12/2010 until 10/12/2010, during a time span in which usually the traffic volume is high, i.e., from 12 am until 4 pm, as shown in [22]. The statistics extracted from the traces are presented in Tab. IV.

Table IV
TRAFFIC PROFILE FROM TRACES PER FRAME

Trace ID	Day	Hours	Average Size [Bytes]	Average Arrival Rate [Mb/s]
1	06/12/2010	12:00-16:00	971	2.7
2	07/12/2010	12:00-16:00	846	2.33
3	08/12/2010	12:00-16:00	961	2.53
4	09/12/2010	12:00-16:00	1399	23.38
5	10/12/2010	12:00-16:00	1030	4.44

Simulations are run by utilizing one source per service. Based on the services subscribed by the ONU, the sources are selectively activated. Statistics are collected on a per-service base.

The metrics used for the performance evaluation are: the average frame queueing delay $E[W_q]$, defined as the average time spent by a frame in the OLT queue and the energy efficiency η as defined by Eq.(10). All the results are reported with the confidence interval at the 95% confidence level.

B. Results

Tab. V reports the results obtained with different service compositions and synthetic traffic. The values of the inter arrival times and of T_{sl} are obtained as described in section II-C, without considering IPDV constraints. Tab. V shows that by setting T_{sl} as a function of the most stringent constraint for the average delay such constraint is satisfied for all the services. However, if some services have constraints on IPDV, such constraint is not met. In Tab. V the IPDV computation is based on [13] Appendix X, assuming that the IPTD is Gaussian distributed. The achieved energy efficiency is high, slightly below 90%.

Tab. VI shows results that are similar to the one presented in Tab. V but with a different IPDV computation. In this case the IPTD of each frame is collected during simulation and then the definition in Eq. 11 is applied to the cumulative distribution function (CDF) of IPTD. Results show that the obtained IPDV is less than the one reported in Tab. V. The main motivation is that the IPTD is uniformly and not Gaussian distributed. Simulation results and Fig. 4 confirm that.

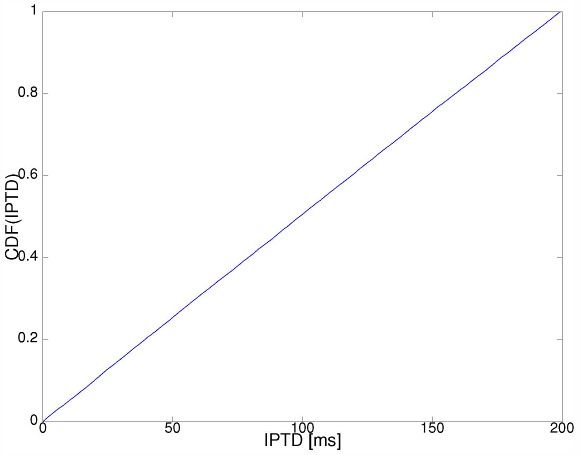

Figure 4. IPTD cumulative distribution function (CDF).

Tab. VII shows some performance results when T_{sl} is computed by taking into account the IPTD and the IPDV constraints of VoIP service. As it can be observed, both IPTD and IPDV constraints are satisfied for all the traffic classes. Furthermore the obtained energy efficiency is less than the one obtained when only IPTD constraints are

Table V
Simulation results with T_{sl} as a function of IPTD; IPDV is computed as in [13].

T_{sl}=22.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
1	125	125	12.3(± 0.04)	20(± 0.03)	81.2 (±2.4X10^{-6})
2	50	125	12.4(± 0.14)	20(± 0.12)	
3	125	125	12.3(± 0.01)	20.1(± 0.01)	
4	12.5	6.25	12.3(± 0.01)	20.1(± 0.01)	

T_{sl}=97.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	D [ms]	$IPDV$ [ms]	η [%]
1	125	125	49.8(± 0.03)	81.2(± 0.02)	87.8 (±0)
2	50	125	49.9(± 0.14)	81.3(± 0.15)	

T_{sl}=97.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
2	50	125	49.8(± 0.08)	81.2(± 0.1)	87.8 (±1.4X10^{-6})
3	125	125	49.8(± 0.01)	81.2(± 0.01)	

Table VI
Simulation results with T_{sl} as a function of IPTD; IPDV is computed based on IPTD collected statistics.

T_{sl}=197.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
1	125	125	12.3(± 0.04)	24.6(± 0.01)	81.2 (±2.4X10^{-6})
2	50	125	12.4(± 0.14)	24.5(± 0.02)	
3	125	125	12.3(± 0.01)	24.6(± 0.01)	
4	12.5	6.25	12.3(± 0.01)	24.6(± 0.01)	

T_{sl}=97.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
1	125	125	49.8(± 0.03)	99.4(± 0.03)	87.8 (±0)
2	50	125	49.9(± 0.14)	99.5(± 0.02)	

T_{sl}=97.2 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
2	50	125	49.8(± 0.08)	99.4(± 0.3)	87.8 (±1.4X10^{-6})
3	125	125	49.8(± 0.01)	99.4(± 0.001)	

Table VII
Simulation results with T_{sl} as a function of IPDV; IPDV is computed based on IPTD collected statistics.

T_{sl}=3.8 ms					
Input parameters			Simulation Results		
Service	$IPTD$ [ms]	$IPDV$ [ms]	$E[W_q]$ [ms]	$IPDV$ [ms]	η [%]
1	125	125	3.1(± 0.01)	6.2(± 0.001)	55.4 (±4.6X10^{-6})
2	50	125	3.1(± 0.03)	6.2(± 0.01)	
3	125	125	3.1(± 0.003)	6.2(± 0.002)	
4	12.5	6.25	3.1(± 0.002)	6.2(± 0.0004)	

considered for the computation of T_{sl} but it is still above 55%.

In the end, the results for the average frame queueing delay obtained by importing the traces into the simulator (i.e., D_{trace} and η_{trace}) and by generating synthetic Poisson arrival process traffic with the data rate and average frame size computed from the traces (i.e.,$D_{Poisson}$ and $\eta_{Poisson}$) are compared. The T_{sl} is computed as a function of IPTD (where $IPTD_{max}$ = 12.5 ms). Tab. VIII shows that the obtained results are similar. Thus, the Poisson assumption well approximates the system behavior.

IV. Conclusion

This paper proposed a method based on queueing theory to compute the ONU sleep time for cyclic sleep with service-based variable sleep period in energy efficient PONs. Such method computes the maximum allowed sleep time based on the services subscribed by the ONU and their frame delay constraints. Both average frame transfer delay and frame delay variation were considered. Results showed that the proposed cyclic sleep with service-based variable sleep period achieves high energy savings while satisfying the delay constraints requested by the considered services.

Table VIII
Simulation results with traffic from traces and related Poisson traffic

T_{sl}=22.2 ms				
Trace ID	D_{trace} [ms]	η_{trace} [%]	$D_{Poisson}$ [ms]	$\eta_{Poisson}$ [%]
1	12.3	81.2	12.3($\pm$ 0.01)	81.2($\pm$3.5X10^{-5})
2	12.3	81.2	12.3($\pm$ 0.004)	81.2($\pm$3.4X10^{-5})
3	12.3	81.2	12.3($\pm$ 0.01)	81.2($\pm$4.2X10^{-5})
4	12.4	81	12.4($\pm$ 0.003)	81.2($\pm$6X10^{-5})
5	12.4	81.2	12.3($\pm$ 0.003)	81.2($\pm$7X10^{-6})

Acknowledgment

This work has been sponsored in part by COST Action Number IC0804 “Energy efficiency in large scale distributed systems”.

References

[1] J. Baliga and et al., “Energy consumption in optical IP networks,” *IEEE/OSA JLT*, vol. 27, no. 13, 2009.
[2] “GPON power conservation,” ITU-T G-series Recommendations - Supplement 45 (G.sup45), 2009.
[3] “10-Gigabit-capable passive optical networks (XG-PON): Transmission convergence (TC) specifications,” ITU-T G-series Recommendations, 2010.
[4] S.-W. Wong, L. Valcarenghi, S.-H. Yen, D. R. Campelo, S. Yamashita, and L. Kazovsky, “Sleep mode for energy saving PONs: advantages and drawbacks,” in *GLOBECOM Workshops, 2009 IEEE*, 2009, pp. 1–6.
[5] B. Skubic and D. Hood, “Evaluation of ONU power saving modes for gigabit-capable passive optical networks,” *Network, IEEE*, vol. 25, no. 2, pp. 20–24, march-april 2011.
[6] L. Shi, B. Mukherjee, and S.-S. Lee, “Energy-efficient PON with sleep-mode ONU: progress, challenges, and solutions,” *Network, IEEE*, vol. 26, no. 2, pp. 36 –41, march-april 2012.
[7] A. Dhaini, P.-H. Ho, and G. Shen, “Toward green next-generation passive optical networks,” *Communications Magazine, IEEE*, vol. 49, no. 11, pp. 94 –101, november 2011.
[8] R. Kubo, J.-I. Kani, Y. Fujimoto, N. Yoshimoto, and K. Kumozaki, “Adaptive Power Saving Mechanism for 10 Gigabit Class PON Systems,” *IEICE Transactions on Communications*, vol. 2, no. E93.B, pp. 280–288, 2010.
[9] M. Fiammengo, A. Lindstrom, P. Monti, L. Wosinska, and B. Skubic, “Experimental evaluation of cyclic sleep with adaptable sleep period length for PON,” in *Optical Communication (ECOC), 2011 37th European Conference and Exhibition on*, sept. 2011, pp. 1 –3.
[10] M. Marsan, A. Anta, V. Mancuso, B. Rengarajan, P. Vasallo, and G. Rizzo, “A simple analytical model for energy efficient Ethernet,” *Communications Letters, IEEE*, vol. 15, no. 7, pp. 773 –775, july 2011.
[11] A. R. Dhaini, P.-H. Ho, and G. Shen, “Energy Efficiency in Ethernet Passive Optical Networks: For How Long Can ONU Sleep?” University of Waterloo, Tech. Rep., March 2011.
[12] “End-user multimedia QoS categories,” ITU-T Recommendation G.1010, nov. 2001.
[13] “Network performance objectives for IP-based services,” ITU-T Recommendation Y.1541, Dec. 2011.
[14] H. Takagi, *Analysis of Polling Systems*, ser. Research Reports and Notes Computer System Series. The MIT Press, 1986.
[15] “Ethernet frame transfer and availability performance,” ITU-T Recommendation Y.1563, Jan. 2009.
[16] J. i. Kani, “Power saving techniques for optical access networks,” OFC Tutorial, Mar. 5 2012.
[17] “Internet protocol data communication service-IP packet transfer and availability performance parameters,” ITU-T Recommendation Y.1540, Mar. 2011.
[18] D. He, L. Liang, and H. Cruickshank, “Internet QoS and traffic modelling,” in *IEE Proceedings - Software*, sept. 2004.
[19] “Yahoo messenger protocol v9,” available at libyahoo2.sourceforge.net/ymsg-9.txt.
[20] P. Seeling, M. Reisslein, and B. Kulapala, “Network performance evaluation using frame size and quality traces of single-layer and two-layer video: A tutorial,” *Commun. Surveys Tuts.*, vol. 6, no. 3, pp. 58–78, Jul. 2004. [Online]. Available: http://dx.doi.org/10.1109/COMST.2004.5342293
[21] “Voice over IP - per call bandwidth consumption,” http://www.cisco.com/en/US/tech/tk652/tk698/technologies_tech_note09186a0080094ae2.shtml#topic1.
[22] X. Liu, N. Ghazisaidi, L. Ivanescu, R. Kang, and M. Maier, “On the Tradeoff Between Energy Saving and QoS Support for Video Delivery in EEE-Based FiWi Networks Using Real-World Traffic Traces,” *Journal of Ligthwave Technology*, vol. 29, no. 18, 2011.