A GI/G/1 Model for 10Gb/s Energy Efficient Ethernet Links
Sergio Herrería-Alonso, Miguel Rodríguez-Pérez, Manuel Fernández-Veiga, Member, IEEE, and Cándido López-García

Abstract—The IEEE 802.3az standard provides a mechanism to build energy efficient Ethernet interfaces via a low power idle mode that they can enter when there is no data to transmit. Several competing algorithms have appeared that make use of this mode to minimize energy consumption with little disruption to the traffic. Two algorithms stand out among those because of their simplicity and performance: frame transmission and burst transmission. Although these algorithms have been shown to be very efficient in simulated scenarios, there is a lack of general analytical models for their behavior. In fact, to this date, the only analyzed traffic patterns have been variants of Poisson traffic. In this paper we provide a general GI/G/1 model for energy consumption and traffic delay for both algorithms. We then develop specializations of the general model for Poisson and deterministic traffic. Finally, we validate the model with the help of both synthetic traffic and real Internet traffic traces.

Index Terms—Power awareness, IEEE 802.3az, energy efficient ethernet, modeling.

I. INTRODUCTION

SINCE the start of the century, the importance of a more efficient usage of power resources has been constantly gaining mind share. In the IT field, the efforts to reduce energy consumption began in naturally power-constrained devices, like laptops, smartphones and, generally, all kinds of portable gadgets. However, not only battery powered devices have been the focus of energy efficiency. Public awareness of environmental concerns, rising electricity costs and even heat dissipation issues have spurred research for energy efficiency in all IT fields, from big data centers to domestic equipment.

To cut the squandering of energy in IT equipment, the first efforts targeted the greatest power consumers in the devices: processors, monitors, network radios in portable devices, etc. Although general purpose networking equipment was a relatively minor energy consumer, networking speed improvements have exponentially increased the power demands of Ethernet links, from less than 0.5 watts for 100Mb/s devices, up to 5 watts for 10Gb/s cards [1], [2]. This, and the fact that the greatest energy consumers had greatly reduced their power demands, made the energy needs of Ethernet links a significant part of the energy budget of IT equipment.

The IEEE formed a study group in late 2006 to develop greener extensions to Ethernet physical interfaces (PHYs). After recognizing that most Ethernet links sit unused most of the time, with load factors well below 30% even for busy servers [3], [4], the group opted to define a new operating state in PHYs to be used when there is no data to transmit. When a PHY is in this new low power idle (LPI) state, it only draws a small fraction of the power needed for normal operation. Although exact power needs may vary among specific PHY and card models, it is expected that power savings of up to 90% are normal in the LPI mode. The detailed operations of this mode were published in the IEEE 802.3az standard, also known as Energy Efficient Ethernet (EEE) [5]. As the PHY is unable to send traffic to its attached link while in the LPI mode, the standard also entails a protocol to coordinate the transitions to and from the LPI mode between the two interfaces attached to a link.

The IEEE did not mandate any algorithm for Ethernet interfaces to decide under which conditions the LPI mode should be entered or left. As a consequence, several alternatives have been proposed to make the best use of the low power mode. Two algorithms stand out among the rest because of their simplicity and performance: frame transmission [6] and burst transmission, also known as packet coalescing [7], [8]. The first is just a straightforward application of the LPI mode. When the interface is in this mode, the PHY is awoken as soon as a new frame is ready for transmission. Correspondingly, when there is no more traffic to be sent, the PHY is put back into the LPI mode again. As the length of the transitions to and from the LPI mode are not negligible and the interface consumes energy while executing them, this algorithm does not perform well when packet interarrival times are in the order of the transition duration. The second algorithm, burst transmission, solves this problem by queuing a few data frames in the transmission buffer before proceeding to awake the device.

The original proposals for both frame and burst transmission were based on exhaustive simulation results. However, the need to properly tune burst transmission and compare its efficiency against frame transmission for general traffic spurred the development of analytical models. Yet, due to the complexity of the system, all models were restricted to Poisson arrival processes and most of them ignored the effects on packet delay [9], [10], [11], [12], [13], [14], [15]. Nevertheless, these models have proved very valuable for assessing qualitative comparisons between the algorithms and to gain a better understanding of the effects of the burst transmission configuration parameters.
In this paper we present new analytical models for both frame and burst transmission algorithms. Our models capture the effects of these algorithms not only on energy consumption but also on packet delay. Moreover, our models are valid for any uncorrelated arrival process and produce accurate results for both the power savings and the added delay to packets. We first derive a general solution to the problem, and particularize it later to Poisson and deterministic interarrival processes. We also validate the model with the help of real traffic traces from the CAIDA project [16] with good results.

The rest of this paper is organized as follows. Section II presents some of the work in the field that preceded this paper. Section III shows the power model on which we will build. In Section IV we validate our analysis through simulation. Finally, Sections V and VI we develop analytic models for both the frame and burst transmission algorithms respectively. Then we particularize our models for Poisson (Section VII) and deterministic (Section VIII) interarrival processes. In Section IX we validate our analysis through simulation. Finally, our conclusions are laid out in Section X.

II. RELATED WORK

The initial efforts to reduce consumption in the networking infrastructure can be traced back to [17] where authors identify several profitable targets to focus on for greening the Internet. Physical links were among those targets. The first proposals to reduce the amount of energy needed in Ethernet links were quick to appear [18], [19], [20], [21]. These proposals could be classified in two separate fields: those advocating the use of a new low power mode at periods with no traffic to be sent, and those that tried to accommodate the Ethernet link speed to the traffic load. Eventually, the IEEE 802.3az standard settled the issue when it introduced a new low power mode incapable of sending traffic, but with very modest power needs for 100Mb/s, 1Gb/s and 10Gb/s Ethernet physical interfaces. The short times needed to enter and leave the low power idle (LPI) mode and the relatively long times needed to renegotiate different link speeds were probably the main reasons behind this decision.

Several different algorithms had been proposed to make use of a new low power mode in Ethernet devices [7], [19]. As the exact capabilities of the finally standardized low power mode were unknown, most proposals worked on the assumption that a sleeping interface could not even sense the presence of incoming traffic in the link. Thus, the algorithms mainly focused on the appropriate selection of sleeping intervals: starting time and sleeping duration, as the low power mode had to be coordinated between both ends of the link. Once the standard was finalized, it became clear that NICs would have the ability to sense incoming traffic while in the low power mode and most of the ideas associated with deciding the length of the sleeping period became no longer necessary.

The frame transmission algorithm is just a direct application of the standardized LPI mode [6]. It consists on immediately entering the LPI mode just after sending the last buffered packet and turning back to the normal mode as soon as there is new traffic to transmit. In spite of its simplicity, it is able to get great power savings for low traffic loads, but it is not very efficient for higher loads or non bursty traffic. The burst transmission algorithm, also known as packet coalescing, trades a small increment in packet delay for much greater energy savings [7], [8]. Instead of exiting LPI as soon as there is new traffic to transmit, it waits to first accommodate some packets in the transmission buffer, thus avoiding inefficiencies with non bursty traffic. To avoid delaying the traffic excessively, there is a timeout for the time the oldest packet waits for new traffic before exiting the LPI mode.

There are already some published analytical models for the aforementioned algorithms. For instance, in [10] there are power and delay models for three pre-IEEE 802.3az algorithms, the last one being very similar to burst transmission. The work considers Poisson traffic arrivals and a general distribution for the packet lengths. An analytical model for frame transmission appears in [11]. It considers Poisson arrivals of packet trains, instead of single packets, to capture the bursty nature of Internet traffic. An alternative model for the energy consumption of frame transmission considering Poisson traffic appears in [12]. The proposed model is an approximation as it splits the time in discrete time intervals and assumes that the transition times are multiples of the frame transmission time. There is also a simple power model for burst transmission in [9] and an analytical comparison between the energy savings that can be achieved when using frame or burst transmission algorithms in [13], but both analyses are again restricted to Poisson traffic. Finally, there are only a few models addressing the trade-off between energy consumption and packet delay [14], [15]. Unfortunately [14] considers Poisson arrivals of packet batches, each packet demanding the same amount of service time, and [15] just deals with Poisson traffic. To the best of our knowledge there is no model considering an arbitrary distribution for the arrival process and the effects of sleeping algorithms on packet delay.

III. EEE OPERATIONS

Fig. 1 illustrates the operations required to put an EEE-capable interface into the low power mode [5]. Let $T_r$ be the time employed to coordinate the status change to the LPI mode and $T_w$ the time needed to awake the interface. There is also a periodic transmission of a refresh signal of length $T_r$ to avoid a complete renegotiation of the link parameters when returning to the normal power mode and thus keep $T_w$ small. However, this parameter is usually ignored in EEE studies since its effects, mainly an increase of power consumption while in the LPI mode, can be taken into account in the average power demands of this mode.

Clearly, for maximizing power savings, interfaces should only be active when there is some data ready for transmission, so they should be put to sleep every time that the transmission

![Fig. 1. EEE-capable interface operations.](image-url)
queue gets empty. However, interfaces should not transition back to active as soon as the first frame arrives, like in the frame transmission method: if the number of frames queued for transmission is too small, it will be likely put back to sleep in too short a time, making the transition to active and back to sleep costly in excess. Therefore, to avoid this, the burst transmission algorithm queues some traffic to transmit large data bursts. With this method, a sleeping interface will resume its normal operation after the amount of queued frames waiting for service reaches \( Q_w \).

On the other hand, assembling frames into bursts increases the delay. To avoid excessive delays, the maximum time an interface can be in the low power mode continuously is limited to \( T_{\text{max}} \) since the first frame is buffered for transmission.

Finally, we would like to point out that we will focus on 10Gb/s links since EEE can achieve considerable energy savings on them and, therefore, they are more interesting from a practical point of view. Thus, as stated by the EEE standard, we assume that transitions from active to the LPI mode cannot be interrupted although a new frame arrives to the interface and that transitions can occur in both directions of the link independently.\(^1\)

IV. POWER EFFICIENCY AND DELAY ANALYSIS

In the following sections we will develop analytic models for both frame and burst transmission methods that can be used to study the influence of configuration parameters and line card physical characteristics on the expected performance. More specifically, we will quantify the achievable energy savings and the impact on average frame delay.

We assume that frame arrivals follow a general distribution with independent interarrival times \( T_n \), \( n = 1, 2, \ldots \), and an average arrival rate \( \lambda \). We also assume that the sequence of service times \( S_n \), \( n = 1, 2, \ldots \), demanded by successive frames is a set of random variables with a common, although arbitrary, distribution function and mean service rate \( \mu \). Obviously, the utilization factor \( \rho = \lambda / \mu \) must be less than 1 to assure system stability. Finally, we assume that the interface has a transmission buffer of infinite capacity.\(^2\)

Fig. 2 illustrates several definitions that will be used in the model. It shows an example of how the transmission buffer of infinite capacity. 2

![Fig. 2. EEE busy cycle when using burst transmission.](image)

are buffered in the transmission queue of the interface. The time interval when the interface is transmitting frames is the busy period \( T_{\text{on}} \). Finally, an inactive period followed by a busy period forms a busy cycle \( T_{\text{cycle}} \).

A. Energy Consumption

Let \( E[P] \) be the mean power consumption for a given interface that never enters a low power mode and \( E[P_{\text{EEE}}] \) the mean power consumption for an EEE-capable interface. \( E[P_{\text{EEE}}] \) depends on the proportion of time the interface spends in each possible state:

\[
E[P_{\text{EEE}}] = \rho_{\text{off}} E[P_{\text{off}}] + \rho_{\text{tra}} E[P] + \rho_{\text{on}} E[P] = \rho_{\text{off}} E[P_{\text{off}}] + (1 - \rho_{\text{off}}) E[P], \tag{1}
\]

where \( E[P_{\text{off}}] \) is the mean power consumed in the LPI mode and \( \rho_{\text{tra}} \) and \( \rho_{\text{on}} \) are, respectively, the fractions of time in which the interface is sleeping, transitioning between states and awake. Note that it is assumed that the interface consumes the same power during transitions as in the active state since many components of the transceivers have to be operative during the state changes. Immediately, the energy consumed on an EEE-capable interface compared with that consumed on a power-unaware interface that is always active is given by

\[
\varphi = \frac{E[P_{\text{EEE}}]}{E[P]} = 1 - (1 - \varphi_{\text{off}}) \rho_{\text{off}}, \tag{2}
\]

where \( \varphi_{\text{off}} = E[P_{\text{off}}]/E[P] \) is the portion of the active mode energy consumption demanded when the interface is sleeping. \( \rho_{\text{off}} \) can be calculated as the ratio between the mean duration of sleeping periods and the mean duration of busy cycles:

\[
\rho_{\text{off}} = \frac{E[T_{\text{off}}]}{E[T_{\text{cycle}}]} = \frac{E[T_{\text{off}}]}{E[T_{\text{off}}] + T_s + T_w + E[T_{\text{on}}]}, \tag{3}
\]

In addition, we know that \( \rho = E[T_{\text{on}}]/E[T_{\text{cycle}}] = E[T_{\text{on}}]/(E[T_{\text{off}}] + T_s + T_w + E[T_{\text{on}}]) \) since, as long as the interface is awake, it is transmitting queued frames. Therefore, rearranging terms, \( E[T_{\text{on}}] \) can be expressed as

\[
E[T_{\text{on}}] = \frac{\rho}{1 - \rho} (E[T_{\text{off}}] + T_s + T_w), \tag{4}
\]

and, substituting this into (3), \( \rho_{\text{off}} \) is now given by

\[
\rho_{\text{off}} = (1 - \rho) \frac{E[T_{\text{off}}]}{E[T_{\text{off}}] + T_s + T_w}. \tag{5}
\]
Finally, substituting $\rho_{\text{off}}$ into (2), we get

$$\phi = 1 - (1 - \rho_{\text{off}})(1 - \rho) \frac{E[T_{\text{off}}]}{E[T_{\text{off}}] + T_s + T_w}.$$  

(6)

In short, to compute the energy consumption of an EEE-capable interface it is only necessary to obtain the average length of sleeping periods $E[T_{\text{off}}]$. Clearly, this value depends on the selected sleeping algorithm, so in the following sections we will particularize $E[T_{\text{off}}]$ for both frame and burst transmission methods.

B. Frame Delay

With a power management scheme in action, frames suffer an additional delay. To compute the average frame delay $E[W]$ in this scenario we can directly apply the result obtained in [22] for those queuing systems with vacations in which the first frame in every busy period suffers a random delay before its transmission begins. In this paper Marshall shows that, when the first frame in each busy period is delayed a random time $W_{\text{first}}$, the mean wait in queue is given by

$$E[W] = \frac{-E[U^2]}{2E[U]} + \frac{E[W_{\text{first}}^2] - E[T_{\text{first}}^2]}{2(E[W_{\text{first}}] + E[T_{\text{first}}])} + \frac{\text{cov}(W, I)}{E[U]},$$  

(7)

where $U_n = S_n - I_n$ and $\text{cov}(W, I) = \text{cov}(W_n, I_n)$ for frame $n$.

Accordingly,

$$E[U] = E[S - I] = E[S] - E[I] = \frac{1}{\mu} - \frac{T_s}{\lambda} = -\frac{1 - \rho}{\lambda},$$  

(8)

and

$$E[U^2] = \sigma_I^2 + E[U]^2 = \sigma_I^2 + \left(\frac{1 - \rho}{\lambda}\right)^2,$$  

(9)

since frame lengths are assumed to be independent from the arrival process and hence $\text{cov}(S_n, I_n) = 0$. Therefore, substituting (8) and (9) into (7), we get

$$E[W] = \frac{\lambda \left(\sigma_I^2 + \sigma_l^2\right)}{2(1 - \rho) \lambda} + \frac{1 - \rho}{2\lambda} + \frac{E[W_{\text{first}}^2] - E[T_{\text{first}}^2]}{2(E[W_{\text{first}}] + E[T_{\text{first}}])} - \frac{\lambda \text{cov}(W, I)}{1 - \rho}.$$  

(10)

Note that the average waiting time of the first frame in each busy period $E[W_{\text{first}}]$ (and $E[W_{\text{first}}^2]$) depends on the sleeping algorithm, while the average empty period $E[T_e]$ (and $E[T_e^2]$) just depends on the distribution followed by frame arrivals.

V. Frame Transmission

In this section, we obtain the average duration of sleeping periods and the average waiting time of the first frame in each busy cycle when the frame transmission method is used. Recall that these values are required to compute the expected energy savings and the average frame delay respectively.

A. Average Duration of Sleeping Periods

As previously explained, with frame transmission the interface is put to sleep every time the transmission queue gets empty and it is immediately awoken upon data arrival. Therefore, for calculating the mean duration of sleeping periods, we have to take into account two different scenarios depending on the arrival time of the first frame after putting the interface to sleep. For the first case, when the first frame in the busy cycle arrives to the interface in less than $T_s$ seconds, the transition to sleep is completely useless since the interface will be reactivated without entering the low power mode at all ($T_{\text{off}} = 0$). However, when the first frame arrives after $T_s$ seconds, then the interface will remain in the sleeping state for $t - T_s$ seconds, where $t$ is the time elapsed since the interface is put to sleep until the arrival of the first frame. Therefore, $E[T_{\text{off}}]$ can be calculated as

$$E[T_{\text{off}}]_{\text{frame}} = \int_{T_s}^{\infty} (t - T_s) f_{T_e}(t) \, dt,$$  

(11)

where $f_{T_e}(t)$ is the probability density function of the length of empty periods or, equivalently, the probability density function of the time elapsed since the beginning of the busy cycle until the arrival of the first frame. Obviously, this function depends on the distribution function followed by frame arrivals. However, if it cannot be calculated or it is simply unknown, we can use the mean interarrival time $1/\lambda$ as a coarse approximation of the arrival time of the first frame and then $E[T_{\text{off}}]_{\text{frame}}$ can be approximated by

$$E[T_{\text{off}}]_{\text{frame}} \approx \max \left(\frac{1}{\lambda} - T_s, 0\right).$$  

(12)

B. Average Frame Delay

Clearly, the first frame in each busy cycle will just wait for the interface to transition to awake since it is precisely the arrival of this frame what wakes up the interface when using this algorithm. Therefore, $E[W_{\text{first}}]_{\text{frame}} = T_w$ and, assuming that the transition time is constant, $E[W_{\text{first}}^2]_{\text{frame}} = T_w^2$.

On the other hand, it can be easily seen that, with this algorithm, $\text{cov}(W, I) = 0$ since the waiting time of a given frame does not depend on the arrival time of the next frame. Therefore, substituting these values into (10), the average frame delay is now given by

$$E[W]_{\text{frame}} = \frac{\lambda \left(\sigma_I^2 + \sigma_l^2\right)}{2(1 - \rho)} + \frac{1 - \rho}{2\lambda} + \frac{T_w^2 - E[T_e^2]}{2(T_w + E[T_e])}.$$  

(13)

VI. Burst Transmission

A sleeping interface using the burst transmission method will be awoken when $Q_w$ frames are queued for transmission, unless the traffic load is so low that the $T_{\text{max}}$ timer expires before the queue reaches the minimum burst size. This means that two different operation regimes can be distinguished depending on which wake-up condition is first met. In this section, we obtain the average duration of the sleeping periods and the average frame delay for both operating scenarios.
A. Average Duration of Sleeping Periods

1) Low Load: If the frame arrival rate is low, it is very likely that reaching the minimum queue length $Q_w$ would take an excessively long time, so, in order to avoid large delays, the sleeping period will be bounded by $T_{\text{max}}$ since the arrival of the first frame in the busy cycle. Then, for a properly configured interface with $T_{\text{max}} > T_s$, we get

$$E[T_{\text{off}}]_{\text{burst,low}} = \int_0^\infty (t + T_{\text{max}} - T_s) f_{T_s}(t) \, dt,$$

(14)

since, in this case, every transition to sleep will guarantee some time in the LPI mode, no matter when the first frame arrives. Again, if $f_{T_s}(t)$ is unknown, the arrival time of the first frame can be approximated by the mean interarrival time $1/\lambda$ and

$$E[T_{\text{off}}]_{\text{burst,low}} \approx \frac{1}{\lambda} + T_{\text{max}} - T_s,$$

(15)

2) High Load: If the load is high enough, $Q_w$ frames will arrive to the upstream interface before the $T_{\text{max}}$ timer expires. Therefore, in this case, the mean duration of sleeping periods is given by

$$E[T_{\text{off}}]_{\text{burst,high}} = \int_{T_s}^\infty (t - T_s) f_{Q_w}(t) \, dt,$$

(16)

where $f_{Q_w}(t)$ is the probability density function of the time elapsed since the beginning of the busy cycle until the arrival of the $Q_w$-th frame. Generally, there is no simple closed form for this function,

4Since we are assuming independent interarrival times, $\{T_s\}$ is a renewal process and $f_{Q_w}(t) = f_{T_s}(t) \ast f_{T_s}(t) \ast \cdots \ast f_{Q_w}(t)$, where $\ast$ is the convolution operator.

but, if the burst size $Q_w$ is large enough, we can approximate the previous expression by

$$E[T_{\text{off}}]_{\text{burst,high}} \approx \frac{Q_w}{\lambda} - T_s.$$

(17)

Note that, if $Q_w$ is too small so that $Q_w/\lambda < T_s$, then it will be very likely to receive $Q_w$ or more frames while transitioning to sleep thus causing the immediate awakening of the interface without entering the low power mode at all. Reasoning in this way, and to assure that this condition is met for all the possible arrival rates, we suggest to set $Q_w > \mu T_s$ so that the likelihood of an useless transition is sufficiently small for any $\lambda$.

3) Operating Scenario: The real operating conditions can be directly determined from the average traffic load. In fact, there exists an arrival rate threshold $\lambda^*$ that delimits whether the system is under low or high load conditions. Thus, with such an average arrival rate $\lambda^*$, it must hold that

$$E[T_{\text{off}}]_{\text{burst,low}} = E[T_{\text{off}}]_{\text{burst,high}}.$$

Then, equating (15) and (17), and solving for $\lambda^*$, we get

$$\lambda^* \approx \frac{Q_w - 1}{T_{\text{max}}}. \quad (18)$$

B. Average Frame Delay

1) Low Load: If the traffic load is low enough so that the sleeping timer $T_{\text{max}}$ expires before $Q_w$ frames arrive to the interface, then the first arriving frame in the busy cycle will simply wait $T_{\text{max}} + T_w$ since the sleeping timer is precisely started on its arrival. Consequently, $E[W_{\text{first}}]_{\text{burst,low}} = T_{\text{max}} + T_w$ and $E[W_{\text{first}}^2]_{\text{burst,low}} = (T_{\text{max}} + T_w)^2$.

In addition, under these load conditions, $\text{cov}(W, I) = 0$ since waiting times and arrival times are certainly independent. Therefore, substituting these values into (10), we have

$$E[W]_{\text{burst,low}} = \frac{\lambda (\sigma^2 + \sigma_I^2)}{2(1 - \rho)} + \frac{1 - \rho}{2\lambda} \left( T_{\text{max}} + T_w \right)^2 - E[T_c^2] \left( T_{\text{max}} + T_w + E[T_c] \right).$$

(19)

2) High Load: When the traffic load is high enough, we can assume that at least $Q_w$ frames will arrive to the interface before the $T_{\text{max}}$ timer expires. Under these conditions, the first arriving frame in the busy cycle must wait for the arrival of the remaining $Q_w - 1$ frames and then, it must wait for the interface to transition to awake. So, on average, $E[W_{\text{first}}]_{\text{burst,high}} = (Q_w - 1)/\lambda + T_w$ and

$$E[W_{\text{first}}^2]_{\text{burst,high}} = \text{Var}[W_{\text{first}}]_{\text{burst,high}} + (E[W_{\text{first}}]_{\text{burst,high}})^2 = (Q_w - 1)\sigma^2 + \left( \frac{Q_w - 1}{\lambda} + T_w \right)^2.$$\n
(20)

On the other hand, for the first $Q_w - 1$ frames in each busy cycle, the waiting time of a frame depends on the following interarrival times, so $\text{cov}(W, I)$ is nonzero. In [23] it is proved that, for single-server queues that wait until $Q_w$ customers are present before starting service again, this covariance term is given by

$$\text{cov}(W, I) = \frac{(1 - \rho)(Q_w - 1)\sigma^2}{Q_w - 1 + \lambda E[T_c]}.$$

(21)

Therefore, under high load conditions, we obtain

$$E[W]_{\text{burst,high}} = \frac{\lambda (\sigma^2 + \sigma_I^2)}{2(1 - \rho)} + \frac{1 - \rho}{2\lambda} \left( T_{\text{max}} + T_w \right)^2 + \frac{\lambda^2 (\sigma^2 (Q_w - 1) - E[T_c^2]) + (\lambda T_w + Q_w - 1)^2}{2\lambda (Q_w - 1 + \lambda (T_w + E[T_c]))} - \frac{\lambda (Q_w - 1)\sigma^2}{Q_w - 1 + \lambda E[T_c]).}$$

(22)

VII. POISSON TRAFFIC

In this section we will obtain the expected energy savings and average frame delay with both frame and burst transmission methods assuming that the number of frames that arrive in a given interval of time to the interface follows a Poisson distribution with average arrival rate $\lambda$. Although it is well-known that frame arrivals on LAN networks do not really follow a Poisson distribution [24], this approximation can serve as a reasonable balance between the best (TCP introducing large bursts of packets into a link) and worst (TCP acknowledgements evenly spaced-out) case scenarios. In any case, Poisson traffic is useful in order to provide a valid approximation to real traffic in sub-second timescales [25] and to aggregated traffic in the Internet core [26].

A. Frame Transmission

As shown in Section V-A, to obtain the average duration of sleeping periods we have to calculate the probability density function of the length of empty periods $f_{T_s}(t)$. However, due
to the memoryless property of Poisson traffic, this function is equivalent to the probability density function of interarrival times. Therefore, \( f_{T_e}(t) = \lambda e^{-\lambda t} \), \( t \geq 0 \), and solving integral (11), we get \( E[T_{\text{off}}|I_n=\text{Exp}] = e^{-\lambda T_i}/\lambda \). Finally, substituting this value into (6), we obtain the expected energy savings for the frame transmission algorithm:

\[
q_{I_n=\text{Exp}}^{\text{frame}} = 1 - (1 - \phi_{\text{off}})(1 - \rho)\frac{e^{-\lambda T_i}}{e^{-\lambda T_i} + \lambda(T_s + T_w)}.
\]  

(23)

As expected, this result coincides with those already obtained in [9], [11].

To obtain the average frame delay, we have to calculate the mean duration of empty periods \( E[T_e] \). Using again the memoryless property of Poisson traffic, it can be easily seen that \( E[T_e|I_n=\text{Exp}] = 1/\lambda \). In addition, \( E[T_e^2|I_n=\text{Exp}] = \sigma^2_I + (E[T_e|I_n=\text{Exp}]^2) = 1/\lambda^2 + 1/\lambda^2 = 2/\lambda^2 \), so substituting these values into (13), the average frame delay is given by

\[
E[W_{I_n=\text{Exp}}^{\text{frame}}] = 1 + \frac{\lambda^2 \sigma^2_I}{2\lambda(1 - \rho)} + \frac{1 - \rho}{2\lambda} + \frac{(\lambda T_w)^2 - 2}{2\lambda(1 + \lambda T_w)}.
\]  

(24)

**B. Burst Transmission**

1) **Low Load**: Substituting \( f_{T_e}(t) = \lambda e^{-\lambda t} \) into (14), we get \( E[T_{\text{off}}|I_n=\text{Exp}] = 1/\lambda + T_{\text{max}} - T_s \), so the expected energy savings in this particular regime are given by

\[
q_{I_n=\text{Exp}}^{\text{burst,low}} = 1 - (1 - \phi_{\text{off}})(1 - \rho)\frac{1 + \lambda(T_{\text{max}} - T_s)}{1 + \lambda(T_{\text{max}} + T_w)}.
\]  

(25)

On the other hand, substituting \( E[T_e|I_n=\text{Exp}] = 1/\lambda \) and \( E[T_e^2|I_n=\text{Exp}] = 2/\lambda^2 \) into (19), we obtain the expected average frame delay:

\[
E[W_{I_n=\text{Exp}}^{\text{burst,low}}] = 1 + \frac{\lambda^2 \sigma^2_I}{2\lambda(1 - \rho)} + \frac{1 - \rho}{2\lambda} + \frac{\lambda^2(T_{\text{max}} + T_w)^2 - 2}{2\lambda(1 + \lambda(T_{\text{max}} + T_w))}.
\]  

(26)

2) **High Load:** With Poisson traffic, interarrival times for all the frames are identically exponentially distributed. Therefore, the arrival time of the \( Q_w \)-th frame is Erlang-\( Q_w \) distributed and \( f_{Q_w}(t) = \lambda Q_w t^{Q_w - 1} e^{-\lambda t}/(Q_w - 1)! \). Then, using this to solve integral (16), we get

\[
E[T_{\text{off}}|I_n=\text{Exp}]^{\text{frame, high}} = \frac{\Gamma(Q_w + 1, \lambda T_s) - \lambda T_s \Gamma(Q_w, \lambda T_s)}{\lambda(Q_w - 1)!},
\]  

(27)

where \( \Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt \) is the upper incomplete gamma function. Finally, substituting (27) into (6), we get the expected energy savings for this scenario:

\[
q_{I_n=\text{Exp}}^{\text{burst, high}} = 1 - (1 - \phi_{\text{off}})(1 - \rho)\frac{\Gamma(Q_w + 1, \lambda T_s) - \lambda T_s \Gamma(Q_w, \lambda T_s)}{\Gamma(Q_w + 1, \lambda T_s) - \lambda T_s \Gamma(Q_w, \lambda T_s) + \lambda(T_s + T_w)(Q_w - 1)!}.
\]  

(28)

VIII. **DETERMINISTIC TRAFFIC**

In this section we will particularize the model for the case in which the interarrival times for all the frames are the same and equal to \( 1/\lambda \).

A. **Frame Transmission**

As shown in Fig. 3, when using frame transmission under low load conditions, the sequence of events that triggers a transition to active is the following: a frame arrival at the sleeping interface will wake it and \( T_w \) seconds later the interface will be able to transmit the arriving frame. Then, if the time until the next frame arrives is sufficiently large (\( 1/\lambda > 1/\mu + T_s + T_w \)), the frame will be transmitted and the interface will be directly put to sleep again. Therefore, with deterministic traffic under these load conditions, it can be easily seen that \( E[T_{\text{frame}}|I_n=\text{Det}] = 1/\lambda - 1/\mu + T_s \) and \( E[T_{\text{off}}|I_n=\text{Det}] = 1/\lambda - 1/\mu + T_s - T_w \). Finally, substituting the average length of sleeping periods into (6), the expected energy savings can be written as

\[
q_{I_n=\text{Det}}^{\text{frame}} = 1 - (1 - \phi_{\text{off}})(1 - \rho)\frac{1}{1/\lambda - 1/\mu + T_s - T_w}.
\]  

(30)

Note that, with higher arrival rates (\( 1/\lambda \leq 1/\mu + T_s + T_w \)), the interface will not ever be in the low power mode and \( q_{I_n=\text{Det}}^{\text{frame}} = 0 \).

As regards the average frame delay, from the previous reasoning we have \( E[T_{\text{off}}|I_n=\text{Det}] = 1/\lambda - 1/\mu + T_s \) and \( E[T_{\text{frame}}|I_n=\text{Det}] = \sigma_I^2 + \sigma_s^2 - 2\text{cov}(I, S) + E[T_{\text{frame}}|I_n\text{Det}]^2 = \sigma_I^2 + (1/\lambda - 1/\mu + T_s)^2 \), so, substituting these values into (13), we get

\[
E[W_{I_n=\text{Det}}^{\text{frame}}] = T_s.
\]  

(31)

if \( 1/\lambda > 1/\mu + T_s + T_w \) as expected.

However, with higher loads (\( 1/\lambda \leq 1/\mu + T_s + T_w \)), the first frame in the busy cycle will arrive shortly after beginning the transition to sleep. Since this transition cannot be interrupted,
transmission under low load conditions. As shown, 1/\lambda - 1/\mu is a good approximation to the duration of the empty periods since it is very likely that some frames arrive to the interface when it is active and transmitting other previously arrived frames, so \[ T_{\text{frame}} \approx 1/\lambda - 1/\mu \] and \[ E[T_{\text{frame}}^{\text{burst,high}}_{h_{n}} - \text{Det}] \approx \sigma_S^2 + (1/\lambda - 1/\mu)^2. \] Then, substituting all these values into (10), we get
\[
E[W]_{h_{n}} - \text{Det} \approx \frac{\lambda^2(T_T + T_w)\sigma_S^2}{2(1-\rho)(1-\rho + \lambda(T_T + T_w))} + \frac{T_T + T_w}{2}. \quad (32)
\]

B. Burst Transmission

1) Low Load: Fig. 4 shows a typical busy cycle with burst transmission under low load conditions. As shown, 1/\lambda - (T_{\text{max}} - \lambda T_{\text{max}})/\lambda - T_w - [\lambda T_{\text{max}}]/\mu is a good approximation to the arrival time of the first cycle. Therefore, solving integral (14) with this approximation, we get
\[
E[T_{\text{off}}^{\text{burst,low}}_{h_{n}} - \text{Det}] \approx 1/\lambda - T_T - T_w - [\lambda T_{\text{max}}]/\mu + [\lambda T_{\text{max}}]/\lambda
\]
and then, the expected energy savings are given by
\[
\phi_{h_{n}} - \text{Det} \approx 1 - (1 - \phi_{\text{off}})(1-\rho) \times \frac{1 - \lambda(T_T + T_w) - \rho[\lambda T_{\text{max}}] + [\lambda T_{\text{max}}]}{1 - \rho[\lambda T_{\text{max}}] + [\lambda T_{\text{max}}]}. \quad (33)
\]
Note that, if the traffic load is so low that \(\lambda T_{\text{max}} < 1\), then
\[ E[T_{\text{off}}^{\text{burst,low}}_{h_{n}} - \text{Det}] \approx 1/\lambda - T_T - T_w - E[T_{\text{off}}^{\text{frame}}_{h_{n}} - \text{Det}] \]
the same amount of power will be saved with both frame and burst transmission algorithms.

To compute the average frame delay, we can use the same approximation again, so
\[
E[T_{\text{e}}^{\text{burst,low}}_{h_{n}} - \text{Det}] \approx 1/\lambda - (T_{\text{max}} - \lambda T_{\text{max}})/\lambda - T_w - [\lambda T_{\text{max}}]/\mu \quad \text{and} \quad E[T_{\text{e}}^{\text{burst,low}}_{h_{n}} - \text{Det}] \approx \sigma_S^2 + (E[T_{\text{e}}^{\text{burst,low}}_{h_{n}} - \text{Det}])^2 \approx \sigma_S^2 + (1/\lambda - T_{\text{max}} - [\lambda T_{\text{max}}]/\lambda - T_w - [\lambda T_{\text{max}}]/\mu)^2. \]
Finally, substituting these values into (19), we have
\[
E[W]_{h_{n}} - \text{Det} \approx \frac{\lambda([\lambda T_{\text{max}}] + \rho[\lambda T_{\text{max}}] - [\lambda T_{\text{max}}] - 1)^2}{2(1-\rho)([\lambda T_{\text{max}}] + \rho[\lambda T_{\text{max}}] - 1)} + T_{\text{max}} + \frac{[\lambda T_{\text{max}}] - 1}{2\mu}. \quad (34)
\]

2) High Load: Under high load conditions, we employ (17) to approximate the arrival time of the \(Q_w/\lambda - 1/\mu\)th frame in the busy cycle by \(Q_w/\lambda - 1/\mu\), then,
\[
E[T_{\text{off}}^{\text{burst,high}}_{h_{n}} - \text{Det}] \approx Q_w/\lambda - 1/\mu
\]
and
\[
\phi_{h_{n}} - \text{Det} \approx 1 - (1 - \phi_{\text{off}})(1-\rho) \frac{Q_w - [\lambda(T_T + 1/\mu)]}{Q_w + \lambda(T_T - 1/\mu)}. \quad (35)
\]
Finally, we obtain the average frame delay substituting
\[
E[T_{\text{e}}^{\text{burst,high}}_{h_{n}} - \text{Det}] \approx 1/\lambda - 1/\mu \quad \text{and} \quad E[T_{\text{e}}^{\text{burst,high}}_{h_{n}} - \text{Det}] \approx \sigma_S^2 + (1/\lambda - 1/\mu)^2 \quad \text{into (22)}:\n\]
\[
E[W]^{\text{burst,high}}_{h_{n}} - \text{Det} \approx \frac{\lambda(\lambda T_{w} + Q_{w} - 1)\sigma_S^2}{(2(1-\rho)(\lambda T_{w} + Q_{w} - \rho))} + \frac{Q_{w} - 1 + T_{w}}{2\lambda}. \quad (36)
\]

IX. Validation

To validate our models for both frame and burst transmission algorithms, we conducted several simulation experiments on the ns-2 simulator [27]. The simulated interface has a 10Gb/s capacity and we vary the average arrival rate from 10Mb/s to 10Gb/s. Frame size is set to 1000bytes.

We chose popular 10GBASE-T interfaces for our experiments, so \(T_e = 2.88\mu s, T_w = 4.48\mu s\) and \(\phi_{\text{off}} = 0.1\) according to several estimates provided by different manufacturers during the standardization process of the IEEE 802.3az standard.

We configured burst transmission with \(T_{\text{max}} = 100\mu s\) and \(Q_{w} = 20\)frames (> \(\mu T_{s} = 3.6\)frames as suggested in Sect VI-A).

We ran each simulation for 10 seconds and measured both the energy consumption and the average frame delay. Each simulation experiment was repeated ten times using a different seed value when initializing the random number generator. Then, an average of the measured parameter was taken over all runs.7

We conducted three different sets of experiments to validate our models for both Poisson and deterministic traffic and our general model when there is no information about the distribution followed by interarrival times.

A. Poisson Traffic

Fig. 5 plots both the measured and the predicted results obtained with Poisson traffic when using the frame transmission method. As shown, EEE reaches important energy savings even with a simple method like this, just increasing frame delay in a hardly perceivable way. In addition, our model predictions are very accurate for both the energy consumption and the average queueing delay.

Fig. 6 shows both the measured and the predicted results when using burst transmission. Estimations under both low and high load conditions are depicted in the graphs. Additionally, frame transmission results are also depicted for comparison. Once more, our model produces very accurate predictions for both the energy savings and the average queueing delay that match the values obtained by simulation. Furthermore, it can be observed that the theoretical arrival rate threshold obtained in (18) delimits in a quite precise way the valid load regions for each of our models (\(\lambda^* \approx 1.5\)Gb/s in this scenario). Finally, it is worth noting that, as expected, sending bursts provides greater energy savings at the expense of slightly increasing the average queueing delay. Note that, in any case, this increment is upper bounded by \(T_{\text{max}} + T_{w}\).

7Although 95% confidence intervals have been also calculated, they are not represented in the graphs since they were consistently lower than \(\pm 1\%\) and just clutter the figures.
B. Deterministic Traffic

Fig. 7 plots both the measured and the predicted results obtained with deterministic traffic when using the frame transmission method. As predicted by our analysis, if \( \lambda > (1/\mu + T_s + T_w)^{-1} \), that is, \( \lambda > 980 \text{Mb/s} \) for the simulated scenario, no power savings can be achieved with this algorithm. Again, our model produces accurate predictions except for those rates slightly higher than the given rate threshold. That is because our model has not taken into account that, at these rates, it is still possible that the interface occasionally enters into the low power mode during a few cycles.

Fig. 8 shows both the measured and the predicted results with the burst transmission method. Note that, as predicted by our model, when the traffic load is so low that \( \lambda T_{\text{max}} < 1 \), that is, \( \lambda < 80 \text{Mb/s} \) in this scenario, the same amount of power is saved with both frame and burst transmission algorithms, but frames suffer the maximum possible delay \( T_{\text{max}} + T_w \).

C. Real Traffic

Finally, we validated our model using real world traffic traces publicly available from the CAIDA archive [16]. The analyzed CAIDA traces were collected during 2011 on a 10Gb/s backbone Ethernet link of a Tier1 ISP between Chicago, IL and Seattle, WA. Fig. 9 shows both the measured and the predicted results for these traces. In this case, we cannot make any assumption about the distribution followed by interarrival times, so we have to rely on the general equations derived in Sections IV-A and VI-B. Thus, to estimate energy consumption we have employed equation (6) so we had to measure the average duration of sleeping periods for the simulated traces. As shown in Fig. 9(a), our model provides very accurate predictions for this parameter. To estimate frame delay we have used equation (13) for frame transmission and equations (19) and (22) for burst transmission, so we measured the average length of empty periods and its square. For this parameter, the model predictions are still quite good, but there is a slight bias for medium traffic loads when using burst transmission, probably because traffic is not 100% uncorrelated as assumed by our analysis.

X. Conclusions

This paper provides a general model for the behavior of the two most popular sleeping algorithms proposed for application in EEE interfaces: frame transmission and burst transmission. This model can be used to analyze the influence of EEE configuration parameters and interface physical characteristics on the expected performance. In contrast to those models already present in the literature, the provided model does not place any additional constraints on the arrival process distribution. To the best of our knowledge, this is the first model to do so. Moreover, our model does not only predict the average power savings, but also the effects of the algorithms on packet delay.

---

The oscillations shown in Fig. 8 in the low load regime are directly caused by the floor and ceiling functions used in equations (33) and (34) that quantify, respectively, the energy savings and the average frame delay achieved when using burst transmission with deterministic traffic.

---

The oscillations shown in Fig. 8 in the low load regime are directly caused by the floor and ceiling functions used in equations (33) and (34) that quantify, respectively, the energy savings and the average frame delay achieved when using burst transmission with deterministic traffic.
To further test the validity of the general model, we also provide two specializations for Poisson and deterministic traffic arrivals. Finally, we have proved the model through simulation with the help of both synthetic and real traffic. The obtained results show the correctness of the models for Poisson and deterministic traffic. Moreover, the tests with real traffic assess the validity of the general model for predicting the effects of EEE algorithms, and their configuration parameters, in the Internet.

ACKNOWLEDGMENTS

This work was supported by the “Ministerio de Educación y Ciencia” through the project TEC2009-12135 of the “Plan Nacional de I+D+i” (partly financed with FEDER funds). Support for CAIDA’s Traces is provided by the National Science Foundation, the US Department of Homeland Security, and CAIDA Members.

REFERENCES

Fig. 9. Results for CAIDA traces.

rate (ALR): analysis of a buffer threshold policy,” in Proc. 2006 IEEE
GlobeCom, pp. 1–6.
energy consumption of Ethernet with an adaptive link rate (ALR),” IEEE
[22] K. T. Marshall, “Bounds for some generalizations of the GI/G/1 queue,”
policy for a class of single-server queues,” Operations Research, vol. 16,
no. 6, pp. 1138–1146, Nov. 1968.
self-similar nature of Ethernet traffic (extended version),” IEEE/ACM
Poisson view of Internet traffic,” in Proc. 2004 IEEE INFOCOM, vol. 3,
pp. 1558–1569.
at short time-scales in a small buffer core network?” in 2009 IEEE
edu/nsman/ns/

Sergio Herrera-Alonso received his M.Sc. and Ph.D. degrees in Telecommunication Engineering from the University of Vigo (Spain) in 2001 and 2006 respectively. He is currently an associate professor in the Dept. of Telematics Engineering at the University of Vigo and affiliated member of the co-located Networking Laboratory. His research interests include quality of service in Internet and energy-efficient networking.

Miguel Rodríguez-Pérez received his M.Sc. and Ph.D. degrees in Telecommunication Engineering from the University of Vigo (Spain) in 2001 and 2006 respectively. He is currently an associate professor in the Dept. of Telematics Engineering at the University of Vigo and affiliated member of the co-located Networking Laboratory. His research interests include congestion control and traffic engineering with a strong focus on energy efficiency.

Manuel Fernández-Veiga received his M.Sc. and Ph.D. degrees from the University of Vigo (Spain) in 1992 and 2001, both in Telecommunication Engineering. He is currently associate professor at the University of Vigo and affiliated member of the co-located Networking Laboratory. His research interests include communication theory, performance analysis of computer networks and wireless networking. He has authored over 50 papers in international conferences and journals.

Cándido López-García received his M.Sc. and Ph.D. degrees in Telecommunication Engineering from the Universidad Politécnica de Madrid (Spain) in 1988 and 1995 respectively. He is currently associate professor at the University of Vigo (Spain) and affiliated member of the co-located Networking Laboratory. His main research interests are in the area of performance evaluation in communication networks. He has authored over 60 papers in international conferences and journals.