On the Impact of the TCP Acknowledgement Frequency on Energy Efficient Ethernet Performance

Pedro Reviriego, Alfonso Sanchez-Macian and Juan Antonio Maestro

Universidad Antonio de Nebrija, Calle Pirineos 55,
28040 Madrid, Spain
{previrie, asanchep, jmaestro}@nebrija.es

Abstract. With the adoption of the recently approved Energy Efficient Ethernet standard, frame scheduling will have a significant impact on Energy Consumption. Acknowledgements are large percentage of frames in Transmission Control Protocol sessions. In this paper the impact of different TCP acknowledgement policies on the energy consumption of Ethernet Networks is analyzed.

**Keywords:** Ethernet, Energy Efficiency, IEEE 802.3az, 10GBASE-T.

1 Introduction

Acknowledgements play an important role in the Transmission Control Protocol (TCP). Acknowledgements (ACKs) are used to detect lost segments, to trigger the transmission of new segments and to determine the sending rate. For those purposes, ideally an ACK should be generated each time a segment is received at the destination. However, to reduce the burden on the network nodes, other acknowledgement policies have been proposed for TCP [1]. Reducing the number of ACKs is especially useful in asymmetric networks or when the cost of sending and ACK is high [2]. Those policies reduce the number of ACKs so that, for example, an ACK is sent only for every second segment that is received. More aggressive schemes that further reduce the number of ACKs have also been proposed, but there has been a concern that an aggressive reduction in the number of ACKs could impact TCP

performance. A recent study, however, suggests that good performance can be achieved even when the number of ACKs is significantly reduced [3].

The efficient use of energy in communications is an area of growing interest [4] and Ethernet is a good example where significant savings can be achieved. Those savings have been estimated to be in the order of TWh [5]. To improve energy efficiency in Ethernet, the IEEE 802.3az Task force (Energy Efficient Ethernet) has defined energy efficiency enhancements, a work that has produced a standard in September of 2010 [6].

Essentially, current Ethernet standards require both transmitters and receivers to operate continuously on a link, thus consuming energy all the time, regardless of the amount of data exchanged. This consumption depends on the link speed and ranges from around 200 mW for 100 Mbps [7] to about 4 W for 10 Gbps [8]. Energy Efficient Ethernet (EEE) aims to make the consumption of energy over a link more proportional to the amount of traffic exchanged [6]. To this end, EEE defines a low power "sleep" mode. The physical layer is put into this sleep mode when no transmission is needed, and woken up very quickly upon data arrival without changing its speed. The low-power mode freezes the elements in the receiver, and wakes them up within a few microseconds. Such sleep/active operation requires minor changes to the receiver elements since the channel is quite stable.

Figure 1 shows a state transition example of a given link following the IEEE 802.3az standard [6]. Here, Ts refers to the sleep time (time needed to enter the low-power mode), and Tw denotes the wake-up time (time required to exit the low-power mode). The transceiver spends Tq in the quiet (energy saving) period. Finally, the standard also considers the scheduling of periodical short periods of activity Tr to refresh the receiver state in order to ensure that the receiver elements are always aligned with the channel conditions.
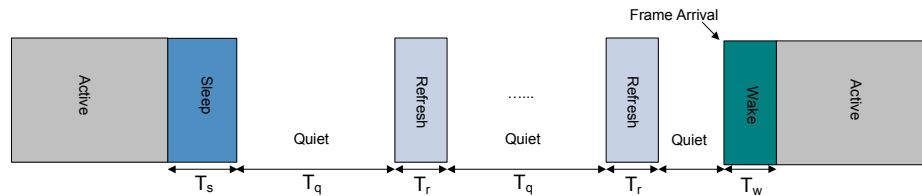
Fig. 1. Transitions between active and low-power modes in Energy Efficient Ethernet.

Concerning energy efficiency of EEE, there is significant energy consumption only during Tw, Ts and Tr periods, and a small fraction of it (about 10%) during Tq, with Tq>>Tr. The minimum and maximum values for timers Tw, Ts, Tq and Tr are specified in the IEEE 802.3az standard for 100BASE-TX, 1000BASE-T and 10GBASE-T. The transition times in all cases are in the order of microseconds, which are comparable or even larger than frame transmission times for some speeds. This results in a significant energy overhead when a link is activated to send a single frame. The performance of EEE in different scenarios has been recently studied in [10] showing that in many cases the transitions cause a large overhead. This is especially harmful for ACKs as they are very small and therefore the relative

overhead is larger. This situation is similar to the one experienced on some wireless networks [2] where there is a large cost for sending a frame regardless of its size.

In this paper, the impact of the acknowledgement policy used in TCP on the energy savings obtained in EEE is evaluated. The results show that reducing the ACK frequency can provide significant energy savings in many situations and therefore brings another incentive to further study the performance of TCP when the number of ACKs is reduced.

2 Impact of the Acknowledgement Frequency on Energy Efficient Ethernet

For this initial evaluation, 10GBASE-T links have been selected. These high speed links are of interest, as they are the ones that consume more power and they are increasingly used in servers. The transition times for 10GBASE-T, as per the latest standard, are approximately 5 μs to wake a link and close to 3 μs to put it into the low power mode [10]. This compares to a transmission time of 1. 2μs for a 1500 byte frame and less than 0.1 μs for the transmission of an ACK.

For this evaluation, the Network Simulator (NS) version 2.33 has been used. The wiredPhy and 802.3 MAC were modified to implement the transitions between active and low power modes as per the EEE standard. The time that a link stays in active mode, low power mode and transitions are logged such that an analysis of the energy consumption can be done.

To estimate the energy consumption from the link states, two alternatives are used. In the first, the energy consumption during transitions is assumed to be the same as in the active state, as in [10]. In the second, the energy consumption during transitions is assumed to be 50% of the active mode to account for the fact that some elements in the transmitter and receiver may be idle during the transitions. In both estimates, the energy consumption in low power mode is assumed to be 10% of the active mode. The first estimate will be referred to as upper bound estimate and the second as lower bound estimate, as they are meant to provide the range of consumptions that can be found in actual implementations.

For this initial study, two simple ACK generation policies are evaluated: i) sending an ACK for every segment that is received (AE), and ii) sending either an ACK for every second segment that is received or a delayed ACK if another frame does not arrive shortly after the first one (DAL2) [3]. In all the experiments, the timer that triggers an ACK transmission in DAL2 if another frame is not received has been selected to be 0.1ms except in the third experiment where 5ms were used. These values ensure that data transfers are completed in almost the same time in AE and DAL2. That means that in all the experiments the network performance from an application perspective is the same regardless of the TCP acknowledgement policy used. This is consistent with previous studies that show how the ACK frequency can be reduced without impacting TCP performance [3]. Therefore in the following, the analysis of the results will focus on the energy consumption.

To evaluate the impact of the ACK frequency on the energy consumption a number of experiments have been done. For the first experiment, the scenario

considered is illustrated in Figure 2 and consists of a server connected to a switch by a 10GBASE-T EEE link that sends data to a server connected to the same switch on a 1 Gbps link. The data is sent over an FTP connection.
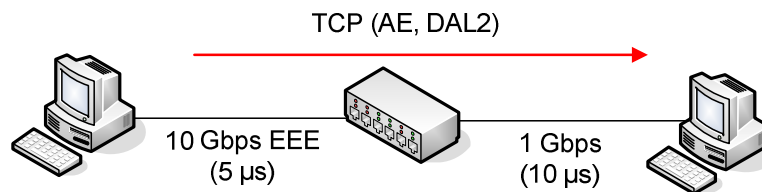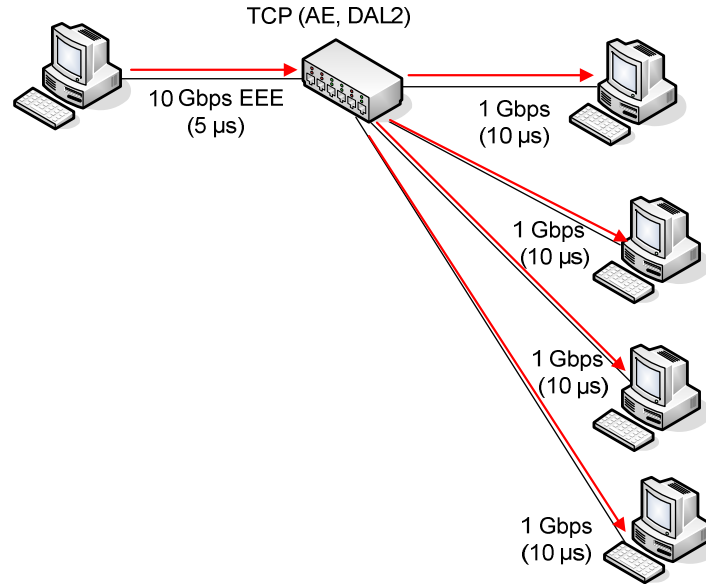

Fig. 2. Scenario for the first experiment.

The results for the first scenario are shown in Table I for the link directions in which the data and ACKs are sent. From the results it can be observed that when AE is used, the energy consumed to send the ACKs is similar to that needed to send the data. This is due to the large transition overheads, as mentioned before. In both cases (AE,DAL2) the link utilization is close to 10%, limited by the 1 Gbps connection. If the energy consumption of AE and DAL2 are compared, it can be seen that DAL2 provides a significant reduction in both directions. This is explained as in this case the 1 Gbps link spaces the ACKs by 12 μs in the AE case and therefore each ACK needs to activate the link. Additionally, each ACK will trigger a frame transmission upon its arrival at the source, and since ACKs are spaced by 12 μs, so will frame transmissions. This means that the link would also have to be activated for each frame. When DAL2 is used, the number of ACKs is reduced to one half. Also, each ACK triggers the transmission of two frames that can share the overhead of activating the link. This explains the reduction in the direction in which the data is sent.

The observed performance in which DAL2 reduces the transition overhead to one half would occur when frames are spaced due to a lower speed connection or if the link load is so low that each frame requires a link activation.

In the second experiment, a scenario in which there are multiple TCP connections is considered. The scenario is illustrated in Figure 3. In this case, there are four simultaneous FTP connections, each limited by a 1Gbps link. The connections in this and the next experiments are started at random instants in time at the beginning of the simulation, in order to avoid source correlation effects.

The results for the second experiment are also shown in Table I. In this case the load of the link is higher and reaches 40% in the data direction. This results in a larger energy consumption and the link is either active or in transitions most of the time. Even for this high load, the use of DAL2 reduces the transition overhead and therefore the energy consumption significantly in both directions.
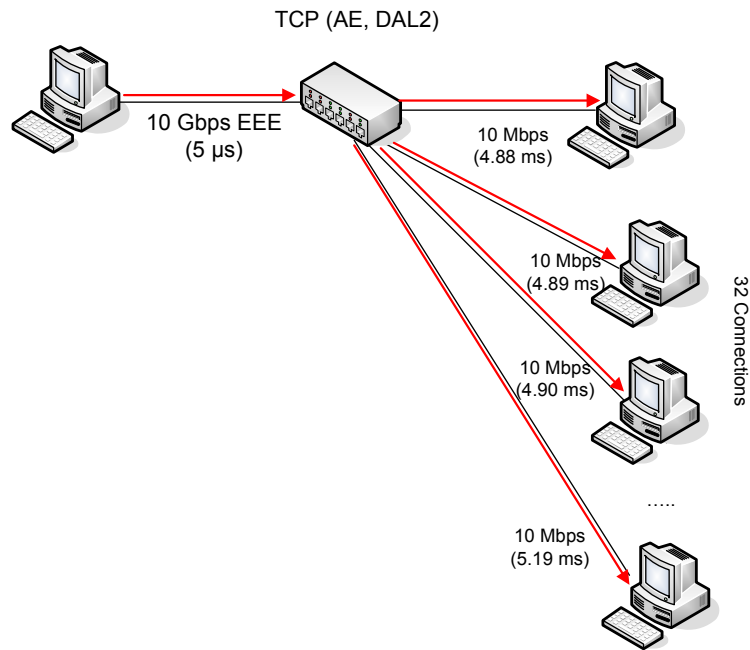
Fig. 3. Scenario for the second experiment.

Table 1. Results for the NS-2 experiments.

ACK Policy	Direction	Percentage of time active	Percentage of time in transitions	Energy consumption Upper bound	Energy Consumption Lower bound
FIRST EXPERIMENT					
AE	Data	9.99	70.60	82.54	47.24
DAL2	Data	9.99	35.28	50.74	33.10
AE	ACKs	0.49	70.61	73.99	38.69
DAL2	ACKs	0.25	35.28	41.97	24.33
SECOND EXPERIMENT					
AE	Data	39.99	47.67	88.92	65.09
DAL2	Data	39.96	35.69	78.08	60.24
AE	ACKs	1.99	94.15	96.51	49.45
DAL2	ACKs	1.01	70.95	74.75	39.28
THIRD EXPERIMENT					
AE	Data	3.13	19.15	31.42	21.09
DAL2	Data	3.13	10.37	24.06	17.81
AE	ACKs	0.16	19.33	28.99	18.52
DAL2	ACKs	0.08	10.49	21.46	15.13
FOURTH EXPERIMENT					
AE	Data	42.95	11.42	59.82	53.62
DAL2	Data	42.95	11.97	60.26	53.81
AE	ACKs	2.14	48.92	56.93	31.93
DAL2	ACKs	1.13	50.14	57.06	31.48

In the third experiment, a situation in which there is a larger number of TCP connections is considered. The scenario is shown in Figure 4 and in this case, 32 connections are used each limited by a 10 Mbps link.

The results are shown in Table I. In this case, the link load is around three per cent, lower than in the previous experiments. However even for this low load, significant energy savings are achieved when DAL2 is used. The relative percentage of time spent in transitions in AE and DAL2 is similar to that of the first experiment, as in most cases each frame transmission requires a link activation when AE is used.
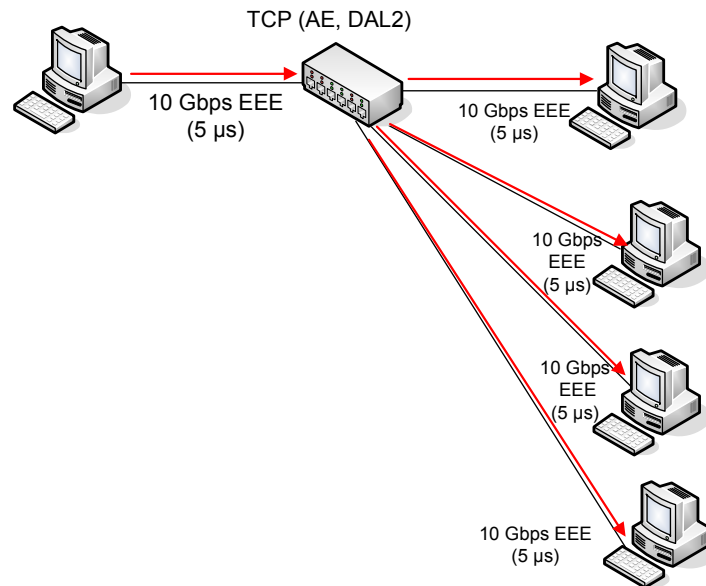

Fig. 4. Scenario for the third experiment.

In the last experiment, a situation in which multiple servers are connected by 10 Gbps EEE links is considered. The scenario is shown in Figure 5 and the analysis focuses on the link of the server that sends data to the other four. In this case, sources have on/off periods determined by an exponential distribution with a sending rate of 10 Gbps.

The results are shown in Table I. In this case, there are no energy savings from using DAL2, as frames are transmitted in bursts and ACKs are spaced by either 1.2 or 2.4 μs, which is less than the time needed to wake and sleep the link. In fact for bursts with an odd number of frames the delayed ACKs can cause additional link transitions thus slightly increasing the energy consumption.

From the experiments presented, it becomes apparent that reducing the ACK frequency will help reduce the energy consumption in EEE in many cases. This is due to the reduced number of ACKs itself and to its effect on grouping data frame transmissions in bursts. More aggressive ACK schemes, such as those evaluated in [3] should provide additional energy savings to those reported here.

Fig. 5. Scenario for the fourth experiment.

3   Conclusions

The reduction of the acknowledgement frequency in TCP has been considered for different reasons in the past. In this paper, a new situation in which the reduction of ACKs is beneficial has been described: Ethernet Networks that implement the Energy Efficient Ethernet standard. The initial results presented show that in many cases significant reductions in energy consumption can be achieved by reducing the acknowledgement frequency. This finding should motivate further research on the reduction of the acknowledgement frequency in TCP to maximize energy savings while preserving performance.

Another alternative to reduce the energy consumption in Ethernet is to use larger frame sizes. For example, if 8,000-byte frames were used instead of 1,500-byte ones, the number of acknowledgements would be automatically reduced by a factor of more than five. This may be an interesting option for networks in which traffic is internal and there are no compatibility issues with legacy equipment that do not support the use of large frames.

Finally, although the experiments have concentrated on 10GBASE-T, similar results are expected for the other Ethernet speeds for which the reduction of the acknowledgement frequency will decrease the energy consumption significantly in many cases.

References

1. M. Allman "On the Generation and Use of TCP Acknowledgments" ACM SIGCOMM Computer Communication Review, vol 28, No. 3, pp. 4–21, Oct. 1998.
2. R. de Oliveira and T. Braun "A Smart TCP Acknowledgement Approach for Multihop Wireless Networks" IEEE Transactions on Mobile Computing, Vol 6, Issue 2, February 2007 pp 192-205.
3. S. Landström and L. Larzon "Reducing the TCP Acknowledgment Frequency" ACM SIGCOMM Computer Communication Review, vol 37, No. 3, pp 7-16, July 2007.
4. M. Gupta and S. Singh "Greening of the Internet" Proc. of ACM SIGCOMM pp 19-26, August 2003.
5. K. Christensen, K., P. Reviriego., B. Nordman.,M. Bennett, M. Mostowfi, and J.A. Maestro, "IEEE 802.3az: The Road to Energy Efficient Ethernet", IEEE Communications Magazine, vol. 48 , issue 11, November 2010, pp. 50-56.
6. IEEE Std 802.3az: Energy Efficient Ethernet-2010.
7. Chou, J. "Low-Power Idle based EEE 100BASE-TX" IEEE 802.3az Task Force Presentation, March 2008.
8. Kohl, B "10GBASE-T Power Budget Summary" IEEE 802.3az Task Force Presentation, March 2007.
9. L.A. Barroso and U. Holzle "The Case for Energy-Proportional Computing" IEEE Computer, vol 40, issue 12, Dec. 2007 pp:33 – 37.
10. P. Reviriego, J.A. Hernández, D. Larrabeiti, J.A. Maestro, "Performance Evaluation of Energy Efficient Ethernet", IEEE Communications Letters , Vol. 13, No 9, Sept. 2009, pp. 697-699.