Increasing the MTU size for Energy Efficiency in Ethernet

P. Reviriego*, A. Sánchez-Macián*, J.A. Maestro* and C.J. Bleakley**

** Universidad Antonio de Nebrija, Spain*

email:{previrie, asanchep, jmaestro}@nebrija.es

*** University College Dublin, Ireland*

email: chris.bleakley@ucd.ie

Abstract— **The commonly used Maximum Transfer Unit (MTU) on the Internet has remained unchanged for many years at around 1500 bytes due mainly to backward compatibility issues. This is in contrast with link data rate, which has increased by several orders of magnitude. In this paper, a new advantage of using larger MTUs is introduced, namely Energy Efficiency. In wire-line environments, the link power consumption is generally roughly independent of the number of frames that are transmitted resulting in a poor energy efficiency. This will change with the development of standards like IEEE 802.3az, Energy Efficient Ethernet. This new standard allows a link to enter a low power mode when there are no frames to transmit therefore making power consumption almost proportional to the link load. In this context the use of larger MTUs minimizes the number of transitions between the active and low power modes thereby improving energy efficiency. The benefits of using larger MTUs in terms of energy efficiency are analyzed in this paper.**

Keywords – **Ethernet, Energy Efficiency, MTU.**

I INTRODUCTION

The Maximum Transfer Unit (MTU) commonly used in the Internet has remained unchanged for many years at around 1500 bytes. This corresponds to the value of the MTU for the original Ethernet standard [1]. Increasing the MTU has been shown to improve performance and reduce processing overhead in routers and switches [2]. The main drawbacks of using larger MTUs are the additional delay and increased probability of a frame suffering an error. The first drawback is minimized by the fact that link speeds have increased by several orders of magnitude such that a 9,000-byte frame takes only 7.2us to transmit over a 10Gbps link, a delay that is acceptable for most applications. The second drawback is not an issue either, as the newest Ethernet standards require low values for the bit error rate (BER). For example, in 10GBase-T, a BER below 10^{-12} is specified [3]. This means that for a 9,000-byte frame the probability of error is below $7.2*10^{-8}$, which is again acceptable for most applications. This leaves backward compatibility as the main issue preventing the widespread use of larger MTUs in Ethernet LANs.

Energy Efficiency in wire-line communication networks has been disregarded for many years but it is now gaining attention given the large amount of energy consumed in network devices [4]. One of the areas where significant savings can be achieved is in the physical layer, as, in many cases, the transmitters and receivers operate continuously even when there is no data to transmit. This is the case in most existing Ethernet standards [5] and has motivated the development of the IEEE 802.3az, Energy Efficient Ethernet (EEE) standard [6]. The importance of Ethernet in terms of energy consumption stems from its wide adoption in LANs. Over a hundred million Ethernet ports are shipped each year and the installed based could be an order of magnitude higher. The energy consumption of all these devices has been estimated to be in the order of a few TWh [4],[5] even though the individual power consumption is small. A 100Mbps device consumes less than 200mW, while at 1Gbps it can reach 500mW and at 10Gbps, 5Watts. This shows the trend to greater power consumption as speed increases in each new Ethernet generation. Consequently, as new technologies are adopted, the energy consumption of Ethernet networks will grow.

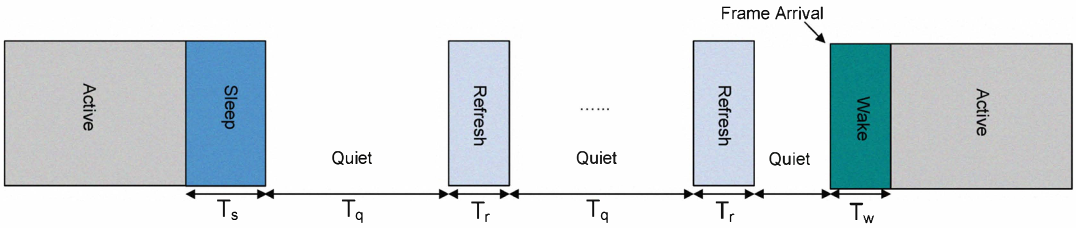

Figure 1 Transitions between the active and low-power modes in Energy Efficient Ethernet.

The EEE standard defines a low power mode such that when there are no frames to transmit, the link enters low power mode, saving a significant amount of energy. The operation of the EEE standard is illustrated in Figure 1. It can be observed that periodic refreshes are scheduled to ensure that the transmitter and receivers stay aligned to the channel conditions. The transition times for switching between active and low power modes are short, as shown in Table I, and therefore should be acceptable to most applications.

Table I Proposed wake and sleep times for different link data rates

Protocol	Min T_w (usec)	Min T_s (usec)
100Base-TX	30.5	200
1000Base-T	16.5	182
10GBase-T	4.48	2.88

In EEE, there is significant power consumption in active mode and during transitions, whereas large savings are achieved in low power mode. However, when isolated frames are transmitted frequently, there is a significant overhead due to mode transitions since the frame transmission times can be shorter than the frame transition times themselves [7]. For example, a 1,500-byte frame takes 120us, 12us or 1.2us to be transmitted at 100Mbps, 1Gbps and 10Gbps, respectively. These transmission times are all shorter than the corresponding mode transition times given in Table I. Hence, the overhead of mode transition is large. The situation is much worse for smaller frames, such as TCP acknowledgements. If larger MTUs were to be used then the number of frames required to transmit a given amount of data would be reduced and so too would be the overhead due to mode transition. This idea is analyzed in more detail in the rest of the paper.

Although it is not the objective of this paper, it is interesting to note that the use of larger MTUs would also reduce the energy consumption in other elements of both switches and routers. This is because many operations are applied per packet, independent of packet size (for example destination lookup). Existing systems already exhibit some reduction of the power consumption when larger frames are used [8] and those savings are likely to increase as more effort is put into designing more energy efficient systems.

II EFFECTS OF A LARGER MTU ON ENERGY EFFICIENT ETHERNET

Increasing the MTU has a double effect on the number of frames transmitted. For example, if an MTU of 9,000-bytes is used instead of the traditional 1,500-bytes, the number of packets required to send a given amount of data would be reduced by a factor of six. Additionally, as most data transfers use TCP and most TCP implementations send an acknowledgement (ACK) packet every time a packet is received, the number of ACKs would also be reduced by a factor of six. This means that both the number of large (data) and small (ACK) frames would be substantially reduced. This reduction would in turn minimize the transition overhead in EEE.

To evaluate the impact of raising the MTU on EEE energy consumption, simulations have been run using 1,500-byte and 9,000-byte frames. In the first set of simulations, Poisson arrivals are assumed as an initial approximation. The results are illustrated in Figures 2 to 4 and show how the use of larger MTUs can significantly reduce energy consumption, especially at higher data rates (1Gbps and 10Gbps). In the simulations the energy consumption in the low power mode was assumed to be 10% that during active mode and during mode transitions [7].

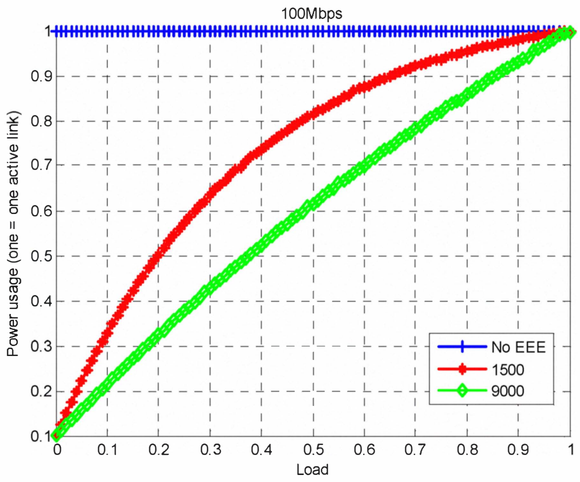

Figure 2 Energy Consumption versus load for 1,500 and 9,000-byte frames for 100Base-TX.

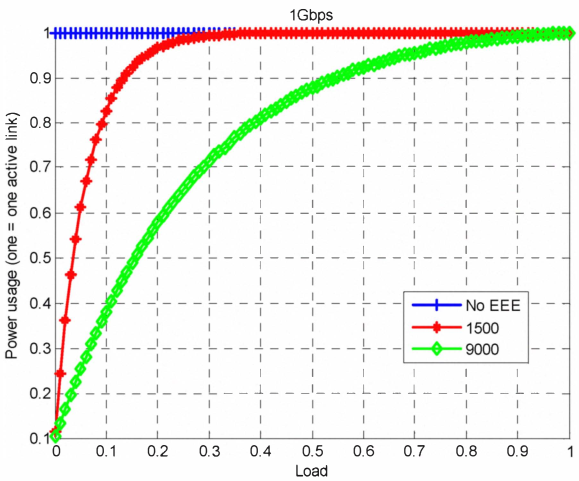

Figure 3 Energy Consumption versus load for 1,500 and 9,000-byte frames for 1000Base-T.

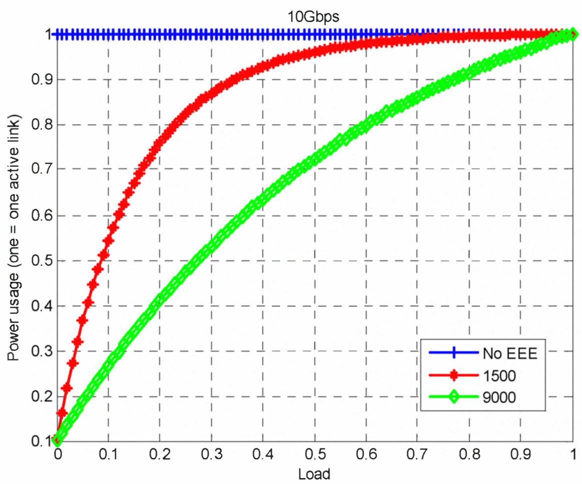

Figure 4 Energy Consumption versus load for 1,500 and 9,000 byte frames for 10GBase-T.

Although the results from these simulations show a large potential benefit in using larger MTUs, it is well known that the traffic in Local Area Networks does not follow a Poisson process [9]. Therefore, to better evaluate the impact of using larger MTUs, simulations using NS-2 [10] have also been done. The EEE transition times have been implemented in NS-2 as a new PHY type, such that the energy consumption can be calculated from the time a link is active, in transition or in low power mode. The use of NS-2 allows us to evaluate the energy consumption in experiments that resemble real scenarios.

In the first experiment the simple scenario shown in Figure 5 is considered. A large FTP is done over a TCP connection and the speed of the transfer is limited by the 1Gbps link. The results are shown in Table II. It can be observed that large energy savings are achieved in both directions. This is because with an MTU of 1,500-bytes the frames are spaced by 12us by the 1Gbps link and so are the ACKs in the reverse direction. This, in turn, causes new data frame transmission to be spaced by 12us such that the link is activated, the frame is transmitted (in 1.2us) and then set back to low power mode before a new frame is transmitted. In this case, most frames require a link transition leading to a large energy consumption overhead. With the largest MTU the number of transitions is greatly reduced.

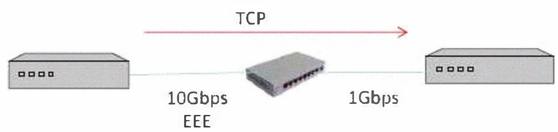

Figure 5 Scenario for the first experiment.

Table II Results for the first experiment

Scenario 1	Direction	Power Consumption
MTU = 1500	Data	82.54
	Ack	73.99
MTU = 9000	Data	31.52
	Ack	22.44

In the second experiment, the scenario shown in Figure 6 is considered in which parallel FTP transfers take place. The results are shown in Table III. In this case, smaller savings are obtained due to the larger link load (around 40% in the data direction) but still the improvement is significant, especially in the ACK direction. This can be explained by noting that with a 40% load, the link will be active a significant fraction of the time in any case. While in the ACK direction, the load is below 1% and reducing the number of ACKs is more beneficial.

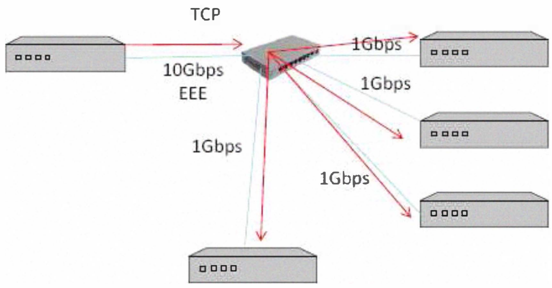

Fig. 6 Scenario for the second experiment.

Table III Results for the second experiment

Scenario 2	Direction	Power Consumption
MTU = 1500	Data	89.26
	Ack	92.49
MTU = 9000	Data	72.10
	Ack	49.57

Finally, in the last experiment, a larger number of lower speed connections are considered, as illustrated in Figure 7. The results are shown in Table IV and again significant energy savings in relative terms are obtained when using a larger MTU. In this case, the link load is around 3% in the data direction. This shows that the use of larger MTUs can have a sizeable impact even at low loads.

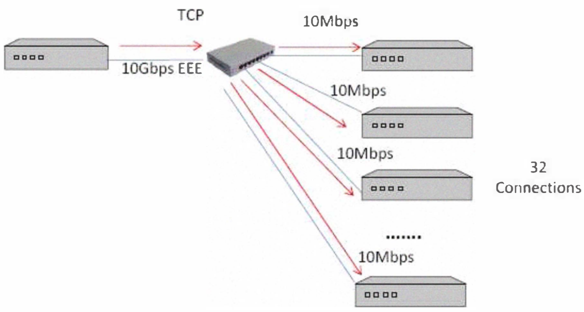

Fig. 7 Scenario for the third experiment.

Table IV Results for the third experiment

Scenario 3	Direction	Power Consumption
MTU = 1500	Data	31.42
	Ack	28.99
MTU = 9000	Data	18.49
	Ack	15.88

The benefits of using larger MTUs depend on whether typical applications can use larger frames or not. Intuitively, use of larger frames seems reasonable as most data transfers exceed 1,500-bytes. However, to corroborate this assumption, several packets traces from a 10Gbps Internet link have been analyzed. These traces are available through CAIDA [11]. In most traces, a significant percentage of the packets are close to 1,500-bytes, meaning that larger MTUs could be used in many connections. As an example, the distribution of packet sizes in one of the traces is shown in Figure 8. It can be observed that nearly 25% of the packets are close to 1,500bytes. If the corresponding ACKs for those frames are added then nearly 50% of the frames would benefit from the use of a larger MTU.

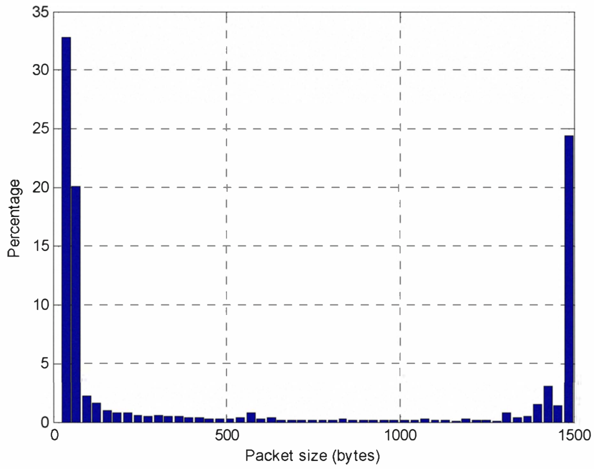

Fig. 8 Distribution of packet sizes in a packet trace of a high speed Internet link.

III CONCLUSIONS

In this work, a new advantage of using larger frames in computer networks and more specifically in Ethernet LANs has been discussed. In particular, it has been shown that significant energy savings can be obtained by using larger frames in LANs that implement the Energy Efficient Ethernet (802.3.az) standard. The simulations show that an MTU of 9,000-bytes would greatly improve the performance compared to an MTU of 1,500-bytes commonly used today. This is due to a reduction in the number of data frames and ACKs in TCP sessions. This reduction minimizes the number of transitions between active and low power modes in Energy Efficient Ethernet thus enabling greater energy savings. The use of larger frames would also have a positive effect on the power consumption of other switch or router elements, as much of the processing is done per packet.

ACKNOWLEDGEMENTS

Support for CAIDA's Internet Traces is provided by the National Science Foundation, the US Department of Homeland Security, and CAIDA Members.

This work was partly funded by a Google Research Award.

REFERENCES

[1] R. Seifert and J. Edwards, "The All-New Switch Book, the Complete Guide to LAN Switching Technology", John Wiley and Sons, 2008.

[2] W. Rutherford, L. Jorgenson, M. Siegert, P. Van Epp and L. Liu, "16000–64000bits MTU experiments with simulation: The case for super jumbo frames", Optical Switching and Networking, Volume 4, Issue 2, June 2007, pp. 121-130.

[3] "IEEE 802.3an: 10GBase-T", published as part of the IEEE 802.3 standard.

[4] M. Gupta and S. Singh, "Greening of the Internet" Proc. of ACM SIGCOMM pp 19-26, August 2003.

[5] C. Gunaratne, K. Christensen, B. Nordman and S. Suen, "Reducing the Energy Consumption of Ethernet with Adaptive Link Rate (ALR)", IEEE Transactions on Computers, vol. 57, issue 4, April 2008 pp.: 448 – 461.

[6] IEEE P802.3az Energy Efficient Ethernet Task Force (http://grouper.ieee.org/groups/802/3/az/public/index.html)

[7] P. Reviriego, J.A. Hernández, D. Larrabeiti, and J.A. Maestro, "Performance Evaluation of Energy Efficient Ethernet", IEEE Communications Letters, Vol. 13, No 9, September 2009, pp. 697-699.

[8] D. Kharitonov, "Time-Domain Approach to Energy Efficiency: High-Performance Network Element Design", IEEE GLOBECOM Workshops, 2009.

[9] W. E. Leland, M. S. Taqqu, W. Willinger and D. V. Wilson, "On the Self-similar Nature of Ethernet Traffic", IEEE/ACM Transactions on Networking (TON), Vol.2 No.1, Feb. 1994, pp.1-15.

[10] L. Breslau, D. Estrin, K. Fall, S. Floyd, J. Heidemann, A. Helmy, P. Huang, S. McCanne, K. Varadhan, Y. Xu, and H. Yu. "Advances in Network Simulation" IEEE Computer, Vol. 33, No.5, pp. 59-67, May, 2000.

[11] C.Walsworth, E. Aben, K.C. Claffy and D. Andersen, "The CAIDA Anonymized 2009 Internet Traces – 15th January 20009", available in the following webpage: http://www.caida.org/data/passive/passive_2009_dataset.xml