On IEEE 802.3az Energy Efficiency in Web Hosting Centers

Vincenzo Mancuso[+] and Angelos Chatzipapas[+,*]
[+]Institute IMDEA Networks, Madrid, Spain, [*]University Carlos III, Madrid, Spain
Email: {vincenzo.mancuso, angelos.chatzipapas}@imdea.org

***Abstract*—This letter presents results from our measurement campaign conducted in a web hosting center in Madrid. We collected traffic data to evaluate the potential power saving that could be achieved by replacing the company's gigabit wired links with the newly released Energy Efficient Ethernet connections (IEEE 802.3az standard). Using traffic traces collected with high precision timestamps, we feed a simulator and use an analytical model to compute the potential power saving achievable for the monitored link. We reveal that at least 80% of the power, consumed by links and network interfaces, can be saved more than 40% of the time. Furthermore, we show the importance of precise arrival time measurements by post-processing the collected traces to introduce random noise to the original timestamps.**

I. Introduction and Motivations

This letter presents traffic measurements that have been taken at InterHost, a web hosting center located in Madrid, Spain. InterHost allowed us to install our traffic measurement tools in front of one of their firewalls, which protects part of InterHost's customer web servers. The goal of our measurements is to collect enough data from a 1000Base-TX link to characterize the traffic behavior of a real commercial installation, and to estimate the power saving that might be achieved at the hosting center by replacing existing Ethernet links with IEEE 802.3az Energy Efficient Ethernet links, also known as EEE links [1]. In this letter, savings are estimated by means of a simulator, using real traces as input, and by means of a model previously proposed for EEE links [2]. Note that the IEEE 802.3az standard specifies that $1\ Gbps$ EEE links can go to low power state only when no traffic is present in both link directions. Therefore, the correlation of traffic in the two link directions is important for estimating the achievable power saving. Unfortunately, such correlation is also difficult to model. Indeed, there are no models available in the literature that consider bidirectional traffic on $1\ Gbps$ EEE links. Note also that, unless coalescing techniques are used, $1\ Gbps$ EEE links save power by introducing negligible delay (few μs) to wake up interfaces in power saving mode [1], therefore, μs accuracy is important in traffic measurements.

Similarly to [2], we use real traces to evaluate the potential EEE power saving, but we focus exclusively on $1\ Gbps$ links and on the impact of uplink/downlink traffic correlation on such saving. With our measurements, we show that EEE might save more than 40% of the link power most of the time, with peaks of 90% or more during night hours. We also unveil that high precision timestamps are key to achieve high accuracy estimations via simulation, and to enable the use of simplified analytical computations. In particular, noisy measurements severely impact the quality of EEE power saving estimates as soon as the maximum timestamp deviation due to noise reaches a few milliseconds, which is below the typical timestamp accuracy of non-dedicated network hardware, i.e., of inexpensive but imprecise driver timestamping. This justifies using specialized high accuracy timestamping hardware.

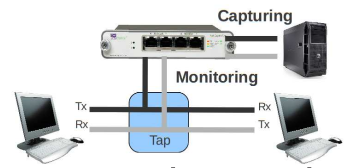


Fig. 1: Measurement architecture with passive tap.

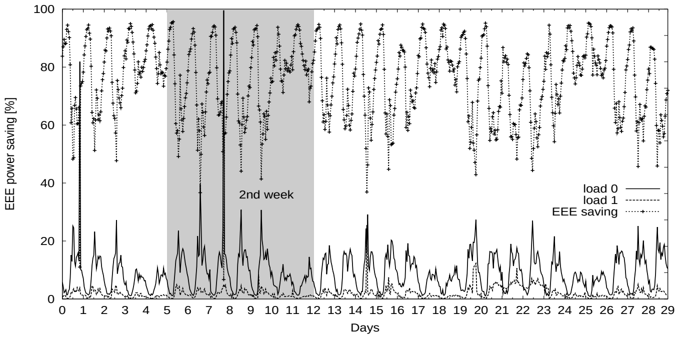
A few number of EEE performance evaluations have appeared recently. In [2] the authors model the behavior of EEE links with independent traffic in the two link directions, and validate their model by means of real traces for $1\ Gbps$ and $10\ Gbps$ links. In [3] the authors measure the power consumption of real EEE cards and report 30% of power saving for $100\ Mbps$ links, and 70% for $1\ Gbps$ links. They observe similar power consumption during state transitions and during the active state of the link. In [4] the authors first simulate the behavior of EEE links with independent traffic in the two directions, and then estimate the power consumption in different utilization scenarios by capturing real traces. They prove that, even in case of low traffic load, there is a consistent energy cost due to transitions to and from the low power state of the EEE link. In [5] the authors evaluate the impact of packet bursts on power saving and packet delay. They show that packet bursts save significant power with negligible packet latency due to burstification. Indeed, according to [6], packet coalescing does not impact the end-to-end packet delay along the Internet as long as coalescing is limited up to 100 packets.

In this work we use the model in [2], which considers pure EEE operations with no coalescing, and provides a good tradeoff between simplicity and accuracy. Few other EEE models recently appeared in literature. The model in [7] allows to compute the EEE power saving from traffic statistics in case of very low load. The model in [8] is more generic but focuses mostly on packet latency and coalescing.

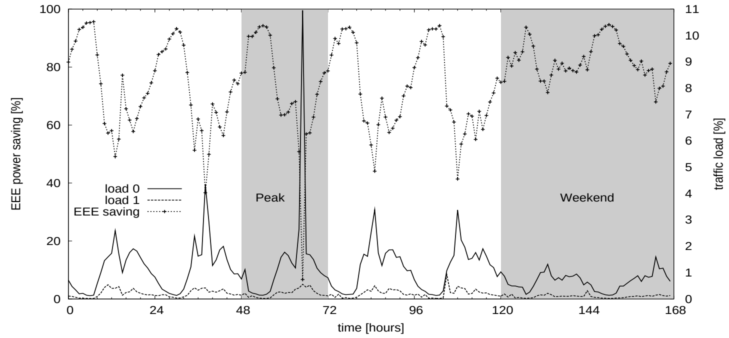
We add to the existing literature a unique evaluation of real traces with bidirectional flows over $1\ Gbps$ links. Furthermore, we are the first to evaluate the importance of precise timestamping on the trace-based performance evaluation of EEE.

II. Unintrusive Measurements at InterHost

Using the model in [2], we estimate the EEE power saving by means of the average and standard deviation of the packet interarrival time, the average packet size, and the offered load. The model uses queuing theory for M/G/1 queues, and models EEE state transitions as a renewal process. Noticeably, the model is robust to non-Poisson arrivals [2]. The analysis yields the average time that the link spends in the four different EEE

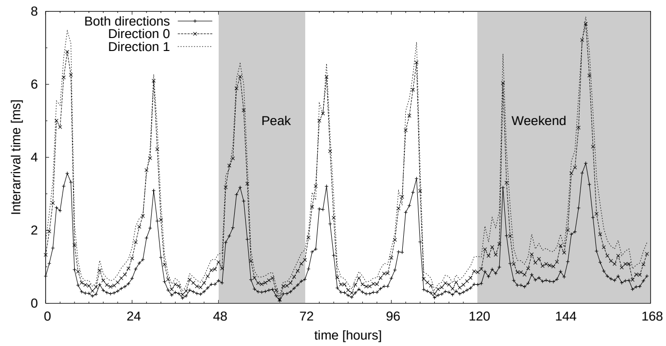


(a) Power saving during the month of February 2012.

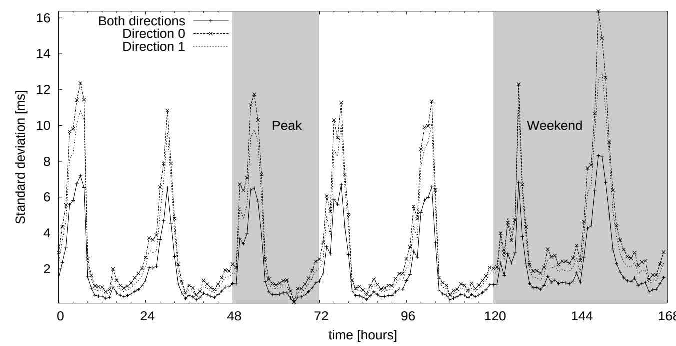


(b) Power saving during the second week of February 2012.

Fig. 2: Monthly and weekly plots for traffic and power saving (February 2012).



(a) Average interarrival time of packets (2^{nd} week of February 2012).



(b) Standard deviation of the interarrival times (2^{nd} week of February 2012).

Fig. 3: Weekly plots for the average and standard deviation of packet interarrival times (February 2012).

states: Active, Low Power Idle (LPI), Sleep, and Wake Up. The time that the link remains in each state, times the power consumed in the corresponding state, gives us the total power consumption of the EEE link. For further details on the model and on EEE please refer to [2] and [1] respectively.

We can also simulate EEE operations by using real traces for packet arrivals. With the *C++* simulator, we compute the exact EEE power saving, and compare this value to the estimate yielded by the model. The simulator uses trace files containing, for each packet, the arrival timestamp and the packet size.

We deployed a monitoring server to capture and store the traffic flowing through a link, and a *tap* to sniff and duplicate real packets without affecting the traffic, as shown in Fig. 1. We use a NetOptics passive device which is inserted in a 1000Base-TX Ethernet link and duplicates each and every signal over the link [9]. The NetOptics device, as shown in the figure, is also able to replicate uplink and downlink traffic over two separate cables connected to the monitoring server. The two monitoring ports of the tap are connected to a digital capture card. Specifically, a high accuracy two-port Endace DAG card [10] is mounted on the monitoring server. The DAG card is a capture device with dedicated CPU and memory, able to capture 100% of the traffic at up to $10\ Gbps$ over each port. Furthermore, the DAG card has a unique timestamping engine that guarantees clock synchronization to the nanosecond over the two monitoring ports. The DAG is activated once per hour to collect at most 100 bytes per packet for 200 seconds. We post-process the traces with the *tshark* packet analyzer and create simplified and anonymous trace files containing only arrival timestamps and packet sizes.

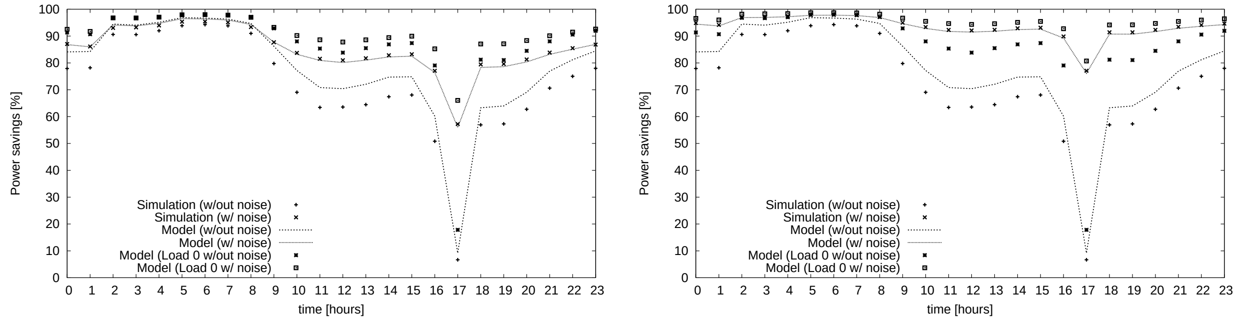
III. Trace-based Simulation and Analysis

We run EEE simulations based on the traces captured by the monitoring server. For the simulation, we use the same *C++* simulator used in [2]. We also extract the statistical parameters needed to run the EEE model in [2], and estimate the EEE power saving through the model as well. In particular, the model uses the average frame size, the average load, and the average and standard deviation of frame interarrival times, to compute the expected energy consumption of EEE links. In contrast, without EEE, the power consumption is constant.

We have captured traffic traces since November 2011, but we only show results for February 2012, which are representative of the rest of our measurements. We observed a weekly periodic behavior, and a daily typical maximum traffic of about 4% in the most loaded link direction, with the exceptional values of 9% and 11% on February 1st and 8th, respectively.

In Fig. 2a we plot the monthly load in each link direction (rightmost y-axis), labeled as “load 0” for one link direction and “load 1” for the other link direction. The figure also shows the power saving that might be achieved by means of EEE links, computed through simulation (leftmost y-axis). Our first observation is that we have a maximum traffic load of about 11% in the most loaded direction, whereas during weekends and overnight the peak load is below 2%. In the figure, traffic patterns are quite regular, showing higher traffic activity over weekdays, followed by lower traffic intensity over the weekend. It is also evident that overnight traffic is very low.

Fig. 2b illustrates results for the second week of February. We choose this week because it shows a traffic spike on February 8th. Daily spikes occurred at about 1 and 6 PM. This traffic distribution over time clearly depends on the nature of the websites hosted at InterHost premises, about which we have no information. However, the measured traffic patterns are qualitatively in line with other patterns reported in literature (e.g., see [11] and references therein). Processing the collected data with the EEE simulator reveals that overnight



(a) Simulation and model with noise in [-5, 5] *ms*.

(b) Simulation and model with noise in [-50, 50] *ms*.

Fig. 4: Evaluation of EEE power saving with model and simulation, with and without noisy timestamp measurements.

and during the weekend, EEE might save 70-90% of the power with respect to legacy gigabit Ethernet. Noticeably, the power saving exceeds 80% in more than 40% of our samples, and during weekdays the power saving is larger than 40% in 99% of our samples.

Figs. 3a and 3b give more information about the traffic arrival characteristics, namely the average interarrival time of the packets and the standard deviation of interarrival times, which are needed in order to run the model in [2] and analytically estimate the EEE power saving. In each of the two subfigures, two of the lines correspond to the traffic in each direction independently, whereas the third line corresponds to the overall link traffic. Both subfigures cover the same time interval covered by Fig. 2b. In Fig. 3a we observe that the packet interarrival pattern is similar for each day but the third day of the week, when we observe a traffic spike (see Fig. 2b), which corresponds to lower interarrival times. Similarly, the pattern observed for the standard deviation of the interarrival time in Fig. 3b is quite regular, and approaches zero in correspondence to the traffic spike observed during the third day of the week. Note that, in both Figs. 3a and 3b, we observe higher values in the weekend, which corresponds to lower traffic activity. As expected from the analysis in [2], Figs. 2b, 3a, and 3b confirm that the higher the average and standard deviation of the interarrival time, the higher the power saving. For instance, the valley in the power saving plot highlighted in Fig. 2b, corresponds to the minimum of interarrival time (Fig. 3a) and of its standard deviation (Fig. 3b). Similarly, peaks of power saving correspond to peaks of average and standard deviation of the interarrival time.

As shown in Fig. 2b, power saving in the Ethernet is often larger than 40%, with peaks exceeding 90%. Considering that a gigabit Network Interface Card (NIC) consumes about 2 W (e.g., see Intel 1G datasheets [12]), the estimated power saving for a single link with two NICs connected sums up to 1.1 to 2.6 $KWh/month$, these numbers being relevant for large data centers or web hosting companies running tens of thousands of links. For example, consider a typical large data center (more than 4650 m^2 nowadays [13]) with $\sim 40,000$ servers (like Google's data centers [14]). According to [15] each server has three network ports on average. Therefore, in total, the data center runs about 120,000 network ports, each of which will typically be very low loaded. Thereby, the power saving could be close to the peak most of the time. Considering the average price of electricity in Europe, i.e., 0.15 $Euros/KWh$, using EEE links might generate a potential economy of more than $280,000$ $Euros/year$. The case of 10 $Gbps$ (and 100 $Gbps$) links is even more promising, since they consume even more power, i.e., $4.5 - 20$ W according to [12].

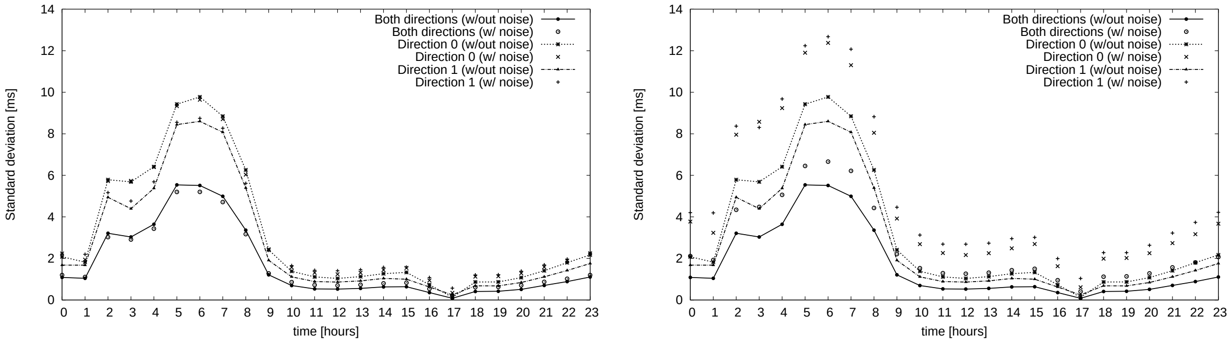
IV. Impact of Noisy Measurements

Here we use modified traffic traces to show the impact of noisy measurements on the quality of EEE power saving estimates. In particular, we picked the day with the maximum traffic recorded over the entire measurement campaign, which corresponds to February 8th, 2012, and perturbed the originally collected timestamps by adding zero-mean uniformly distributed noise. For each packet trace, we plot the EEE power saving based on simulation and model. However, since the model runs for unidirectional traffic only, we feed the model with either the traffic of each link direction separately, or with a single trace representing both traffic directions. Specifically, when the model is computed over the aggregate traffic measured over the two link directions, we merge the traffic traces obtained for the two directions. Consequently, the model cannot completely capture the correlation of traffic in the two directions. In fact, the IEEE 802.3az standard says that link interfaces can enter the LPI state only when *both link directions* are idle. Instead, we use a simplified model to cope with a unique link direction, resulting from the merging of the two real link directions.

Figs. 4a and 4b depict the EEE power saving computed via simulation and model. Input traces were used both with the original high precision timestamp and with modified timestamps. In Fig. 4a, original timestamps have been altered by adding uniformly distributed noise in the range $[-5, 5]$ ms. In Fig. 4b, noise ranges in $[-50, 50]$ ms.

Each plot contains the power saving estimates obtained with the model computed on the aggregate link traffic (merging of the two traffic directions, labeled as "Model" in the figures), with and without the artificially added noise. The figures also report the EEE power saving as estimated by the model when considering only the traffic in the most loaded link direction ("load 0" in the figures), with and without noise. Finally, the figures include the EEE power saving computed through simulations, with and without noise. Therefore, in each plot, the values obtained without noise are always the same, thus representing the benchmark for the experiment.

From the figures, we note that model and simulator report similar trends, with close but distinct power saving estimates. Differences are well explained considering that the model does not capture the bidirectional nature of the EEE power saving mechanism. Limited differences are due to the low



(a) Standard deviation of interarrival times with noise in [-5, 5] *ms*. (b) Standard deviation of interarrival times with noise in [-50, 50] *ms*.

Fig. 5: Standard deviation of packet interarrival times with and without noisy timestamp measurements.

load measured on both link directions. The model which only considers the most loaded link is obviously the one reporting the highest power saving. Similarly, the model considering the aggregate traffic yields the lower power saving, since it does not consider that packets belonging to opposite traffic directions might be served in parallel.

Let's now consider the impact of noise. In Figs. 4a and 4b, we can see that savings estimated with noisy measurements tend to be higher. In fact, adding noise to timestamps contributes to break the traffic correlation between the two directions, so that (i) simulator and model yield very similar results, and (ii) power saving occurs with roughly the product of probabilities of having each link direction idle. As a result, $\sim 80\%$ power saving can be estimated even for the peak hour.

We remark that noisy measurements cause erroneous power saving estimates, and conclude that timestamp errors larger than few *ms* are not tolerable to achieve accurate estimates. Considering that *ms* accuracy in time-stamping is barely achievable with ordinary operating systems and driver-operated time-stamping, which depends on system interrupts, we also conclude that dedicated traffic measurement tools are needed, as the ones that we have used for our measurements.

Note that timestamping noise does not affect load, interarrival average and packet size. It changes only the standard deviation of the interarrival time. Therefore, we show in Figs. 5a and 5b the standard deviation of interarrival times with and without noise, for the same time interval used in Figs. 4a and 4b. Fig. 5a illustrates the standard deviation with noise uniformly distributed in [-5, 5] ms, and Fig. 5b refers to uniform noise in [-50, 50] ms. The figures report separately the interarrival statistics for each traffic direction, plus the statistics for the overall arrival process (i.e., considering interarrivals between packets accessing the link, independently from their flow direction). We observe that the presence of noise in the timestamps can induce to deal with packets separated by a short interval as if they were sent back-to-back (note that we do not allow timestamp noise to induce packet sequence reordering), this error reducing the standard deviation of interarrival times at particular measurement epochs. Most importantly, we note that small errors in the estimate of the standard deviation of the interarrival time cause large errors in the estimate of the EEE power saving.

V. CONCLUSIONS

In this letter we summarized the results of our traffic measurements that have been collected at InterHost's Web Hosting premises in Madrid. The goal of our work was to collect and analyze *unique* data on bidirectional traffic in a real network, and estimate the potential savings that can be achieved by adopting the recently released IEEE Standard 802.3az instead of the legacy Ethernet. Overall, we observed traffic patterns yielding the possibility to save at least 40%, and up to more than 90%, in each gigabit link. Such a power saving would represent a non-negligible operational cost reduction for a data center, in the order of several hundreds of thousands of Euros per year. Moreover, we analyzed the importance of high precision traffic measurements on the power saving estimation. Our analysis unveils that precise timestamping is needed to use analytical models, which allows EEE power saving estimation with no need for time consuming simulations.

REFERENCES

[1] IEEE Standard 802.3az, "Energy Efficient Ethernet," 2010.

[2] M. Ajmone Marsan, A. Fernandez Anta, V. Mancuso, B. Rengarajan, P. Reviriego, and G. Rizzo, "A Simple Analytical Model for Energy Efficient Ethernet," *IEEE Comm. Letters*, no. 99, pp. 1–3, June 2011.

[3] P. Reviriego, K. Christensen, J. Rabanillo, and J. A. Maestro, "Initial Evaluation of Energy Efficient Ethernet," *IEEE Communication Letters*, vol. 15, no. 5, pp. 578–580, May 2011.

[4] P. Reviriego, J.A. Hernandez, D. Larrabeiti, and J.A. Maestro, "Performance Evaluation of Energy Efficient Ethernet," *IEEE Communication Letters*, vol. 13, no. 9, pp. 697–699, Sept. 2009.

[5] P. Reviriego, J.A. Maestro, J.A. Hernandez, and D. Larrabeiti, "Burst Transmission for Energy Efficient Ethernet," *IEEE Computer Society*, July 2010.

[6] K. Christensen, 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, no. 11, pp. 50–56, Nov. 2010.

[7] P. Reviriego, V. Sivaraman, Z. Zhao, J.A. Maestro, A. Vishwanath, A. Sanchez-Macian, and C. Russell, "An Energy Consumption Model for Energy Efficient Ethernet Switches," *OPTIM 2012 (Workshop)*, Madrid, Spain, July 2012.

[8] S. Herreria-Alonso, M. Rodriguez-Perez, M. Fernandez-Veiga, and C. Lopez-Garcia, "Optimal Configuration of Energy Efficient Ethernet," *Computer Networks*, Mar. 2012.

[9] NetOptics website: http://www.netoptics.com/products/network-taps.

[10] Endace website: http://www.endace.com.

[11] F. Malandrino, M. Kurant, A. Markopoulou, C. Westphal, and U.C. Kozat, "Proactive Seeding for Information Cascades in Cellular Networks," in *Proceedings of IEEE INFOCOM'12*, Mar. 2012.

[12] R. Sohan, A. Rice, A.W. Moore, and K. Mansley, "Characterizing 10 $Gbps$ Network INterface Energy Consumption," in *35th IEEE Conference on Local Computer Networks*, Oct. 2010.

[13] S. Bapat, "The Future of Data Centers (... and the Stuff That Goes In Them)," http://www.e3s-center.org/events/09/e3s-symposium.htm, June 2009.

[14] Google, "Google Data Center Video Tour," http://www.google.com/about/datacenters/events/2009-summit.html#tab0=4, Apr. 2009.

[15] U.S. Environmental Protection Agency, Energy Star Program, "Report to Congress on Server and Data Center Energy Efficiency, " *Public Law 109-431*, Aug. 2007.