

Performance Evaluation of Energy Efficient Policies for Ethernet Switches

Beatrice Paillassa, Riadh Dhaou, Gentian Jakllari, Thaenchaikun Chakadkit
IRIT-ENSEEIH, University of Toulouse, France
{beatrice.paillassa,riadh.dhaou@enseeiht.fr, jakllari,chakadkit.thaenchaikun}@enseeiht.fr

Abstract—Energy efficiency has emerged as a defining scientific and engineering challenge of our time. Ethernet, as the dominant wireline technology, has been a focus of intense research and development efforts with the goal of significantly decreasing its energy consumption. In 2010 the IEEE Std 802.3az, which uses a Low Power Idle (LPI) mode to reduce the energy consumption of a link when there is no data traffic, was approved. Recently, the major manufacturers have brought to market the first switches that implement the new energy efficient standard.

In this paper, we make use of the first generation of hardware that support the IEEE 802.3az to get a better understanding of the behavior of energy efficient mechanisms for ethernet. Based on measurement on DLINK switches that are IEEE 802.3az compliant, we build a power model that reflects more accurately the power used by real hardware in practice. We use the measurement-driven power model to analyze the behavior of *Lazy Start*, the state of the art algorithm for energy efficient ethernet, in the ns-3 simulator. Based on our analysis, we provide recommendations for improvements and potential directions for future work.

Index Terms—Energy Efficient Ethernet, IEEE 802.3az, Green Networking

I. INTRODUCTION

Energy efficiency has emerged as a fundamental engineering and scientific challenge for the 21st century. As is the case for most modern industries, the energy consumption of the Information and Communication Technology (ICT) industry has been steadily increasing in the last decade. With Ethernet being the dominant wireline communications technology for LANs, efforts have been made by the research community as well as the industry to find ways to reduce its energy consumption [2], [3]. In [2], Gupta and Singh were among the first to propose an *On/OFF* algorithm for reducing the energy consumption of Ethernet. The basic idea is to turn off the transceivers, which according to the authors can consume up to 20% of the total switch power budget, when there is no data traffic. This approach is generally referred to as Low Power Idle (LPI). The challenge, however, is deciding when to turn off the transceivers. To this end, Gupta and Singh propose a threshold based approach. As long as the queue size is estimated to be under a threshold the transceivers are allowed to go to sleep, otherwise they are turned back on. In [3], the authors identify some weaknesses in the Gupta-Singh algorithm and propose a new algorithm to address them. The new algorithm, referred to as *Lazy Start*, introduces two

key changes. First, the transceivers go to sleep only when the queue is empty, as opposed to when the queue size is under a threshold. Second, the devices go back on only when, either the queue has exceeded a threshold or, a maximum allowed amount of time has elapsed. The latter is necessary to bound the delay when the traffic is too low to fill the queue fast enough. Through ns-2 simulations, the authors showed that *Lazy Start* offered superior performance in terms of energy savings compared to the Gupt-Singh algorithm.

In parallel with the research community and the industry, the IEEE 802.3 Working Group created in 2006 a new task force, IEEE 802.3az, in charge of standardizing the methods for improving the energy consumption of the Ethernet devices. The resulting standard, namely IEEE Std 802.3az-2010¹, was ratified by IEEE in September 2010 [4], [5]. The IEEE 802.3az standard adopts the LPI mode of operation to save energy, however, it does not define a specific policy for deciding when to go to sleep and when to stay active.

While a lot of progress has been made in terms of algorithms and protocols for energy efficient ethernet, there are still unexplored areas that deserve further study. We identify two important areas we aim to address in the current work. First, a lot of the schemes proposed, including the *Lazy Start* [3], which is considered the state of the art, were validated through simulations before the first commercially available IEEE 802.3az devices were available. Note that, the standard was only ratified in September of 2010 and the major manufacturers, like CISCO and DLINK, have only recently made available switches that are IEEE 802.3az compliant. As such, the validations so far are generally based on power models that were not based on the behavior of real IEEE 802.3az devices. Instead, the studies so far have used values recommended by the manufacturers or the IEEE standard [3]. However, we believe that a power model based on real life IEEE implementations would be a better estimation of what savings one is expected to get in practice. Second, because switches implementing IEEE have only recently been made available, their performance is still poorly documented.

In this paper, we address both these areas. We perform measurements with IEEE 802.3az compliant switches, specifically the DLINK DGS 1100-24. Our experimental results serve a dual purpose. First, based on the measurements, we make

¹The IEEE 802.3az is also known as EEE, for Energy Efficient Ethernet. We use the names IEEE 802.3az and EEE interchangeably for the rest of the paper.

new recommendations on how to implement the spanning tree protocol in EEE compliant switches, which are counter to the recommendations based on non-EEE switches [6]. Second, we develop a power model based on measured power values to be used by performance evaluation studies. To showcase the value of our measurement-based power model, we use it to analyze the performance of the *Lazy Start* in ns-3. In our simulations we find that a small change in the power values can have a dramatic impact on the energy efficiency of *Lazy Start*. Therefore, the accuracy of the power model is paramount. In addition, in our simulations we find that one of the important parameters to affect the energy consumption is the packet size. Thus, for small packets aggregating packets at the transport level may be particularly useful. Finally, we observe that the deployment of EEE switches has an impact in terms of delay. Therefore, care must be taken to avoid disruption delay sensitive applications, such as VOIP, in the name of energy efficiency.

The rest of the paper is organized as follows. In Section II we present our experimental analysis for identifying the correct power model. In Section III, we present our simulation study based on the power model identified in Section II. We conclude our study and discuss directions for future work in Section IV.

II. EXPERIMENTATIONS ON POWER MODEL

The objective of the work is to establish the power consumption model for the latest enterprise switches that implement the IEEE 802.3az, the IEEE standard for energy efficient ethernet. Previous studies in the green networking area have focused on measurement of non-IEEE 802.3az switches and routers, and of mechanisms capable to improve their energy efficiency [6], [7]. We consider the DLink switch (DGS 1100-24), which implements IEEE 802.3az and supports data rates varying from 10Mbps to 1Gbps. To measure the power levels we use the HM8115-2 power analyzer [8]. Our measures focus on the power states of the switches. There are three states a switch can be in and, therefore, three states for which we need power values; namely, the idle state, the transmission state, and the sleeping state.

A. Idle State

We start by studying the idle state by performing measurements on a switch with no data traffic. Note that, the IEEE 802.3az is only implemented on Gbps ports. Compared to previous studies about non-IEEE 802.3az switches [6], energy values presented in Figure 1 are minimized (around 50W in [6] and 5W here) but the global interpretation is similar. Energy is consumed not only for transmission but also during idle periods. As stated by previous studies, the power consumption increases with the number of active ports. By itself, the switch consumes around 5W (value with 0 active ports). We can deduce a maximum consumption per port by approximating it with the maximum measured value. Therefore the power per port values in idle state are: 0.05W at 10 Mbps, 0.15 for 100 Mbps and 0.3 at 1Gbps.

B. Sleeping State

Switch manufacturers do not at this point provide the capability to users to set the switches to sleep manually and that is the case for our DLINK DGS 1100-24. Therefore, measuring the power level for the sleep state directly is not practically possible. Instead, we resort to the following procedure to get a good estimate of the sleep state power level and compare it to that of idle state. First, we connect two DLINK DGS 1100-2 switches to each other with a single cable and leave them idle. As there is no data traffic, the respective ports on both the switches will be mostly in sleep as per IEEE 802.3az. Therefore, the measured power levels under this scenario are a very good estimate of the power level used at sleep state. We get a value for the global power gain equal to 5%, and conclude that: the power per port in sleeping state for 1 Gbps is 0.03 Once this experiment is done, we connect a DLINK DGS 1100-2 to a PC that uses a legacy, i.e. non-IEEE 802.3az ethernet card. Doing so, will lead to the switch going to backward compatibility mode and disable the IEEE 802.3az. Therefore, the power measured under this scenario is representative of the idle state since there is no data traffic and the switch does not go to sleep mode. We observe that the power level for sleep state represents 10% of that for the idle state. This value is in conform with the IEEE 802.3az objective of having the power level of sleep state be 10% of that of the idle state.

Something worth stressing is that this case is ideal for IEEE 802.3az: since there is no traffic, the sleeping time is maximized. But, in the case of data transmissions, the power consumption at 1Gbps would be higher than that at 10 Mbps (depending upon the manufacturer). Therefore, rather than deactivating all the unused ports by management as proposed in [6], it seems preferable to make a difference between the real idle ports and those redundant. These ports are useful to increase the network robustness. They are generally chosen by a spanning tree protocol in order to avoid the topology loops. We propose to negotiate for them, when they are not included in the active topology. The speed leading to the lower consumption between the connected switches should be chosen, rather than the higher one as it is automatically done by the auto-negotiation function currently in place. By doing so, spanning tree protocols could advantageously influence the bit rate negotiation.

C. Active Mode

To put forward the LPI process, we connect two DLINK DGS 1100 switches together and then to two PCs respectively. As mentioned earlier, the PCs are equipped with non-IEEE 802.3az cards. Therefore, in this setting, the two links from the switches to the respective PCs do not use the IEEE 802.3az, while the connection between the two switches does. In Figure 2, we varied the inter arrival time of the UDP packet generation. When this time is superior to T_s (transition time from active to sleeping states) the LPI mechanism can be activated and sleeping states occur. When it is superior to the threshold specified in [5] for a given data bit rate and

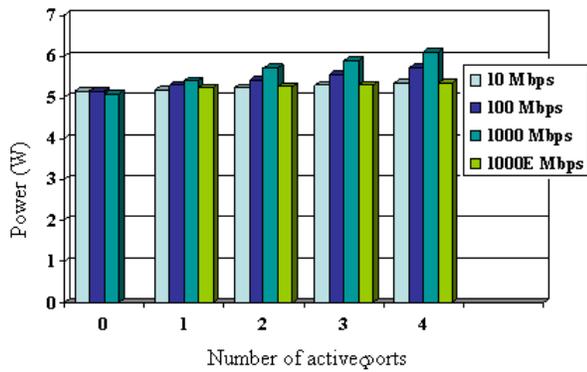


Fig. 1. The idle state power consumption.

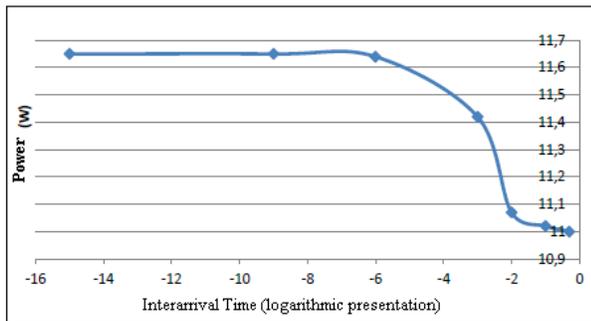


Fig. 2. Power function of time between packet arrivals.

coding (1000Mbps, $T_s=0.202$) the consumed power tends to be constant ($P=11.65$). We approximate the power formulation by first the 2 switch powers for two active ports (1 with IEEE 802.3az and 1 without: $10.22 + 0.3 \cdot 2 + 0.03$) plus, secondly the traffic management consumption (i.e. transmission, memorization) onto 3 ports. So, the traffic cost without IEEE 802.3az per port (inter arrival time less than 10-6 s. on Figure 2) is maximized as : power per port in transmit state for 1 Gbps is 0.03

In [9] a model of traffic consumption is proposed, without considering the LPI details and so the parameters and their influences. To get more understanding about the IEEE 802.3az switches we need to vary the parameters that are related to some queuing management that are vendor specific. From simulation studies, the right parameters could be defined (i.e. those optimizing objective as power, delay ..). Next section presents such results. Even if each manufacturer would result in different power consumed values, we assert that the global tendency would be the same.

III. SIMULATIONS

A. The Impact of EEE policies parameters

We use ns-3 tool to show the impact of the policies parameters. We simulate the buffering management policy proposed by Herreira and al. in [3] based on the parameters that we obtained in the previous section. The chosen management

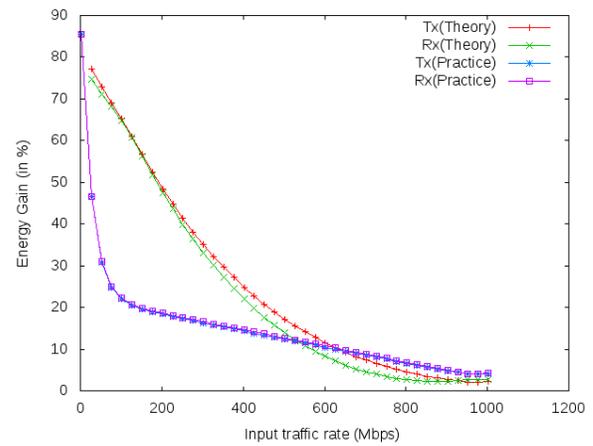


Fig. 3. Energy Gain for UDP traffic (in the direction S to D).

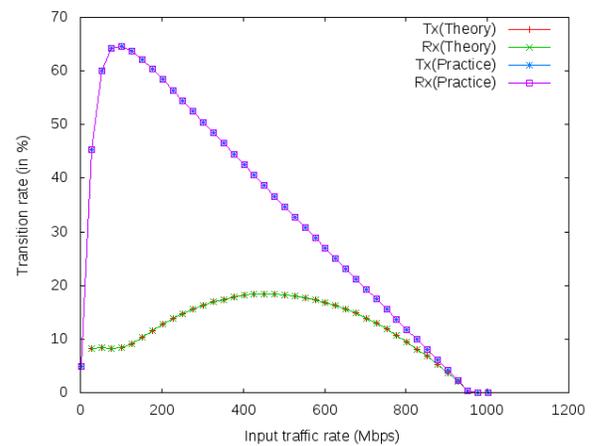


Fig. 4. Transition rate (in the direction S to D).

policy improves the Gupta Singh algorithm by introducing an empty queue sleep condition, and also a timer condition that is function of the power consumption. There is a minimum profitable sleeping interval that has to be greater than the transition time, and has to generate a power consumption smaller than the one that would result if the port was in the active mode without any transmission.

The theoretical results are based on the parameters values that were obtained from [2]. In the following, we refer to the simulation results based on the parameters values derived from our experiments.

We assume that the SLEEP mode consume 10% of the consumption of the READY mode, i.e. the active mode without any transmission, (in theory respectively 0.1 and 1W [3] and in practice respectively 0.03 and 0.3). The same quantity of energy is consumed during the states transmission (TX BUSY) and during the transitions between active and sleep modes (TX UP and TX DOWN). The corresponding power is (in theory, 2W for transmission and 1.3W for the reception, and in practice 0.3W for both values, from our measures). A simple

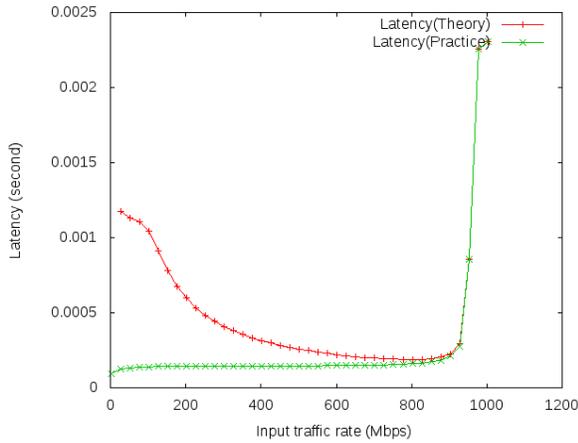


Fig. 5. Delay with EEE scheme (in the direction S to D).

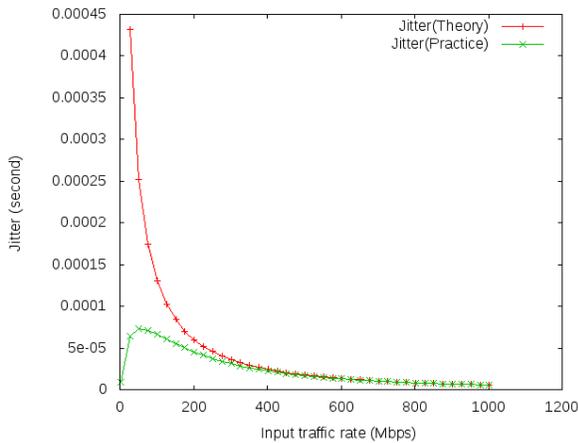


Fig. 6. Jitter with EEE scheme (in the direction S to D).

Ethernet EEE link connecting a source (S) and a destination (D) is studied. We measure the energy consumption due to the transmission and the reception of the Ethernet frames for both directions (S to D and D to S). Firstly, we measure the energy gain as a function of the input traffic load. Then we evaluate the number of transitions between states. Finally, we study the impact of the parameters of the EEE schemes on the energy gain and on the QoS metrics such as delay and jitter. The impact of these parameters is analyzed for different kind of traffics (we present here results for UDP and for TCP traffic flows).

B. Energy saving obtained with EEE schemes for UDP Traffic

1) *Impact of the input data rate:* We trace on the Figure 3, for a simple Poisson UDP traffic flow the obtained energy gain by using of EEE schemes. The capacity of the transmission queue is limited to 300 packets (1000 bytes for each packet).

Initially the transmitter is in the LPI (Low Power Idle) mode. The transition from the LPI to the active mode is triggered either when the queue size reaches a threshold value

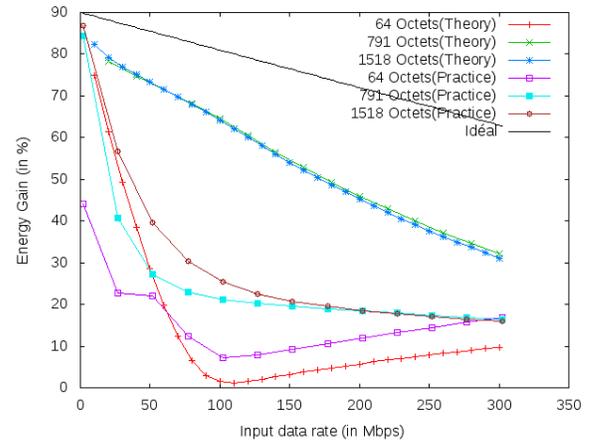


Fig. 7. Impact of the packet size on the energy gain.

(10% of the capacity) or at the expiration of the TX Sleep timer (this timer is triggered upon the arrival of the first packet and expires 2,75ms later). The theoretical curves obtained by simulation are not linear (as it is expected theoretically) but they are conform to the results obtained in previous works [4]. The results based on practice values are worst but are linear starting from 50Mbps/s. The gap between the simulation results and the expectations is due to the number of transitions between states. The same quantity of energy is assumed to be consumed during the transition and during the active state. The maximal number of transitions is obtained when the data rate is around 450Mbps as shown in the Figure 4. In practice, the transition rate is higher (up to 65%) for low data rates (about 50Mbps/s). The input rate impacts also the packets delay and jitter (see Figures 5 and 6). We notice a significant increase of packet jitter, mainly for low data rates. The average value of the obtained jitter with EEE is 0.5ms while it is less than 0.1 ms without the EEE schemes. When the link is highly loaded the LPI scheme is less often used and the delay and jitter decrease. Practical results for delay and jitter confirm these trends. Finally, we notice that the packet size has an important impact on the energy saving. The larger the packets are the better the energy gains are obtained. The Figure 7 shows that small packets (64 bytes) gives bad results in terms of energy consumption compared to large ones.

2) *Impact of the queue size.:* To show the impact of the queue size, we keep the same parameters values than previously while fixing the data rate to 80 Mbits/s. The increase of the queue capacity (while keeping the same threshold value to 10%) impacts only the packets loss. The increase of the capacity reduces the energy consumption, as the number of transitions between states will be reduced. As shown on the Figure 8, the increase of the queue capacity, above 225 packets, does not enhance the energy consumption performance. This value drops to 50 in practice.

The increase of duration in the idle state and the decrease of the number of transitions between states are the main factors

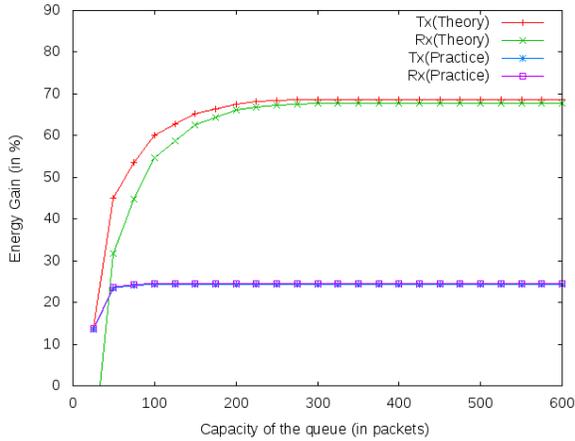


Fig. 8. Energy gain (in the direction S to D).

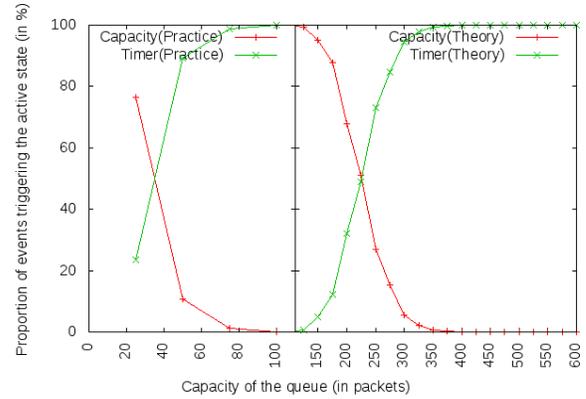


Fig. 9. Proportion of events (in the direction S to D).

to enhance the energy efficiency. The Figure 9 shows the proportion if each events triggering the active mode according to the input data rate. We can notice that above a capacity of 225 packets the active mode is triggered by the expiration of the TX sleep timer. We obtain similar results when varying the value of this timer while keeping the queue capacity constant.

3) *Energy saving obtained with EEE schemes for TCP Traffic:* The same experiment is performed with transmission of TCP packets (536 bytes length). Two main results have to be highlighted. The first result concerns the value the TX Timer timer that has no impact on the energy consumption. The second result concerns the impact of the capacity of the queue in both directions (TCP data and ACK). As shown on the Figure 10, when the capacity of the queue is less than 120 packets there is congestion and some packets are lost.

The congestion control mechanism of TCP contributes to the decrease of the energy gain. Above a capacity of 120 packets the curves have are conform to the trend obtained for UDP traffic. The obtained gain on the return path is, independent of the capacity, and is around 80%.

IV. LESSONS AND PERSPECTIVES

The experimental study confirms the results obtained previously. However, as the thresholds largely depend on the power values, a slight change in the values may have a dramatic impact on the energy efficiency. However, very few products currently implement EEE functions and the lack of flexibility in the use of those implemented functions and the possible settings of these led us to evaluate the performance of the EEE policies via simulation. The analysis of the results shows an opportunity to adapt the parameters of the policy as a function of the load. The maximal energy consumption gain is obtained for lightly loaded systems. One of the important parameters in the energy consumption is the packet size. Aggregating packets at the transport level may be particularly useful for small packets.

However, we observed that the deployment of EEE switches has an impact in terms of delays. The impact of these delays,

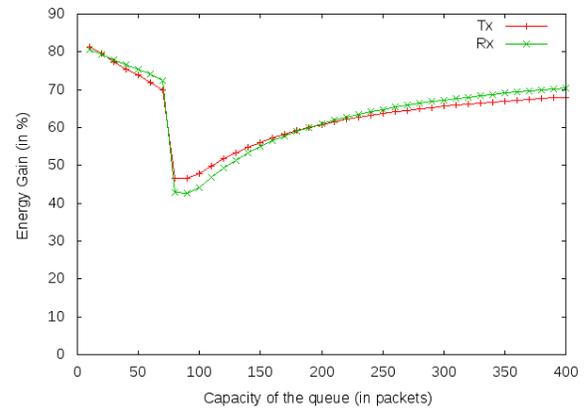


Fig. 10. Energy Gain for TCP traffic (in the direction S to D).

using buffering mechanisms, is important to analyze. This is especially important for time constrained traffics (such as VoIP, for example). For this type of traffic it would be interesting to compare the use of EEE mechanisms at the switch level with an aggregation of packets at the source level.

The second observation we have outlined in this paper is on the impact of congestion control mechanisms on the gains in terms of energy consumption. These schemes contribute to the mitigation of energy gains. The parameters of the EEE policies must adapt according to QoS metrics (such as as packet loss or delivery delay) that can be measured locally. Looking ahead, we can deepen the analysis to find effective ways to correlate the adaptation parameters to QoS metrics. Finally, it would also be interesting to see if the adaptation must be considered for each flow or an for an aggregate flow.

ACKNOWLEDGEMENTS

We thank Maxime Lecanu and Ammari Abdelkarim for their help in generating the data presented in this paper.

REFERENCES

- [1] M. Gupta and S. Singh, "Greening of the internet," in *Proceedings of the 2003 conference on Applications, technologies, architectures, and protocols for computer communications*, ser. SIGCOMM '03. New York, NY, USA: ACM, 2003, pp. 19–26.
- [2] —, "Using Low-Power Modes for Energy Conservation in Ethernet LANs," in *INFOCOM 2007. 26th IEEE International Conference on Computer Communications. IEEE*, may 2007, pp. 2451–2455.
- [3] S. Herrería-Alonso, M. Rodríguez-Pérez, M. Fernández-Veiga, and C. López-García, "Opportunistic power saving algorithms for ethernet devices," *Comput. Netw.*, vol. 55, no. 9, pp. 2051–2064, Jun. 2011. [Online]. Available: <http://dx.doi.org/10.1016/j.comnet.2011.02.012>
- [4] K. Christensen, P. Reviriego, B. Nordman, M. Bennett, M. Mostowfi, and J. Maestro, "IEEE 802.3az: the Road to Energy Efficient Ethernet," *Communications Magazine, IEEE*, vol. 48, no. 11, pp. 50–56, november 2010.
- [5] "IEEE Standard for Information Technology–Telecommunications and information exchange between systems–Local and metropolitan area networks–Specific requirements Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications Amendment 5: Media Access Control Parameters, Physical Layers, and Management Parameters for Energy-Efficient Ethernet," *IEEE Std 802.3az-2010 (Amendment to IEEE Std 802.3-2008)*, pp. 1–302, 27 2010.
- [6] P. Mahadevan, S. Banerjee, and P. Sharma, "Energy proportionality of an enterprise network," in *Proceedings of the first ACM SIGCOMM workshop on Green networking*, ser. Green Networking '10. New York, NY, USA: ACM, 2010, pp. 53–60. [Online]. Available: <http://doi.acm.org/10.1145/1851290.1851302>
- [7] J. Chabarek, J. Sommers, P. Barford, C. Estan, D. Tsang, and S. Wright, "Power awareness in network design and routing," in *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, april 2008, pp. 457–465.
- [8] in <http://www.hameg.com/0.147.0.html>.
- [9] P. Reviriego, V. Sivaraman, Z. Zhao, J. Maestro, A. Vishwanath, A. Sanchez-Macian, and C. Russell, "An energy consumption model for energy efficient ethernet switches," in *High Performance Computing and Simulation (HPCS), 2012 International Conference on*, july 2012, pp. 98–104.
- [10] M. Gupta and S. Singh, "Dynamic ethernet link shutdown for energy conservation on ethernet links," in *Communications, 2007. ICC '07. IEEE International Conference on*, june 2007, pp. 6156–6161.