Implications of energy efficient Ethernet for hubs and switches

P. Reviriego* and J.A. Maestro

Universidad Antonio de Nebrija,
C/Pirineos, 55 E–28040 Madrid, Spain
Fax: +34-914521110
E-mail: previrie@nebrija.es
E-mail: jmaestro@nebrija.es
*Corresponding author

C.J. Bleakley

University College Dublin,
Belfield, Dublin 4, Ireland
Fax: +353-12695396
E-mail: chris.bleakley@ucd.ie

Abstract: The efficient use of energy in communications is an area of growing interest. Until recently energy efficiency received little attention in most wireline communications standards and implementations. In many cases, the transmitter and receiver operate at full power, even when no data is being sent. This is the case in most wireline Ethernet standards that results in a considerable waste of energy. Efforts are now underway to develop new standards, such as energy efficient Ethernet, with the aim of reducing energy consumption. The changes introduced by energy efficient Ethernet have different implications for each network element. The implications for hubs are different to those for switches. These implications are analysed in this paper. It is shown that the adoption of the new standard will make hubs less energy efficient than switches. The implications studied in this paper illustrate the potential impact of energy efficient Ethernet on Ethernet networks.

Keywords: Ethernet; energy efficiency; hubs; switches; LANs; energy efficient Ethernet.

Reference to this paper should be made as follows: Reviriego, P., Maestro, J.A. and Bleakley, C.J. (2011) 'Implications of energy efficient Ethernet for hubs and switches', *Int. J. Communication Networks and Distributed Systems*, Vol. 6, No. 1, pp.3–11.

Biographical notes: Pedro Reviriego received his MSc and PhD (Honors) in Telecommunications Engineering from Technical University of Madrid, Madrid, Spain, in 1994 and 1997, respectively. From 1997 to 2000 he was an R&D Engineer at Teldat, Madrid, working on router implementation. In 2000, he joined Massana to work on the development of 1000BaseT transceivers. During 2003, he was a Visiting Professor at University Carlos III, Leganés, Madrid. From 2004 to 2007, he was Distinguished Member of Technical Staff with LSI Corporation working on the development of Ethernet transceivers. He is currently with Universidad Antonio de Nebrija, Madrid. His research interests are fault tolerant systems, performance evaluation of communication networks and the design of physical layer communication devices. He has authored numerous papers in international conferences and journals. He has also participated in the IEEE 802.3 standardisation for 10GBaseT.

Juan Antonio Maestro received his MSc in Physics and PhD in Computer Science from Universidad Complutense de Madrid, Madrid, Spain, in 1994 and 1999, respectively. He has served both as a Lecturer and Researcher at several universities, such as Universidad Complutense de Madrid, Universidad Nacional de Educación a Distancia (Open University), Madrid, Saint Louis University, Madrid, and Universidad Antonio de Nebrija, Madrid, where he currently manages the Computer Architecture and Technology Group. His current activities are oriented to the space field, with several projects on reliability and radiation protection, as well as collaborations with the European Space Agency. He is the author of numerous technical publications, both in journals and international conferences. Beside these, he has worked for several multinational companies, managing projects as a Project Management Professional, and organising support departments. His areas of interest include high level synthesis, signal processing, fault-tolerance and reliability.

Chris Bleakley received his BSc (Honours) in Computer Science from Queen's University, Belfast, Northern Ireland in 1990 and PhD in Electronic Engineering from Dublin City University, Dublin, Ireland, in 1995. In 1995, he joined Broadcom Eireann, a joint venture of Ericsson, Eircom and Trinity College Dublin, as a Senior Researcher. From 1997–2003, he worked for Massana, a research-lead fabless semiconductor company focused on integrated circuits for physical-layer communications. From 2001, his role was Vice-President of Engineering. He joined University College Dublin, Dublin, Ireland, as a Lecturer in 2003. His current research interests include fault tolerant design, embedded systems and digital signal processing for ubiquitous computing applications.

1 Introduction

The efficient use of energy in communications is of growing concern for the industry. A large number of communications devices in use today, combined with further expected growth means that communications systems could become increasingly large consumers of energy on a global scale. The energy consumption of the communications equipment in the core of internet alone is estimated to be over 6 TWh per year (Gupta and Singh, 2003).

Much of the problem has been due to the lack of focus on energy efficiency in the design of wireline communication systems. One example considered in this work is Ethernet. Currently, Ethernet devices consume energy when the link is established even if there is no data being transmitted (Gunaratne et al., 2008). It has been estimated that this design flaw leads to the waste of over 3 TWh per year (Gunaratne et al., 2008). The root cause of the problem is the lack of energy efficient criteria in the specifications of original Ethernet standards. These standards require receivers and transmitters to operate continuously even in the absence of data. This issue is now being addressed by the IEEE 802.3az task force (energy efficient Ethernet). The task force aims to introduce energy efficiency enhancements to the existing Ethernet standards and complete the new standard by the end of 2010 (Bennet et al., 2010).

When energy efficient Ethernet is widely adopted, the energy consumption of Ethernet physical layer devices will change substantially. Today, energy consumption is almost independent of whether data is being transmitted or not and depends only on the length of time for which the link is established. This will change to a situation in which energy consumption is dependent on the length of time for which data is actually being transmitted over the link. This change has wide implications for Ethernet network nodes and also for upper layer protocols. In this paper, the implications for hubs and switches are analysed.

The remainder of the paper is structured as follows. In Section 2, current work on the energy efficient standard is reviewed. A brief description of Ethernet hubs and switches is provided in Section 3. In Section 4, the implications of energy efficient Ethernet on the network are discussed. Finally, conclusions are presented in Section 5.

2 Energy efficient Ethernet

The energy consumption of a conventional wireline Ethernet link is currently roughly proportional to the length of time that the link is established. For example, a link to a PC would normally be established whenever the PC is on. The other factor that affects power consumption is the link speed – higher speed means greater power consumption. For example, the power consumption of a 1 Gbps link is in the range of 1–2 watts while for 10 Gbps power consumption is much greater.

The main idea behind energy efficient Ethernet is to put wireline Ethernet physical layer devices in a low power mode when no data is being transmitted. This technique has the potential to provide large energy savings since links are normally lightly loaded (Bennet et al., 2007).

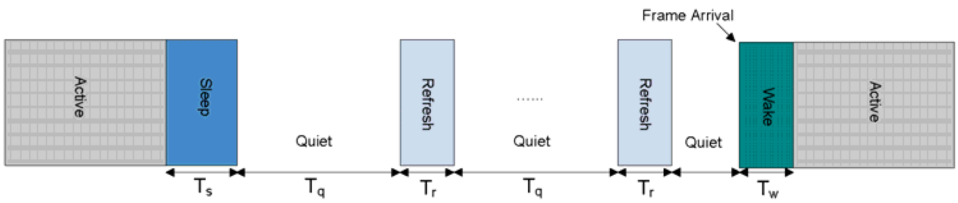
A number of alternative approaches can be used to implement the low power modes. The most obvious is to reduce the speed of the link when there is little traffic (Gunaratne et al., 2008). This can be achieved through auto-negotiation (Spurgeon, 2000) which is already part of the IEEE 802.3 Ethernet standards. Auto-negotiation is currently used when the link is first established in order to determine the link speed to be used. Normally, the highest link speed that the devices at both ends of the link support is selected. However, one of the ends could reduce the link speed by restarting the auto-negotiation process and advertising a lower speed. This would cause the link to operate at a lower speed. Unfortunately, the auto-negotiation process takes between a few hundred milliseconds and a few seconds to complete, which is excessive for many applications and would impact on the user experience. Consequently, alternative approaches, such as rapid PHY selection (RPS) (Christensen, 2007), have been proposed to accelerate link speed change. In RPS, frame exchange is used to agree a link speed change without the need for restarting the auto-negotiation process. Thus, speed changes can take place in a much shorter time.

The advantage of using speed change mechanisms for energy reduction is that they can be implemented with minimal changes to the existing standards. There are, however, two major drawbacks. The first is that the link has to be re-established at the new speed. Even for RPS, this takes a few hundred milliseconds. This time is needed to adjust all of the elements in the receiver, such as the equalisers, cancellers and timing circuits, to the channel conditions. These elements are used differently at different link speeds.

Therefore a change in speed requires adjustments in the receiver elements. During this time, the link is down and no traffic can be exchanged. This means that the total time it takes for a speed change is still on the order of hundreds of milliseconds. The second drawback is that, although reducing the link speed reduces energy consumption, transmitter and receiver still operate continuously, only at a lower speed.

An alternative approach is to introduce changes in the standards such that the physical layer devices support low power modes that can put a device to sleep and wake it up very quickly (in the order of micro-seconds) without a speed change. This is the option chosen by the IEEE 802.3az task force. They analysed mechanisms to support the use of low power modes at each Ethernet speed, for example 100 Mbps, 1 Gps and 10 Gbps. They propose that the receiver elements be frozen when the device enters a low power mode. When the receiver is woken, it is expected that only minor adjustments are needed since the channel is quite stable. These adjustments can be performed in a few microseconds, in contrast to the milliseconds that are needed to re-establish a link. To ensure that the receiver elements are aligned with the channel, periodic short periods of activity are scheduled to refresh the receiver state while in a low power mode. As an example, the proposed state transitions in the IEEE 802.3az draft (Bennett et al., 2009) are illustrated in Figure 1. The minimum and maximum allowed values for the timer parameters are specified in the standard draft for 100Base-TX, 1000 Base-T and 10 GBase-T. Wake up times of the order of a few microseconds are supported much less those required for speed change. The energy savings achieved by implementing the low power modes are substantial. For example, in Chou et al (2008), the savings for a 100 Base-TX device were shown to be over 75% of the total power when the PHY is in the low power mode compared to the active state.

Figure 1 Illustration of EEE operation (see online version for colours)

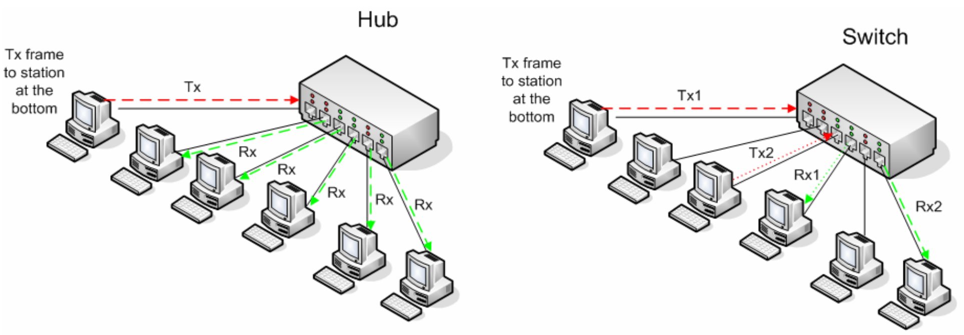


Once energy efficient Ethernet is adopted, the energy consumption of an Ethernet link will depend more on the length of time that the link is actually carrying data. The time that the link is established is then divided into the time with data transmission (link active) and the time without data transmission (link inactive). When the link is active, power consumption is similar to that of conventional Ethernet. When the link is inactive, power consumption is significantly reduced as the physical layer devices are in a low power mode. The exact time a link stays in low power mode will depend on the arrival times of the data frames and the algorithm used to enter the low power mode. In discussing the implications for hubs and switches, we will consider that the time spent in low power mode is roughly equal to the time when no data is being transmitted on the link.

3 Hubs and switches

Hubs and switches are traditional elements in Ethernet networks (Spurgeon, 2000). Both are used to interconnect nodes in a network. Examples of six port hub and six port switch topologies are shown in Figure 2. A hub broadcasts each frame received on all other ports, as shown in Figure 2. This creates a logical bus between all nodes connected to a hub. A switch applies more sophisticated processing to frames (Seifert and Edwards, 2008). Each frame is only transmitted on the port via which the frame destination can be reached, as shown in Figure 2. While a hub operates at the physical layer, a switch operates at the link layer. In a switch, after each frame is received, the destination address is used to select the output port for that frame. Normally, switches buffer frames such that a frame can be delayed until earlier frames are transmitted on an output port. In a hub, no buffering is performed. Frames arriving simultaneously cause collisions and the loss of both frames.

Figure 2 Illustration of hub and switch topologies (see online version for colours)



Historically, hubs have been widely used due to their low cost as hub devices are simple since no switching logic is needed. Switches have the advantage of providing more capacity for the same topology since simultaneous frames can be transmitted on different ports, as shown in Figure 2. In the past decade, the cost of digital electronics has reduced dramatically. Thus, the cost difference between hubs and switches has diminished leading to the wide adoption of switches (Spurgeon, 2000). In addition to cost reductions, the past decade has seen a substantial increase in the level of system integration leading to solutions that implement an eight or 16 port switch in a single device (Seifert and Edwards, 2008). Depending on application, there are many types of Ethernet switches, ranging from desktop to campus or enterprise scale. Desktop switches are used to connect users and normally have a low port count (up to 24 ports). These switches only support low speeds, in contrast campus or enterprise switches have a much larger port count (up to hundreds) and use higher speeds to aggregate traffic (Seifert and Edwards, 2008).

4 Implications of energy efficient Ethernet

Before the adoption of energy efficient Ethernet, the energy consumption of physical layer devices in a hub and switch was similar and proportional to the time for which their links were established. In a switch, additional power is needed to implement the switching logic. Therefore, all other factors equal, one would expect hubs to be more energy efficient than switches. This advantage of hubs will disappear with the introduction of energy efficient Ethernet, as we will now explain.

In energy efficient Ethernet the situation changes because a hub sends data on all ports. Therefore the links will be active for longer and so the hub will consume more energy than an equivalent switch. For example, an eight port hub will transmit an incoming frame on seven ports whereas a switch would transmit the frame on only one port. This will result in significantly higher energy consumption for the hub. In fact, the situation for a hub is worse since there is no easy way to put the ports in low power mode. Since no buffering is performed in a hub, if a port is in a low power mode then any frames sent to the port will be lost, as illustrated in Figure 3. Therefore it only makes sense for a hub to enter low power mode when the remote ends of all of the ports request low power mode, as is shown in Figure 4. That is the only case for which it is guaranteed that no data will be received destined for a port in low power mode and therefore be lost. As soon as one port receives a request to wake up from the remote end, then all the other links must be woken up so as to avoid potential data loss. This means that, as long as one port is active, all ports must be active. This will substantially reduce the savings achieved by implementing energy efficient Ethernet. Another consequence of this limitation is that, unless all of the devices connected to a hub implement energy efficient Ethernet, the hub will not be able to set any of the ports to low power mode since the legacy device could transmit data at anytime leading to potential data loss. Since the adoption of existing Ethernet standards has taken many years (Bavel et al., 2005) it is likely to take some time before energy efficient Ethernet hubs can benefit from the new standard. This is not the case in switches as each port can be put in a low power mode independently and traffic is buffered. Finally, if hubs are interconnected then the limitation will apply to all ports on all of the interconnected hubs. That is, unless the remote devices on all ports on the interconnected hubs request low power mode, the hubs cannot put any of their ports in low power modes without risking data loss.

Figure 3 Illustration of data loss in a hub when a port is in low power mode (see online version for colours)

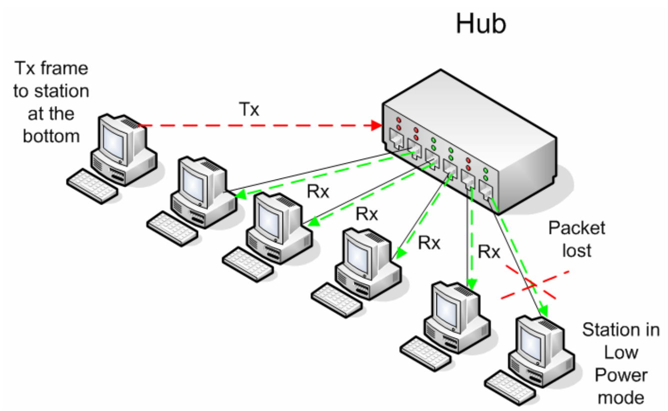
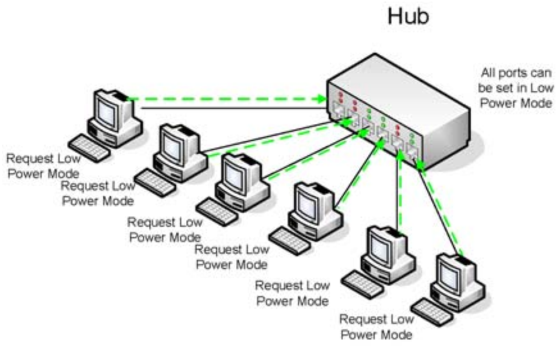
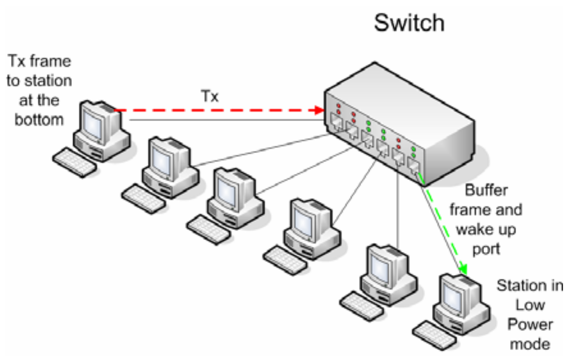


Figure 4 Illustration of low power mode in a hub when low power mode is requested on all ports (see online version for colours)



For switches, the situation is different. Firstly, data is sent only on the port through which the destination is reached and not on all ports. Secondly, each port can enter a low power mode independently of the other ports with no risk of data loss. This is because the switch will identify the output port and, if it is in low power mode, will buffer the frame until the port is woken up, as shown in Figure 5.

Figure 5 Illustration of data for a port that is in low power mode in a switch (see online version for colours)



To illustrate the differences in energy consumption, a case study follows. Let us assume that a 16 port energy efficient Ethernet hub or switch is used to connect PCs and that all ports are used. The average on time for the PCs is assumed to be eight hours with a 2% traffic load. This is a typical scenario for an office with activity during working hours only. Herein we consider 100 Base-TX connections since detailed power consumption figures are available (Chou et al., 2008). The first configuration that is studied is the use of a 16 port hub and we further assume that at least one of the PCs does not implement energy efficient Ethernet. This means that all ports will be active for the eight hours that the PCs are on. Therefore daily energy consumption will be 16 ports times eight hours times the per port power consumption which is taken to be 250 mW as in (Chou et al., 2008). The second configuration consists of a hub but we assume that all connected PCs implement energy efficient Ethernet. Assuming for simplicity that the activity on all ports is independent then all ports will be idle roughly 72% of the time. This is obtained from the probability of all ports being inactive at a given point in time calculated as $(1\text{-}0.02)^{16}$. That means that the physical layer devices would be in low power mode 72% of the time. Thus, the inactive energy consumption is sixteen ports times 72% of eight hours times the per port power consumption in low power mode which is taken to be 60 mW as in Chou et al. (2008). The consumption in active mode has to be added which is 16 ports by 28% of eight hours times the per port conventional power consumption of 250 mW. In the final configuration considered a switch is used. A similar calculation can be performed but, in this case, the active time is only 2% of eight hours. The results are summarised in Table 1 and show that the difference in energy consumption between hubs and switches is substantial. In this case study, depending on whether all of the remote devices implement energy efficient Ethernet or not, a hub will need almost two to four times the energy of a switch. Although these results are just an example, they serve to illustrate the benefits of using switches instead of hubs when energy efficient Ethernet is implemented.

Table 1 Energy consumption for the case study

Configuration	*Daily energy consumption*
Hub with legacy PCs	32 Wh
Hub with energy efficient PCs	14.4 Wh
Switch	8.2 Wh

Finally, it is worth noting that these figures are for the energy used in the physical layer devices only. In the case of switches, additional energy will be used in the switching logic. However the energy consumed in the switching logic will be also reduced when energy efficient Ethernet is implemented, as low power modes can also be used for the switching logic (Dove, 2008). Since switching is done with digital circuitry, in contrast to physical layer devices which have substantial analogue circuit components, we would expect larger reductions in power consumption for the switching logic as microelectronics technology scales to finer process geometries.

Switch designers have more options for reducing power consumption than simply adopting energy efficient Ethernet. In switches, some processing is performed only once per frame, for example identifying the output port for the frame. In many cases, this processing uses content addressable memory (CAM) (Seifert and Edwards, 2008). CAM tends to consume a large amount of energy. If larger frames are used and the quantity of data remains constant then the frame processing overhead can be reduced and energy consumption is reduced with it. This is an incentive to use the so called jumbo frames (Spurgeon, 2000), which are larger than traditional Ethernet 1500-byte limit. For example, jumbo frames of 9,000 bytes would reduce the number of frames needed to carry data by up to six times. Unfortunately, the use of jumbo frames poses some compatibility issues and their adoption is not straightforward. Switches also check the cyclic redundancy code (CRC) in each frame to ensure that no errors have occurred. Performing this check consumes energy. In most Ethernet standards, the bit error ratio (BER) required for the device is quite low. For example, in 1000Base-T, BER has to be below 10^{-10} that means, for 10,000 bit frames, one frame in a million, at most, will suffer a bit error. Since so few frames have errors, an option to consider is disabling CRC checking in switches. If more energy is used to perform CRC checking on all frames than in re-sending the small number of frames in error then this will reduce overall energy consumption. In any case, the error will be detected at the destination. A careful analysis of the potential benefits in energy savings and the implications for mean time to false packet acceptance (MTTFPA) would have to be performed to ensure that reducing CRC checking is beneficial for a given network.

In summary, with the adoption of energy efficient Ethernet, the energy consumption of switches will be significantly reduced, whereas the energy reduction in hubs will be limited. As a result, switches will become much more energy efficient than hubs.

5 Conclusions

In this paper, the implications of energy efficient Ethernet on hubs and switches have been discussed showing that the adoption of the new standard will make switches much more efficient than hubs. The discussion illustrates the implications of the new standard on network nodes. The authors expect that energy efficient Ethernet will also have an impact on the scheduling of the frames on the ports and on upper layer protocols. Once the new standard is adopted at the physical layer, changes in both will enable greater energy savings. The implications of frame size and CRC checking for the energy consumption of switches were also discussed. These examples illustrate the authors' contention that many network design choices will have to be reviewed in the context of the emergence of energy efficiency as a first class design requirement.

References

Bavel, N.V., Dove, D., Flatman, A. and McConnell, M. (2005) 'Short haul 10 Gbps Ethernet copper PHY call for interest' Presentation at *IEEE 802.3 November meeting*.

Bennett, M. et al. (2007) 'Server bandwidth utilization plots' Presentation at *IEEE 802.3 March meeting*.

Bennett, M. et al. (2009) *IEEE P802.3az Energy Efficient Ethernet Standard Draft D2.0*.

Bennett, M. et al. (2010) *IEEE P802.3az Energy Efficient Ethernet Task Force*, available at http://grouper.ieee.org/groups/802/3/az/public/index.html.

Chou, J., Kuo, A., Tseng, D., Lin, J. and Huang, K. (2008) 'Proposal of low-power idle: 100 Base-TX, Presentation at *IEEE 802.3 January meeting*.

Christensen, K. (2007) 'Rapid PHY selection (RPS): emulation and experiments using PAUSE', Presentation at *IEEE 802.3 March meeting*.

Dove, D. (2008) 'Energy efficient Ethernet: switching perspective', Presentation at *IEEE 802.3 January meeting*.

Gunaratne, C., Christensen, K., Nordman, B. and Suen, S. (2008) 'Reducing the energy consumption of Ethernet with adaptive link rate (ALR)', *IEEE Transactions on Computers*, Vol. 57, No. 4, pp. 448–461.

Gupta, M. and Singh, S. (2003) 'Greening of the internet', *Proc. of ACM SIGCOMM*, pp.19–26.

Seifert, R. and Edwards, J. (2008) *The All-New Switch Book: The Complete Guide to LAN Switching Technology*, August, John Wiley & Sons.

Spurgeon, C.E. (2000) *Ethernet: The Definitive Guide*, February, O'Reilly (ISBN 10: 1–56592–660–9).