Contemporary Engineering Sciences, Vol. 7, 2014, no. 24, 1373 - 1381
HIKARI Ltd, www.m-hikari.com
http://dx.doi.org/10.12988/ces.2014.49169

Traffic Load-Based LPI Control Mechanism in Ethernet Access Network

Young-Hyun Kim

Department of Computer Engineering, Sunchon National University
255 Jungang-ro, Sunchon-si, Jeollanam-do, Republic of Korea

Sung-Keun Lee

Department of Multimedia Engineering, Sunchon National University
255 Jungang-ro, Sunchon-si, Jeollanam-do, Republic of Korea

Jin-Gwang Koh

Department of Computer Engineering, Sunchon National University
255 Jungang-ro, Sunchon-si, Jeollanam-do, Republic of Korea

Abstract

IEEE802.3az working group has recommended LPI based Energy Efficient Ethernet standard in order to efficiently reduce unnecessary waste of energy that might occur while there is no data to be transmitted. This paper proposes an enhanced LPI control mechanism to improve energy efficiency between terminal equipment and Ethernet switch in an EEE environment. The proposed mechanism measures the amount of traffic coming to the terminal during specific period of time. According to the measured traffic amount, ON and OFF duration for the following basic cycle will be decided. The performance evaluation results indicate that the suggested mechanism improves the overall performance in a way of significant reduction of energy consumption while it maintains an average packet delay to similar levels compared to the traditional method.

Keywords: Traffic Load Based LPI Control, Energy Efficient Ethernet, Ethernet Switch, IEEE802.3az

1 Introduction

As the demands of supporting multimedia services over Internet get higher recently, data transmission speed of network device has increased consistently. However, high speed network device tends to consume more energy. Therefore there are a number of studies actively conducted on the improvement of networking mechanism enabling to consume less energy as well as studies on the improvement of network device towards high speed capability [1]. Ethernet is the most representative technology in access networking. There is a significant potential for improving energy efficiency in an access network since the period during which data transmission is actually performed is substantially short [2]. To enhance energy efficiency for Ethernet, IEEE working group has started working on EEE (Energy Efficient Ethernet) in 2007, and confirmed the standard, 802.3az LPI(Low Power Idle) in September of 2010[3]. Regarding the details about operation of LPI client sub-layer which actually controls LPI mode, however, there is no specification made. The mechanism is recommended to be implemented by the manufacturers. Therefore the studies on exploring an optimized mechanism in controlling LPI have emerged as critical issue. This paper proposes an enhanced LPI control mechanism enabling to improve energy efficiency between terminal equipment and Ethernet switch.

2 Related Works

LPI is a mechanism enabling to save unnecessary waste of energy by switching the Ethernet link to idle mode when there is no frame to be transmitted through the link [3]. According to the performance evaluation on EEE conducted by previous studies, LPI mechanism shows highly varying energy efficiency depending on traffic characteristics of transport layer [4]. The performance of the LPI mechanism can improve by applying packet coalescing capability [4]. For achieving better energy efficiency for small-scale Ethernet switch, a synchronized coalescing mechanism has been suggested [5]. However, the synchronized coalescing mechanism has problem to support real-time multimedia service. To overcome such a drawback, an adaptive synchronized coalescing mechanism has been proposed [6]. It is determined whether the following state will remain in ON state or switch to OFF state depending on the amount of traffic arriving during ON state, rather than alternation between fixed period of ON and OFF [6]. An enhanced synchronizing packet coalescing mechanism measures the amount of incoming traffic during a certain period. Based on the measurement, the method predicts traffic characteristics of following period and adjusts threshold value suitable for the forecasted traffic load [7].

3 Traffic load-based LPI control mechanism

The proposed mechanism is applied to the operation between Ethernet switch and Ethernet terminal. The two systems connecting through the Ethernet interface stay in either of two distinctive states, namely ON or OFF. In ON state, the system can send and receive data properly. In OFF state, the system cannot send nor receive data because it blocks the power supplied to most of components in the system. As for the Ethernet switch, state transition between ON and OFF takes place repeatedly in a cycle basis. Basic cycle is set to a constant value (T_{cycle}). T_{cycle} consists of sum of two variable values, T_{on} and T_{off}. T_{on} denotes time duration during which port stays in ON state without changing its state, while T_{off} denotes time duration during which ports stays in OFF state. The Ethernet switch comprises of one up-link and multiple down-link ports. Basic cycle for all the links starts exactly at the same time. For each link, durations of T_{on} and T_{off} might be different depending on the amount of traffic delivered to the corresponding link. In other words, the values of T_{on} and T_{off} for each link is to be determined appropriately in accordance to the amount of traffic coming to the link. The prediction of an incoming traffic amount is carried out by an Ethernet terminal node. This paper defines several LPI client states for the terminal and the switch as shown in Table 1. There are three states commonly applied to terminal and switch: IDLE, OFF and ON state. I_wait state is used only by the switch.

Table 1. State definition for LPI protocol

State	Description
IDLE	Link has not been activated.
OFF	Link has been activated. However the power supplied to most of components are blocked. Hence, data cannot be transferred nor received. Only refresh signal can be transferred.
ON	All components operates in normal condition. Data can be transferred and received properly.
I_wait	Wait for information about Ton of basic cycle sent by a terminal

In this paper, there are a number of timers used. The types and roles of the timers are shown in Table 2.

Table 2. Timer definition for LPI protocol

Timer	Description
Tx_Timer	Timer to control data transmission cycle.
ON_Timer	Timer to specify the duration during which it operates in ON state.
OFF_Timer	Timer to specify the duration during which it operates in OFF state.
Wait_Timer	Timer to specify the maximum duration which it waits until getting information about Ton.

Initially the basic cycle consists of sum of T_{on_def} and T_{off_def}. After the initial cycle is completed, the cycles from then on will be determined by two variables of T_{on} and T_{off} that reflect current traffic conditions. Figure 1 illustrates how the switch and the terminal node operate to achieve better energy efficiency in the Ethernet respectively.

The terminal node operates as follows. In IDLE state, once the link gets activated, the terminal node carries out synchronization process with the switch. In an initial state, the value of basic timer is set by negotiation process with the switch. Therefore the terminal node changes the link state to active without sending information about T_{on}. Subsequently, Tx_Timer and ON_Timer are triggered. After that, it changes its state to ON. If it receives packets from the upper layer in IDLE state, it stores the packet in the transmit queue with maintaining its state as IDLE. In ON state, when Tx_Timer expires, a certain amount of packets stored in the transmit queue are sent out to the switch in accordance with the transmission capacity of the link. Subsequently, Tx_Timer is restarted, which is a timer for next transmission cycle. In ON state, when a set of packets arrive from the upper layer, the packets are stored in transmit queue.

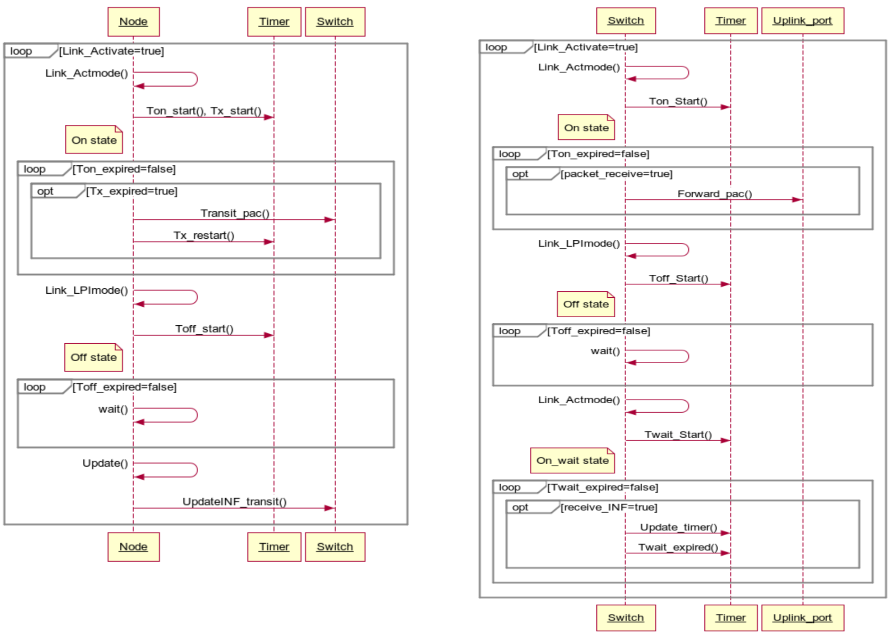

Figure 1. Sequence diagram of the terminal and switch

When ON_Timer expires, which means ON state within the cycle has been terminated, Tx_Timer is aborted and the link enters LPI mode. Then it starts OFF_Timer and turns into OFF state. In OFF state, when the terminal node receives a number of packets sent by the upper layer, it stores the packets in the transmit queue and updates the traffic characteristics information. Once OFF_Timer expires, new T_{on} and T_{off} value for the following cycle are calculated based on the traffic characteristics information stored in queue. Now it activates the link by switching to Active mode. Then it sends Ton information to the switch. It starts Tx_Timer and ON_Timer. After that, it enters ON state. The switch operates as follows. In IDLE state, when the link is activated, the switch performs synchronization process with the terminal node. Through negotiation process, the values of basic parameters and basic timers are determined. Initially, basic timer is already set to default value. It means it is not necessary to wait for receiving T_{on} information from the terminal. Therefore, the switch changes the link state to active mode without waiting for T_{on} information. Subsequently, it starts ON_Timer. Then it turns to ON state straight away. Now it is ready to receive the packets sent by the terminal. When receiving the packets sent by the terminal node in ON state, it stores the packets in the up-link transmit buffer of core network. In OFF state, when OFF_Timer expires, the corresponding link is switched to Active mode. Subsequently, Wait_Timer is triggered and then it enters T_{on_Wait} state. Each terminal node evaluates T_{on} and T_{off} value for the following cycle by measuring the amount of packets stored in transmit queue, which have arrived from the application layer during OFF period of previous basic cycle. Considering the total amount of packets, the transmission capacity of the tx link and overhead, the terminal node evaluates the expected time how long it will take to send the packets to the switch. Based on this evaluation, Ton is decided. Care should be taken when deciding T_{on}. T_{on} needs to be integer multiples of T_{on_def} without exceeding half of basic cycle duration. The terminal node sends the packets stored in transmit queue only during T_{on} period. When T_{on} period expires, data transmission should be aborted. Then it turns to LPI mode, which leads to saving energy consumption. The switch is also allowed to send out data to an internet core network only during T_{on} period. Once T_{on} expires, data transmission is aborted. Then it goes to LPI mode while T_{off} period so as to save energy consumption.

4 Performance Evaluation

The performance of the proposed mechanism was evaluated by simulation. CSIM 20 simulator was used for simulation. Network architecture used by previous study [6] was applied to our simulation. Ethernet link capacity is set to 10 Gbps. T_{cycle}, T_{on_def} and T_{off_def} are set to 100msec, 5msec and 95msec respectively.

Simulation has been carried out using two modelling approaches: (1) aggregate traffic modelling for the traffic through all links connecting to the switch, and (2) independently separate traffic modelling for each link.

Incoming traffic arrival pattern is assumed to follow a Poisson distribution with fixed packet size of 1,500 bytes. Script has been coded to decide an average arrival rate according to the varying overload. For an adaptive PPSE mechanism with two threshold values (1,000 packets and 5,000 packets) and the proposed mechanism, the performance evaluation was performed by measuring the energy consumption rate and the average packet delay by varying traffic load ranging from 1% to 30%. Energy consumption rate in ON state is assumed to 100% since all the devices fully operate. For OFF state, the energy consumption rate is assumed to 10%. Packet delay denotes the time duration to be taken for the packet to be delivered to the network interface card from the host terminal.

Simulation result for an aggregate traffic modelling is shown in figure 2.

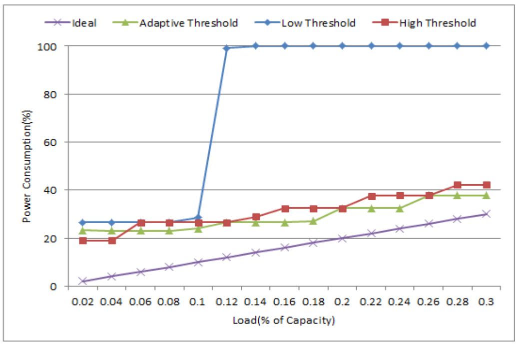

Figure 2. Energy consumption

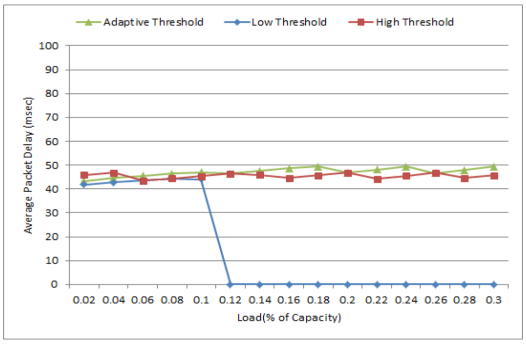

Figure 3. Average packet delay

In this figure, energy consumption rates are compared among three different cases: ideal case, adaptive PPSE mechanism and the proposed mechanism. The existing switch remains in ON state all the time regardless of traffic load. Therefore it shows 100% of energy consumption rate. In ideal case, energy consumption rate is linearly proportional to network load. The performance evaluation results for low and high threshold are shown to be highly consistent with previous studies. The new mechanism shows lower energy consumption rate than the adaptive PPSE mechanism. While the traditional mechanism shows 26.5% of the energy consumption rate, the proposed mechanism shows 23%, which indicates there is a significant improvement in energy efficiency by applying the proposed mechanism. Figure 3 is a graph that shows the simulation result in terms of an average packet delay. The average delay reaches down to the lowest level when the threshold is low. When the traffic load gets more than 12% of link capacity, link stays in ON state all the time. Therefore the average delay becomes zero. For most cases of different traffic loads, the proposed mechanism shows similar or lower average packet delay than the adaptive PPSE with high threshold.
When the traffic load drops down below 5%, the proposed mechanism shows 44.8msec of the average packet delay, which is lower than the average delay of

the traditional mechanism with high threshold. In range of 6~10% that is an average load observed in access network, the proposed mechanism shows higher average delay than the case with fixed threshold. When the traffic load exceeds 15%, the average delay tends to gradually increase. The performance evaluation results indicate that the mechanism contributes to significant reduction of energy consumption rate while it maintains the average packet delay to similar levels. Hence, the proposed mechanism is considered to enhance the overall performance. For each of four links, we carried out the simulation by varying the amount of load. Fixing the load of link 1 and varying the load of the rest of links, we measured the energy consumption rate and the average packet delay for each link. Performance evaluation results for an individual traffic are shown in table 3.

Table 3. Performance evaluation for individual traffics

Input Load				Energy Consumption (%)				Average Packet Delay (msec)			
1	2	3	4	1	2	3	4	1	2	3	4
0.02	0.02	0.02	0.02	14.5	14.5	14.5	14.5	46.1	46	46	46
0.02	0.04	0.06	0.08	14.5	14.5	19	19	46	47	43.1	44.1
0.02	0.06	0.08	0.1	14.5	19	19	20	46	43.1	44	44.1
0.02	0.1	0.12	0.14	14.5	20	23.4	23.4	46	44.1	41.1	42
0.02	0.12	0.14	0.16	14.5	23.4	23.4	25.5	46	41.1	42.1	41.8
0.02	0.14	0.16	0.18	14.5	23.4	25.5	27.9	45.9	42.1	41.7	39.2
0.02	0.16	0.18	0.2	14.5	25.5	27.8	27.9	45.9	41.8	39.2	43.9
0.04	0.02	0.02	0.02	14.5	14.5	14.5	14.5	46.9	46	45.9	46.1
0.04	0.04	0.06	0.08	14.5	14.5	19	19	46.8	47	43.1	44
0.04	0.06	0.08	0.1	14.5	19	19	20	47	43.1	44	44.1
0.04	0.1	0.12	0.14	14.5	19.8	23.4	23.4	46.8	44.2	41.1	42.1
0.04	0.12	0.14	0.16	14.5	23.4	23.4	25.5	46.9	41.2	42.1	42
0.04	0.14	0.16	0.18	14.5	23.4	25.5	27.9	47	42	41.6	39.2
0.04	0.16	0.18	0.2	14.5	25.5	27.9	27.9	46.9	41.8	39.2	41.5
0.06	0.02	0.02	0.02	19	14.5	14.5	14.5	43	46	46.1	46.1
0.06	0.04	0.06	0.08	19	14.5	19	19	43.1	47	43.1	44
0.06	0.06	0.08	0.1	19	19	19	20	42.9	43.1	44	44.3
0.06	0.1	0.12	0.14	19	19.8	23.4	23.4	43	44.1	41.1	42
0.06	0.12	0.14	0.16	19	23.4	23.4	25.5	43	41.1	42.1	41.8
0.06	0.14	0.16	0.18	19	23.4	25.6	27.9	43	42.1	41.7	39.1
0.06	0.16	0.18	0.2	19	25.4	27.9	28	43.1	41.9	39.1	42.9
0.08	0.02	0.02	0.02	19	14.5	14.5	14.5	43.9	45.9	46	46.1
0.08	0.04	0.06	0.08	19	14.5	19	19	44	47	43.1	44
0.08	0.06	0.08	0.1	19	19	19	19.8	44	43.2	44.1	44.2
0.08	0.1	0.12	0.14	19	19.9	23.4	23.4	43.9	44.1	41.1	42.1
0.08	0.12	0.14	0.16	19	23.4	23.4	25.4	44	41.1	42.1	42.2
0.08	0.14	0.16	0.18	19	23.4	25.6	27.9	44	42.1	41.6	39.2
0.08	0.16	0.18	0.2	19	25.5	27.9	27.9	44	41.6	39.1	41.2
0.1	0.02	0.02	0.02	19.8	14.5	14.5	14.5	44.3	46.1	46	46.2
0.1	0.04	0.06	0.08	19.8	14.5	19	19	44.3	46.9	43.1	44
0.1	0.06	0.08	0.1	19.8	19	19	19.8	44.3	43.1	44	44.3
0.1	0.1	0.12	0.14	19.9	20	23.4	23.4	44.2	44	41.1	42
0.1	0.12	0.14	0.16	19.7	23.4	23.4	25.4	44.3	41	42.1	42.1
0.1	0.14	0.16	0.18	19.9	23.4	25.4	27.9	44	42.1	42.1	39.1
0.1	0.16	0.18	0.2	19.8	25.4	27.9	27.9	44.2	42	39.1	40.5

It is found that the energy consumption is differentiated for the traffic load of each link. Since T_{on} period for each link is adjusted by traffic load, energy consumption rate decreases in case of low link load, while energy consumption increases in case of high link load. Packet delay is shown differently depending on the energy consumption rate. In case of low energy consumption rate, T_{on} period gets shorter. As a result, packet delay increases. However, the overall packet delay is less than 50msec so that the amount of increment cannot affect the quality of real-time application service. In an adaptive PPSE mechanism, state transition between ON and OFF for all the ports is determined depending on the condition of a specific port that has the highest traffic load. Therefore the drawback of this method has been pointed out that it does not exactly reflect the characteristics of all the traffics coming to the switch. The suggested mechanism has been designed to overcome the downside of adaptive PPSE.

5. Conclusion

The proposed mechanism measures the amount of incoming traffic arriving at the terminal. Based on the measured data, ON and OFF durations are determined for the following cycle. Furthermore, the ON and OFF durations are delivered to the switch, so that synchronized data transmission can be carried out. Since the switch operates LPI mode depending on the traffic load, the proposed mechanism is capable of overcoming the drawback of the adaptive synchronized coalescing mechanism effectively. The simulation results indicate that the performance improves by and large since the proposed method contributes reduction of energy consumption rate without significant increasing of an average packet delay. According to the traffic load for each link, the energy consumption shows to be differentiated. Based on the traffic load, the new mechanism applies different T_{on} duration to an individual link, so that it results to high energy consumption rate with high load and low energy consumption rate with low load. Therefore the proposed mechanism is considered to effectively overcome the problem of the existing method that is not capable of reflecting the characteristics of all traffics incoming to the switch.

Acknowledgements. This research was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the CITRC support program (NIPA-2014-H0401-14-1008) supervised by the NIPA

References

[1] Report to Congress on Server and Data Center Energy Efficiency," EPA ENERGY STAR Program, U.S. Public Law 109-431, (2007).

[2] C. Gunaratne, K. Christensen, et al., Reducing the energy consumption in Ethernet with ALR, *IEEE Trans. Computers*, **57.4** (2008), 448-461.

[3] K. Christensen, P. Reviriego, et al., IEEE 802.az: the road to energy efficient ethernet, *IEEE Communications Magazin*e, **48.11** (2010), 50-56.

[4] Reviriego, J. Hernandez, et al., Performance Evaluation of Energy Efficient Ethernet, *IEEE Communications Letters*, **13.9** (2009), 1-3.

[5] P. Reviriego, J. Hernadez, D. Larrabeiti, J. Maestro, Burst Transmission in Energy Efficient Ethernet, *IEEE Internet Computing*, (2010).

[6] M. Mostowfi, K. Christensen, Saving energy in LAN switches: New methods of packet coalescing for Energy Efficient Ethernet, *Green Computing Conference and Workshops(IGCC)*, (2011), 1-8.

[7] Y.H Kim, S.K Lee, et al., Enhanced synchronizing packet coalescing mechanism for improving energy efficiency in Ethernet switch, *International Journal of Smart Home*, **7.3** (2013), 375-384.

Received: August 30, 2014