Photon Netw Commun (2015) 30:71–84
DOI 10.1007/s11107-015-0503-5

Adaptive state transition control for energy-efficient gigabit-capable passive optical networks

Steven S. W. Lee[1] · Kuang-Yi Li[1]

Received: 6 July 2014 / Accepted: 26 March 2015 / Published online: 5 April 2015

Abstract In this paper, we investigate power management problems that affect gigabit-capable passive optical networks (GPONs). In GPONs, the basic principle for power reduction is to keep optical network units (ONUs) in the Power Saving state, wherein some of the hardware and software functions are turned off. Current research focuses on scheduling and determining the length of the sleep periods for ONUs that are in the Power Saving state. Our investigation indicates that keeping ONUs in the Power Saving state is not necessarily the most energy-efficient practice. The Power Saving state and the Full Power state must be jointly considered. Our study also reveals that traffic distribution is a critical factor. Considering only the average is insufficient. The variance of packet arrival also must be included when designing a green GPON. We have analyzed the power consumption in a GPON and determined the optimal load threshold for triggering a state transition from the Power Saving state to the Full Power state. For the reverse direction, we propose a neural network-based adaptive control scheme to achieve near optimal control of the transition from Full Power to Power Saving. We also propose a burst transmission scheme to determine the sleep period for an ONU in the Power Saving state. Unlike the proposal of ITU-T, which uses a fixed length for the sleep period, the state sojourn time in our approach is dynamically adjusted. We have carried out extensive simulations to evaluate the performance of the proposed scheme. Simulation results show that the total energy consumption of the proposed scheme is almost equal to the optimal control scheme.

✉ Steven S. W. Lee
ieeswl@ccu.edu.tw

1 Department of Communications Engineering, National Chung Cheng University, Chiayi, Taiwan

Keywords Gigabit-capable passive optical network · Energy-efficient PON · Neural network · Power management

1 Introduction

Because of the need to reduce greenhouse gas emissions and reduce environmental pollution, energy efficiency has become an important issue in the design of communication networks. In recent years, standard organizations including IEEE and ITU-T have proposed ways to reduce power consumption for communication networks. To further enhance eco-consciousness, the EU has proposed a code of conduct to express the maximum allowed power consumption for broadband communication equipment [1]. These new rules have turned energy efficiency into a practical issue. They will encourage industrial engineers and academic researchers to focus on designing greener networks.

The Passive Optical Network (PON) is one of the major technologies used to provide broadband access. A PON allows end users to connect their CPE to the Internet through a common Ethernet interface. To study power management in a PON, the user behavior on Ethernet is the best reference.

The estimated energy consumption of the Ethernet network interface controllers in desktop PCs and other network edge devices in the USA was approximately 5.3 TWh/year in 2005 [2]. However, one study [3] shows that the average utilization of Ethernet is only 1–5 % for desktop links. Based on this low network utilization, we have proposed a Fixed Bandwidth Allocation (FBA) scheme for saving energy in an EPON [4]. Several other studies on EPONs have been proposed in recent years. Zhang and Ansari have proposed sleep-aware traffic scheduling schemes for an energy-efficient 10G EPON [5]. Kubo et al. have proposed sleep and adaptive link rate control functions to reduce power consumption for sleep mode ONUs [6]. In another study [7], three ONU receiver architectures were presented, and the energy saving performance for ONUs using sleep mode on those architectures was evaluated. Dhaini et al. [8] proposed a green bandwidth allocation scheme that uses batch-mode transmissions in both upstream and downstream directions to save energy consumption for a given amount of sleep time. They continued this work by proposing a new sleep-time sizing and scheduling framework. Each ONU is considered to be an M/G/1 system [9]. A closed-form model for the proposed sleep-time model is addressed. Other studies on Power Saving techniques and mechanisms for PONs can be found in the tutorial paper [10].

ITU-T has proposed a standard for power management [11–13]. In this proposal, an ONU can enter the Power Saving mode with permission from the OLT to reduce energy consumption. The permission is conveyed in a newly defined Physical Layer OAM (PLOAM) command. Two operation modes are defined: the doze mode and the cyclic sleep mode. In the doze mode, the receiver of the ONU is always working, regardless of the downstream traffic condition. The transmitter can go to sleep when the upstream load is light. Because the receiver is always working an OLT can wake the ONUs when needed. In contrast, an ONU operating in the cyclic sleep mode can turn off both its receiver and transmitter when it sleeps. All G.987.3-compliant implementations are expected to support the doze mode. Support for the cyclic sleep mode is optional. In this paper, we focus our design on doze mode and leave cyclic sleep mode to future studies.

Although ITU-T has made proposals for power management in a GPON system, the detailed control policy is considered to be implementation-dependent. Therefore, the control policy to determine the timing for an ONU to enter and leave the Power Saving state is still not clear. Only our previous work [14] and another previous study [15] take this issue into account. To the best of our knowledge, other works in the literature only focus on ONUs operating in the Power Saving state. The Full Power state is called the Active state in [15]. In that paper, an analytical model based on Poisson process is proposed. In this paper, we take self-similarity traffic into consideration and proposed a neural network-based control scheme for the system. Traffic generated by Poisson process is called short-range dependent, while traffic following self-similarity distribution is called long-range dependent. The internet traffic has been proved to be long-range dependent. Our study shows that there is a big gap on power consumption between short-range dependent traffic and long-range dependent traffic. Purely taking short-range dependent into consideration is not optimal for Power Saving.

When an electronic device is turned on, there is a surge current in a short period of time. This is called the inrush current phenomenon [16]. Inrush current causes large energy consumption as the optical transceiver warms up. Hence, an ONU should wake up its transceiver as infrequently as possible. Because of the large energy cost of warming up a transceiver, keeping an ONU in the Power Saving state is not necessarily the most energy-efficient option. Whether an ONU should be operated in the Power Saving or Full Power state depends not only on the first moment (average load) but also on the second moment (load variance) of the arrival traffic.

We have previously analyzed the power consumption in an energy efficient GPON and derived the optimal condition for performing the state transition from the Power Saving state to the Full Power state [14]. However, it is difficult to determine the optimal strategy for making state transitions from Full Power to Power Saving. When an ONU is operating in the Power Saving state, it can count the number of times the transceiver is turned on and off. This number can be used to determine an optimal traffic load for triggering a state transition. In the Full Power state, the transceiver is always on. Without the transceiver on/off count, the state transition is determined based on the traffic load recorded the last time the ONU was in the Power Saving state [14]. Because traffic conditions change from time to time, estimation will be inaccurate if it does not take new traffic condition into account.

Artificial neural networks have been widely used to control complex systems. In a previous study [17], a neural network-based scheme was proposed that allocates upstream bandwidth in a passive optical network. In a complex system, there are many parameters that need to be taken into account. One major benefit of using a neural network is that it can learn how to react to the environmental inputs through a learning process. The inputs can be generated offline by simulation programs. There are many engineering problems that have no exact model to describe the behavior of the environmental inputs, so it is much easier to generate the input data through simulations. We make use of the same concept in this paper and propose a neural network-based scheme for controlling the state transition. A large set of input traffic patterns with a variety of mean loads and different levels of burstiness were generated and used as the training data for our neural network. Hence, our approach is able to handle a wide range of input traffic.

When an ONU is in the Power Saving state, the lengths of the sleeping cycle and working cycle as defined in the ITU-T specification are fixed. In this paper, we apply a burst transmission scheme and propose a control policy for ONU operating in the Power Saving state. We perform simulations to analyze the performance of the proposed burst transmission scheme and make performance comparisons to the IEEE 802.3az control scheme [18]. The efficiency of power saving and average packet delay are also discussed in the paper.

The remainder of this paper is organized as follows. In Sect. 2, we introduce the state diagrams used in this work. In Sect. 3, we present several adaptive control schemes for state

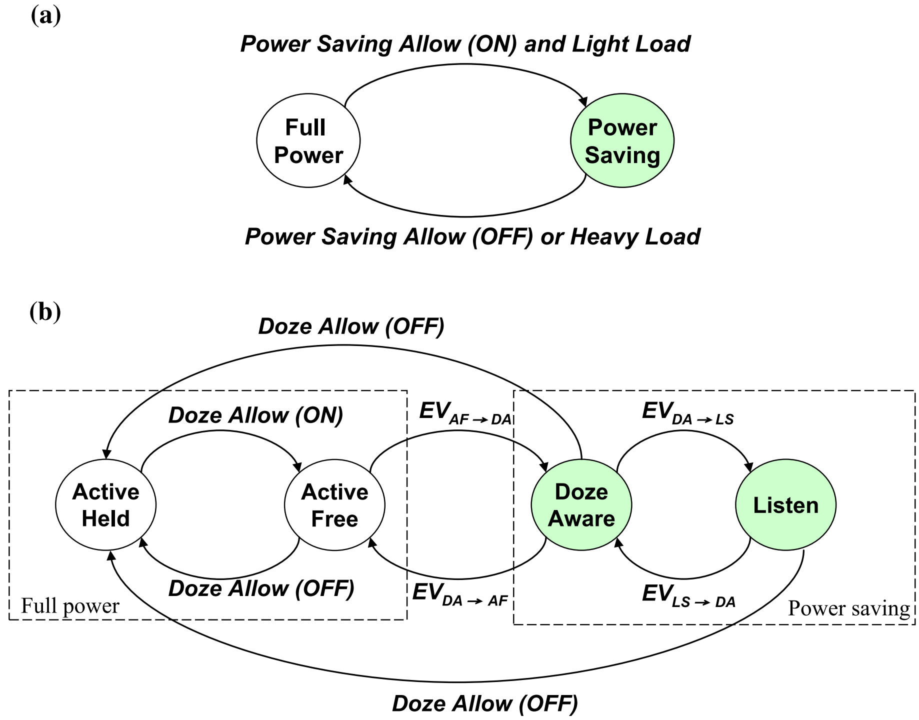

Fig. 1 ONU state diagram. **a** GPON ONU state diagrams, **b** ONU doze mode state transition diagram

transitions between Full Power and Power Saving states. In Sect. 4, we present control policies for determining the state transitions between sub-states within the composite Power Saving state. In Sect. 5, we show the simulation results. Concluding remarks are provided in the final section.

2 ONU state diagrams

The state diagram is basically the same as that proposed in our previous study [14]. To make the paper self-contained, we describe the state diagram in more detail in this section. The state diagrams used in this paper are shown in Fig. 1. We first use two composite states, as shown in Fig. 1a, to give a high-level view of energy-efficient operations in an ONU. A detailed state diagram is shown in Fig. 1b. As in Fig. 1a, there are two states: Full Power and Power Saving. The OLT issues a Power Saving Allow message with permission parameter ON or OFF to tell an ONU whether it is allowed to enter the Power Saving state or not. In the Full Power state, all components are performing normally without applying any power saving functions. In the Power Saving state, the ONU tries to reduce its power consumption by turning off unnecessary components or forcing some components into power down mode. In a GPON system, an ONU is allowed to enter the Power Saving state only when it receives a Power Saving Allow (ON) PLOAM sent from the OLT. By issuing a Power Saving Allow (ON/OFF) message, an OLT is able to control the operating state of any ONU in its network.

There are two power saving modes suggested in ITU-T. In the doze mode, when an ONU is operating in the Power Saving state, the transmitter and its corresponding driver circuit are turned off to reduce power consumption. The receiver is always working regardless of whether the ONU is in the Power Saving state or in the Full Power state. Thus, the ONU is able to receive downstream traffic and control messages from the OLT at any moment. In the cyclic sleep mode, both the transmitter and the receiver are turned off when the ONU is in the Power Saving state. Because a sleeping ONU cannot receive any message from the OLT, the OLT cannot wake an ONU by sending a wake-up command. Usually, an ONU is woken when the sleep time goes over a pre-determined value or when a local event occurs.

In this work, we consider the power saving design for a GPON operating in doze mode. We follow ITU-T G.987 to divide the composite Full Power state and the Power Saving state in Fig. 1a into four sub-states shown in Fig. 1b. Those sub-states are named Active Held, Active Free, Doze Aware, and Listen. Although we use the same four sub-states as they appeared in G.987, the schemes to control state transitions are different.

In Fig. 1b, there are four total sub-states: Active Held, Active Free, Doze Aware, and Listen. State transitions are triggered either by a Doze Allow (ON/OFF) PLOAM or by events generated by the ONU itself. In Fig. 1b, Doze Allow is a PLOAM message that comes from the OLT. An event is labeled $EV_{x \to y}$ where x is the source state and y is the destination state. We say that an ONU is in the Full Power state or the Power Saving state when there is no ambiguity and no need to specifically identify the sub-state.

In doze mode, the power savings come from turning off the transmitter when there is no upstream data. The Active Held state is the initial state whenever an ONU is powered on. In the Active Held state, the ONU is fully responsive, forwarding downstream traffic and responding to all bandwidth allocations. A state transition does not occur until the ONU receives the Doze Allow (ON) permission from the OLT. Doze Allow (ON) is a GPON PLOAM message that allows the ONU to enter the Active Free State.

As in the Active Held state, an ONU in the Active Free state is fully responsive, forwarding downstream traffic and responding to all bandwidth allocations. The ONU keeps its transmitter and receiver working in this state. The only difference between Active Held and Active Free is that in the latter the ONU is allowed to enter the Power Saving state freely. Whether to perform power saving is a local decision for the ONU.

Event $EV_{AF\rightarrow DA}$ and Event $EV_{DA\rightarrow AF}$ are used to control the state transitions between the Active Free state and the Doze Aware state. Generally speaking, both events are activated when traffic conditions in the upstream direction change. ONU performs state changes to reduce power consumption. More specifically, when an ONU is in the Active Free state, low upstream load activates $EV_{AF\rightarrow DA}$, which triggers the ONU to enter the Doze Aware state. In contrast, when there is a large amount of upstream traffic waiting for transmission, it enables $EV_{DA\rightarrow AF}$, which triggers a state transition from the Doze Aware state to the Active Free state. We will present several control schemes for determining the activating conditions for both $EV_{DA\rightarrow AF}$ and $EV_{AF\rightarrow DA}$ in Sect. 3.

As an ONU is operating in the Doze Aware state, its receiver and transmitter remain powered on. This state persists until the ONU has transmitted all its upstream data to the OLT. In our design, $EV_{DA\rightarrow LS}$ is the upstream queue empty indication that triggers a state transition from the Doze Aware state to the Listen state. In the Listen state, only the transmitter is turned off. The receiver still works. The downstream data can still be received by the ONU. The ONU remains in the Listen state until Event $EV_{LS\rightarrow DA}$ happens. We will evaluate two policies for generating $EV_{LS\rightarrow DA}$. The details are presented in Sect. 4.

The Listen state is the actual one to make ONU save energy. An ONU only remains in the Doze Aware state for a short time. If the ONU stays in the Doze Aware state for a long time, it indicates that the upstream traffic is heavy. Thus, the system should go to the Full Power state to avoid turning on and off the transmitter frequently. If the load of upstream traffic is not heavy enough to force the ONU to return back to the Full Power sate, the ONU remains in the Power Saving state. An ONU continues its power saving operations until it receives a Doze Allow (OFF) from the OLT. That forces the ONU back to the Active Held state, and all communication functions turn on.

Please note that the state transitions are controlled by a pair of timers in the G.987. $EV_{DA\rightarrow LS}$ and $EV_{LS\rightarrow DA}$ are triggered by the expirations of local timers T_{aware} and T_{lowpower}, respectively. The ONU periodically changes between the Listen and the Doze Aware states. They do not take buffer status and traffic condition into consideration. The $EV_{AF\rightarrow DA}$ and $EV_{DA\rightarrow AF}$ are out of scope of G.987. How to decide both events are not specified in the Recommendation. In this work, our proposal clearly defined all of the triggering events. Since our scheme takes traffic and buffer condition into consideration, our system is more energy efficient.

3 State transition control between full power state and power saving state

In this section, we present schemes to determine state transitions between the Full Power state and the Power Saving state. These schemes use the same condition $EV_{DA\rightarrow AF}$ (taken from our previous study [14]) to trigger state transitions from the Power Saving state to the Full Power state. For the reverse direction, the newly developed simulation-based scheme and neural network-based schemes are used to control the state transition from the Full Power state to the Power Saving state. We first show that the triggering condition $EV_{DA\rightarrow AF}$ used in these schemes is optimal.

3.1 Optimal state transition from power saving state to full power sate

Symbols used in our analysis are listed below.

T	Observation duration
$P_{\text{warm up}}$	Mean power consumption during transmitter warm-up
$t_{\text{warm up}}$	Time required for a transmitter to finish a single warm-up
$T_{\text{total warm up}}$	Total transmitter warm-up time within observation time T; $T_{\text{total warm up}}$ is the product of the number of times a transmitter is turned on and the transmitter warm up duration $t_{\text{warm up}}$;
P_{data}	Mean power consumption for transmitting upstream data;
T_{data}	Total data transmission time in observation time$T$;
P_{idle}	Mean power consumption when ONU is idle. The ONU is idle when it is ready to transmit data and is waiting for the OLT to allocate upstream bandwidth.
$T_{idle(PS)}$	The total transmitter idle time in the observation time T when the ONU is in the Power Saving state.
$T_{idle(FP)}$	The total transmitter idle time in the observation time T when the ONU is in the Full Power state.

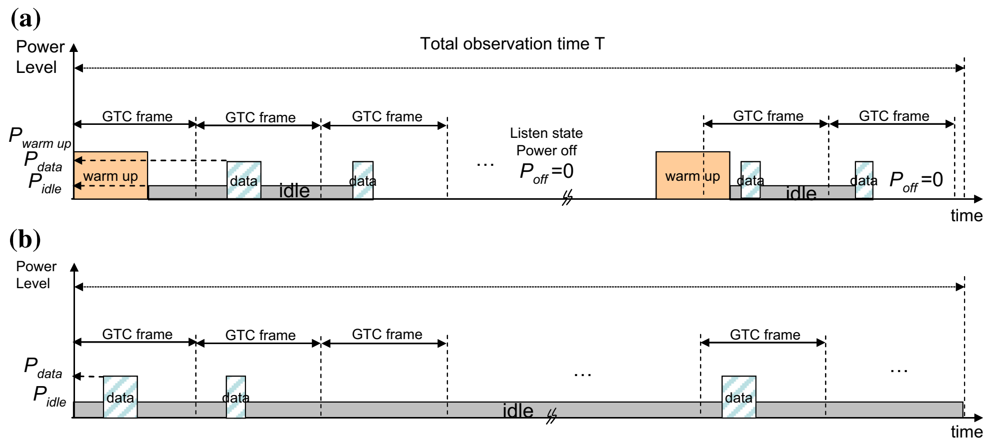

Fig. 2 ONU power consumption and timing in doze mode. **a** ONU in Power Saving states (i.e., Doze Aware and Listen States). **b** ONU in Full Power States (i.e., Active Held and Active Free States)

P_{off}	The power consumption is zero when the transmitter is off.

In GPON, the upstream consists of consecutive GTC frames. Each GTC frame duration is 125 µs. Upstream GTC frames are made up of bursts from different ONUs. The start time and stop time for a burst is determined by the dynamic bandwidth allocation (DBA) agent inside the OLT [19]. Figure 2 shows the power consumption and timing diagrams for a transmitter operating in the Full Power state and the Power Saving state. Due to the rush current effect, the average power consumption during the circuit warm-up period might be higher than during the data transmission and idle periods.

State transitions from the Power Saving state to the Full Power state are determined by comparing the power consumption for both states during a period of observation time, T. We dynamically determine the best load threshold Th_DZ to minimize power consumption. By comparing the power consumption in both states, we can derive the dynamic threshold Th_DZ.

As an ONU enters the Power Saving state, it starts to observe the traffic condition in an interval T to obtain the total warm-up time $T_{\text{total warm up}}$ and the total upstream data transmission time T_{data}. To keep the ONU in the Power Saving state, the power consumption for the ONU to operate in the Full Power state has to be larger than the power consumption for the ONU to operate in the Power Saving state. Thus, if the following condition (1) is true, the system should remain in the Power Saving state; otherwise, it should enter the Full Power state:

$$P_{\text{warm up}} \times T_{\text{total warm up}} + P_{\text{idle}} \times T_{\text{idle}(\text{PS})} < P_{\text{idle}} \times T_{\text{idle}(\text{FP})} \tag{1}$$

Assuming the system is properly maintained, there is no packet loss caused by buffer overflow. Thus, the total time for data transmission should be the same regardless of whether the ONU is transmitting upstream data in the Full Power state or in the Power Saving state. That is, the total data transmission time T_{data} is the same in Fig. 2a, b. Thus, the power consumption used for data transmission does not appear in (1). Furthermore, because the transmitter is either idle or transmitting data, Eq. (2) is always true.

$$T_{\text{idle}(\text{FP})} = T - T_{\text{data}} \tag{2}$$

Based on (1) and (2), we can obtain (3).

$$\begin{aligned}(P_{\text{warm up}} \times T_{\text{total warm up}} + P_{\text{idle}} \times T_{\text{idle}(\text{PS})})\\ /(P_{\text{idle}} \times T) < 1 - T_{\text{data}}/T\end{aligned} \tag{3}$$

Using $Load = T_{\text{data}}/T$, we can rewrite (3) as below.

$$\begin{aligned}Load < 1 - (P_{\text{warm up}} \times T_{\text{total warm up}} + P_{\text{idle}} \times T_{\text{idle}(\text{PS})})\\ /(P_{\text{idle}} \times T)\end{aligned} \tag{4}$$

We define Th_DZ to be the right hand side of (4), that is,

$$\begin{aligned}Th\_DZ = 1 - (P_{\text{warm up}} \times T_{\text{total warm up}} + P_{\text{idle}} \times T_{\text{idle}(\text{PS})})\\ /(P_{\text{idle}} \times T)\end{aligned} \tag{5}$$

In our system, the event $\text{EV}_{\text{DA}\rightarrow\text{AF}}$ becomes true as the traffic load is higher than Th_DZ. Please note that Th_DZ is not a constant value. It depends on traffic condition. To implement the algorithm, the device manufacture can record P_{warmup}, P_{idle}, and transmitter warm-up time in nonvolatile memory so the ONU can access those parameters to calculate Th_DZ. Another approach is to download those two parameters through GPON OMCI [20]. In operating, ONU

MAC only needs to collect the total amount of upstream data transmission time T_{data} and the number of transmitter being turning on in a period T. Because $T_{\text{idle(PS)}}$ is equal to $T - T_{\text{data}} - T_{\text{total warm up,}}$ all of the parameters in (5) are known for the ONU. Therefore, if load is larger than Th_DZ, the ONU should leave the Power Saving state and enter the Full Power state. Th_DZ is then stored in the system that will be used as a reference for the Mean Load Scheme described in the next section.

3.2 Adaptive control schemes

In this subsection, we present three control schemes for managing state transitions between the Full Power state and the Power Saving state. These schemes use load threshold Th_DZ for state transition control from the Power Saving state to the Full Power state. For the reverse direction, each scheme has its own triggering condition (i.e., $\text{EV}_{\text{AF}\rightarrow\text{DA}}$) for the state transition from the Full Power state to the Power Saving state.

As mentioned in Sect. 3.1, an ONU operating in the Power Saving state knows the total number of transceiver warm-up times. Thus, all of the terms in (5) are known. However, in the reverse case, when the ONU is operating in the Full Power sate, it has no idea how many warm-up times would happen if the ONU were operating in the Power Saving state. Consequently, an ONU operating in the Full Power state cannot derive the power consumption for the Power Saving state through the computation used in Sect. 3.1 Therefore, Eq. (5) cannot be applied for state transition control.

Intuitively, because the network load is light, letting ONUs operate in the Power Saving state consumes less energy than leaving them in the Full Power state. However, as the load increases, keeping the ONU in the Full Power state becomes a better choice due to the elimination of unnecessary power consumption due to frequent transmitter warm-up processes in the Power Saving state. Although traffic load is the dominating factor, it is not the only one relevant for power consumption. For two traffic patterns having the same average load, the one with higher burstiness needs fewer warm up times because the average amount of continuous upstream data is larger than the other. Therefore, in designing a green GPON system, traffic burstiness also needs to be taken into account. We propose three control schemes for determining the state transition from the Full Power state to the Power Saving state.

3.2.1 Mean load scheme

The Mean Load Scheme only takes traffic load into consideration. The benefit of the mean load scheme is its simplicity. In Sect. 3.1, we have shown that threshold Th_DZ is the optimal reference for determining whether to perform a state transition from the Power Saving state to the Full Power state. It is natural to use this threshold as a reference to determine state transitions from the Full Power state to the Power Saving state. We have used this threshold as the trigger event $\text{EV}_{\text{AF}\rightarrow\text{DA}}$ in our previous work [14]. Using this approach, the ONU only needs to calculate the average upstream load. If the arrival traffic has an average load lower than Th_DZ, the ONU goes from the Active Free state to the Doze Aware state.

To avoid frequent state transitions between the two states, the actual threshold can be set to be $\alpha \times Th\_DZ$. More specifically, the thresholds for the state transitions from the Doze Aware state to the Active Free state and vice versa are Th_DZ and $\alpha \times Th\_DZ$, respectively. The purpose of α is to provide a double threshold to avoid system oscillation between the Power Saving state and the Full Power state. The ratio $\alpha$is smaller than 1. A small value of α reduces the chance of a state transition. However, it might waste more power. The network operators can determine the observation time T and the ratio α to optimize the tradeoff between transition frequency and power saving ratio. The use of a double threshold in energy efficient Ethernet design has been analyzed previously [2].

3.2.2 Simulation-based scheme

In the second approach, the ONU uses the exact power consumption values to determine whether it should leave the Full Power state. The implementation is basically like a simulation program. The traffic traces of the current period T serve as the input to the simulation program running in the ONU. Although the ONU is operating in the Full Power state, the program behaves like it is in the Power Saving state. By doing so, the program can obtain the total number of warm-up times required in current period T. At the end of an observation time, we simply make a comparison between the power consumption in the Full Power state and the Power Saving state. If the value for the latter case is smaller, the ONU should leave the Full Power state and jump to the Power Saving state; otherwise, it remains in the Full Power state. This simulation-based scheme can use power consumption information for both states. However, due to the high computation cost, this scheme is not a candidate for practical implementation. Nevertheless, it provides a good reference for performance comparisons.

3.2.3 Neural network scheme

Neural networks have been used to provide adaptive control in many applications. The major benefit of a neural network is its ability to provide adaptive control in real time. Figure 3 depicts the neural network used in our design. There are seven hidden layers in the network, and each hidden layer

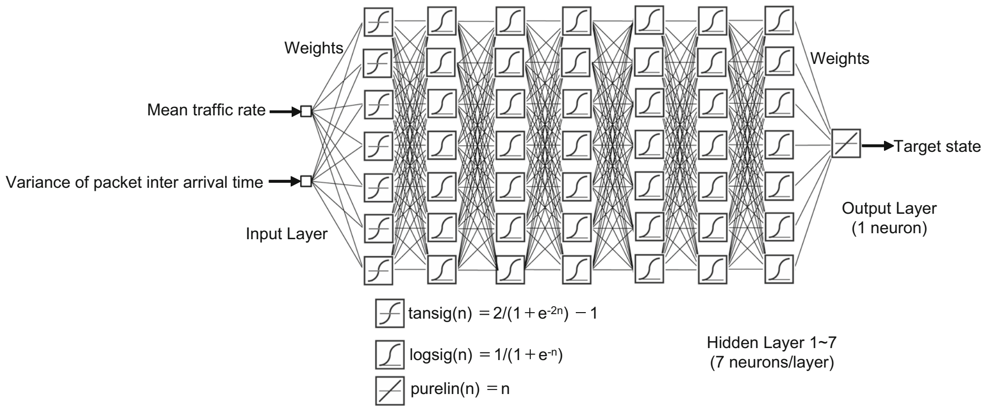

Fig. 3 Neural network architecture

has seven neurons. Each neuron in the first layer uses *tansig* as the activation function. The neurons from layer two to layer seven use the *logsig* function. The function for the final output stage is *purelin*.

The link weight system $\boldsymbol{w}$ in the neural network is obtained through a learning process. In the process, a set of training data is fed into the neural network. There is an extensive study showing that most network traffic flows can be characterized by self-similarity and long-range dependence. A reference list has been previously published [21]. A flow of long-range dependent traffic can be generated by aggregating multiple substreams following a Pareto distribution. The shape of a Pareto distribution depends on the parameter $\boldsymbol{\alpha}$. The relationship between the Hurst parameter H and the shape parameter α is $H = (3 - \alpha)/2$ [22]. The Hurst parameter is approximately 0.8 for common network traffic [23]. Traffic with larger H has larger burstiness. To obtain an accurate and realistic performance analysis, we use long-range dependent traffic as the training data. Our training data contain a wide range of traffic distributions with a variety of mean loads and variances. They cover typical traffic distributions found in real GPON networks. Each training data unit is a vector (m, b, s), where m, b, and s are the mean rate, burstiness, and the desired state $s \in 0/1$ (Full Power/Power Saving), respectively.

We use a back-propagation technique to adjust the link weights to minimize the error between the output of the neural network and the target output state. As the learning process finishes, the final link weight system is provisioned into the neural network. Usually, a larger sized neural network can help to reduce the error gaps between the desired output given by the training data and the real output of the neural network. However, larger sized neural network requires more physical resources. We have tried different sized neural networks and found that seven layers and seven neurons are small enough to provide accurate results. If the size of the neural network is reduced, the error gaps between the training data and the neural output are not acceptable.

The neural network can be installed in the ONU (or OLT) and used to control the transitions from the Full Power state to the Power Saving state. Please note that because the learning process is performed off-line and the weight system is obtained before the neural controller is equipped in the ONU (or OLT), the computation time is not an issue in the proposed neural network-based scheme. This is in contrast to directly applying the simulation-based scheme shown in Sect. 3.2.2

4 Transition control within the power saving state

In this section, we present an IEEE 802.3az-based scheme and the proposed burst transmission scheme for state transition control within the composite Power Saving state. More specifically, for an ONU operating in doze mode, our policies are used to wake up the ONU system and force it to leave the Listen state and to enter the Doze Aware state. The awoken ONU will stay in the Doze Aware state until all backlog data in the upstream direction are transmitted to the OLT or until the ONU receives a Doze Allow (OFF) command.

4.1 IEEE 802.3az

The first control scheme is directly borrowed from IEEE 802.3az. An ONU in the Listen state is activated and jumps to the Doze Aware state immediately when a packet comes into the system. When all of the packets for the upstream direction have been transmitted and the buffer becomes empty, the ONU returns to the Listen state. The major difference between GPON and Ethernet is that a GPON has to perform bandwidth allocation in the upstream direction. Therefore, some additional delay time for an ONU is inevitable.

4.2 Burst transmission

The second control scheme is inspired by the burstification process used in Optical Burst Switching (OBS) networks

Table 1 Simulation parameters

Line rate	2.5 Gbps
Frame size	1505 bytes (including 5 bytes GEM header)
Upstream burst overhead	25 bytes (including guard time 108 bits + preamble 8 bytes + delimiter 3 bytes)
Upstream burst header	3 bytes (BIP 1 byte + ONU_ID 1 byte + Ind 1 byte)
Bandwidth allocation	32 Mbps
ONU buffer threshold	1 Mbytes
Distance between ONU and OLT	20 Km
$P_{\text{warm up}}$	0.594 W
P_{data}	0.495 W
P_{idle}	0.132 W

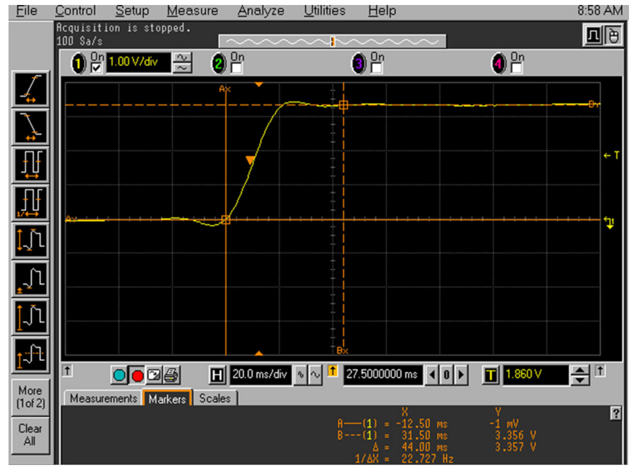

Fig. 4 Transceiver warm-up time

[24]. We have used the idea in our earlier work on energy-efficient GPONs [14]. We also find that a similar idea has been called a coalescing scheme in an energy-efficient EPON [25]. For an ONU, an upstream packet that comes into an ONU in the Listen state is stored in the ONU buffer. If the packet happens to be the first packet in the buffer, a timer is activated. The ONU remains in the Listen state until either the timer reaches the maximum Queue Head Waiting Time (QHWT) or the queue length is greater than a pre-determined Queue Length Threshold (QLT). QHWT provides the maximum value for packet waiting in the ONU. The QLT is used to restrict the amount of data stored in the ONU before it is woken. To avoid buffer overflow, the following equation has to be satisfied:

$$(\text{N-QLT})/r > t_{\text{warm up}} + \delta \qquad (6)$$

In (6), N is the buffer size, r is the maximum input rate of the UNI interface, and δ is the maximum upstream bandwidth allocation latency. Based on (6), $N - r(t_{\text{warm up}} + \delta)$ is the maximum QLT that guarantees an ONU free from packet loss.

5 Simulation results

The detailed simulation settings are summarized in Table 1. In the simulations, the optical line rate is set to 2.5 Gbps for both the upstream and downstream directions. Each packet contains an Ethernet frame and a 5-byte GEM header. The Ethernet frames are generated using the traffic generator program [26] with frame size fixed at 1500 bytes. The power consumption is obtained from the transceiver module in the GPON system developed by ITRI. We also include the warm-up timing diagram based on a real experiment in Fig. 4. It might take over 15 ms depending on the design of the driver circuit. In the simulations, the total simulation time is divided into multiple time windows. Each time window is 5 min.

To focus on the energy efficiency improvements contributed by different schemes and to eliminate the effect of dynamic bandwidth allocation (DBA), we use a fixed 32 Mbps static bandwidth assignment in the upstream direction. More specifically, the OLT constantly allocates 32 Mbps of upstream bandwidth for each ONU. Without loss of generality, we observe only one specific ONU. The load shown in all of the figures in this section is defined to be the ratio of the mean rate of the arrival traffic to the 2.5 Gbps optical line rate. In other words, the load is 0.0001 when the average arrival rate at the observed ONU is 0.25 Mbps.

The neural network shown in Fig. 3 is used in the simulations. We generate 9024 sets of training data. Each set of training data consists of the mean load of the arrival traffic, variance of packet inter arrival time, and the desired target state. The target state s is either the Full Power state or the Power Saving state.

For the training data, the mean loads of the arrival traffic are further classified into two parts. The first part ranges from 0.00005 to 0.0001, and the second part ranges from 0.0001 to 0.001. Because the loads in the first part are smaller than those in the second part, we use different granularity for the training data generation. The sample interval is increased by 0.000001 in the first part, and the interval is increased by 0.00001 in the second part. The arrival packets are obtained from the long-range dependent traffic generator reported previously [26]. By changing the Hurst parameter H, we obtain traffic with different burstiness. For each mean load, eight patterns of burstiness are generated by changing the H parameter from 0.55 to 0.9 with an interval of 0.05.

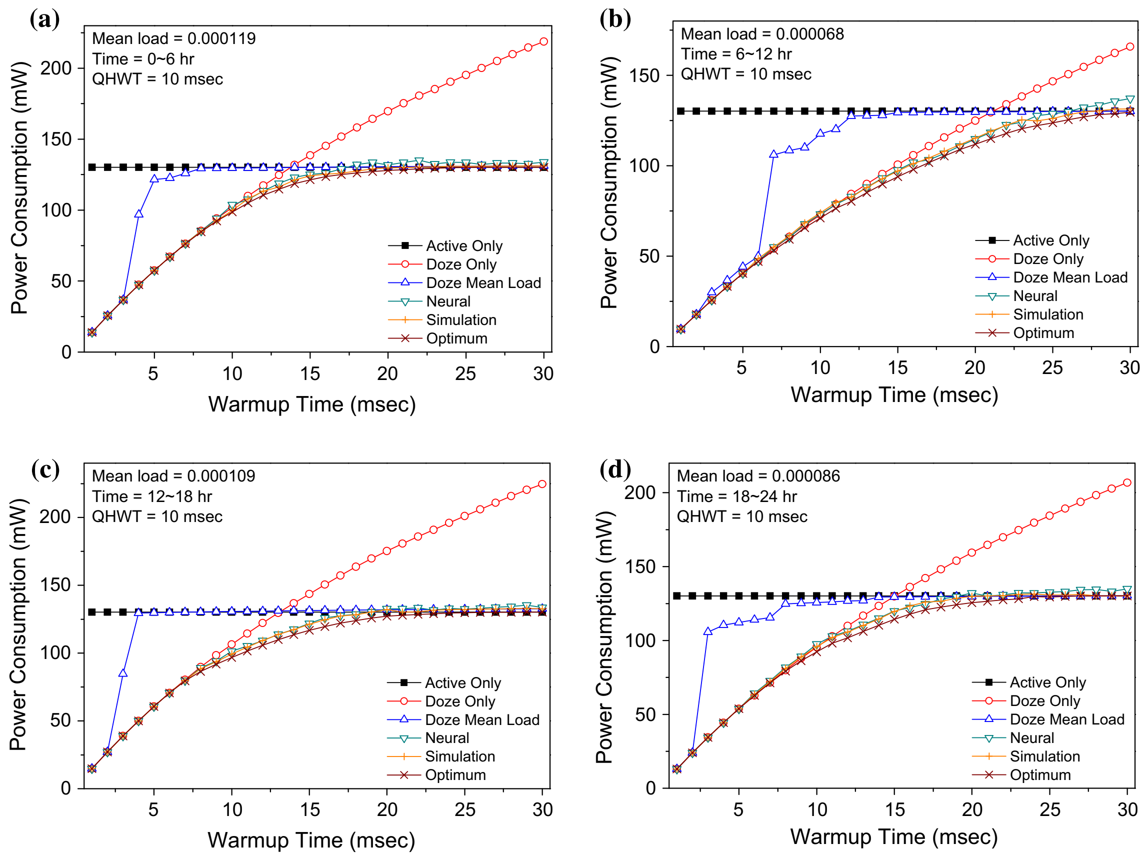

Fig. 5 Power consumption using real network traffic. **a** Midnight (0:00–6:00), **b** morning (6:00–12:00), **c** afternoon (12:00–18:00), **d** evening (18:00–24:00)

5.1 Adaptive control for state transition between full power state and power saving state

In the first set of simulations, we evaluate the performance of different adaptive control schemes for state transition control between the Full Power state and the Power Saving state. We use both real traffic and programmatically generated traffic in the simulations. The traffic in the former case is obtained from the access point of our laboratory in National Chung Cheng University during a typical day. In addition to the real traffic traces, we conducted simulations using long-range dependent and short-range dependent traffic and made performance comparisons.

The three proposed schemes (Mean Load, Simulation, and Neural) described in Sect. 3.2 are used to determine the state transitions between the Full Power state and the Power Saving state. In addition to the three control schemes, two other schemes, *Active Only* and *Doze Only*, are also considered in our simulations. In the *Active Only* and*Doze Only* schemes, the ONU remains unchanged in the Full Power state and the Power Saving state, regardless of the traffic conditions. The power consumption of the *Active Only* scheme is the same as a typical non-energy aware ONU. In contrast, the ONU using the *Doze Only* scheme only operates in the Power Saving state. Because most approaches in the literature only consider an ONU operating in the Power Saving state, we use this simulation to identify the importance of adaptive control.

In addition to taking the five schemes into consideration, we calculate the power consumption using optimal control in our simulations. Given the whole set of traffic traces, we can determine the most energy-efficient state for the ONU and its total power consumption. In practical situations, the arrival traffic is a random process, and we cannot know the exact arrival traffic for the next observation period. The optimal power consumption values are used for performance comparison purposes. They are also used to determine the target state s in the training process for our neural network.

The simulation results for the real traffic cases are shown in Fig. 5. In this set of simulations, there are four time periods in the day. Each time period has a slightly different mean load distribution. The average load is light (approximately 297.5 Kbps) even in the most busy time period. The results meet the low utilization behavior of Ethernet [3]. Observing Fig. 5, we find that the neural network-based approach gives a power consumption value very close to the optimal solution. Purely using mean load to determine the state transitions

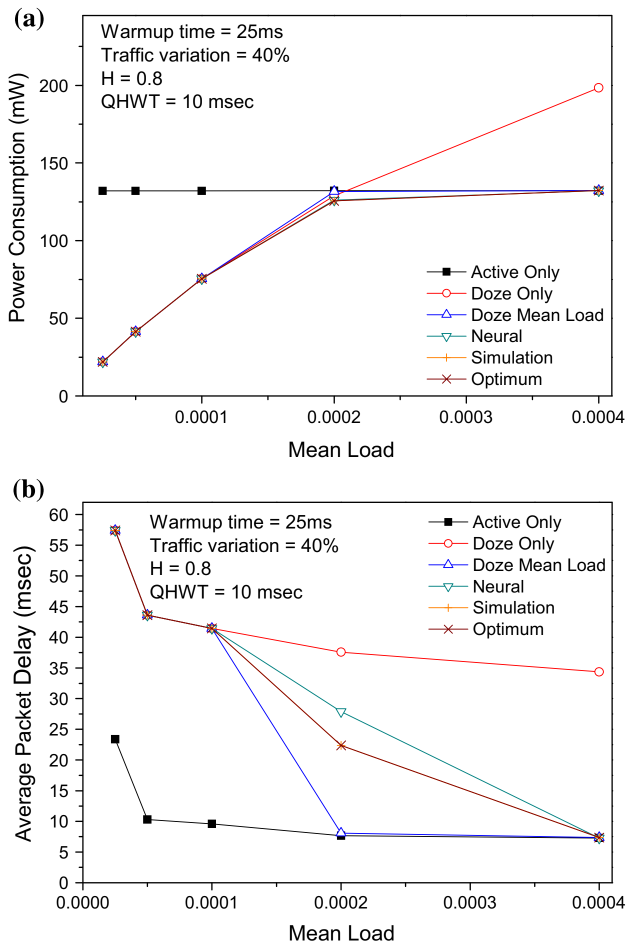

Fig. 6 Power consumption and average packet delay in different loads. **a** Total power consumption versus traffic loads, **b** average packet delay versus traffic loads

between the Full Power state and the Power Saving state has good performance for transceivers requiring large or small warm-up times. However, the performance is not satisfactory when the warm-up time is in the middle range. For an ONU with small warm-up time, *Doze Only* consumes low power. However, as the traffic load increases, keeping ONU active is a better choice, and we will see this situation later.

To evaluate the network over a wider range of traffic distributions, we generate traffic using a long-range dependent model that has been used in PON for many studies [26]. We apply traffic with $H = 0.8$ in this set of simulations. The mean load is regenerated at the beginning of each time window. Those mean loads are with 40 % variation centered at a nominate load. In other words, given load r, the actual mean load used in each time window is randomly selected from $1.4r$ to $0.6r$, following a uniform distribution. The power consumption and average delay under different loads r are shown in Fig. 6.

Figure 6a displays the power consumption under loads ranging from 0.000025 to 0.0004. When the system load is very light, the power consumption for Doze Only is very close to the optimal solution. In contrast, when the system load is heavy, an ONU operating in the Full Power state is the best choice. This means that the Active Only scheme has superior performance. However, when the load is in the middle range, keeping the ONU in the Power Saving state or in the Full Power state is not good enough. The ONU should dynamically adjust the state according to the traffic conditions.

Observing the two adaptive schemes, we find that the mean load scheme consumes more energy than the proposed neural network scheme. Although traffic load is the most important factor for power consumption, the results indicate that it is insufficient to consider only the average load. The power consumption of the neural network approach is very close to the optimal solution regardless of the system load.

Figure 6b shows the average delay under different traffic loads. It is clear that Active Only and Doze Only have the smallest and largest delays, respectively. The delay obtained using the neural network scheme is similar to the optimal solution. We also find that the delay decreases as the load increases in the neural network scheme. This is because, as the load increases, there is a higher probability that the ONU will stay in the Full Power state. Even for an ONU that is in the Power Saving state, higher loads increase the sojourn time for an ONU to stay in the Doze Aware state. Thus, the delay time decreases as the load increases.

In the next set of simulations, we evaluate the impact of traffic burstiness on system performance. Three H parameters 0.7, 0.8, and 0.9, are used. Traffic with larger H has larger burstiness. Figure 7 indicates that power consumption is strongly influenced by traffic burstiness. For two traffic patterns with the same average load, the one with smaller burstiness consumes more power because high burstiness means a greater probability of generating a longer train of arrival packets, which leads to a longer average time in the state. Thus, the power consumption is reduced because the transceiver is less frequently turned off and on. The results also show that only taking the first moment (average load) of traffic distribution into consideration is not enough. Because the proposed neural network approach takes both the first moment and second moment (variance) of the traffic into consideration, its performance closely matches the optimal results.

We further apply Poisson and Pareto distributions in our study to evaluate the impact of short-range dependent traffic and long-range dependent traffic on system performance. Figure 8a presents the results. We observe that the power consumption for the long-range dependent Pareto traffic consumes less energy compared to the short-range dependent Poisson traffic. For two traffic arrivals with the same

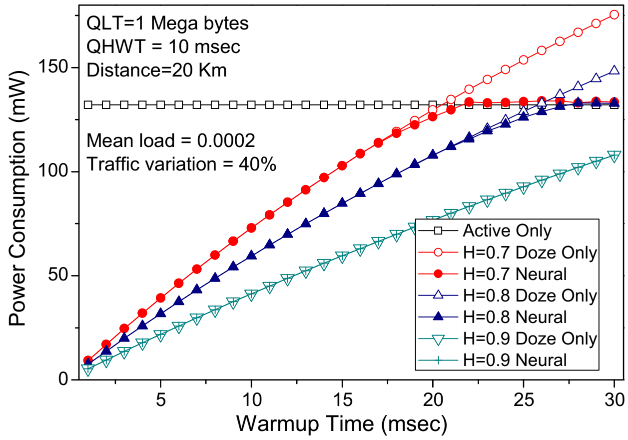

Fig. 7 Impact of traffic burstiness on power consumption

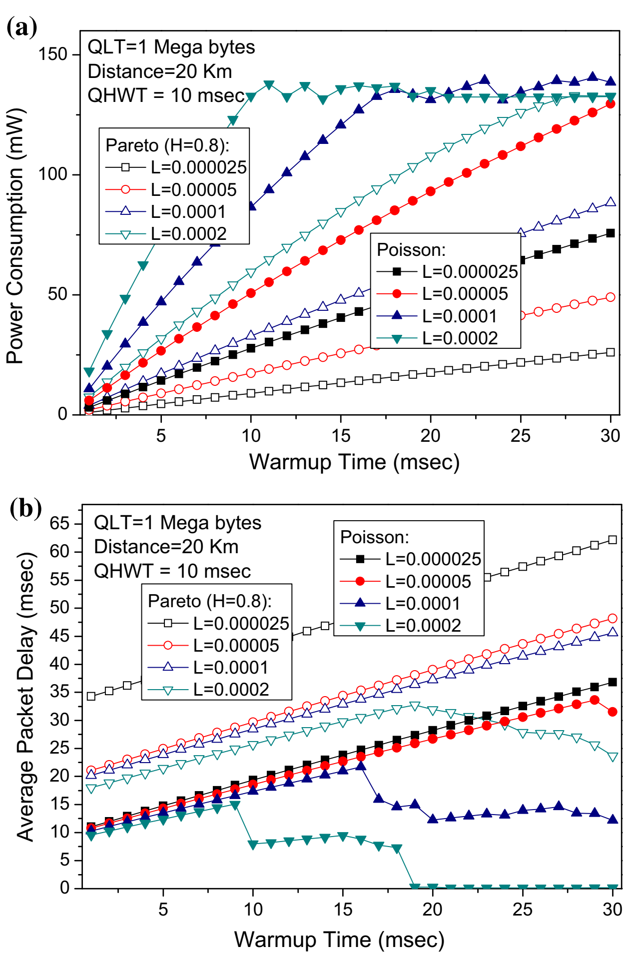

Fig. 8 Performance comparisons between long-range dependent and short-range dependent traffic. **a** Power consumption, **b** average packet delay

average arrival rate, the bursty one has fewer transitions between the Full Power and Power Saving modes. This reduces the power consumption by reducing the number of transceiver warm-ups.

Figure 8b compares the average packet delay between the Poisson and Pareto distributions. Because Pareto traffic has longer burstiness in arrival traffic, the queueing delay of Pareto traffic is larger than those with Poisson traffic. This makes the total average delay for Poisson traffic shorter than that of Pareto traffic. Please note that the curves bend down at certain points as the warm-up time increases. This is because, when the warm-up time increases, to reduce total power consumption, the ONU has a higher and higher probability of remaining in the Full Power state. This reduces the average delay.

5.2 Adaptive control for state transition inside power saving state

In the second set of simulations, we study the system performance for an ONU operating in the Power Saving state. We evaluate the IEEE 802.3az-based control schemes and make performance comparisons with the burst transmission scheme. The difference between both schemes is the timing for generating event $EV_{LS \to DA}$. Just like the IEEE 802.3az operating in Ethernet, in the simulations, the ONU always stays in the Power Saving state. It never returns to the Full Power state.

The system denoted *802.3az optimum* uses optimal control to make state transitions between the Full Power state and the Power Saving state. The optimal control is implemented by examining the input traffic in advance so an optimal decision can be made to determine whether the ONU should stay in the Full Power state or the Power Saving state. The 802.3az optimum is only used to understand the limitation of applying 802.3az in GPON. Because we cannot predict future traffic, it is impossible to achieve this optimum in reality.

In this simulation, we set QLT to 1 MByte and QHWT to 10 ms. Figure 9a shows the results for power consumption. Comparing the burst transmission scheme with the IEEE 802.3az scheme, we observed that the former has superior performance in power saving. This is because frequent transitions between the Listen state and Doze Aware state waste power in the waking up processes.

We also observe that the power consumption of 802.3az optimum approaches a constant value. That value is actually the power consumption observed when system using an Active Only scheme. The simulation results reveal that 802.3az cannot provide extra power savings over the Active Only scheme. Because 802.3az optimum is actually a lower bound on power consumption and uses optimal control, the power consumption of any other version of an 802.3az-based control scheme would consume more power than the Active Only scheme. We have examined the results carefully and realize that the extra power consumption comes from the inadequate state control when 802.3az is applied in the GPON system. When the ONU is in the Listen state, any incoming

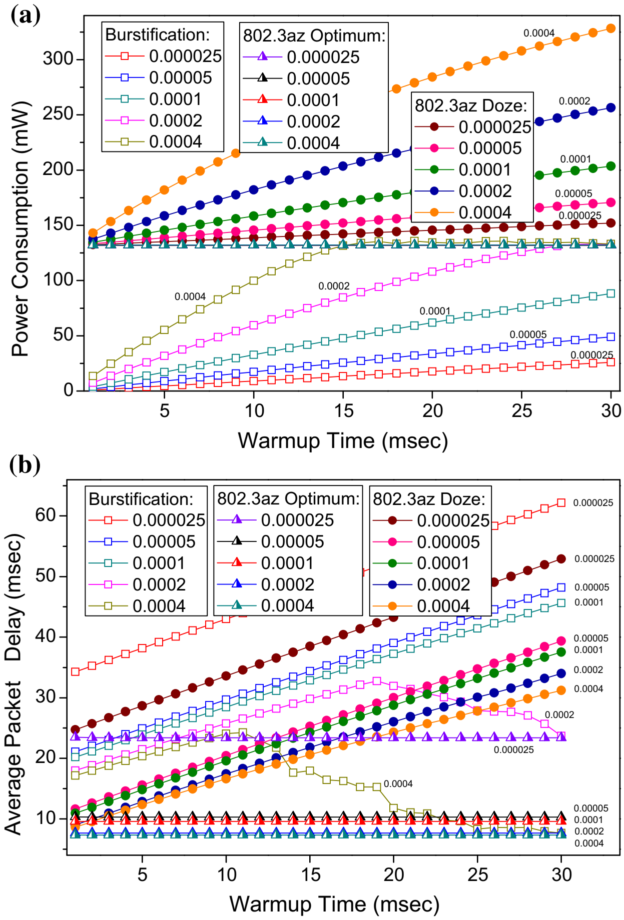

Fig. 9 Performance comparisons between the burst transmission scheme and the IEEE 802.3az-based schemes. **a** Total power consumption versus transceiver warm-up time, **b** average packet delay versus transceiver warm-up time

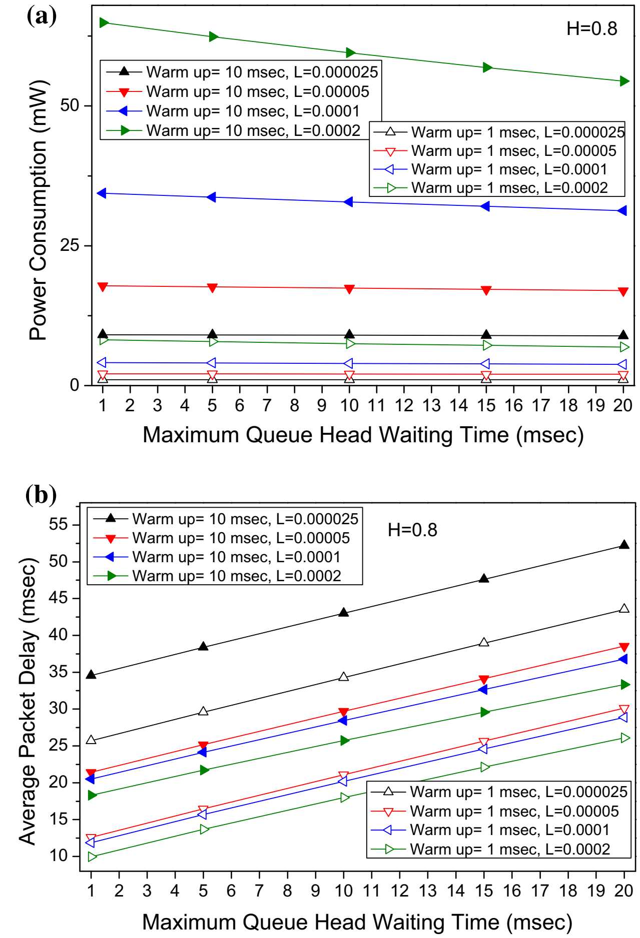

Fig. 10 Power consumption and average delay under different QHWT. **a** Total power consumption versus QHWT, **b** average delay versus QHWT

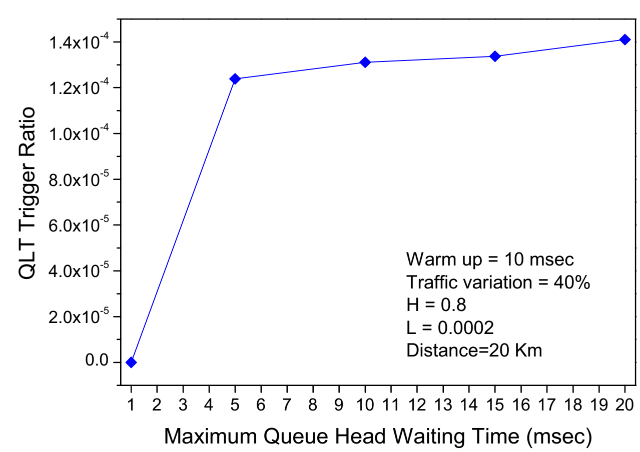

Fig. 11 Ratio of state transitions from Listen state to Doze Aware state caused by QLT

frame triggers the ONU to jump to the Doze Aware state. As the upstream buffer becomes empty, it goes back to the Listen state immediately. Such control generates a lot of unnecessary state transitions between the Listen state and the Doze Aware state. Due to the rush current effect, frequent transition results in a lot of power consumption.

The average delays for the burst transmission scheme and the IEEE 802.3az are presented in Fig. 9b. The average delay of the IEEE 802.3az increases as the load decreases. The 802.3az optimum has the lowest delay among those schemes. Because the 802.3az optimum system is operating in the Full Power state most of the time, the delay is very close to that obtained for the Active Only scheme.

To determine the effect of QHWT on system performance, we draw the curves for power consumption under different QHWT values and show the results in Fig. 10a. In this set of simulations, we use two warm-up times: 10 and 1 ms. The input traffic follows a Pareto distribution with $H = 0.8$. Because traffic with larger QHWT has a better traffic burstification effect, a larger QHWT results in greater power savings. We have observed that the power consumption decreases almost linearly as the QHWT increases. Traffic with a larger average load has a larger slope. When the load is smaller than 0.00005, the burstification effect is diminished to the point that the results become a flat line.

We also observe the impact of QHWT versus packet delay, and the simulation results are displayed in Fig. 10b. The delay in the GPON system results from a combination of QHWT, OLT, transceiver warm-up time, and queueing delay. To evaluate the impact of QLT, we record the number of state transitions from Listen to Doze Aware caused by QLT and show the results in Fig. 11. In this figure, the vertical axis is defined to be the ratio of the number of transitions caused by QLT to the total number of state transitions caused by QLT plus QHWT. The results show that, when QLT = 1 Mbyte, most triggers come from QHWT, not QLT. However, the buffer threshold is still important. It is largely used for avoiding a buffer overflow caused by sudden input bursts, not for reducing power consumption in the ONU.

Because we use long-range dependent traffic in the simulation, long trains of packets would come into the system even when the network load is light. That results in a large queuing delay. We have shown earlier in Fig. 8b that the delay for short-range dependent traffic is much smaller than for long-range dependent traffic. Although long-range dependent traffic is used, the QHWT successfully provides an upper bound for the average delay. Please note that to focus on the performance of different schemes for power saving, we use fixed upstream bandwidth allocation for each ONU in our simulation. The delay can be further reduced when a dynamic bandwidth allocation is used.

6 Conclusion

In this paper, we address energy management problems in GPON networks. Unlike most existing studies, which consider energy control only within the Power Saving state, our approach takes both the Power Saving state and the Full Power state into consideration. In this work, we also reveal that traffic distribution is another critical factor for energy consumption. Considering only the average load of input traffic is not enough. The variance of packet arrival is also important for designing green GPON systems.

We have proposed an adaptive control scheme for doze mode GPON systems. The proposed approach uses the optimal load threshold to determine the state transitions from the Power Saving state to the Full Power state. For the reverse direction, we have proposed a neural network-based scheme to determine state transitions from the Full Power state to the Power Saving state. We also propose a burst transmission scheme for controlling state transitions within the composite Power Saving state. Simulation results indicate that the proposed neural network-based adaptive control scheme can achieve power consumption that is very close to the theoretical optimum.

Acknowledgments The authors would like to thank the anonymous reviewers for their helpful and constructive comments, which greatly contributed to improving the final version of this paper. This work was supported in part by the Hsinchu Science and Industrial Park of Taiwan, under Grant 99IA06. The authors would like to thank OPNET Telecommunication Inc., Taiwan and ITRI, Taiwan for help in performing this research.

References

[1] Code of Conduct on Energy Consumption of Broadband Equipment, version 4, European Commission, Institute for Energy, Renewable Energy Unit, 10 February (2011)
[2] Gunaratne, C., Christensen, K., Nordman, B., Suen, S.: Reducing the energy consumption of ethernet with adaptive link rate (ALR). IEEE Trans. Comput. **57**(4), 448–461 (2008)
[3] Christensen, K.J., Gunaratne, C., Nordman, B., George, A.D.: The next frontier for communications networks: power management. Comput. Commun. **27**(18), 1758–1770 (2004)
[4] Lee, S.S.W., Chen, A.: Design and analysis of a novel energy efficient ethernet passive optical network. In: Proceedings of ICN'10, April (2010)
[5] Zhang, J., Ansari, N.: Toward energy-efficient 1G-EPON and 10G-EPON with sleep-aware MAC control and scheduling. IEEE Commun. Mag. **49**(2), S33–S38 (2011)
[6] Kubo, R., Kani, J., Ujikawa, H., Sakamoto, T., Fujimoto, Y., Yoshimoto, N., Hadama, H.: Study and demonstration of sleep and adaptive link rate control mechanisms for energy efficient 10G-EPON. IEEE OSA J. Opt. Commun. Netw. **2**(9), 716–728 (2010)
[7] Wong, S.-W., Valcarenghi, L., Yen, S.-H., Campelo, D.R., Yamashita, S., Kazovsky, L.: Sleep mode for energy saving PONs: advantages and drawbacks. In: Proceedings of the IEEE GreenCom (2009)
[8] Dhaini, A.R., Ho, P.-H., Shen, G.: Toward green next-generation passive optical networks. IEEE Commun. Mag. **29**(11), 94–101 (2011)
[9] Dhaini, A.R., Ho, P.-H., Shen, G., Shihada, B.: Energy efficiency in TDMA-based next-generation passive optical access networks. IEEE/ACM Trans. Netw. **22**(3), 850–863 (2014)
[10] Kani, J.: Power saving techniques and mechanisms for optical access networks systems. J. Lightwave Technol. **31**(4), 563–570 (2013)
[11] ITU-T G.987.3: 10-Gigabit-capable passive optical networks (XG-PON): transmission convergence (TC) layer specification, January (2014)
[12] ITU-T G.sup45: GPON power conservation, ITU-T G-series recommendations-supplement 45, May (2009)
[13] Skubic, B., Hood, D.: Evaluation of ONU power saving modes for gigabit-capable passive optical networks. IEEE Netw. **25**(2), 20–24 (2011)
[14] Lee, S.S.W., Chen, A.: Adaptive control and performance evaluation of burst transmission in energy efficient GPON. In: Proceedings of IEEE ISCC (2012)
[15] Newaz, S.H.S., Cuevas, A., Lee, G.M., Crespi, N., Choi, J.K.: Evaluating energy efficiency of onus having multiple power levels in TDM-PONs. IEEE Commun. Lett. **17**(6), 1248–1251 (2013)
[16] Laughton, M.A., Say, M.G.: Electrical Engineer's Reference Book. Butterworth, London (1990)
[17] Peng, J.-W., Wu, H.-Y., Chang, C.-J.: Dynamic bandwidth allocation method of ethernet passive optical network, US patent, US20100254707 A1

[18] Media Access Control Parameters, Physical Layers, and Management Parameters for Energy-Efficient Ethernet, IEEE Standard 802.3az-2010
[19] ITU-T G.984.3: Gigabit-capable Passive Optical Networks (G-PON): Transmission convergence layer specification, January (2014)
[20] ITU-T G.984.4 ONT management and control interface (OMCI) specification
[21] Willinger, W., Taqqu, M., Erramilli, A.: A bibliographical guide to self-similar traffic and performance modeling for modern high-speed networks. In: Kelly, F.P., Zachary, S., Ziedins, I. (eds.) Stochastic Networks, pp. 339–366. Oxford University Press, Oxford (1996)
[22] Willinger, W., Taqqu, M., Sherman, R., Wilson, D.: Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level. In: Proceedings of the ACM SIGCOMM, pp. 100–113, August (1995)
[23] Leland, W., Taqqu, M., Willinger, W., Wilson, D.: On the self-similar nature of ethernet traffic. IEEE/ACM Trans. Netw. **2**(1), 1–15 (1994)
[24] Battestilli, T., Perros, H.: An introduction to optical burst switching. IEEE Commun. Mag. **41**(8), s10–s15 (2003)
[25] Rodriguez-Perez, M., Herreria-Alonso, S., Fernandez-Veiga, M., Lopez-Garcia, C.: Improving energy efficiency in upstream EPON channels by packet coalescing. IEEE Trans. Commun. **60**(4), 929–932 (2012)
[26] Kramer, G.: Synthetic self-similar traffic generation. http://glenkramer.com/trf_research.shtml

Steven S. W. Lee received his Ph.D. degree in electrical engineering from National Chung Cheng University, Taiwan, in 1999. From 1999 to 2008, he was with Industrial Technology Research Institute (ITRI), Taiwan, where he was the leader of Intelligent Optical Networking project and a section manager of Optical Communications Department. He was also a Research Associate Professor of National Chiao Tung University, Taiwan, during 2004–2007. In 2008, he joined National Chung Cheng University, Taiwan, where he is currently an Associate Professor of the Department of Communications Engineering. Dr. Lee has published over 70 papers in international journals and conferences in the areas of broadband communications. His research interests include optical and broadband networking, network optimization, software defined networking, and green network design.

Kuang-Yi Li received his B.S. degree in information management from Saint John's University, Taiwan, in 2010. He is now working towards his Ph.D. degree in communications engineering at National Chung Cheng University, Taiwan. His research interests include network optimization, software defined networking, and green network design.