Service Class Resource Management for Green Wireless-Optical Broadband Access Networks (WOBAN)

Maha Ahmed, *Student Member, IEEE*, Iftekhar Ahmad, *Member, IEEE*, and Daryoush Habibi, *Senior Member, IEEE*

***Abstract*—Exponential growth in the volume of wireless data, boosted by the growing popularity of mobile devices, such as smartphone and tablets, has forced the telecommunication industries to rethink the way networks are currently designed, and to focus on the development of high-capacity mobile broadband networks. In response to this challenge, researchers have been working toward the development of an integrated wireless optical broadband access network. Two major candidate technologies, which are currently known for their high capacity as well as quality of service (QoS) for multimedia traffic, are passive optical networks (PON), and fourth generation (4G) wireless networks. PON is a wired access technology, well known for its cost efficiency and high capacity; whereas 4G is a wireless broadband access technology, which has achieved broad market acceptance because of its ease of deployment, ability to offer mobility, and its cost efficiency. Integration of PON and 4G technologies in the form of wireless-optical broadband access networks, offers advantages, such as extension of networks in rural areas, support for mobile broadband services, and quick deployment of broadband networks. These two technologies however, have different design architectures for handling broadband services that require quality of service. For example, 4G networks use traffic classification for supporting different QoS demands, whereas the PON architecture has no such mechanism to differentiate between types of traffic. These two technologies also differ in their power saving mechanisms. In this paper, we propose a service class mapping for the integrated PON-4G network, which is based on the M/G/1 queuing model. We also propose a class-based power saving mechanism which significantly improves the sleep period for the integrated optical wireless unit, without compromising support for QoS. Results indicate that our proposed class-based power saving scheme reduces power consumption by up to 80%, and maintains the QoS within the requirements of the service level agreement.**

***Index Terms*—Power saving, quality of service, service class, wireless-optical broadband access networks.**

I. Introduction

BROADBAND access networks have become an essential part of worldwide communication systems because of the exponential growth of broadband services such as video on demand, high definition TV, internet protocol TV and video-conferencing.

These bandwidth-intensive services and their demand for QoS are the key reasons for the declining popularity of back-haul technologies such as T/E-carrier and digital subscriber line [1], and the emergence of optical networks and wireless broadband access networks. Optical networks are known for their high capacity and reliability, and are widely used in backbone communication networks [2]. In recent years, increasing demand for higher data rates in the access network, has motivated telecommunication engineers to focus on developing an access network version of optical networks. One such version is the passive optical networks (PON) which is a highly efficient technology with low cost and power consumption attributes. The major components of PON are the optical line terminal (OLT) which resides at the central office, optical network unit (ONU) which resides at or near the end user premises and a splitter which connects ONUs to the OLT. The OLT controls the PON and connects it to the backbone network.

Another popular broadband access technology is the fourth generation (4G) wireless technology, which is robust, cost-efficient, highly scalable, and supports mobility and multimedia traffic. 4G technology sustains a high downstream data rate up to 1 Gb/s and 500 Mb/s upstream for a fixed connection. Currently two variants of 4G technology are available; Worldwide Interoperability for Microwave Access (WiMAX), based on the IEEE 802.16 m standard, and long-term evolution advanced (LTE-A).

Wireless-Optical Broadband Access Networks (WOBAN) integrate PON and wireless networks, to cost effectively deliver broadband services to the end users. This integration relies on 10 Giga-Ethernet PON (10 G-EPON) as the back-haul for the 4G network and offers benefits such us extended coverage, reduced service delivery cost and mobility support [3], [4]. WOBAN however, faces numerous challenges including maintaining QoS in the integrated network, caused as a result of the differences in QoS interpretation in the constituent technologies. 4G differentiates the services in several specific classes, for example, five in WiMAX whereas LTE-A defines eight QoS Class Identifiers (QCI) [5]. In contrast, PON does not specify classes, but it can support multiple queues for traffic differentiation [6]. As the traffic moves between wireless and the optical interface in WOBAN, service class mapping at the interface becomes a major challenge. In this paper, we propose a solution for the service class mapping problem, between 4G classes and the 10G-EPON queues, based on an M/G/1 queuing model. We investigate the queue behavior of these classes and develop a model to calculate the required length for different queues which ensures that QoS for various end applications is not compromised. To our best knowledge, this is the first work of its kind that addresses the

Manuscript received April 7, 2014; revised July 23, 2014 and September 5, 2014; accepted October 20, 2014. Date of publication October 26, 2014; date of current version December 16, 2014.

The authors are with the School of Engineering, Edith Cowan University, Joondalup, WA 6027, Australia (e-mail: mahaa@our.ecu.edu.au; i.ahmad@ecu.edu.au; d.habibi@ecu.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JLT.2014.2365254

service class mapping, scheduling and queue modeling problems at the ONU in WOBAN.

As the Internet continues to expand, its power demand also escalates, e.g., the annual increase in energy consumption by the telecommunication networks grew from 1.3% to 1.8% of world consumption between 2007 and 2012 [7]. Recent studies suggest that the Information and Communications Technology (ICT) sector consumes 8% of world power consumption and unless power saving mechanisms are implemented, this may rise to 14% by 2020 [8]. About half this power is consumed by devices that are powered ON, but unused [9]. Despite PON being energy efficient [10], the evolution from EPON to 10G-EPON to increase the data rate ten fold still increases power consumption. In WOBAN, reducing power consumption is a research topic of major significance. In the 4G interface within the WOBAN architecture, sleep mode has been introduced to reduce power consumption [11]. In the PON part, the international telecommunication union has introduced sleep and dose modes in Gigabit PON (GPON) [12]. In sleep mode, the ONU's transceiver is turned OFF for substantial periods while in dose mode, only the transmitter is turned OFF when possible. Different mechanisms have been proposed to apply sleep mode in PON technology [13]–[17]. No study, however, was identified in the literature which considered the QoS classes in sleep management in standalone PON or WOBAN. In this paper, we extend the concept of service class mapping and propose a class based dynamic bandwidth allocation scheme for power savings with a specific sleep period for each class in WOBAN. The proposed scheme significantly increases the sleep period and reduces the overhead between the OLT and ONU. Moreover, it maintains the QoS by considering the tolerated delay of each type of traffic when calculating the sleep cycle.

The rest of this paper is organized as follows. In Section II, we overview the integration between the PON and wireless networks, class mapping and the relevant common techniques for saving power. In Section III, we characterize WOBAN architecture. In Section IV, we propose the QoS mapping and queue management scheme. In Section V, we describe our proposed power saving scheme. In Section VI, we present the performance evaluation that emphasizes the merits of our solution for deciding the initiation and duration of the relevant sleep cycle. In Section VII, we summarize these innovations.

II. Background and Related Works

A. Traffic Management at the ONU

In WOBAN (see Fig. 1), the ONU architecture includes single or multiple queues, which accommodate the traffic until it is dispatched. Major challenges exist with QoS mapping between different classes of traffic and these ONU queues. Researchers have attempted to address this problem from different viewpoints to arbitrate the appropriate number of queues in ONU. In [6], Gangxiang *et al.* suggested eight priority queues at the ONU to accommodate traffic from the BS. They however, did not propose any model for calculating the appropriate length of each queue and scheduling packets from these queues. In [17], Shi *et al.* developed a model for accommodating two services, real-time and best effort traffic, to control sleep periods for the ONU. They used one preemptive queue for both of the services. Another study [1] proposed a centralized scheduling mechanism at the OLT, to reduce the end-to-end delay, based on the Hybrid ONU-BS architecture. While this included QoS mapping between eight priority queues and WiMAX traffic, they did not specify any class mapping or queue modeling. For bandwidth allocation, Yang *et al.* in [3] suggested a virtual ONU-BS unit to achieve QoS-aware dynamic bandwidth allocation. Their proposed model involves more than one priority queue at ONU. Although their study considered the signaling of bandwidth allocation, it did not specify the queues apart from highlighting the bandwidth requested for each queue. In [18], Assi *et al.* used multiple priority queues for developing an efficient bandwidth allocation scheme, which redistributed excess bandwidth with an arbitrary large queue size (10 MB), but only considered the EPON QoS.

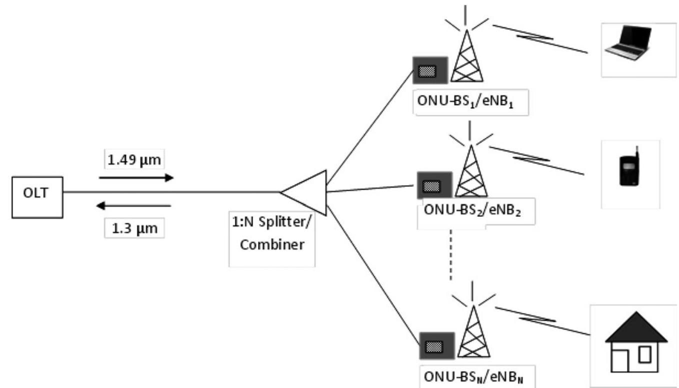

Fig. 1. Architecture of integrated network of 10G-PON and 4G (WOBAN).

Since queue length affects the integrated network's performance, through queuing delay and the packet dropping rate, queue management for ONU-BS complements QoS mapping. Despite the importance of the queue management at ONU-BS on the QoS, only few studies have addressed this issue. In [19], Kramer *et al.* investigated queue behavior and size only for the EPON network and did not consider the different classes of traffic and the QoS. Obele *et al.* in [20] studied queue behavior in a WiMAX-NGPON network by deploying one-to-one mapping of only four services between the two technologies. Their complex approach solves the embedded Markov chain when estimating queue waiting time for each service.

B. Traffic Management and Power Savings

While broadband services demand high data rates and QoS from WOBAN, the increased power consumption becomes a major challenge. Accordingly, Wong *et al.* proposed two ONU architectures to support sleep mode and minimize the overhead of clock recovery when ONU switches from sleep mode to active mode [21]. In [13], Zhang *et al.* proposed four levels of power saving and the transition between these power states depended on traffic conditions. In [14], the authors proposed an adaptive sleep duration based on measuring the traffic situation periodically. In their study, the negotiable sleep time is implemented through a three way handshake mechanism between OLT and ONU. In [16], the author proposed a scheme that computes sleep time based on the negotiation between OLT and ONU. In [22], Dhaini *et al.* considered a green bandwidth allocation scheme which involves enabling fast and deep sleep of the ONU and dynamic bandwidth allocation based on an upstream-centric strategy.

In [13], [14], authors advocated basing decisions, for transition to sleep time and the duration of sleep time, on traffic conditions. In [13], they proposed switching to sleep state based on the absence of traffic and repetitive wake up after each sleep duration to check for arriving traffic. That study initially made the sleep duration equal to the duration of no traffic, then there was a controlled increase in duration up to 50 ms. However, that study did not consider the traffic classes. In contrast, a sleep cycle was proposed in [14], based on classes of traffic for the downstream transmission, in which the arrival of high priority traffic wakes up the ONU. These researchers sought additional power savings by considering the concepts of sleep mode and dose mode from different viewpoints. Overall, there are challenges associated with the sleep mode such as when to initiative the sleep time cycle, how long should the sleep time duration be, what is the effect of the recovery time for transition from sleep to wake up, and how should the registration of the ONU be maintained.

Although researchers have attempted to address QoS in PON, none of these studies addresses both service class mapping and queue modeling at the same time at the ONU in WOBAN. There is also no available model that deals with the power savings issue and QoS at the ONU. In this work, we have addressed the above mentioned problems and the proposed solutions are described in Section IV and V.

III. Preliminaries

A. Architecture of an Integrated 10G-PON and WiMAX Network

Fig. 1 shows the architecture of an integrated 10G-PON and 4G system. The integrated unit is the ONU-BS which is comprised of the 10G-EPON's ONU and the BS that represents the base station/e-NodeB (BS/eNB) from the 4G system. The 10G-EPON system is a point-to-multipoint system in which the operation, administration and maintenance function is based on the Multi-Point Control Protocol (MPCP). This is based on five 64-byte messages that govern two processes namely: auto discovery and upstream bandwidth allocation. The former allows a new ONU to join the system while the later process is responsible for allocating bandwidth to upstream traffic. 10G-EPON is a full duplex system that supports two channels (for upstream and downstream transmission) with different wavelengths. The downstream transmission (from OLT to ONUs) is broadcast to all ONUs while the upstream transmission (from ONUs to OLT) is shared between the ONUs. The OLT uses a poll and grant mechanism based on MPCP messages to control access for ONUs to the upstream channels. The access mechanism is based on TDMA; the OLT during each polling cycle, distributes access time to each ONU. The ONU requests bandwidth from an OLT for delivering the packets, waiting in its queues, by sending the required bandwidth in the REPORT message, while the OLT assigns the bandwidth to the ONU by sending the GATE message which contains the grant length and its starting time for each ONU. The bandwidth allocation scheduling at the OLT is periodic and can be classified as fixed bandwidth allocation or dynamic bandwidth allocation. In fixed bandwidth allocation, the assigned bandwidth is fixed for all ONUs while the assigned bandwidth is variable in dynamic bandwidth allocation depending on the requested bandwidth from the ONUs. One limitation is that the fixed allocation scheme can be under or over-utilized for light and heavy ONU load, respectively. Therefore, the dynamic allocation can allocate bandwidth more efficiently for each ONU.

The wireless part of the WOBAN architecture can be WiMAX or LTE-A. The IEEE 802.16 m is a wireless standard for a metropolitan area network and is known as WiMAX. The cell architecture of the WiMAX consists of one central base station and several subscriber stations (SS). Each SS requests bandwidth and the BS broadcasts the awarded bandwidth. The media access control layer of the WiMAX, which is responsible for distributing resources among the SSs, is classified as connection-oriented where every connection is mapped to one of five predefined QoS parameter service flows; Unsolicited Grant Service (UGS), real-time Polling Service (rtPS), extended real-time Polling Service (ertPS), non-real-time Polling Service (nrtPS), and best effort (BE).

The LTE-A is the other 4G technology and the main components of LTE-A networks are evolved-Node B (eNB) and access gateway. An eNB manages the radio resource allocation and controls the LTE-A air interface. For providing quality of service, LTE-A defines eight different types of services with the highest priority being real-time gaming while TCP based services are given the lowest priority e.g. email which is BE [5].

In the WOBAN architecture depicted in Fig. 1, subscribers are connected to the broadband network through BSs which are connected to the optical network through ONUs. As such, the way an ONU treats incoming and outgoing packets, has a significant impact on the end user's quality of service. This treatment of packets also influences power savings at ONUs. In this paper, we address these two challenges by first proposing a class based queue management scheme for ONUs and then extending this concept for increasing the sleep period at each ONU.

IV. Proposed QoS Management Scheme

As discussed in Section II, ONU and BS adopt dissimilar strategies for handling services that require QoS. For merging the gap in QoS handling, we propose an architecture for the ONU-BS unit as shown in Fig. 2. The ONU-BS unit consists of a QoS mapper unit that maps traffic classes defined in the 4G standard to the appropriate class in the ONU unit. We also model the queue, corresponding to various classes in the ONU unit, and show how to calculate appropriate queue lengths.

A. QoS Mapping

The queuing system in our ONU-BS architecture is modelled as an M/G/1 system, where the arrival of packets from

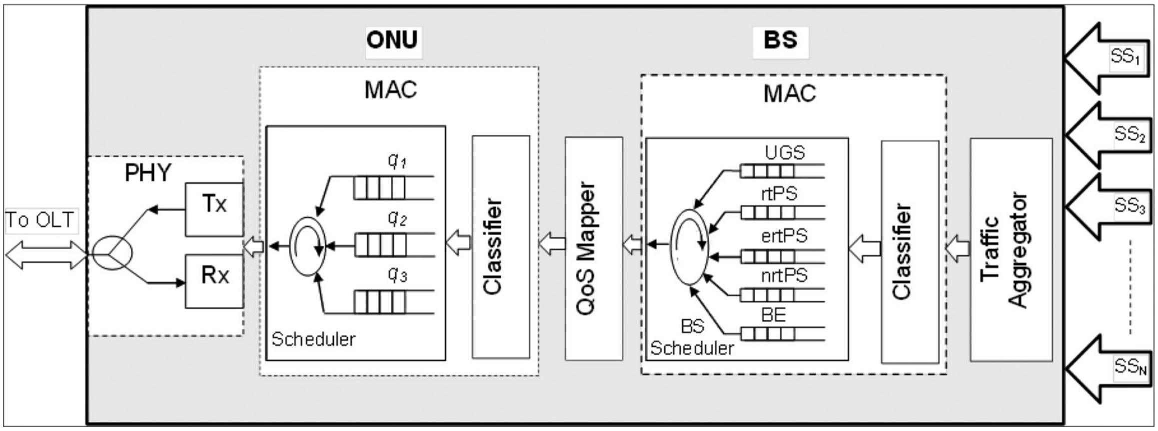

Fig. 2. Proposed ONU-base station (BS) architecture.

subscribers is considered as a Poisson random variable with a mean arrival time λ. As different packet lengths require different serving times, the distribution of the serving time is considered as general. The ONU scheduler is responsible for allocating serving time to different traffic types depending on their priority. Our queue model analysis is based on classifying the incoming traffic from the 4G base station into three priority ONU queues namely q_1, q_2, and q_3 in descending order of priority:

1) q_1 is the queue for *class* 1 packets; these are urgent and delay sensitive, such as Voice over IP (VoIP). Urgent and delay sensitive applications belong to the QCI1 and QCI3 traffic types in the LTE-A, and UGS in the WiMAX 4G standards.
2) q_2 is the queue for *class* 2 packets; these are more delay tolerant, but have a maximum delay limit, such as video streaming. LTE-A defines QCI2, QCI5 and QCI7 traffic types and the WiMAX defines ertPS and rtPS for these types and traffic.
3) q_3 is the queue for *class* 3 packets; these are delay tolerant and do not have a delay limit, such as e-mail, File Transfer Protocol (FTP), chat and web browsing. LTE-A defines QCI4, QCI6 and QCI8 and the WiMAX defines nrtPS and BE traffic types for such services.

In our system, we use three priority queues at the ONU, because additional queues increase the complexity of the scheduling system and power consumption, while fewer queues would harm service differentiation. Every single queue uses a drop-tail discipline and the serving of packets within the same queue operates on a first-come-first-served basis. The scheduling policy at the ONU is non-preemptive in the sense that when packets from a lower priority queue are being served, an arrival of a packet to the higher priority queue does not affect the serving of lower priority packets. For example, during the transmission window at an ONU which is determined by the OLT, the scheduler at the ONU serves the q_1 packets first until it is emptied, then it moves to q_2 and finally, it serves q_3. In our implementation, we used a non-preemptive policy which means that when the scheduler at the ONU has served all q_1 packets and moved to q_2, arrival of new packets at q_1 does not interrupt the serving of packets at q_2. Information about these new packets at q_1 will be included in the REPORT that will be sent to the OLT at the end of the current transmission window. The REPORT will help the OLT to decide the length of the next transmission window for the ONU. This management strategy makes sure that packets in q_3 are not heavily dominated by q_1 and q_2 packets and thereby, creates a scenario where low priority traffic does not suffer from resource starvation.

B. Queue Management at the ONU Unit

The QoS mapping unit is responsible for mapping the packets delivered by the BS unit to the appropriate queue at the ONU unit and vice versa. The next challenge is to compute the optimum length for each queue so that both queuing delay and packet dropping rates, for packets in each queue, are minimized. From a design perspective, a queue with longer than optimum length minimizes the packet dropping rates at the cost of increased delays. This also increases power consumption because of the buffer's demand for more powered hardware [23]. A queue with shorter than optimum length reduces delay at the cost of an increased packet dropping rate. In the next section, we address this issue by first analysing the expected waiting time of packets in each queue, for a given scheduler, and then show how to calculate the optimum queue length for each queue in the proposed M/G/1 model.

1) Expected Waiting Time in Relation to a Packet's Priority at the ONU Unit: An incoming packet at the ONU experiences a delay which depends on the traffic class the packet belongs, and load. The general formula that expresses such delays can be modelled after the Pollaczek–Khinchin formula for the expected waiting time of non-preemptive priority queues and can be given as:

$$E\left(W^k\right) = \frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_k \times E\left(S_k^2\right)}{\left(1-\rho_{k-1}^-\right)\left(1-\rho_k^-\right)} \quad (1)$$

where $E(W^k)$ denotes the expected waiting time of the arriving packets in their corresponding queue k; λ_k denotes the Poisson arrival rate of the packets in the kth queue; ρ_k denotes the traffic intensity of class k, the stability condition of which is $(\sum_{k=1}^{K} \rho_k < 1)$; related to this, $\rho_k^- = \sum_{i=1}^{k} \rho_i$; k denotes the traffic class indexed by $1, 2, 3$; S_k denotes the class k service times with finite expected or average service time $(E(S_k))$ and expected second moment service time $(E(S_k^2))$.

This waiting time is based on the system's status at the instant the packet departs from an M/G/1 queue. The waiting time of a newly arrived packet depends on the class of the packet and the number of packets for other classes waiting in their queues. The expected waiting time of the packets at q_1 is:

$$E\left(W^1\right) = \frac{\frac{1}{2}\lambda_1 \times E\left(S_1^2\right)}{(1-\rho_1)}. \tag{2}$$

Eq. (2) shows that the waiting time of the arrival packets that belong to the highest priority traffic class (*class* 1) depends on the arrival rate of *class* 1 traffic and its expected serving rate. The waiting time of a newly arrived *class* 1 packet thus depends on the number of packets already in the q_1.

The expected waiting time for *class* 2 packets is expressed as:

$$E\left(W^2\right) = \frac{\frac{1}{2}\left(\lambda_1 \times E\left(S_1^2\right) + \lambda_2 \times E\left(S_2^2\right)\right)}{(1-\rho_1)(1-\rho_1-\rho_2)}. \tag{3}$$

The waiting time of newly arrived packets in q_2 depends on the number of packets waiting in q_1 plus the number of packets already entered into q_2. For the lowest priority packets (*class* 3), any newly arrived packet sees all the packets waiting for being served in the q_1, q_2, and q_3. Therefore, the expected waiting time for *class* 3 packets is longer than *class* 1 and 2, and can be given as:

$$E(W^3) = \frac{\frac{1}{2}\left(\lambda_1 \times E(S_1^2) + \lambda_2 \times E(S_2^2) + \lambda_3 \times E(S_3^2)\right)}{(1-\rho_1-\rho_2)(1-\rho_1-\rho_2-\rho_3)}. \tag{4}$$

To calculate the $E(W^k)$, $E(S_k^2)$ must be calculated as:

$$E(S_k^2) = x_k^2 + \sigma^2 \tag{5}$$

where x_k denotes the service time for the packets in q_k and is equal to packet size ×8/bandwidth; σ^2denotes the variance of service time and can be adjusted to better reflect the system by using

$$\sigma^2 = \sum_{k=1}^{3} \alpha_k \times (x_k - T_{\text{avg}})^2 + \beta \times (T_{\text{gap}} - T_{\text{avg}})^2. \tag{6}$$

In Eq. (6), we assume that there is a dummy packet with serving time equal to T_{gap}, the time period between two successive grant periods (see Fig. 3) and this T_{gap}, for any ONU in a system consisting of *N* ONUs, denotes the time of serving the other (*N*-1) ONUs and can be calculated as follows:

$$T_{\text{gap}} = T_{\text{PON}} - T_{\text{grant}} \tag{7}$$

$$T_{\text{gap}} = (N-1) \times T_{\text{grant}} + N \times T_g \tag{8}$$

where T_g is the guard time between grant periods to two successive ONUs, and T_{grant} and T_{PON} denote the allocated time for ONU to transmit and polling cycle time, respectively, and the relationship between them is as follows:

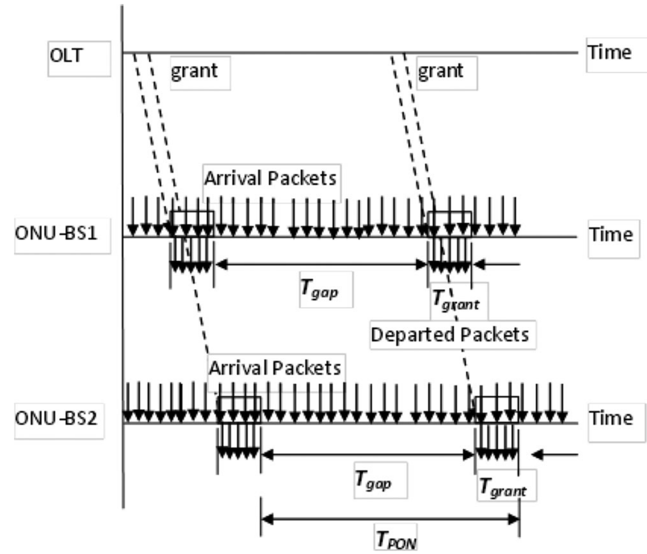

Fig. 3. The time allocation (grant) of two ONU-BSs in TDM-PON.

$$T_{\text{PON}} = \sum_{n=1}^{N} \left[T_{\text{grant}}^{n} + T_g\right]. \tag{9}$$

Here n denotes the nth ONU.

In Eq. (6), T_{avg} denotes the average value of the service time for all types of traffic and it is calculated using Eq. (10):

$$T_{\text{avg}} = \left(\sum_{k=1}^{3} x_k + T_{\text{gap}}\right)/4. \tag{10}$$

In Eq. (6), α_k and β are parameters that are linked to the T_{grant} and T_{PON}, and to the normalized assigned bandwidth (p_k) for queue k (q_k) in the following way:

$$\alpha_k = p_k \times T_{\text{grant}}/T_{\text{PON}} \tag{11}$$

$$\beta = 1 - T_{\text{grant}}/T_{\text{PON}} \tag{12}$$

where p_k denotes the normalised assigned bandwidth to different queues $(\sum_{k=1}^{3} p_k = 1)$;

From Eq. (9), we can see that the length of the polling cycle T_{PON} is based on the granted time for each ONU and the number of ONUs. In dynamic bandwidth allocation, the duration of T_{grant} varies with the requested bandwidth from the ONU. Consequently, the duration of the polling cycle also becomes variable. However, in fixed bandwidth allocation because the granted transmission time for each ONU is fixed, the polling cycle time is fixed and can be calculated as follows:

$$T_{\text{PON}} = N \times (T_{\text{grant}} + T_g). \tag{13}$$

2) Expected Waiting Time Due to the Scheduler Grant: The bandwidth allocation and scheduling mechanisms of the ONUs are centrally controlled by the OLT. Fig. 3 shows that the ONU is only allowed to transmit for the duration of T_{grant} which reflects the allocated bandwidth for the ONU. Because the scheduler at ONU-BS cannot schedule without grant, packets which arrive during the absence of the grant (T_{gap}) stay in their corresponding queues until the grant occurs. This waiting time for grant in the TDM system is additional to the priority-related waiting time. Therefore, the expected total transit time in any queue can be calculated as follows:

$$E(T^k) = E(W^k) + c \times T_{\text{gap}} \tag{14}$$

where $E(T^k)$ denotes the total expected transit time for the packets in kth queue; c denotes a constant ≤ 0.5, which is linked to the differences in waiting time in relation to grant time that is experienced by packets due to differences in arrival time. The average waiting for the grant (T_{grant}) for all the packets is approximately half of T_{gap} (hence 0.5).

3) Queue Length Estimation Based on Expected Waiting Time: According to Little's Law, the ONU-BS unit's queue length is equivalent to the average delay of the packets and can be written as:

$$L_q^k = \lambda_k \times E(T^k) \tag{15}$$

where L_q^k denotes the length of the kth queue; λ_k denotes the mean arrival rate of packets.

By analysing the Eq. (14) and adjusting these equations to our system, we propose:

$$L_q^k = \lambda_k \times \left(\frac{\sum_{i=1}^{k} \lambda_i \times \left(x_i^2 + \sigma^2\right)}{2\left(1 - \rho_{k-1}^-\right)\left(1 - \rho_k^-\right)} + c \times T_{\text{gap}} \right) \tag{16}$$

where ρ_k denotes the utilization intensity and $\rho_k = \lambda_k \times x_k$. In Sections III and IV, we showed how service class mapping can be achieved in an integrated PON-4G network. In the next section, we extend the service class mapping model for developing a power saving mechanism for the ONU-BS unit.

V. Proposed Class-Based Power Saving Model

The class-based power saving model (CBPS) model is based on a DBA scheme which increases the sleep duration of ONUs. This model supports both dose and sleep modes at ONUs. The model uses batch-mode transmission in which packets are buffered at the ONU and OLT for upstream and downstream traffic, respectively. Upstream and downstream traffic are dispatched when the transmitter and the receiver wake up, respectively. In the following sections, we discuss the power states of the ONUs during the operation of the DBA, the sleep time duration and their relationships with queue management as proposed in Section IV.

A. Transition Between Different Power States

In our proposed scheme, we use two different modes for power saving, one for dealing with *class* 1 traffic (see Fig. 4(a)) which is highly time sensitive and requires immediate attention from the transmitter while the other mode (see Fig. 4(b)) is for *class* 2 and 3 traffic. Mode 1 is based on two power states because the transmitter and receiver have the same sleep and awake times. As shown in Fig. 4(a), our proposed scheme switches between the all-wake-up and all-sleep states for *class* 1 traffic. The all-wake-up state stays on for ($T_{\text{grant}} + T_{oh}$) and all-sleep state stays on for ($T_{\text{PON}} - T_{\text{grant}} - T_{oh}$) (Table II). This is because the delay for *class* 1 traffic should not exceed

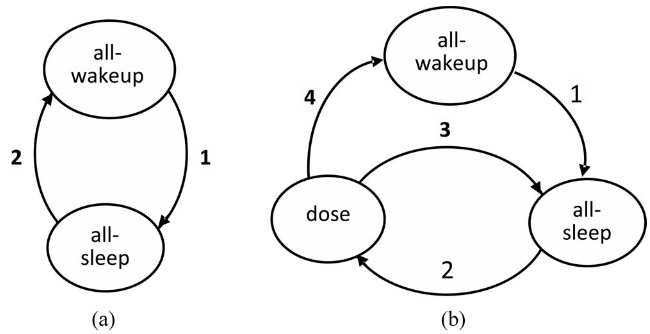

Fig. 4. Different power states of ONU based on; a-existence and b-absence of a class 1 traffic.

one polling cycle duration. The transition from the all-sleep to all-wake-up occurs when the sleep time expires.

In mode 2, we have three power states namely: all-wake-up, all-sleep and dose. These three states arise from the difference between the sleep time duration of the transmitter and the receiver. Because classes 2 and 3 tolerate longer delays, we use a longer sleep time for the transmitter than for the receiver. This is because the OLT transmits downstream traffic and control messages (GATE) for the ONUs even when the transmitter is sleeping. Therefore, for maintaining the QoS of downstream traffic, the receiver needs to wake up in every cycle. Firstly, the decision for sleep time duration in our scheme is based on the upstream traffic condition. Secondly, there is no mechanism to inform the ONU about the priority of downstream classes, so we bounded the delay of downstream traffic within one polling cycle. The transition between all-wake-up to all-sleep occurs in similar way to mode 1 (see Fig. 4(a)). Because the transmitter is switched to sleep for ($N_{\max} \times T_{\text{PON}} - T_{\text{grant}} - T_{oh}$) duration (see Table II) and the receiver wakes up in every polling cycle, the transition between the all-sleep to dose state occurs. T_{oh} depicts the overhead time required for the ONU to synchronize and warm up after it wakes up; $N_{\max}$ depicts the maximum number of polling cycles over which an ONU can continuously sleep, and is calculated as ($N_{\max} = \frac{T_s^{\max}\ (\text{Table I})}{T_{\text{PON}}}$). The reverse transition occurs when the transmitter still has sleep time and the receiver has finished receiving the downstream data, so it can switch to sleep mode again. The transition between dose to all-wake-up occurs with the expiration of sleep time for the transmitter and receiver and the approximation of the starting time if bandwidth is allocated. Algorithms 1 and 2 describe the transition between the power states of mode 1 and 2, respectively. Importantly, to allow the all-sleep state, the OLT needs to buffer the downstream traffic of the sleep ONUs and schedule concurrently with the upstream traffic scheduling for the intended ONU. The decision whether to use mode 1 or 2 at the ONU is not static, but depends on the status of its q_1, q_2 *and* q_3 occupancies which is likely to change with time. Thus the ONU makes this decision every time before it goes to sleep.

For ONU's transmitter, the sleep time (T_s) and wake up time (T_a) can be re-expressed in Eq. (17) and (18), respectively.

$$T_s = \begin{cases} T_{\text{PON}} - T_{\text{grant}} - T_{oh}, & \text{class } 1 \\ N_{\max} \times T_{\text{PON}} - T_{\text{grant}} - T_{oh}, & \text{class } 2, 3 \end{cases} \tag{17}$$

$$T_a = T_{\text{grant}} + T_{oh}. \tag{18}$$

Algorithm 1 The transition between all-wake-up and all-sleep states.

start: If OLT allocates grant for the ONU_i
    ONU_i sends data to the OLT;
    ONU_i sends the REPORT to the OLT;
    set sleep_time = $(T_{PON} - T_{grant} - T_{oh})$;
    ONU_i switches_to_ sleep_ mode;
    after sleep_time_duration
    ONU_i _Rx is wake_up;
    If ONU_i _Tx transmitted BW requests for q_1
        ONU_i _Tx is wake_up;
        go to line start;
    Else
        go to line k; (Algorithm 2)
    End if
End if

Algorithm 2 The transition from all-sleep to all-wake-up through dose state.

k: If ONU_i _Tx transmitted BW request for q_2 or q_3
    s = 1;
    set ONU_i _Tx_sleep_time =
    $N_{max} \times T_{PON} - T_{grant} - T_{oh}$;
    a: after s$\times T_{PON}$
        ONU_i _Rx is wake_up and read the downstream traffic;
        If ONU_i _Tx_sleep_time > (s$\times T_{PON}$)
            ONU_i _Rx enter sleep;
            s = s + 1;
            go to line a;
        Else
            ONU_i _Rx is wake_up;
            ONU_i _Tx is wake_up;
        End if
End if

In the dynamic bandwidth allocation, the length of polling cycle ($T_{\text{PON}} = \sum_{n=1}^{N}(T_g + T_{\text{grant}}^n)$) varies depending on the load conditions. Because T_{PON} depicts the polling cycle when all the ONUs are polled and transmitting their upstream traffic (see Fig. 3), we need to calculate the maximum duration of the polling cycle to estimate the maximum sleep time.

The maximum length of the polling cycle ($T_{\text{PON}}^{\max} = N \times (T_g + T_{\text{grant}}^{\max})$) and ($T_{\text{grant}}^{\max} = \frac{\sum_{k=1}^{3} q_k^{\max} \times P_k \times 8}{R}$), where $T_{\text{grant}}^{\max}$ depicts the maximum transmission window which should be equal to the maximum requested bandwidth by any ONU; P_k depicts the packet size for class kth; $q_k^{\max}$ depicts the maximum queue size for class kth; R depicts the transmission rate.

B. Sleep Time Duration Policies

The sleep time duration differs in mode 1 and 2 (see Table II) only for the transmitter. Moreover, the suggested transmitter sleep time duration depends on the class of traffic. Sleep time was fixed for every traffic class, so that the OLT can infer

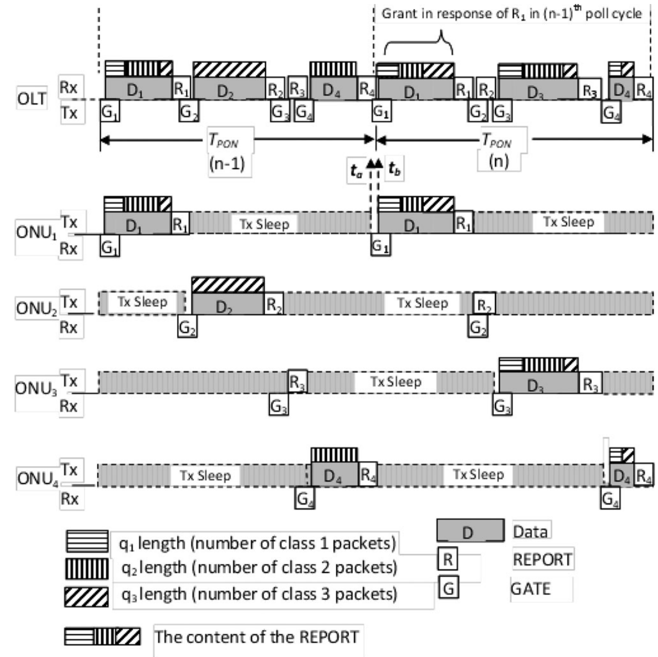

Fig. 5. Time diagram of CBPS operation for four ONUs.

TABLE I
MAXIMUM SLEEP TIME OF DIFFERENT SLEEP ASSUMPTIONS FOR *class* 2 AND *class* 3 TRAFFIC

Assumption for Sleep Time	Max. Sleep Time ($T_s^{\max}$)
Aggressive assumption	45 ms
Fine assumption	35 ms
Conservative assumption	25 ms

this wake up time from the REPORT message (as described in Fig. 5). For instance, when the ONU requests bandwidth, it puts information regarding its q_1, q_2 and q_3 occupancies in the REPORT message. So, if it reports packets in q_1, the OLT knows that the transmitter will wake up at the end of one polling cycle ($T_{\text{PON}} - T_{\text{grant}} - T_{oh}$). Similarly, if ONU sends a REPORT with 0 for the q_1 and requests bandwidth for the q_2, the OLT infers that the transmitter's wake up time is after T_s for *class* 2 and so on. For q_3, even though this traffic tolerates a sleep time duration longer than that for *class* 2, the ONU's transmitter needs to be triggered every 50 ms because of the restriction of the 802.3 av (the standard for 10G-EPON). This requires a maximum gap between two REPORTs of 50 ms to maintain the ONU's registration with the OLT [13], [22].

For *class* 1, we set the sleep time as less than one polling cycle ($T_{\text{PON}} - T_{\text{grant}} - T_{oh}$). However, if the polling cycle duration is not longer than the double length of overhead time then it will not activate the sleep state of the transceiver. In this work, we consider the maximum polling cycle as a fixed cycle even with dynamic allocation of the bandwidth. Moreover, in our experiments, we used only aggressive and conservative sleep times with 3 and 2 ms overhead times, respectively (see Table I).

TABLE II
MAXIMUM SLEEP TIME OF THE ONU'S TRANSMITTER FOR DIFFERENT TRAFFIC

Class	Priority	Tx Sleep Time (T_s)	Mode Type
1	High	T_{PON} -$T_{\mathrm{grant}} - T_{oh}$	1
2	Medium	$N_{\max} \times T_{\mathrm{PON}} - T_{\mathrm{grant}} - T_{oh}$	2
3	Low	$N_{\max} \times T_{\mathrm{PON}} - T_{\mathrm{grant}} - T_{oh}$	2

C. CBPS Based ONU

The essence of the CBPS scheme is that the ONUs transit to sleep mode after transmitting data and the REPORT. Despite the higher power saving of the sleep state, compared to dose mode, the ONU transits to sleep mode only for a short time, and stays in dose state for longer. As previously stated, the transition to sleep and dose states depends on the class priority of the available traffic in the ONUs. Fig. 5 illustrates the operation of the CBPS mechanism in cooperation with dynamic bandwidth allocation. We show only the sleep time of the ONU's transmitter because as stated previously the sleep time of the receiver is bounded to $(T_{\mathrm{PON}} - T_{\mathrm{grant}} - T_{oh})$. We can see the OLT periodically polls 4 ONUs and assigns bandwidth for them (we show only two polling cycles). For instance, in polling cycle (n-1), we see that ONU_1 receives a GATE message (G_1) and infers the starting time and length of its allocated grant. ONU_1 dispatches its waiting packets (D_1) and requests bandwidth by reporting the occupancy of its queues in REPORT (R_1) for the next polling cycle; this report includes bandwidth request for q_1, q_2 and q_3. Therefore, ONU_1 enters sleep for $(T_{\mathrm{PON}} - T_{\mathrm{grant}} - T_{oh})$ as it requests bandwidth for q_1 in response to *class* 1 traffic. The allocated grant in the polling cycle (*n*) is calculated based on the REPORT R_1 for the (n-1) polling cycle. Similarly ONU_3 and ONU_4 also request bandwidth for q_1. ONU_2's transmitter is switched to sleep mode for $(N_{\max} \times T_{\mathrm{PON}} - T_{\mathrm{grant}} - T_{oh})$ because it requested bandwidth for q_2 and q_3 in the previous REPORT. In this scheme, the OLT grants bandwidth, to all ONUs. For ONUs in sleep mode, a $T^{\min}_{\mathrm{grant}}$ of 64 bytes is required to allow the transmission of REPORTS and detection of downstream traffic. Algorithm 3 summarizes all the operations of our proposed class based dynamic bandwidth allocation.

The total energy consumption of the transmitters for N ONUs is:

$$E = \sum_{n=1}^{N} [a \times E^n_{cls1} + (1-a) \times E^n_{nocls1}] \quad (19)$$

where E_{cls1} and E_{nocls1} denote the energy consumption with (a = 1) and without (a = 0) *class* 1 traffic, respectively.

By introducing T_s and T_a for *class* 1 and other classes (*non-class 1*), we generate

$$E = \sum_{n=1}^{N} (a \times (T^n_s(cls\,1) \times P_s + T^n_a(cls\,1) \times P_a) + (1-a) \times (T^n_s(nocls\ 1) \times P_s + T^n_a(nocls\ 1) \times P_a)). \quad (20)$$

Eq. (20) shows that each ONU's transmitter has both awake and sleep energy consumption.

By specifying the T_s (Eq. (17)) and T_a (Eq. (18)) of each class, the following can be shown:

$$E = \sum_{n=1}^{N} (a \times ((T_{\mathrm{PON}} - T^n_{\mathrm{grant}} - T_{oh}) \times P_s + (T^n_{\mathrm{grant}} + T_{oh}) \times P_a) + (1-a) \times ((N_{\max} \times T_{\mathrm{PON}} - T^n_{\mathrm{grant}} - T_{oh}) \times P_s + (T^n_{\mathrm{grant}} + T_{oh}) \times P_a)). \quad (21)$$

In the above equation, E denotes the total energy consumption by the transmitters of the ONUs; P_s denotes transmitter power consumption in sleep state; P_a denotes transmitter power consumption in active state. From Fig. 5, $T_{oh} = t_a - t_b$, where t_a denotes wake up time and t_b denotes the transmission start time for the data and REPORT.

VI. PERFORMANCE EVALUATION

In this section, we present the simulation results to validate our proposed queue management and power saving schemes. We chose the Network Simulator 2.29 [24] and created a model to simulate the 10 G-EPON and WiMAX. We chose WiMAX model as a preferable option because of its compatibility with the simulator. The simulation parameters for both the 10G-EPON and WiMAX are configured as shown in Tables III & IV, respectively. For traffic modelling, we followed a traffic profile [18], [25] where high priority traffic such as VoIP is modelled by a constant bit rate source, video applications are modelled using a variable bit rate source, and BE traffic arrives dynamically. Similar traffic profile was used in other relevant studies [18], [25]. The UGS (*class* 1) traffic was modelled after the VoIP applications, using a constant bit rate with a packet size of 100 bytes, and intervals of 0.1 and 0.2 s for modelling high and low load VoIP applications. For simulating the video streaming (*class* 2) applications, we chose variable bit rate traffic with packet sizes uniformly distributed between 500 and 1500 bytes. To simulate the BE traffic, we used a TCP/ Newreno type [24] with a packet size of 60 bytes. We changed the arrival rates of applications to investigate the impact of various loads on network performances. It can be noted that because we are using 10G-EPON with a maximum capacity of 10 Gb/s, WiMAX system (802.16 m) which provides the access network connections in WOBAN, works as the bottleneck.

We maintained the same scenario and traffic profile when comparing our proposed model against existing models. We compared our proposed approach against a model published recently [17], which classifies the traffic into two classes, expedited forwarding (EF) and best effort (BE). The high priority class is expedited forwarding, which is delay sensitive (e.g., VoIP and video streaming). The low priority class is BE and is not delay sensitive (e.g., FTP, email). We refer to this model as EF_BE and our proposed model as Diff_srv. The EF_BE model assigns a queue of infinite length to each traffic class. If there is no packet in the queue corresponding to EF traffic, the

Algorithm 3 Class Based Dynamic Bandwidth Allocation (DBA).

OLT:
start new polling cycle: Loop $\forall ONU$
If state != wake up
Allocated BW $= T_{grant}^{min}$
Else
Allocated BW $= T_{grant}$ based on ONU's previous REPORT message.
End if
send data (downstream packets)
collect upstream packets
collect REPORT
End Loop
OLT considers the REPORTs received from all ONUs and calculates the wake up time of Rx and Tx for all *ONUs* based on each *ONU's* power state (Mode 1 or 2);
OLT calculates (T_{grant}) for each *ONU* (Eq. 9)
OLT generates GATE messages and schedules them for each
ONU based on its wake up time
go to start new polling cycle:
ONU:
In every polling cycle:
If Rx_sleep_time expired
Rx wakes up
Rx receives GATE message
Rx receives downstream packets
End If
If Tx _sleep_time expired
Tx wakes up
If (T_{grant}) is approached
ONU sends data (upstream packets)
ONU sends a REPORT message to OLT indicating its q_1, q_2 and q_3 occupancies
If q_1 is not empty
ONU moves into mode 1 (Algorithm 1)
Else
ONU moves into mode 2 (Algorithm 2)
End If
ONU decides the next wake up time for Rx and Tx
End If
End If.

TABLE III
SYSTEM PARAMETER OF SIMULATION IN 10G-EPON

Parameter	Value
Sync time	2, 3 ms
PON cycle time	1, 5, 10, 20, 30, 40 ms
Number of ONUs	16
Guard time	1 μs
Mean packet length/q_1 q_2 q_3	100 750 60
Bandwidth quota/q_1 q_2 q_3	0.3 0.5 0.2

TABLE IV
SYSTEM PARAMETER OF SIMULATION IN WIMAX

Parameter	Value
WiMAX frame	20 ms
Number of UGS connections	0-10
Number of rtPS connections	0-12
Number of BE connections	0-2
Modulation	BPSK 1/2
CP	0.25
Channel bandwidth	20 MHz

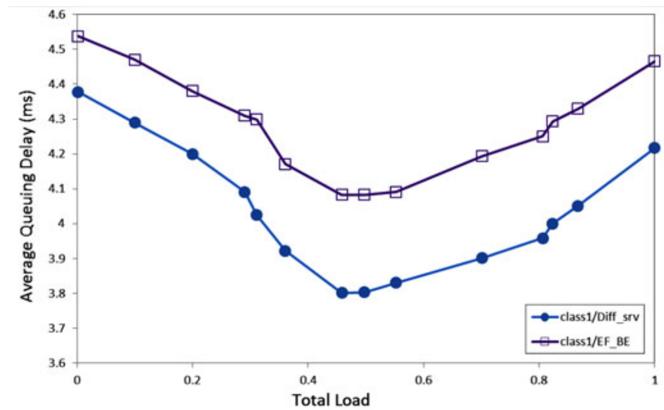

Fig. 6. Comparison between our proposed Diff_srv and EF_BE model for the average queuing delay with *class* 1 packets.

EF_BE model puts ONU to sleep mode and the sleep period is calculated based on a state-transition graph model in response to the load in the queue corresponding to EF traffic. We compared our model against the EF_BE model because this is the only model available in the literature that deals with QoS of broadband services and power savings at the same time.

A. Queue Management

This section compares the performance of the proposed queue management model against the FE_BE model [17]. In the simulation, the length of each queue in the proposed Diff_srv system was calculated using Eq. (16). We then monitored the average queuing delay in both the Diff_srv and EF_BE systems at various loads as shown in Figs. 6 and 7. These figures show that the Diff_srv system provides a much lower average queuing delay for *class* 1 flows under all tested loads. Fig. 6 indicates that our proposed Diff_srv scheme consistently outperforms the EF_BE model in terms of queuing delay under all load conditions. Reduced queuing delay for time critical applications such as *class* 1 traffic, implies reduced end-to-end delay which is a key parameter for measuring QoS. The first reason for this outcome is our proposed system's ability to calculate the appropriate queue length and the second reason is its ability to serve the q_1 immediately after the ONU-BS receives the grant. The proposed model also outperforms the EF_BE model for *class* 2 traffic in terms of average queuing delay under all tested loads (see Fig. 7). Fig. 8 shows the packet dropping rates for *class* 1 and *class* 2 traffic. It is evident that packet dropping rate is zero for *class* 1

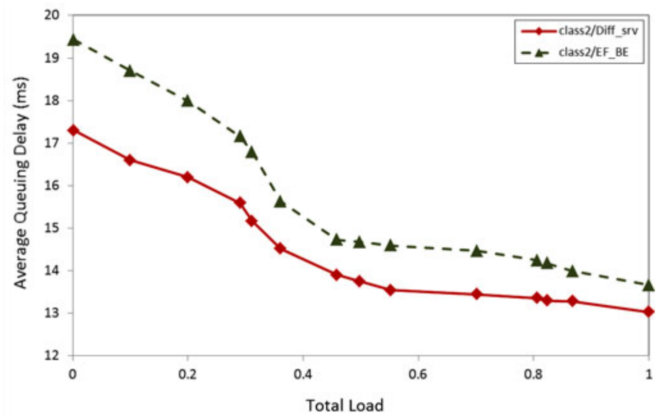

Fig. 7. Comparison between our proposed Diff_srv and EF_BE model for the average queuing delay with *class* 2 packets.

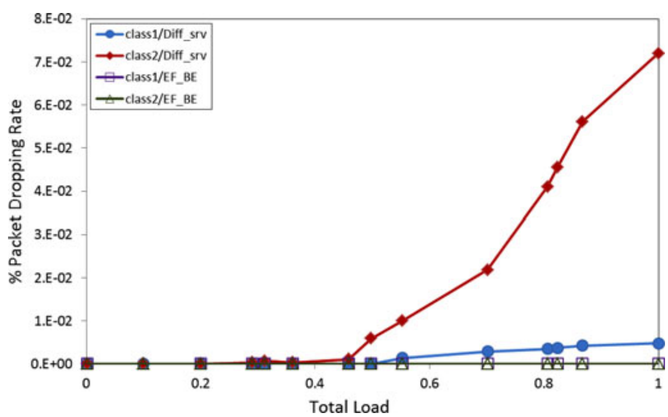

Fig. 8. Packet dropping rate comparison between our proposed Diff_srv and EF_BE model.

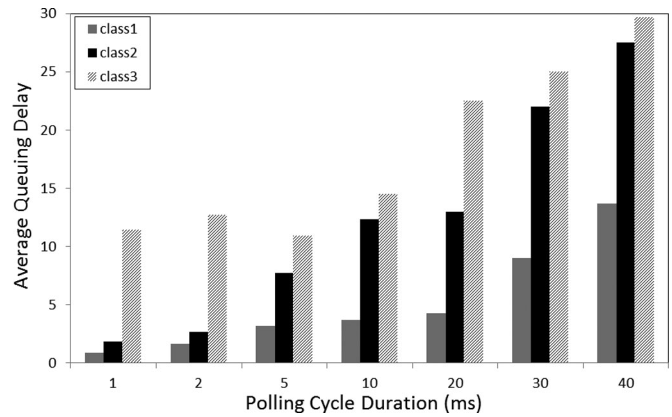

Fig. 9. Average queuing delay versus different polling cycle duration (T_{PON}) for Diff_srv model.

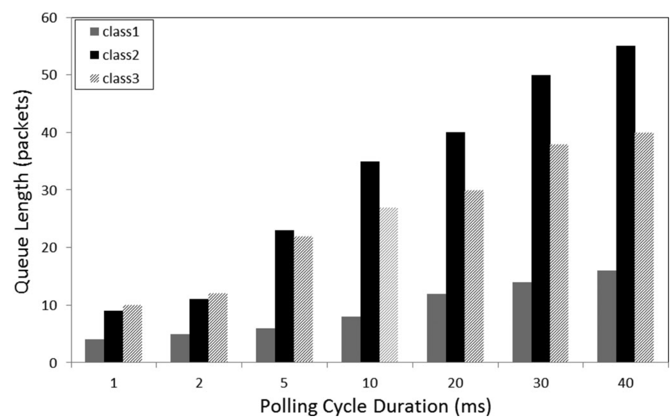

Fig. 10. Queue length versus different polling cycle duration (T_{PON}) for Diff_srv model.

and *class* 2 in our model for low to moderate load. Under heavy load, the system experiences insignificant dropping rates (i.e., 0.0047% and 0.072%, respectively) for *class* 1 and *class* 2 traffic. The packet dropping rate for the EF_BE is zero under all loads because it uses a huge buffer length. For the Diff_srv system, another influence on queuing delay is the polling cycle duration.

Fig. 9 shows the average queuing delay for each traffic class at various polling cycle durations. The delay is linked to the average queue length as shown in Fig. 10. When the polling cycle is very short (e.g., 1 ms, 2 ms), the grant period is also very short for each ONU and when the grant period is short, packets from *class* 1 and 2 traffic start to heavily dominate *class* 3 packets because of their higher priorities. *Class* 3 packets however, are still served at a very low rate because our scheme allows a small fraction of each grant period for *class* 3 traffic so that it does not experience complete resource starvation. When the polling cycle increases (e.g., 5 ms and above), each ONU starts to enjoy much larger grant periods in each polling cycle. When the grant period increases, their relative serving period in each grant period also increases for all traffic classes. This however, comes at the cost of increasing the waiting period (T_{gap}) before the next grant period is made available to each ONU; this is one reason why average queue length increases with increasing polling cycles for all traffic classes. The growth rate of queue length for different classes at various polling cycles also depends on packet arrival rates to the ONU, which is dictated by the wireless BS. In our simulation scenario for the given traffic profile, at 5 ms polling cycle, a combination of parameters such as the serving period for *class* 3, the serving periods for *class* 1 and 2, packet arrival rates for *class* 1, 2 and 3 to the ONU and the waiting period for the next grant period for each ONU, contributes to a scenario where *class* 3 packets were transported with a lower average delay.

B. Power Saving

This section presents the simulation results of our proposed class based power saving (CBPS) model against existing models [13], [17] available in literature. The common concept for power savings in existing models is to adopt a mechanism for avoiding or quitting sleep mode when high priority traffic is present at the ONU queue. We refer to this concept as the Idle Transmitter Time-based Power Saving (ITT-PS) model. The ITT-PS is a framework for calculating the sleep time as a function of the idle transmission period [13]. Initially, the sleep time begins with a specific value. As long as there is no upstream traffic, the

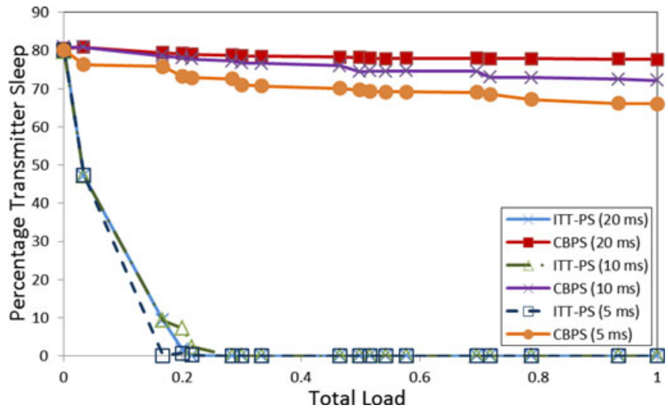

Fig. 11. Comparison between the CBPS and ITT-PS model for percentage power saving.

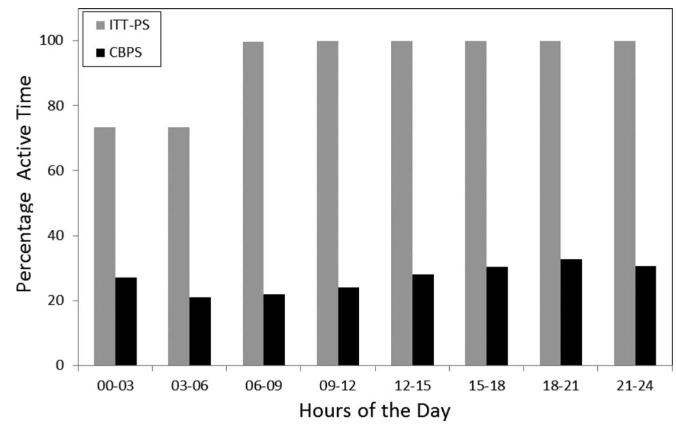

Fig. 12. Power saving comparison between ITT-PS and CBPS over 24 h.

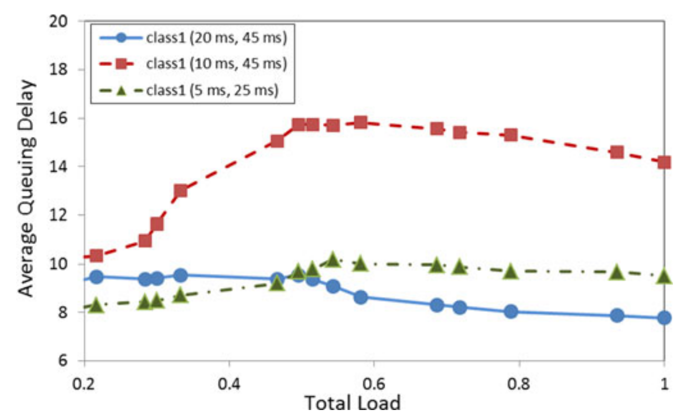

Fig. 13. Average queuing delay of *class* 1 traffic.

sleep time is gradually increased until it reaches the maximum sleep time of 50 ms. The arrival of traffic favors the wake up state for the transmitter. The simulation parameters for both the 10G-EPON and WiMAX are configured as shown in Tables III and IV, respectively. In the simulation, we measure the total sleep time of the transmitter and the average queuing delay experienced by different types of traffic.

The percentage sleep time of the transmitter in ONU is used as a parameter to measure the energy efficiency performance. The more time the transmitter is in sleep mode, the less energy is consumed by the ONU. The results provide the total sleep time used in our proposed CBPS model, compared to ITT-PS, for three values of the PON cycle time, namely, 5, 10 and 20 ms. This highlights the effectiveness of making the sleep time accord with the traffic class. We then show QoS provisioning for different classes of traffic in the CBPS model.

Fig. 11 presents the comparison between the proposed CBPS and ITT-PS models in terms of percentage sleep time for the ONU's transmitter. In the CBPS model the percentage sleep time is not affected much by an increase in the offered traffic load. The CBPS model consistently maintains a saving up to 77% for PON polling cycle of 20 ms. This is because of the low impact of guard (1 μs) and the overhead time (2–5 ms) [21] on the long polling cycle (20 ms). Furthermore, the long idle time for each ONU causes a longer sleep time.

Notably, a shorter polling cycle causes lower power savings; therefore, a polling cycle with 10 ms can save up to 72% while the 5 ms polling cycle shows the lowest power savings (about 66%). This is again due to the higher impact of the guard and overhead time on the shorter polling cycle length. However, in the ITT-PS model the savings in power declined sharply from 80% as high priority traffic increases in the ONU's queues. This occurs because the ITT-PS model ignores the idle time of ONUs even with heavy traffic which contributes to about 93% for a system consisted of 16 ONUs (the transmission time of each ONU is about 1/16 of the total transmission time). Such differences can be attributed to the CBPS efficiency in exploiting the idle time to make the ONU turn its transmitter into sleep mode and save more power. Moreover, there is a difference between the power saving of our proposed scheme for different PON cycle time 20, 10 and 5 ms but almost no differences between the performance curves of the other proposal. The results for CBPS reveal that increasing the polling cycle duration increases the sleeping time and consequently the power saving. Fig. 12 shows the comparison between our model and the ITT-PS model in terms of ONU's active time (%) for a full 24 h period following the traffic profile used in [4]. The traffic profile captures the varying traffic loads during a 24 h period. As indicated in Fig. 12, our proposed model consistently outperforms the ITT-PS model by a significant margin.

Fig. 13 illustrates the influence of CBPS on the behavior of WOBAN by showing the average queuing delay of *class* 1 traffic for different polling cycle durations. We use pairs (T_{PON}, T_s) in our discussion to describe every curve in the figure. As shown, the shortest average delay belongs to the (5, 25 ms) pair when the load is below the half of the full load. Then, with increasing load, (20, 45 ms) exhibits the minimum average delay. The reason is that, with increasing load, the maximum transmission window becomes over-utilized for (5, 25 ms) which causes more unscheduled packets waiting in the queues for the next polling cycle. For the (20, 45 ms) pair, the light-load penalty is clear below 50 percent load and is caused by small grants (small number of packets) with long idle time; then the situation improves with increasing load to produce an average delay of less than 8 ms. The worst performance belongs to (10, 45 ms); this highlights the importance of the ratio between the sleep time and polling cycle and the warm-up time. Usually, after the 55% load, all scenarios illustrate a steady reduction on the average delay due to the fact that with increasing load, the number of packets flowing in the queues also rises. This implies that the transmission of more packets at one grant consequently minimizes the average delay.

VII. Conclusion

Reducing power consumption and improving quality of service are two major concerns in the telecommunication industry. Several studies have investigated and evaluated the potential of achieving a more efficient communication system in terms of power consumption. This paper proposed a new class based power saving model for the 10G PON in WOBAN architecture which integrates 10G-EPON and 4G networks. The proposed model significantly improves the power saving by efficiently exploiting the idle periods within the tolerated delay of different services, to increase the sleep time of the optical part of the ONU. The results reveal that using CBPS obtains power savings of up to 80%. Our proposed class based queue model also provides better quality of service compared to other models.

References

[1] J. Bokrae, C. JungYul, H. Young-Tae, K. Min-Gon, and K. Minho, "Centralized scheduling mechanism for enhanced end-to-end delay and QoS support in integrated architecture of EPON and WiMAX," *J. Lightw. Technol.*, vol. 28, no. 16, pp. 2277–2288, Aug. 2010.

[2] S. Choi and J. Huh, "Dynamic bandwidth allocation algorithm for multimedia services over ethernet pons," *ETRI J.*, vol. 24, no. 6, pp. 465–468, 2002.

[3] K. Yang, S. Ou, K. Guild, and H.-H. Chen, "Convergence of ethernet PON and IEEE 802.16 broadband access networks and its QoS-aware dynamic bandwidth allocation scheme," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 2, pp. 101–116, Feb. 2009.

[4] P. Chowdhury, M. Tornatore, S. Sarkar, and B. Mukherjee, "Building a green wireless-optical broadband access network (WOBAN)," *J. Lightw. Technol.*, vol. 28, no. 16, pp. 2219–2229, Aug. 2010.

[5] M. A. Ali, G. Ellinas, H. Erkan, A. Hadjiantonis, and R. Dorsinville, "On the vision of complete fixed-mobile convergence," *J. Lightw. Technol.*, vol. 28, no. 16, pp. 2343–2357, Aug. 2010.

[6] S. Gangxiang, R. S. Tucker, and C. Chang-Joon, "Fixed mobile convergence architectures for broadband access: integration of EPON and WiMAX [Topics in Optical Communications]," *IEEE Commun. Mag.*, vol. 45, no. 8, pp. 44–50, Aug. 2007.

[7] S. Lambert, W. V. Heddeghem, W. Vereecken, B. Lannoo, D. Colle, and M. Pickavet. (2012 Dec.). Worldwide electricity consumption of communication networks. *Opt. Exp.*, *20(26)*, pp. B513–B524. [Online]. Available: http://www.opticsexpress.org/abstract.cfm?URI=oe-20-26-B513

[8] M. Pickavet, W. Vereecken, S. Demeyer, P. Audenaert, B. Vermeulen, C. Develder, D. Colle, B. Dhoedt, and P. Demeester, " Worldwide energy needs for ICT: the rise of power-aware networking," in *Proc. IEEE Ants: 2nd Int. Symp. Adv. Netw. Telecommun. Syst.*, 2008, pp. 1–3.

[9] LIGHTWAVE. (2014). IEEE ratifies new 802.3az standard to reduce network energy footprint. [Online]. Available:" http://www.lightwaveonline.com/articles/2010/10/ieee-ratifies-new-8023a z-standard-to-reduce-network-energy-footprint–104392129.html.

[10] J. Baliga, R. Ayre, K. Hinton, and R. Tucker, "Energy consumption in wired and wireless access networks," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 70–77, Jun. 2011.

[11] P. Dini, M. Miozzo, N. Bui, and N. Baldo, "A model to analyze the energy savings of base station sleep mode in LTE HetNets," in *Proc. Int. Conf. Green Comput. Commun./IEEE Int. Conf.Internet Things Cyber, Phys. Social Comput.*, Aug. 2013, pp. 1375–1380.

[12] ITU, "G. Sup45: GPON power conservation. (2014, Sept.). [Online]. Available: http://www.itu.int/rec/T-REC-G.Sup45/en

[13] J. Zhang and N. Ansari, "Toward energy-efficient 1G-EPON and 10 G-EPON with sleep-aware MAC control and scheduling," *IEEE Commun. Mag.* vol. 49, no. 2, pp. s33–s38, Feb. 2011.

[14] R. Kubo, J. Kani, Y. Fujimoto, N. Yoshimoto, and K. Kumozaki, "Adaptive power saving mechanism for 10 Gigabit Class PON Systems," *IEICE Trans. Commun.*, vol. E93-B, no. 2, pp. 280–288, 2010.

[15] Y. Yan, S. Wong, L. Valcarenghi, S.-H. Yen, D. Campelo, S. Yamashita, L. Kazovsky, and L. Dittmann, "Energy management mechanism for ethernet passive optical networks (EPONs)," in *Proc. IEEE Int. Conf. Commun.*, May 2010, pp. 1–5.

[16] J. Mandin. (Sep. 2008). EPON power saving via sleep mode. presented at the IEEE P802.3av 10GEPON Task Force Meeting. [Online]. Available:. http://www.ieee802.org/3/av/public/2008_09/3av_0809_mandin_1.pdf.

[17] L. Shi, B. Mukherjee, and S. Lee, "Energy-efficient PON with sleep-mode ONU: Progress, challenges, and solutions," *IEEE Netw.*, vol. 26, no. 2, pp. 36–41, Mar./Apr. 2012.

[18] C. M. Assi, Y. Yinghua, S. Dixit, and M. A. Ali, "Dynamic bandwidth allocation for quality-of-service over ethernet PONs," *IEEE J. Select. Areas Commun.*, vol. 21, no. 9, pp. 1467–1477, Nov. 2003.

[19] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON: Design and analysis of an optical access network," *Photon. Netw. Commun.*, vol. 3, no. 3, pp. 307–319, 2001.

[20] B. O. Obele, M. Iftikhar, S. Manipornsut, and M. Kang, "Analysis of the behavior of self-similar traffic in a QoS-aware architecture for integrating WiMAX and GEPON," *IEEE/OSA J. Opt. Commun. Netw.,* vol. 1, no. 4, pp. 259–273, Sep. 2009.

[21] S. Wong, L. Valcarenghi, S. Yen, D. Campelo, S. Yamashita, and L. Kazovsky, "Sleep mode for energy saving PONs: Advantages and drawbacks," in *Proc. IEEE GLOBECOM Workshops*, 2009, pp. 1–6.

[22] A. R. Dhaini, P. Ho, and G. Shen, "Toward green next-generation passive optical networks," *IEEE Commun. Mag.*, vol. 49, no. 11, pp. 94–101, Nov. 2011.

[23] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester, "Power consumption in telecommunication networks: Overview and reduction strategies," *IEEE Commun. Mag.*, vol. 49, no. 6, pp. 62–69, Jun. 2011.

[24] The network simulator -ns-2. (2014, Sept.). [Online]. Available: http://www.isi.edu/nsnam/ns/

[25] A. R. Dhaini, C. Assi, M. Maier, and A. Shami, "Per-stream QoS and admission control in ethernet passive optical networks (EPONS)," *J. Lightw. Technol.*, vol. 25, no. 7, pp. 1659–1669, Jul. 2007.

Maha Ahmed received the Bachelor of Engineering degree in Electrical from Al-Mustansiriya University, Baghdad, Iraq, and the M.Sc. degree from the University of Technology, Baghdad, in 1991 and 2002, respectively. She is currently working toward the Ph.D. degree at the School of Engineering, Edith Cowan University, WA, Australia. Her current research interests include optical, wireless and hybrid wireless-optical broadband access networks.

Iftekhar Ahmad is currently working as a Senior Lecturer with the School of Engineering, Edith Cowan University, Australia. He received the Ph.D. degree in communication networks from Monash University, Australia, in 2007. His research interests include 5G technologies, green communications, QoS in communication networks, wireless sensor networks and computational intelligence.

Daryoush Habibi (M'95–SM'99) received the Bachelor of Engineering degree in electrical with (First Class Hons.) from the University of Tasmania, Burnie, Tasmania, in 1989 and the Ph.D. degree from the same University in 1994. He was with Telstra Research Laboratories, Flinders University, Intelligent Pixels, Inc., and Edith Cowan University, where he is currently a Professor and the Head of the School of Engineering. His research interests include engineering design for sustainable development, reliability and quality of service in communication systems and networks, smart energy systems, and environmental monitoring technologies. He is a Fellow of Engineers Australia, Electrical College Board member of Engineers Australia, ITEE College Board member of Engineers Australia, Editor-in-Chief of the Australian Journal of Electrical and Electronic Engineering, and President of the Australian Council of Engineering Deans.