2015

Improving energy efficiency and quality of service in an integrated wireless-optical broadband access network

Maha Shihab Ahmed
Edith Cowan University

Improving Energy Efficiency and Quality of Service in an Integrated Wireless-Optical Broadband Access Network

by

Maha Shihab Ahmed

This thesis is presented in fulfilment of the requirements for the degree of

Doctor of Philosophy

SCHOOL OF ENGINEERING

FACULTY OF HEALTH, ENGINEERING AND SCIENCE

EDITH COWAN UNIVERSITY

19th November 2015

USE OF THESIS

Abstract

Exponential growth in the volume of wireless data, boosted by the growing popularity of mobile devices such as smart phones and tablets, is forcing telecommunication industries to rethink network design, and focus on developing high capacity mobile broadband networks. Accordingly, researchers have undertaken developmental work for an integrated wireless-optical broadband access network (WOBAN). Passive optical networks (PONs) and fourth generation (4G) wireless networks are two major candidate technologies for the WOBAN. PON is a wired access technology, well-known for its high capacity, whereas 4G is a wireless broadband access technology, popular for its ease of deployment and ability to offer mobility. Integration of PON and 4G technologies, as a wireless-optical broadband access network, offers advantages such as extension of networks in rural areas, support for mobile broadband services, and rapid deployment of broadband networks. However, these two technologies have different design architectures for handling broadband services which require Quality of Service (QoS), for example, 4G networks use traffic classification for supporting different QoS demands whereas PON does not differentiate between traffic types. This integrated network must also be energy efficient, as a green broadband access network, without hindering QoS. While these technologies both use sleep mode, they differ in their power saving mechanisms.

This thesis first addresses a QoS solution for the incompatibility between these technologies. Service class mapping is proposed in Chapter 3 for the integrated WOBAN, based on the M/G/1 queuing model supported by an innovative priority scheduler. Once class mapping is deployed, a power saving mechanism can be devised by exploiting traffic differentiation. Specifically, a class-based strategy is proposed which helps optimise the sleep period for the terminal units of the optical network, without compromising QoS.

Since the optical network involves control and terminal nodes, both of which consume power, this thesis proposes an energy efficient mechanism that involves both components. In contrast, other published strategies (Chapter 2) have only considered the terminal units. Chapter 4 presents the mechanism for enabling global sleep (control and terminal nodes)

and local sleep (terminal nodes), based on the available traffic's class structure. This mechanism enables sleep for different components within the bandwidth allocation by adapting the switching between predefined polling cycle lengths.

As the WOBAN is comprised of both wireless and optical parts, a dynamic resource management mechanism is needed which responds to changing daily traffic patterns across a green integrated network. Consequently, Chapter 5 proposes a mechanism which dynamically adapts the polling cycles, of the optical and wireless parts of the network, to the changing traffic volume and class composition. Tailored sleep durations for the components of the WOBAN are facilitated within the resource management regime, as these components differ in their ability to function efficiently if management of the sleep periods is not responsive to the changing traffic volumes and class composition.

This dissertation creates new knowledge by seamlessly integrating the two parts of WOBAN and introducing differentiated, class-based sleep for the components of the hybrid network to help realise a green WOBAN.

The declaration page
is not included in this version of the thesis

ACKNOWLEDGMENTS

Firstly, I would like to express my gratitude to my supervisor Professor Daryoush Habibi for his insightful discussions on my research work through our weekly meetings.

I would like to gratefully acknowlege my co. supervisor Dr Iftekhar Ahmad for his enthusiasm and mentoring through out my PhD journey. His knowledge and logical way of thinking have been of great value to me.

My special thanks go to Dr Greg Maguire, the research writing consultant in the Faculty of Health, Engineering and Science for his constructive advice and help.

My thanks also go to Dr Eric Dines for his technical advice including appropriate use of some specialised descriptions in this thesis.

I am also grateful to Dr Viet Phung for his help and suggestions during our valued discussion.

I wish to thank my friends Miss Elizabeth Taylor and Mrs Noura Thomas for their generous personal support.

I would like to thank all the members of wireless research and research lab community.

I would like to thank my parents, brothers and sisters for their love, support and encouragement through my whole life.

Most importantly, I am indebted to my dear husband Qays and my lovely daughters Basma and Sama for their patience, encouragement and everlasting love.

Contents

List of Papers

List of papers from this research:

1. M. Ahmed, I. Ahmad, and D. Habibi, "Service Class Resource Management for Green Wireless-Optical Broadband Access Networks (WOBAN)," *IEEE/OSA Journal of Lightwave Technology*, vol. 33, no. 1, pp. 7-18, 2015.

2. M. Ahmed, I. Ahmad, and D. Habibi, "Green Wireless-Optical Broadband Access Network: Energy and Quality-of-Service Considerations," *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 669-680, 2015.

3. M. Ahmed, I. Ahmad, and D. Habibi, ""Load-Adaptive Resource Management for Green Wireless-Optical Broadband Access Networks (WOBAN)," submitted to the *IEEE/OSA Journal of Lightwave Technology.*

4. M. Ahmed, D. Habibi, and I. Ahmad, "An improved architecture for integrating fourth generation wireless and passive optical networks," TENCON 2012 - 2012 IEEE Region 10 Conference, 1–6.

List of Tables

List of Figures

List of Algorithms

Abbreviations

10G-EPON	10 Giga-Ethernet Passive Optical Network
1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	3rd Generation Partnership Project
4G	Fourth Generation
5G	Fifth Generation
AWG	Arrayed Waveguide Gratings
AGW	Access Gateway
BE	Best Effort
BS	Base Station
CO	Central Office
CWDM	Coarse Wavelength Division Multiplexing
DBA	Dynamic Bandwidth Allocation
DL-MAP	Down Link-MAP
DSL	Digital Subscriber Loop
DWDM	Dense Wavelength Division Multiplexing
eNB	evolved NodeB
EPC	Evolved Packet Core

EPON Ethernet Passive Optical Network

ertPS extended-real-time Polling Services

E-UTRAN Evolved Universal Mobile Telecommunications System Terrestrial Radio Access Network

FBA Fixed Bandwidth Allocation

FCFS First Come First Serve

FDD Frequency Division Duplexing

FSAN Full Service Access Network

FTP File Transfer Protocol

FTTB Fibre To The Building

FTTC Fibre To The Curb

FTTH Fibre To The Home

Gbps Giga bit per second

GBR Guaranteed Bit Rate

GPON Gigabit Passive Optical Network

HDTV High-Definition Tele-Vision

ICT Information and Communication Technology

IEEE Institute of Electrical and Electronics Engineers

IMT-A International Mobile Telecommunications-Advanced

IP Internet Protocol

ITU-T International Telecommunication Union-Telecommunication standardization sector

LTE Long-Term Evolution

LTE-A Long Term Evolution-Advanced

M/G/1 Memoryless/General/single server

MAC	Media Access Control
MAN	Metropolitan Area Network
Mbps	Mega bit per second
MPCP	Multi-Point Control Protocol
nrtPS	non-real-time Polling Service
OCDMA	Optical Code Division Multiple Access
OFDMA	Orthogonal Frequency-Division Multiple Access
OLT	Optical Line Terminal
ONU	Optical Network Unit
PON	Passive Optical Network
QAM	Quadrature Amplitude Modulation
QCI	Quality of service Class Identifier
QoS	Quality of Service
QPSK	Quaternary Phase Shift Keying
rtPS	real-time Polling Services
RTT	Round-Trip Time
SC-FDMA	Single Carrier-Frequency Domain Multiple Access
SLA	Service Level Agreement
SS	Subscriber Station
TCP	Transmission Control Protocol
TDD	Time Division Duplexing
TDM	Time Division Multiplexing
TDMA	Time Division Multiplexing Access
UGS	Unsolicited Grant Services

UL-MAP	Up Link- Map
UMTS	Universal Mobile Telecommunications System
VoD	Video on Demand
VoIP	Voice over IP
WDM	Wavelength Division Multiplexing
WiMAX	Worldwide Interoperability for Microwave Access
WOBAN	Wireless-Optical Broadband Access Network
XG-PON	10 Giga Passive Optical Network

Chapter 1

Introduction

In recent years, broadband access technologies have attracted increasing attention because of the exponential growth of the Internet and the popularity of many broadband services such as high definition TV, video on demand, video streaming, voice over IP and peer-to-peer file sharing, all of which demand a high bandwidth for delivery. Video traffic alone accounts for approximately 84 percent of the Internet traffic in the USA [9], and its bandwidth demands will prove very challenging for current broadband access technologies. Broadband technologies differ in their ability to cost-efficiently provide the above mentioned services. Wireline and wireless are the two categories of broadband access technologies which connect the end user to the service provider and compete to provide the required bandwidth for delivering broadband services. However, each technology features different advantages and limitations regarding bandwidth, coverage, cost, power consumption, quality of service (QoS), mobility, and reliability issues. For instance, wireless technologies allow low cost deployment and offer ubiquity and mobility, but are limited by bandwidth. Optical networks are the most promising wireline solution due to their reliability and high bandwidth, but they are costly to deploy and lack mobility. The telecommunications industry and its consumer groups envisage a green next-generation broadband access network, capable of delivering high bandwidth, QoS, reliability, and mobility. Since no individual technology will satisfy all these requirements, a hybrid access network comprised of multiple complementary technologies can be the solution [7].

Hybrid networks, constructed by integrating an optical network with a wireless broadband network, inherit reliability and high bandwidth from the former and mobility from the latter. In meeting the aforementioned challenges, hybrid wireless-optical networks also help overcome the inequitable digital division between rural and urban users. However, hybrid networks face QoS and energy efficiency challenges that must be addressed.

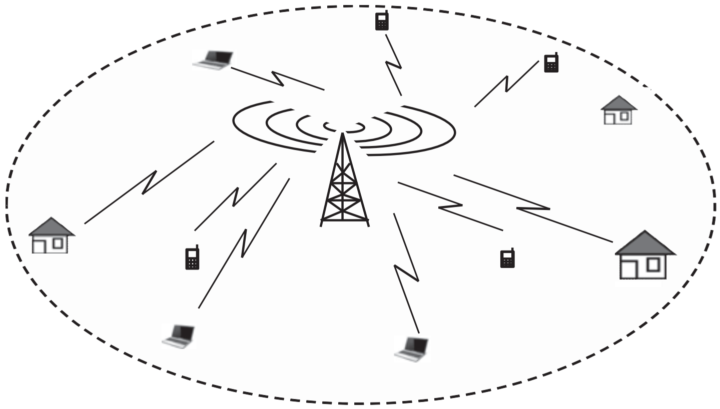

Figure 1.1: Connectivity in wireless network.

1.1 Wireless Communication Networks

Wireless broadband networks use radio or microwave frequencies to communicate between the receiver and transmitter as shown in Fig. 1.1. Wireless technology is becoming increasingly important for people worldwide due to its ubiquity and convenience. By 2019, mobile connected devices are expected to equate to 150% of the world's projected population [10]. Wireless and cellular communication systems are undergoing continuous development to cope with growing demand for better and diverse services. For example, the first generation (1G) wireless system provided poor quality analogue voice service, then the second generation (2G) system improved voice quality by using digital voice, and supported data services. The progression from 2G to third generation (3G) was even more significant. 3G was considered to be a broadband wireless system because it supported the Internet, streaming video, and multimedia messaging services. The current wireless access technology is the fourth generation (4G), which is comprised of two technologies, Worldwide Interoperability for Microwave Access (WiMAX), and Long Term Evolution (LTE). The 4G technology was released in 2012 and its goals were to accelerate data rates and increase the capacity to provide bandwidth-intensive services like HDTV, which requires up to 19.3 Mbps of downstream bandwidth [11] [12]. The 4G communication system utilizes only Internet Protocol (IP) to transfer packets, a standardization which accelerates the switching of packets between access and backbone networks [13]. Despite their design differences, both WiMAX and LTE have surpassed the requirements for a 4G wireless system in the International Mobile Telecommunications-Advanced (IMT-A) standard of 1 Gbps and 100 Mbps download bandwidth, for stationary and mobile wireless access devices respectively [14]. Although 4G has not been fully deployed worldwide, research on fifth

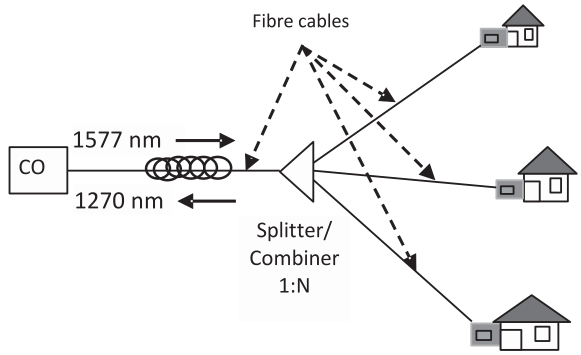

Figure 1.2: Connectivity in PON technology.

generation wireless systems (5G) has commenced and this technology is expected to be available by 2020 [15].

1.2 Optical Networks

Optical networking technology is unique in that its huge bandwidth can theoretically support data rates up to 25,000 Gbps [16]. This capacity and high reliability make optical networking the most attractive choice for backbone networks. It can also extend directly to the end user, via Fibre-To-The-Home (FTTH), Fibre-To-The-Building (FTTB), and Fibre-To-The-Curb (FTTC); using either of two connectivity architectures, Point-to-Point (P2P) and Point-to-Multi-Point (P2MP). Since the laying of fibre optic cable to each subscriber is a highly expensive option, Passive Optical Network (PON), which is based on P2MP connectivity architecture, is an economical alternative as it enables numerous users to share the fibre length and deploys passive components along the fibre path. However, a PON deployment still requires the laying of a fibre cable all the way to, or close to the users' premises (Fig. 1.2).

1.3 Wireless-Optical Broadband Access Network (WOBAN)

The Wireless-Optical Broadband Access Network (WOBAN), a hybrid network which integrates PON and 4G, could develop into a major access network as it offers the advantages of optical and wireless technologies, and suits both urban and rural areas. However, PON and 4G follow standards which were developed to meet different targets; thus a hybrid wireless optical network presents many technical challenges.

As an access network, WOBAN must support QoS while delivering various types of traffic, each of which has particular service challenges. For example, VoIP is affected by the delay and jitter, so the WOBAN must minimise these problems to maintain appropriate QoS. 4G networks have strong QoS credentials as they differentiate traffic delivery into flows with predefined QoS. However, while PON networks have no defined standards for the QoS, they do support queues. Despite the importance of queue management and the scheduler in achieving QoS at the PON networks, they are not yet standardized. Therefore, research into integrating the 4G and PON technologies has become important for achieving a hybrid network of wireless and wireline i.e., a WOBAN, particularly for QoS mapping and queue management. Even though 4G employs its own QoS strategies, it does not guarantee the overall QoS for the entire network. A design is needed for the optical component of the network to deal with a 4G's differentiated traffic. To this end, an efficient QoS mapper at the interface between 4G and the optical part of the network, augmented by robust queue management at the optical part, and a suitable scheduler to deal with different traffic types, are essential. The QoS mapper ensures the required QoS for each type of traffic when it travels between the 4G network and the PON. The queue management and the scheduler then play vital roles in serving each type of traffic, according to the required QoS. Important questions in this context include: i) the number of queues that are needed; ii) the appropriate queue management design for the WOBAN. (Queue management involves the number, length, and type of queues; and how to schedule the traffic from these queues when the network deals with traffic that is highly sensitive to delays, i.e. VoIP, and best effort traffic such as email, which has no QoS requirement); iii) the means for maintaining the QoS requirement and the fairness between the traffic types when dealing with these varieties of traffic.

Another aspect of the viability of an access network is energy efficiency. Developing a green access network means reducing the power consumption without affecting the QoS and the normal functionality of the network. Sleep and doze modes are commonly used to reduce the power consumption; in sleep mode, both the transmitter and the receiver are switched into sleep mode during the idle periods while in doze mode only the transmitter switches to sleep mode as the receiver remains active. The power consumption in sleep mode is much less than that in active mode. The sleep durations should be defined so that the components are ready to serve the traffic when it arrives. Deep and cyclic are the two categories of sleep mode. In deep sleep, the components enter into sleep mode for a long time whereas in cyclic, the components enter into sleep mode for a short duration and wake up to check and serve, and then go back to sleep. The challenges with sleep mode in QoS supportive networks are in deciding when to move into sleep mode, and for how long without affecting the QoS, given that long sleep durations cause longer delays but save more energy. The energy efficiency mechanism should take into account all components that consume power in the WOBAN. However, the roles of the various components and their idle times differ. Therefore, controlling the idle durations for each of the components, based on their role and their traffic types, can be achieved to enable energy savings within the resource management mechanism. Moreover, the resource management mechanism should also consider the daily traffic fluctuations in order to adjust the allocated resources to suit the existing traffic. The gains from responding to changes in the traffic are in improving the channel utilization, reducing the power consumption, and maintaining the QoS. The central components of this thesis revolve around this vital research area.

1.4 Significance

It has recently been revealed that the annual electricity consumption of all Information and Communication Technology (ICT) systems is about 1,500 TWh, which represents 10% of globally generated electricity [17]. Importantly if that electricity was generated from oil, the greenhouse gas emissions would include about 1099 million tonnes of CO_2 [18]. The increased power consumption of ICT in recent times is due to the huge growth of connected devices and global data traffic, which has been boosted by the emergence of wireless broadband access networks for mobile devices, i.e., smart phones and tablets. Furthermore, access networks tend to consume the most power in the ICT sector, and specifically 70% of the Internet's consumption [7]. This increase in power consumption has environmental and economic impacts from increasing both global warming and the operational cost of the networks [19]. These figures, together with increasing awareness of global warming, demonstrate the need for better strategies for reducing power consumption. Thus the ever increasing demand for electricity can be significantly mitigated by using green networks that have efficient resource utilization and lower energy consumption. An overall strategy is required to meet the increasing demand for high bandwidth, cost efficient delivery of broadband services, especially for video services which consume a huge bandwidth, and to support mobility in response to the increasing number of smart phones and tablets. This thesis presents mechanisms for improving QoS and reducing power consumption in the next generation broadband access networks.

1.5 Motivation

The future-proof network should be green, consume less power, and support QoS. The presence of more than one category of access network i.e., wireless and wireline, the diversity of traffic, the dominance of video traffic, the cost, the required coverage, and the distribution of the users, lead to a consideration of what the ideal access network is. This ideal network should maximize customer experience, be affordable, and be more environmentally friendly. As each category has its own advantages and disadvantages, finding the optimum broadband access network requires integration of the wireless and the optical components. Thus it is important to develop a green hybrid network that include an optical network, to obtain high capacity and reliability, and 4G, to obtain the required QoS and mobility; which provides the motivation for conducting this project.

1.6 Objectives

This thesis focuses on the design of future-proof broadband access networks. WOBAN is introduced as a promising solution for broadband networks as it links the merits of low costs and the ubiquity of wireless networks, together with the reliability and high capacity of the optical networks. However, QoS and energy efficiency are the major challenges that must be addressed. The QoS challenge is related to the gap between the two technologies that make up WOBAN. The wireless technology is more precise in defining both the QoS and the difference in the QoS parameters for different types of traffic; however the optical technology has no specific definition for the QoS. To overcome this problem, the fusion between the two part technologies must achieve seamless movement of traffic between the two technologies. The second potential challenge is power consumption. Realising a green WOBAN is about minimizing power consumption whilst maintaining acceptable levels of QoS for the real time traffic and, in particular, to cost efficiently deliver broadband services to the end users. This new hybrid network poses several challenges that motivate the in-depth investigations within this thesis. The key questions in this thesis are:

- How to integrate the two technologies (i.e., PON and 4G) to improve overall network performance and maintain QoS for the network?
- How to develop an energy efficient WOBAN by reducing the power consumption of the PON's terminal and control nodes without affecting the QoS?
- In addition to the above PON node issues, how to make the integrated network more

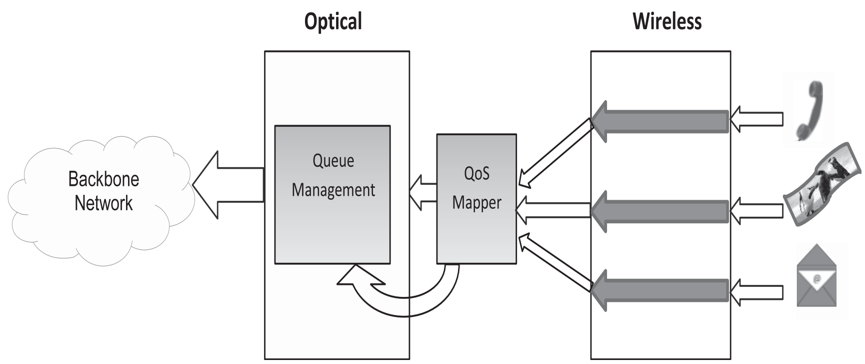

Figure 1.3: Layout of the integration between wireless and optical networks.

efficient by linking power consumption to the available load?

1.6.1 Establishing the Integration and Maintain QoS

The diversity of the wireless technologies and optical technologies makes the selection of the technologies that construct WOBAN a key factor for controlling the available capacity. The first requirement for realizing the integration of the wireless and optical networks is in achieving the fusion of the two technologies. Figure 1.3 shows the layout of the integration between the wireless and optical networks. Preparing the background for the integration requires solving the mapping challenge between the classes of traffic in the wireless network and the queues of traffic of the optical network. Moreover, servicing the queues which hold these different types of traffic is important; therefore, the queue management that suits this type of network is specifically addressed in this thesis. It is understood that this will require a scheduler that respects high priority traffic without disadvantaging low priority traffic. This study proposes a design for queue management and a scheduler that suits the WOBAN, so that QoS for each type of real time traffic is maintained without disadvantaging the non-real time traffic. From this, the relationships between the delays for each class, the arrival rate and the allocated resources, are derived.

1.6.2 Green Resource Management

Different service classes require different QoS parameters. Resource management affects the QoS, as the traffic is in two states; either wait or serve. Different usage of the access network requires different resource management. This thesis proposes a new resource management mechanism that lengthens the idle duration, to extend the sleep durations and increase power savings. The class of traffic is added as another dimension to the

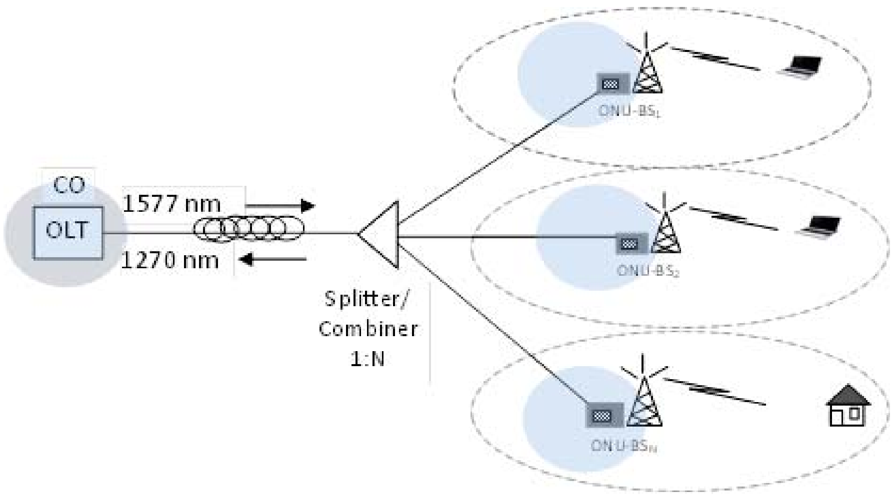

Figure 1.4: Layout of the WOBAN showing the points of power consumption in the optical part.

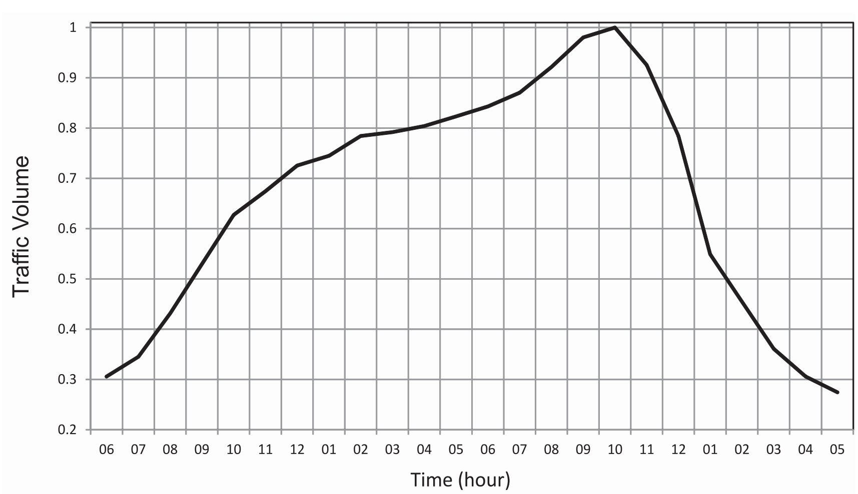

Figure 1.5: Aggregated daily traffic profile of a sample network [3].

resource management in order to maintain the appropriate service level agreement for each class. This mechanism targets the energy efficiency of the key components in WOBAN that consume the most power (shaded parts in Fig. 1.4).

1.6.3 Coordinated Resource Management

This part of thesis deals with the effect of the two resource management stages in WOBAN; specifically at the control nodes of the optical and wireless parts of the network affecting the QoS of all traffic flows. A comprehensive investigation shows that power can be saved and QoS supported if the control nodes are responsive to the composition and volume of traffic. The design of the proposed mechanism aims to make the resource management mechanism responsive to the dynamics of the changing daily traffic. A typical pattern for the relative aggregated traffic volume is shown in Fig. 1.5. In this proposal, either the resource management of the wireless part of the network depends on the changes in the optical part of the network, or it is independently changed. This mechanism provides more comprehensive control over the delay and jitter which affect the service level agreement. It is important to include all the components of the WOBAN (all the shaded parts in Fig. 1.6) when optimising energy efficiency. Strategies which improve energy efficiency through coordinated resource management in the WOBAN should also be evaluated for any effects on QoS.

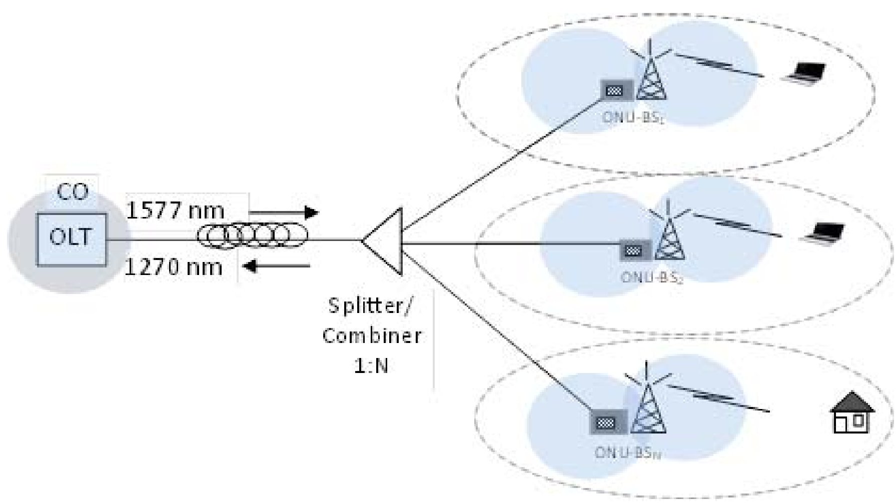

Figure 1.6: Layout of the WOBAN that shows the points of power consumption in both the optical and wireless parts.

1.7 Research Contributions

This research contributes to the development of an environmentally friendly hybrid wireless-optical broadband access network and the QoS provisioning for that network. To facilitate the fusion between the two technologies, it presents the mechanisms that adjust the periodic resource management to the dynamics of daily traffic, so as to improve both energy efficiency and QoS. Specifically, the best durations for the allocated resources and sleep periods are identified to suit the different types of traffic. In this thesis, each chapter contributes to an aspect of these overall goals:

- The first contribution of this thesis is to propose a solution to realize the seamless integration between 4G and PON. It is necessary to cater for different traffic classes that require different QoS. Therefore, a QoS mapper is proposed to assign the differentiated traffic at the 4G network to the appropriate queues at the PON. Following that, a design for queue management is proposed with a scheduler to maintain the priority of the traffic. The mathematical model and performance results from the associated simulation are outlined.

- The energy efficiency of the integrated network depends on reducing the power consumption when the load is low. This research deploys a cyclic sleep for the terminal nodes of the optical part of the network to exploit the idle durations. Differentiated fixed sleep duration is proposed in this thesis. The mathematical models and performance results are compared with those from published alternative strategies.

- An inflexible approach to the dynamics of daily traffic (shown in Fig. 1.5), should be considered as being energy inefficient as it usually reflects peak loads. Thus, a differentiated resource management, based on the instantaneous traffic conditions, is proposed in this thesis. The network switches between predefined durations for allocating resources, to adjust power consumption to the available load. In this study, even the control node is included in this management strategy to boost the power saving. The mathematical model, algorithm, and the performance with comparative studies are outlined.

- Since both the wireless and optical parts contribute to power consumption and affect the QoS, this research presents the interoperability between the resource management at the optical and wireless parts of the network to enhance both the QoS and energy efficiency. This part of the thesis presents the concept of extending the idle durations at the optical and wireless parts of the network to the limit that an efficient traffic flow can tolerate, to maximize power savings. The mathematical model, the algorithms and the simulation results are outlined.

1.8 Thesis Outline

The organization of this thesis is as follows:

- Chapter 1 introduces the broadband access networks, particularly wireless and optical, and the limitations in their deployment. In addition, the importance of developing a green hybrid network which integrates the optical and wireless parts (WOBAN) is presented.

- Chapter 2 presents the technical background of network architecture, QoS, resource management and energy efficiency for the optical, wireless, and WOBAN. A comprehensive review of QoS and energy efficiency is provided. In addition, a critical overview of the advantages and disadvantages of the strategies implemented for energy efficiency and QoS is provided.

- Chapter 3 proposes a solution to bridge the gap between the wireless and optical networks. As the traffic moves between the wireless and optical parts of WOBAN, service class mapping at the interface becomes a major challenge. This work proposes a solution for the QoS mapper at the interface between the optical and wireless technologies. The queue behaviour and the classes of traffic are investigated and a model is developed to calculate the maximum lengths for different queues. This model ensures that QoS for various end applications is not compromised. This queue model is augmented by a priority scheduler which ensures that traffic is serviced based on its class priority and the QoS requirements of each class. However, this scheduler should not disadvantage the low priority traffic. By differentiating the traffic into classes, this thesis proposes a class-based differentiated sleep to reduce power consumption for terminal nodes in the optical part of the WOBAN. This work has been published in *IEEE/OSA Journal of Lightwave Technology*, vol. 33, no. 1, pp. 7-18, 2015.

- Chapter 4 focuses on solving the problems of power consumption and under-utilization of channels. The challenge is in maintaining the QoS of different types of traffic alongside solving the problem of power saving, and furthermore, making the power consumption proportional with the traffic load. This thesis proposes local and global sleep mechanism based on the type and size of the load. This is achieved by introducing sleep mode for the control node and terminal nodes of the optical part of the WOBAN. In this work, a novel dynamic resources management mechanism is proposed to exploit the fluctuation of daily traffic to maximize the energy efficiency of the WOBAN. This is done by making the wake up and the sleep time of the control and terminal nodes of the optical part of WOBAN, relate to the quantity and priority of the traffic. This work is published in *IEEE/OSA Journal of Optical Communications and Networking*, vol. 7, no. 7, pp. 669-680, 2015.

- Chapter 5 proposes the coordination of the resource management for the optical and wireless parts to enable longer sleep periods and maximize energy efficiency for all of the WOBAN. The resource management at the wireless part follows the dynamics in the optical part or relies on monitoring traffic in the local cell. Thus, this work is based on adapting the dynamic resource management at both parts of WOBAN, to traffic changes during the day in order to achieve the goals of supporting QoS and increasing the energy efficiency. This work has been submitted for publication at *IEEE/OSA Journal of Lightwave Technology*.

- Finally, Chapter 6 summarises and concludes the contributions and significance of this research in developing an environmentally friendly WOBAN. Furthermore, several extensions and directions for future research are recommended.

Chapter 2

Background and Literature Review

This chapter largely addresses the problem of realizing the integration between optical and wireless systems, supporting QoS and developing energy efficient mechanisms for this hybrid network. QoS is highly dependent on the designs deployed for QoS mapping, queue management, scheduling and resource management. All of these factors affect the behaviour of the WOBAN which integrates wireline and wireless network technologies. High capacity wireline networks use optical technology, with the passive optical network (PON) being the most popular as it cost effectively achieves high levels of bandwidth, coverage and reliability. For these reasons, PON has replaced T1/E1 and digital subscriber line technologies. The Fourth Generation (4G) wireless network is currently the superior option for wireless broadband access and it is designed to support high data rates for downstream and upstream traffic. WOBAN [6], which integrates the 4G as the last mile network and 10G-EPON as the back haul network, has the advantages of both the optical and wireless networks; however, it faces QoS and energy efficiency challenges.

This chapter provides an overview of the relevant background theory and literature. Firstly, an overview of the access networks, i.e. 4G and PON, which form the WOBAN, is presented, with particular emphasis on architecture, important components and different standards and technologies. Section 2.4 briefly explores resource management strategies for these networks. Section 2.5 presents the state of the art research and related works for developing energy efficiency mechanisms and QoS for those access networks. Finally, Section 2.6 summarises the major ideas presented in this chapter.

2.1 4G Technologies

The prevalence of broadband services and limitations of earlier technologies drove the introduction of 4G systems, which are commonly referred to as fourth generation wireless technology. Orthogonal Frequency Division Multiplexing Access (OFDMA) and Multiple-Input and Multiple-Output (MIMO) technologies heightened the performance of 4G. OFDMA is a robust solution for the intra-cell interference problem and supports multiple subscribers. MIMO Technology [20] is a key strength of 4G, which significantly improves the spectral efficiency. 4G supports both Time Division Duplex (TDD) and Frequency Division Duplex (FDD) as the modes of operation. In TDD, the upstream and downstream channels share the same frequency, but at different intervals, while in FDD the upstream and downstream channels have different frequencies. 4G supports many types of traffic from VoIP to streaming video and online game services, and provides high data rates up to the IMT-Advanced constraints (1 Gbps and 100 Mbps for fixed and mobile use, respectively).

There are two commonly used 4G standards, Worldwide Interoperability for Microwave Access (WiMAX2) [21] and Long Term Evolution-Advanced (LTE-A) [22]. Both LTE-A and WiMAX2 are based on multiple access technologies in downstream and upstream transmissions. LTE-A uses OFDMA and Single Carrier-Frequency Division Multiplexing Access (SC-FDMA) in downstream and upstream transmission, respectively. WiMAX2 uses OFDMA in both upstream and downstream transmissions. The Medium Access Control (MAC) layer protocol deploys TDM and TDMA for the downstream (from base station (BS) to subscriber stations (SSs)) transmission and upstream (from SSs to BS) transmission, respectively. Both TDM and TDMA technologies use time division to distribute the available bandwidth among a number of subscribers. TDMA combines multiple signals from different SSs into a single channel. Whereas TDM multiplexes signals come from the BS into the channel. The upstream channel is shared among the SSs while the downstream channel is a broadcast.

Table 2.1: QoS parameters for IEEE 802.16 [1].

Class of Service	QoS Parameters	Application
UGS	constant bit rate, max. packet dropping rate,	VoIP
	max. jitter and max. latency	
ertPS	guaranteed min. data rate,	Silent suppressed VoIP
	variable bit rate and max. latency	
rtPS	guaranteed min. data rate,	MPEG video,
	variable bit rate and max. latency	video Conference
nrtPS	guaranteed data rate	FTP
BE	no requirements	E-mail, HTTP

2.1.1 WiMAX Technology

The IEEE 802.16 commonly known as WiMAX is a wireless standard for a Metropolitan Area Network (MAN). In 2001, the standard IEEE 802.16a was launched for fixed broadband wireless access, while in 2004 IEEE 802.16d was published. In the following year, the mobile WiMAX (IEEE 802.16e) emerged and was considered as a pre-4G technology. The latest amendment for WiMAX is the IEEE 802.16m (WiMAX2) which was released in 2012 and is considered as a 4G technology. WiMAX2 is designed with a physical layer operated at a frequency less than 6 GHz and is capable of supporting high data rates. The MAC of WiMAX is connection oriented and supports two modes of operation: mesh mode [23], which permits the communication between two subscriber stations, and Point-to-Multi-Point mode (P2MP), which depends on using a base station (BS) to govern the subscriber stations. WiMAX defines different service flows for the connections and each service flow has specific QoS parameters, such as jitter, delay and throughput. Each cell is constructed from several subscriber stations which are connected to a BS. BSs are responsible for resource management and bandwidth allocation. WiMAX can support different applications by using multiple types of service flow. WiMAX supports QoS by differentiating the traffic into five service flows and allocates bandwidth per service flow per connection. The services supported by the IEEE 802.16m [24] are: 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 highest priority is UGS service, which supports a constant bit rate (CBR) and delay sensitive application, i.e., VoIP. In the second priority is rtPS, which is a variable bit rate (VBR) and delay dependent application such as streaming video. Then, ertPS which is characterized as VBR and supports delay sensitive flows such as VoIP with silence suppression. The nrtPS is also VBR, but supports non-real time applications, i.e., FTP. The BE is best effort such as E-mail and web browsing. BE and nrtPS have the lowest priority because they do not require a set quality of service. Table 2.1 illustrates the QoS parameters and applications, for the five service classes which are supported by WiMAX.

2.1.2 Long Term Evolution-Advance (LTE-A) Technology

Long Term Evolution (LTE) is a 3rd Generation Partnership Project (3GPP) standard which evolved from the Universal Mobile Telecommunication System (UMTS). This wireless technology started with Release-8 [25] [26] which provided a data rate and latency of about 300 Mbps and 5 ms, respectively. The physical layer supports different modulation techniques such as QPSK, 16 QAM, 64 QAM and OFDM. LTE is equipped with a two-layer retransmission algorithm to capture the errors in transmissions [27]. This is a high speed wireless technology that enables the migration of an Internet application from fixed to roaming such as VoIP, HDTV and video streaming. Unlike WiMAX which needs to install its own network infrastructure, LTE uses the installed mobile telecommunication system infrastructure to provide real and non-real time services. Therefore, the challenge for LTE is increasing the throughput, reducing the latency and exploiting the existing infrastructure.

Table 2.2: QoS parameters for LTE -A [2].

Service	QoS Parameters	Application
QCI1	constant bit rate, max packet dropping rate, max jitter and max delay	VoIP
QCI2	guaranteed min data rate, max sustained data rate, max packet error rate and max delay	Conversational Video (live streaming)
QCI3	guaranteed min data rate, max sustained data rate, max packet error rate and max delay	Real Time Gaming
QCI4	guaranteed min. data rate, max. sustained data rate,	Non-conversational Video
QCI5	max packet error rate, max delay	IP Multimedia Sub-system signalling
QCI6	max packet error rate, max delay	Video (buffered streaming), FTP and Email
QCI7	max packet error rate, max delay	Voice, Video (live streaming)
QCI8	max packet error rate, max delay	Video (buffered streaming), FTP and E-mail

LTE-A architecture evolved from LTE and provides a variety of coverage including macro cell and indoor. This architecture is typically constructed from three domains: Evolved Packet Core (EPC) as the core network, Evolved-Universal Terrestrial Radio Access Network (E-UTRAN), and User Equipment (UE). UTRAN consists of a base station, called “evolved NodeB (eNB)” which is responsible for radio resource allocation and air interface control. LTE-A is based on bearers as information carriers, and is a virtual connection of predefined performance. To maintain the quality and priority of the data packets, the bearer informs the networks on handling these packets. LTE-A defines eight levels of QoS class identifier (QCI) [28]. These QCIs are classified as either guaranteed bit rate (GBR), which requires a minimum bit rate such as QCI1-QCI4, or Non-GBR such as QCI5-QCI8. For each QCI, the priority, packet delay budget and packet loss rate are defined. When the user equipment establishes a connection with the eNB, it is assigned a non-GBR bearer which supports “always on” connectivity. Because the user can run different applications on the equipment, a dedicated bearer is assigned to each level of QoS. Each user can establish a maximum of three signalling bearers and 8 data bearers [2].

There are different types of services that are supported by LTE-A with the highest priority, based on delay perspective, being real-time gaming (QCI3) which requires very tight end-to-end delay (50 ms) followed by QCI1 and QCI2 which require 100 ms and 150 ms end-to-end delay, respectively and the lowest is TCP based services (QCI8) e.g. email which is BE [28]. Table 2.2 illustrates the QoS parameters and application of the services which are supported by LTE-A.

Table 2.3: PON Standards (Note XG-PON uses X, the Roman numeral for 10.)

Classification	Standard	Bandwidth
APON/BPON	ITU-T G.983	622 Mbps/ 155 Mbps
GPON	ITU-T G.984	2.5 Gbps/ 1.25 Gbps
EPON	IEEE 802.3ah	1.25 Gbps/ 1.25 Gbps
10G-EPON	IEEE 802.3av	10 Gbps/ 10 Gbps
XG-PON	FSAN/ITUT-T Rec. G.987	10 Gbps/ 2.5 Gbps

2.2 Passive Optical Networks (PON)

PON is a class of optical technology for which there are different standards as shown in Table 2.3. The main components of PON are the Optical Line Terminal (OLT) and Optical Network Units (ONU)s as shown in Fig. 2.1. The interface between the PON and

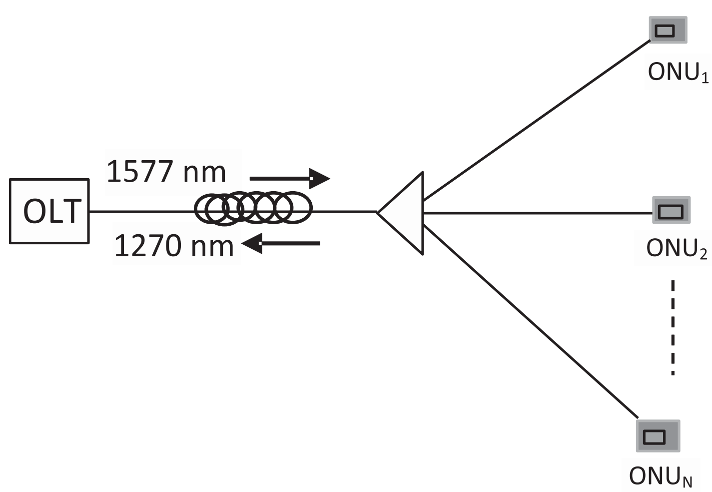

Figure 2.1: PON Architecture

backbone network is the OLT, which for each PON network, is located at the central office of the service provider and is connected to N ONUs through a 1:N splitter/combiner. The split ratio is between 4 and 64, and the maximum distance between OLT and ONU is 20 km [29]; both factors are affected by the power budget. Fibre cables connect the OLT to the splitter, and the splitter to the ONUs which are located at or close to the subscriber premises. In Time Division Multiplexing (TDM) PON, two optical channels are used; downstream and upstream. The data flow from the OLT to the ONUs is called downstream while the reverse flow is called upstream. The OLT's main function is controlling the transmission from and to the ONUs. The OLT broadcasts the data and only the intended ONU obtains the data while the other ONUs discard it. In upstream, the OLT controls the transmission from the ONUs via a Bandwidth Allocation (BA) mechanism. As PON is Point-to-Multi-Point (P2MP), its MAC layer is based on Multi-Point Control Protocol (MPCP) to arbitrate the ONUs access to the upstream channel. The MPCP supports messages namely: REPORT and GATE of 64-byte to control the upstream transmission [30]. The REPORT message carries the ONU's bandwidth request and the GATE message carries the allocated transmission window for the ONUs. The OLT uses Time Division Multiplexing Access (TDMA) to assign each ONU a time slot when they are allowed to transmit. The MPCP also performs auto-discovery, registration and ranging processes [31] [32], in the last of which, the OLT measures the Round Trip Time (RTT) between itself and each ONU and keeps the RTTs of all ONUs for later use in appropriately calculating the allocated time slot for each ONU. It also applies synchronization which prevents potential collisions due to concurrent transmission from multiple ONUs.

To cope with increasing capacity demands, the Institute of Electrical and Electronics Engineers (IEEE) and the Full Service Access Network (FSAN) sought to expand of PON's capacity, so the concept of New Generation (NG) PON standards was proposed. The trends for NG PON include increasing the bandwidth, extending the coverage, and enlarging the number of users. NG PON encompasses the various standards presented below.

2.2.1 New Generation (NG) PON

XG-PON and 10G-EPON (shown in Table 2.3) represent the first approaches to NG PON which use TDM, evolved from Giga (GPON) and EPON, respectively [33]. Thus, XG-PON and 10G-EPON coexist with GPON and EPON on the same fibre plant, respectively. XG-PON provides 2.5 Gbps and 10 Gbps for upstream and downstream respectively, and 10G-EPON can provide a symmetric bandwidth of 10 Gbps [34].

The second approach to the NG PON is Wavelength Division Multiplexing (WDM) PON, which uses multiple wavelengths to increase the available bandwidth for upstream and downstream communication. In WDM PON, each ONU is assigned a pair of dedicated wavelengths, one for downstream and another for upstream to achieve Point-to-Point (P2P) connection. Effectively, several virtual PONs share the same physical infrastructure. In WDM PON, the ONUs are multiplexed to different wavelengths by Arrayed Waveguide Grating (AWG) which replaces the TDM passive splitter used in TDM PON architecture [35] [36]. Coarse WDM (CWDM) PON is the cheaper WDM PON which is constrained by eight channels because it uses a gap of 20 nm between channels [12].However, Dense WDM (DWDM) only utilizes a gap of 0.2 nm to 0.8 nm between channels and thus uses a huge number of wavelengths. CWDM PON and DWDM PON are used in metro networks and long-back haul connection [37]. WDM PON can provide a raw bandwidth of more than 10 Gbps, and is better than TDM PON in three aspects: there is no sharing of wavelength channel capacity because it is dedicated to only one user; synchronization is not required between wavelength channels; and it is more secure because no one can access another's channels. However, it is an expensive choice. Another solution for the NG PON is Optical Code Division Multiple Access (OCDMA) PON, in which, the users share the wavelength by using their allocated optical codes as signatures and coding and decoding are performed in the optical part [12].

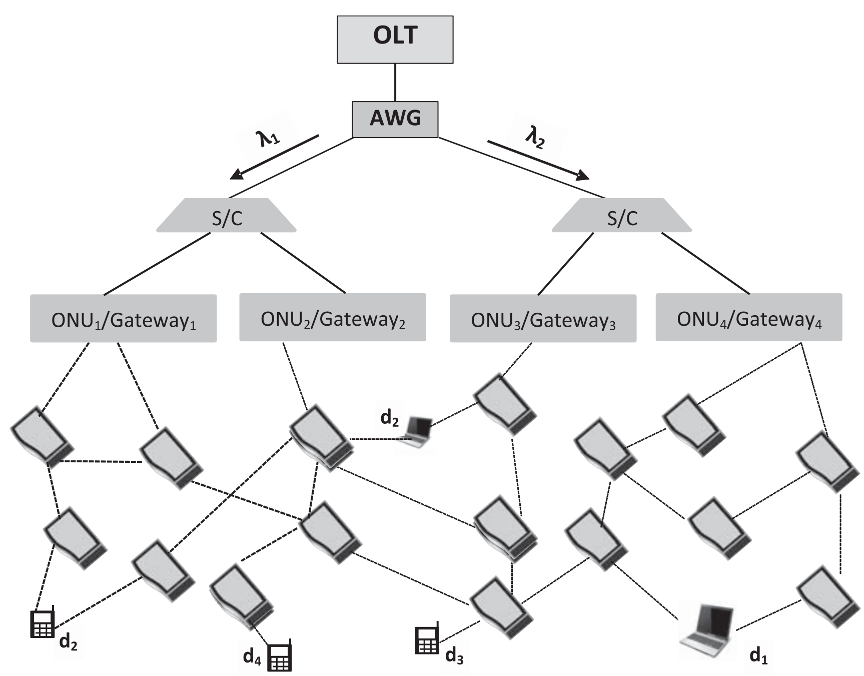

Figure 2.2: The connectivity layout in fibre-Wireless (FiWi) [4].

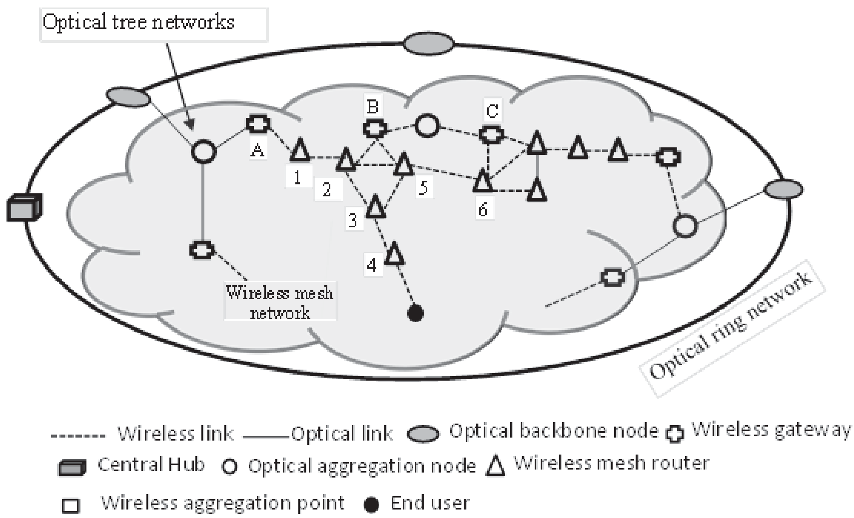

Figure 2.3: The architecture of an integrated network of TDM/WDM PON and wireless mesh proposed by [5].

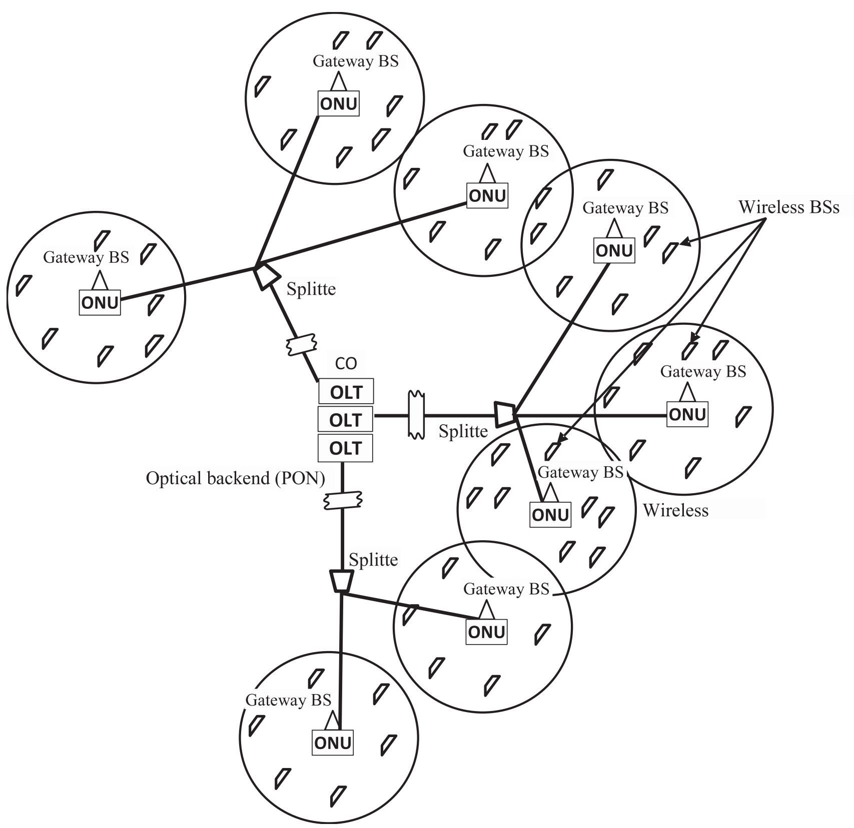

Figure 2.4: WOBAN architecture based on a mesh wireless part [6].

Other approaches to the NG PON involve hybrid PON such as WDM/TDMA PON [38] and the combination of OCDMA with CWDM [12]. In the former hybrid, the bandwidth over the wavelength is shared between various users through TDMA, while the latter can support many more subscribers. Another approach is Long Range PON (LR-PON) which extends the maximum distance to 100 km and uses many splitters which are each connected to 256 ONUs, so this PON can connect a few thousand ONUs. An optical amplifier is used because the losses for long range propagation and for splitting need to be substituted [39]. Using LR-PON will reduce the required number of OLTs and thus minimize capital and

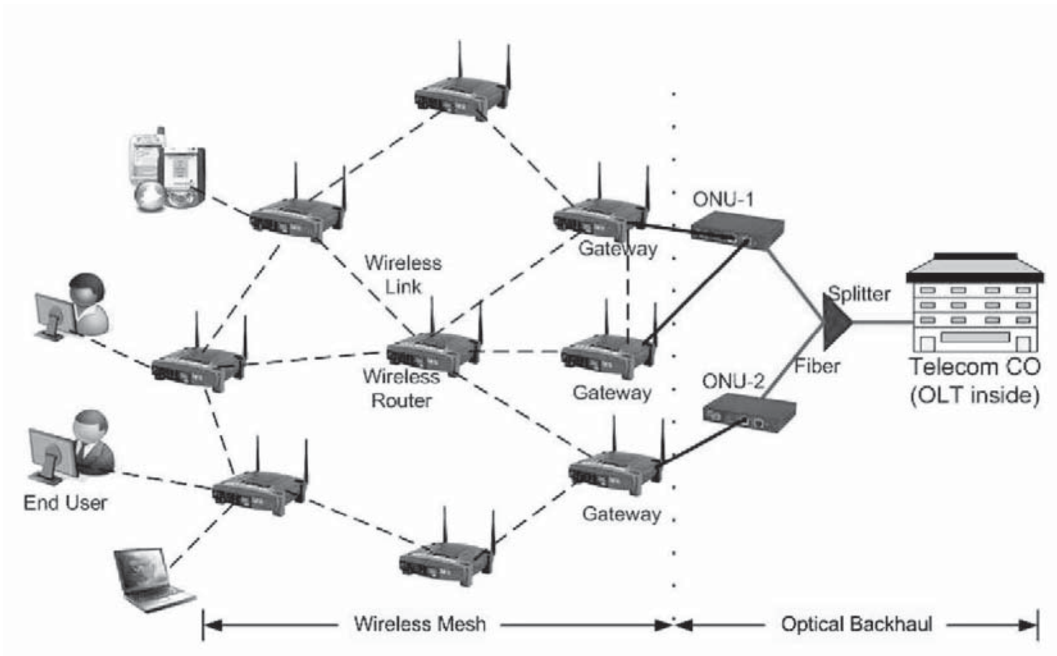

Figure 2.5: WOBAN architecture [7].

operational expenditure while increasing the number of connected users [40] [38].

2.3 Hybrid Access Networks

The motivation for using hybrid networks is to aggregate the advantages of different technologies in one network and the dominant ones are constructed from a Passive Optical Network (PON) and a wireless network, because they provide reliability, high bandwidth and ubiquity. For example, FiWi (Fig. 2.2) [4] integrates WDM/TDM PON with wireless mesh network to reduce the effects of interference in the wireless part by transmitting data between two wireless subscribers through the optical part instead of a multi-hop wireless network. In [5], TDM/WDM PON was integrated with a mesh wireless network as shown in Fig. 2.3; the reconfiguration scheme at the optical part was employed to balance the load and enhance the resource utilization. Figures 2.4 [6] and 2.3 [7] show WOBAN architectures constructed from the PON and mesh wireless network. The concerns of researchers developing this network are the routing mechanisms to minimize delay [41] and network planning [42]. In these studies, traffic differentiation was not considered providing that QoS is not guaranteed, since mesh wireless network is based on WiFi technology (IEEE 802.11 standard).

Another study addressed the challenges of integrating LTE with WDM PON [28]. The latter was chosen to avoid the typical (TDM) PON topology, as that topology did not support a full distribution access network, and to isolate OLT from controlling the upstream transmission, and detecting the failure in fibre cable because it is time costly. Therefore, the authors in [28] proposed an architecture for integrating a network based on deploying ring fibre to connect the distributed eNBs, each of which is integrated with ONU, while the OLT is connected to the EPC's Access Gateway (AGW). The author did not provide details of the queues apart from using one-to-one mapping with eight QCI of the eNB but how to calculate the queue size, the expected delay, and the behaviour of the end-to-end delay, jitter and throughput is missing as they only proposed a methodology.

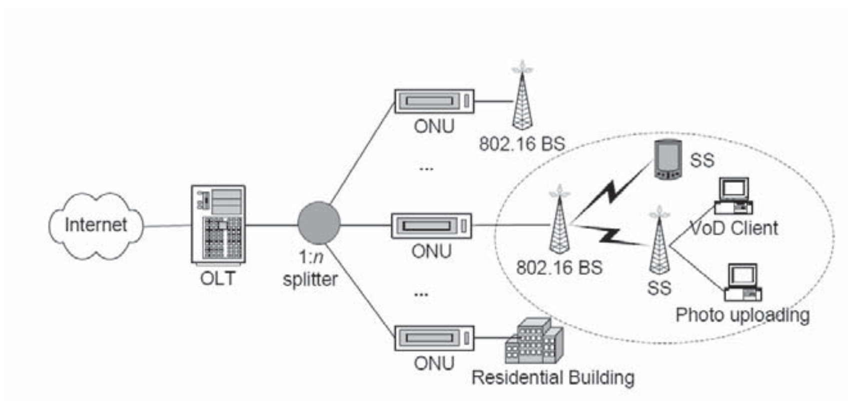

Figure 2.6: WOBAN architecture based on multi-point to point wireless part [8]

EPON can also be integrated with WiMAX which is facilitated by the similarity between the two technologies in terms of bandwidth allocation. Four ways for integrating the aforementioned technologies were proposed by [43] namely: independent, hybrid, unified connection-oriented and Microwave-over-fibre architectures. A WOBAN based on P2MP wireless network such as WiMAX (Fig. 2.6), was proposed in [8] [44], but QoS was a major concern. Jung *et al.* in [44], proposed centralized scheduling at the OLT to schedule the traffic from the SSs at the wireless part, and hybrid integration between the BSs and the ONUs is used. The OLT assigns a transmission window to the ONU based on the aggregated requests, from the SSs associated this ONU, which are forwarded from the BS to the ONU and then the OLT. Yang *et al.* in [8], proposed a two stage bandwidth allocation, based on using a virtual ONU-BS unit whose role was to control the bandwidth allocated for the BS which in turn dictates the allocated bandwidth per connection to the SSs. The proposed mechanism for bandwidth was based on reserving a percentage of bandwidth to schedule the best effort traffic. Both studies [8] [44] used priority queues (the former used three and the later used eight queues) to accommodate the different services of the WiMAX, however, they did not investigate the queue issues such as the queue length and the expected waiting time for packets in these queues which are crucial to maintain the QoS. Their strategies were based on informing the OLT about the requested bandwidth of the SSs to maintain the QoS. These strategies however, increase the overhead by not solving the QoS problem from the queues at the ONU because they are used to send messages between the BSs and the ONUs then from the ONUs to the OLT to inform the OLT how much bandwidth is required by each BSs. To this end, the focus of this thesis is around an architecture based on using P2MP wireless topology connected to the optical network.

2.4 Resource Management

Network performance is greatly affected by resource management with bandwidth being a vital resource that all of the connected users share. The available bandwidth and the allocation of this bandwidth amongst users vary for different technologies. The following sections discuss the resource management for the PON, 4G and WOBAN.

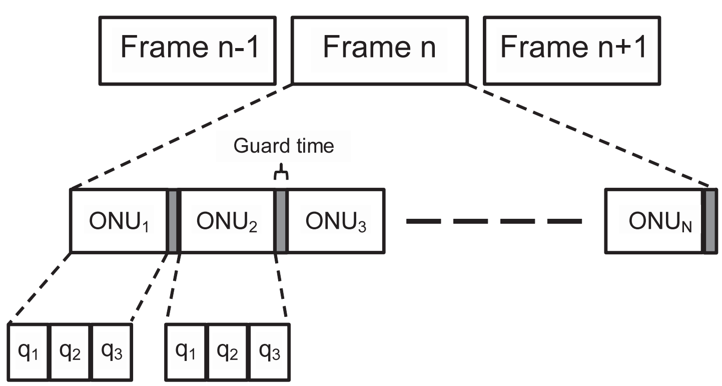

Figure 2.7: Frame structure for 10G-EPON.

2.4.1 Resource Management in PON

In PON, all ONUs share access to the upstream channel. To avoid collision when two ONUs intend to transmit at the same time, the medium access control of PON uses a scheduling mechanism to control the ONUs' access to the upstream channel [33]. In PON, there are two schedulers; one at the OLT (Inter-ONU) and the other at the ONU (Intra-ONU). The Inter-ONU scheduler arbitrates the order of the ONUs to access the upstream channel and the Intra-ONU scheduler arbitrates the traffic from different queues at each ONU, while for the collective scheduling of the ONUs, the time for which each ONU can transmit is important. In TDMA scheme, the OLT assigns a time slot called a grant (transmission window) for each ONU based on the polling mechanism; ONUs request their desired amount of bandwidth and the OLT decides the respective amounts and assigns it. This assignment is periodic and the granted transmission window depends on the length of the polling cycle and the bandwidth requested by individual ONUs as shown in Fig. 2.7. The OLT uses either Fixed Bandwidth Allocation (FBA) or Dynamic Bandwidth Allocation (DBA) [31] [45]. In FBA, fixed grants are assigned for the connected ONUs, so the polling cycle duration is also fixed and the ONUs do not need to request any bandwidth as the OLT assigns the same size transmission window for all ONUs. In this case, the granted transmission window might be underestimated or overestimated for the heavy loaded and light loaded ONUs, respectively. The underestimated transmission window wastes bandwidth while overestimated transmission window causes more delay and packet dropping and both of them affect channel utilization and network performance. DBA can enhance the network performance as the OLT uses statistical multiplexing which considers the bandwidth requested by ONUs, and the available bandwidth. So, in DBA, the grant size varies with the requested bandwidth, and so the polling cycle duration also varies. To prevent the highly loaded ONUs from monopolizing the channel and affecting the service of the other ONUs, the maximum grant size technique is important and will help predict the polling cycle duration. In contrast, lightly loaded ONUs result in more frequent polling which wastes more power. To this end, the bandwidth allocation mechanism plays a crucial role in network performance.

2.4.2 Resource Management in 4G

In 4G networks, the geographical cell is constructed from the BS and the SSs. The BS controls the resource management which uses TDD or FDD to allocate the resources among the SSs. In TDD, the upstream and downstream channels share the same frequency while

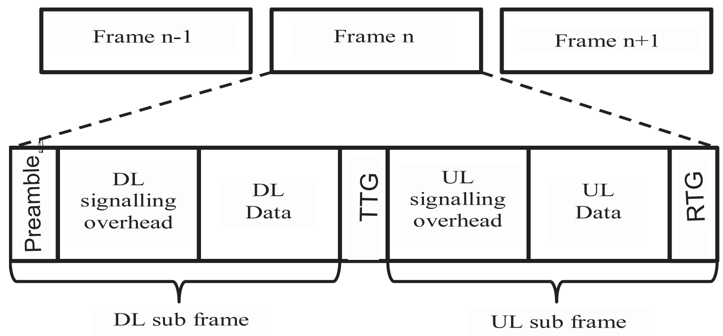

Figure 2.8: TDD frame structure of IEEE 802.16 standard.

in FDD each channel has a different frequency. Figure 2.8 shows the frame structure for the IEEE 802.16 standard. This frame is of fixed duration and composed of two sub-frames; downstream and upstream. In P2MP operation, the transmission is between the BS and SSs. For upstream transmission, BS allocates resources among SSs by slicing the upstream sub-frame, so it assigns time slots for each SS. As noted in section 2.1, the transmission method in the downstream communication, which is from BS to SSs, is based on broadcasting. The BS informs all the SSs about their upstream and downstream time slots by broadcasting DL-MAP and UL-MAP control messages. From these control messages, the SSs learn about their transmission and receiving windows [24]. 4G networks are superior in QoS provision by supporting class differentiation of traffic as they assign each connection to a predefined QoS service. These QoS requirements affect the scheduling and the radio resource allocation. As mentioned earlier, WiMAX supports five types of services and each requires different QoS parameters while LTE supports eight types of services which are also require different QoS.

2.4.3 Resource and Traffic Management in WOBAN

As WOBAN is a hierarchical network, there are two stages of resource management; at the BS and the OLT. The efficiency of both stages contributes to the behaviour of the whole network. In WOBAN (Fig. 2.6), the ONU architecture includes single or multiple queues, which accommodate traffic until it is dispatched. Mapping between different classes of traffic at the wireless part and these ONU queues is challenging. Researchers have attempted to address this problem from different viewpoints to arbitrate the appropriate number of queues in ONU. In [43], Gangxiang *et al.* suggested eight priority queues at the ONU to accommodate traffic from the BS. However, they did not propose any model for calculating the appropriate length of each queue and scheduling packets from these queues. In [46], 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 a single pre-emptive queue for both services, but did not address the queue length estimation problem. Another study [44] 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 modelling. For bandwidth allocation, Yang *et al.* in [8] suggested a virtual ONU-BS unit to achieve QoS-aware dynamic bandwidth allocation and their proposed model involves more than one priority queue at the ONU. Although their study considered the signalling of bandwidth allocation, they did not specify the queues apart from highlighting the bandwidth requested for each queue. In [45], 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 did not address the queue length and its relation with the resource management, which affect the QoS.

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, few studies have addressed this issue. In [47], Kramer *et al.* investigated queue behaviour and size for the EPON network, but did not consider the different classes of traffic and the QoS. Obele *et al.* in [48] studied queue behaviour in a WiMAX-GEPON network, by deploying one-to-one mapping of only four services between the two technologies. These authors had to arbitrate the queue size to solve the embedded Markov chain and thus estimate the delay and the packet dropping rate in the queues. This approach is impractical as it is only applicable for small size queues of a few packets. Moreover, these authors did not consider a bandwidth allocation mechanism which affects the allocated transmission window, when calculating the waiting time at the queues. Also they used strict priority for scheduling the packets and thus low priority packets may suffer excessively long delays and high dropping rates.

2.5 Energy Efficiency

Since climate change and operating cost have become an increasingly important issue, researchers have sought to develop energy efficient access networks. The following sections discuss the recent research in this field.

2.5.1 Energy Efficiency in PON

As power consumption scales up with the transmission rates, one energy conservative approach involves using an Adaptive Link Rate (ALR), to adapt the transmission rate to the traffic load in Ethernet networks [49]. As WDM PON supports multiple wavelengths where each user is assigned a wavelength, the authors in [50] proposed using this technology (WDM PON) in which different transmission rates are carried by different wavelengths. The data rates are based on the applications run by the end users. In TDM PON, the authors in [51] exploit the compatibility between EPON and 10G-EPON to propose the switching between 1 Gbps and 10 Gbps, by incorporating transceivers for 1 Gbps and 10 Gbps into the architecture for the ONU and OLT. However, this proposal involves the complexity of just when to make the decision to switch between these two transmission rates. Those authors in [51] proposed a sleep mechanism for the ONUs that enables sleep and periodic wake up based on a three way handshake protocol between the ONU and OLT. Their work was supposed to use DBA as a mechanism for resource management; however, there is a lack of results that show the effects of the resource allocation parameters on the QoS and power savings. Moreover, the authors proposed different sleep schedules for different classes of traffic, but there was no sleep when high priority traffic occurred. Since the OLT is a significant power consumer, in [52] authors proposed connecting multiple OLTs through optical switches to turn off the OLT under a light load scenario so that another OLT could serve the network. Apart from the delay issue, which is caused by the multiple stages of switching, this approach is costly because it requires additional optical switches and cables. To develop an energy efficient ONU, Wong *et al.* proposed two ONU architectures to support sleep mode and minimise the overhead of clock recovery (from 2-5 ms to 125 μs) when ONU switches from sleep mode to active mode [53]. Since maintaining the ONU in an idle state wastes energy, another approach used by researchers to improve energy efficiency is adopting sleep mode for the ONU when there is no traffic. Mandin [54] proposed to initiate the sleep time after a three way handshake protocol between the ONU and OLT to ensure that there is no upstream or downstream traffic, and sleep could be interrupted by the arrival of traffic. This necessitates a cyclic wake up for ONUs to check for upstream traffic before returning to sleep if there is none. Conversely, if the ONU detects any packets in its queues, it informs the OLT by requesting bandwidth, which eventually cuts off the sleep phase and the ONU resumes normal work.

Ying *et al.* [55] proposed two mechanisms to put the ONU into sleep mode; one is based on the scheduling of upstream traffic and the other on downstream traffic. In this work [55], the OLT schedules the sleep and wake up time for the ONUs using a message similar to the GATE. In [56], the authors proposed a strategy for switching from DBA to FBA when the load becomes low, so that the ONU can enter into sleep mode and periodically wake up to transmit and receive traffic. These authors also mentioned that the transmission rate could change between 1 Gbps, 100 Mbps and 10 Mbps based on the load, but they did not describe how the transition between these different rates can be achieved. In their work, the traffic is not differentiated so if the traffic is low the system will turn to FBA and a low rate is chosen even if the traffic contains voice or video. Hence, performance could be degraded which indicates that QoS was not considered. Furthermore, queue issues such as size and expected waiting time were not addressed despite their impact on the network behaviour especially when long polling cycles that are used. In [57], 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. The authors sought to put the ONU into sleep after it receives the GATE message and is informed about the transmission window details; the ONU wakes up when the transmission window is initiated. In their work, the authors also did not differentiate the traffic. In [58] [59], the researchers based the decisions for transition to sleep time and the duration of sleep time, on traffic conditions. In [58], Zhang *et al.* proposed four levels of power saving and the transition between these power states depending on traffic conditions. In their work, the sleep duration is either within one or more DBA cycles and the maximum sleep period is 50 ms. Despite the use of DBA for resource management, which indicates the authors' concern for QoS, 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 and then there was a controlled increase in duration up to 50 ms. However, the authors [58] only proposed the framework, but did not address the queue issues and did not differentiate the traffic. In contrast, a sleep cycle was proposed in [59], based on classes of traffic for the downstream transmission, in which the arrival of high priority traffic wakes up the ONU. These authors proposed adaptive sleep duration based on measuring the downstream traffic situation periodically. In their study, the negotiable sleep period is implemented through a three way handshake mechanism between the OLT and ONU and the arrival of the upstream traffic which makes the ONU ignore the OLT request for sleep. At the end of each sleep duration, the ONU wakes up, then transmits if there is any traffic or otherwise goes back to sleep. In that research, the arrival of high priority traffic stops the sleep mode and hence, limits power savings in a practical scenario. Overall, there are challenges associated with the sleep mode such as when to initiate the sleep time cycle, the appropriate sleep time duration, the effect of the recovery time for transition from sleep to wake up, and how should registration of the ONU be maintained. Although researchers have attempted to address QoS in PON, none of these studies concurrently addresses both service class mapping and queue modelling at the ONU in WOBAN. There is also no available model that deals with the power saving issue and QoS at the ONU. This thesis has addressed the above mentioned problems and proposes efficient solutions.

In general, the incorporation of sleep mode within resource management is important to maintain the performance of the network and the QoS. The servicing of the traffic from each ONU is informed by the GATE message, therefore, if the ONU is in sleep mode when the OLT advertises the GATE, the ONU will miss its allocated time slot and its traffic must wait. To ensure the servicing of traffic at the ONU, the receiver should wake up to receive all messages and data, then obtain the details of the allocated transmission window.

2.5.2 Energy Efficiency in 4G

As the BSs are the biggest power consumers in the wireless network [60], one of the green approaches in wireless networks is turning off, or putting the lightly loaded BS in sleep and letting the neighbouring BSs take over the traffic [61] [62] [63]. All these techniques are based on a long sleep strategy (turning off); however, the working BSs and SSs still waste power due to the idle durations for listening to the radio channel when there is no traffic to receive. These idle durations only can be controlled within the resource management and specifically, a cyclic sleep strategy is needed to solve the problem of idle listening and to support QoS.

2.5.3 Energy Efficiency in WOBAN

There have been some studies conducted to improve the energy efficiency of WOBAN. Chowdhury *et al.* [7] proposed using routing protocols. These authors sought to green the WOBAN architecture, based on a wireless mesh network, by re-routing the traffic in the wireless part to achieve the two requirements. The first was to increase reliability against fibre failure, ONU failure or failure in the wireless part (routers and gateways). The second was to turn off the ONUs with light loads to save more energy. This approach relies on the assumption that routers require less power than ONUs. However, the wireless part was not considered in that energy efficiency and still wasting power when underutilised. Therefore, the authors in [64]proposed to turn on only the radio interface part of the wireless node to perform the rerouting and turn off the other components of the nodes which are unused parts when the attached ONU is in sleep mode. Another study [65] proposed to use power saving mechanisms for the all integrated network. Power saving mode and sleep mode were proposed for the wireless and optical parts of the integrated network, respectively. These authors coordinated the sleep time of the ONU according to the access point that the ONU was attached to. However, their study considered the downstream traffic only, so the timing signals between the ONU and OLT were ignored.

All the above studies are based on wireless mesh network, not 4G technology and point-to-point connections between the base station and the wireless subscribers. As such the QoS is not considered in these proposals because the used wireless technology has not differentiated and defined the QoS for different services like 4G. Moreover, the amount and traffic profile was not considered in their energy efficiency which represents a major point for energy deficiency. This research addresses these signeficant gaps in the field of supporting QoS.

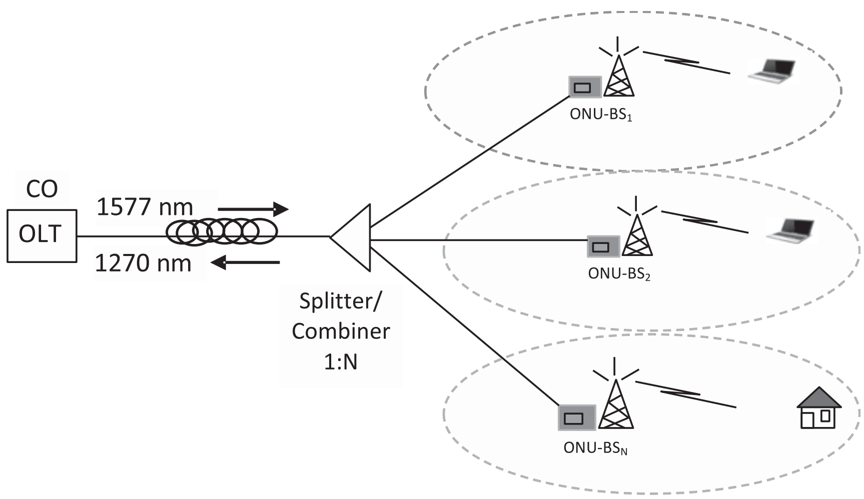

Figure 2.9: A WOBAN architecture.

2.5.4 Research Questions

This thesis targets a WOBAN that integrates 10G-EPON and 4G as shown in Fig. 2.9, since it is collectively superior to 10G-EPON and 4G as previously mentioned. Indeed, the aim of this integration is the efficient delivery of broadband services, including voice, video and best effort, to the end user. In this network, the integrated ONU-BS unit is crucial for realising the integration and provisioning of QoS. All traffic between the wireless users and the OLT passes through this hybrid unit (ONU-BS). In upstream, each BS gathers the traffic from the wireless users and forwards it to the associated ONU. The ONU then sends it to the OLT located at the central office to connect the WOBAN to the backbone network. Therefore, the design of the ONU-BS ultimately affects the performance of the WOBAN and represents a major challenge especially with the diversity of traffic and the QoS requirement.

Clearly solutions are needed to maximize resource utilization and energy efficiency. In the TDM PON, the durations of serving or waiting states for ONUs change with the load. As the waiting durations are inevitable and power consumption becomes a big concern, lengthening the waiting time to enable relatively longer sleep should be considered. The questions are: what is the acceptable limit for the waiting periods and how this waiting time can be controlled? Tolerated delays for the available traffic can be considered to define the boundaries for the waiting periods and this could be achieved through the resource management mechanism. To this end, this dissertation mainly focuses on the realization of WOBAN and developing mechanisms, to better perform resource management, that consider both the QoS and energy saving through cyclic sleep.

The objectives of this thesis are as follows:

- Maintaining QoS in the integrated network when there are different QoS interpretations in the constituent technologies. In WOBAN, BS differentiates the services into several specific classes, for example, five in WiMAX whereas LTE-A defines eight QoS Class Identifiers (QCI). In contrast, the ONU architecture includes single or multiple queues, which accommodate the pending traffic until it is released. The challenge with integrating BS and ONU is mapping the traffic at the BS and the queues at the ONU. Furthermore, 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 influence on the QoS of the integrated design of QoS mapper, queue management and resource management at the ONU-BS, there has not previously been enough research to comprehensively address this issue in PON or WOBAN.
- A common approach to reduce the power consumption is enabling a sleep mode. There is no available model that deals with the power savings and QoS issues and considers the existence of different loads at the ONU in PON and WOBAN. Sleep mode has often been used by researchers as a power saving approach when traffic is

absent. While the idea can be useful in scenarios where FTTH or FTTB technology is used for serving a small number of subscriber stations, this is not useful in WOBAN where the ONU is connected to a BS serving hundreds of subscriber stations because there is very little chance that the ONU queue becomes completely empty when serving a Macro BS. As such, a more intelligent power savings strategy is required and this provides the motivation for this research.

- In order to develop an efficient green WOBAN, QoS sustainability and reduced power consumption are essential considerations. However, the problem of reducing power consumption in the QoS support network is challenging, as the network must be service-ready at all times. Sleep periods that are longer than required (in the context of traffic load) result in more power savings at the cost of increased packet dropping rates, jitter and delay, whereas shorter periods cause more power consumption. QoS demands, often expressed in terms of delay, jitter, packet loss requirements, for various traffic classes (e.g., voice, video, data) must also be considered when calculating the appropriate sleep periods for OLTs and ONUs and the class specific load. For example, if the load from real-time service classes is low, ONUs and, where possible, OLTs can be put into sleep for longer durations since delay is not a major concern for non-real-time and best effort traffic. A fixed power saving strategy is also not sufficient because traffic load varies significantly during various periods in a day. As such, the problem that needs to be addressed is - how to develop an efficient resource management scheme that allows adaptive sleep periods for ONUs and OLTs so that power consumption can be reduced without sacrificing the QoS of end applications.

- The sleep mode approach has been widely adopted for wireless and PON networks. No research has been identified which concurrently consider the energy efficiency of both the wireless and optical parts of WOBAN. In the literature, sleep mode has been proposed for only the ONUs yet there are other components that need to contribute to energy efficiency namely, OLT, BSs and SSs. This research aims to extend the idle durations not only for the optical components but also for the wireless part components to maximize the sleep time. This can be achieved by controlling the idle times to mitigate the effect of the overhead. As no study has so far developed a mechanism for QoS and energy efficiency, this thesis will fill this gap by proposing a dynamic resource management, at the optical and wireless parts that simultaneously adapts to dynamic daily traffic.

2.6 Closing Remarks

This chapter has provided an outline of the research undertaken in the course of this study in realizing a green WOBAN. Since this thesis presents mechanisms to enhance the energy efficiency and QoS of a hybrid network of an integrated 10G-EPON and 4G (WOBAN), it is important to demonstrate the factors that relate to this subject. Therefore, an overview of the PON and 4G, with their standards and technologies is given. An overview for the hybrid network of 10G-EPON and 4G was presented. As the WOBAN's QoS and power consumption are affected by the optical (10G-EPON) and wireless (4G) parts, these variables have been strongly emphasised in this literature review. Specifically, 4G is much better than PON at organising different types of traffic according to their QoS requirements.

This research explores the potential of using sleep mode and its impact on the performance of PON, 4G and WOBAN. An investigation and evaluation of the performance of energy efficiency schemes has been conducted. In particular, this research focuses on how to provide appropriate QoS with minimum power consumption.

Chapter 3

Quality of Service (QoS) Support Model for Green WOBAN

It is commonly understood that there are several important factors that influence the deployment of access networks such as the QoS, energy efficiency, mobility, capacity, cost, coverage, and reliability. While the QoS ensures the success of transmission of diverse traffic according to its requirements, the energy efficiency provides the flexibility for reducing the power consumption according to the presence of traffic, which results in lower bills for end users. Integration of PON and 4G technologies, in the form of wireless-optical broadband access networks, offers advantages such as extension of networks throughout 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 QoS. For example, 4G networks use traffic classification for supporting different QoS demands, whereas the PON architecture has no such mechanism to differentiate between various types of traffic. These two technologies also differ in their power saving mechanisms. This chapter proposes a service class mapping for an integrated 10G-EPON-4G network, based on the M/G/1 queuing model. A class-based power saving mechanism which significantly improves the sleep period for the integrated optical wireless unit, without compromising support for QoS is also proposed.

3.1 Architecture of an Integrated 10G-PON and WiMAX Network (WOBAN)

Figure 2.9 presents a WOBAN architecture which integrates the 10G-EPON with 4G standard system. The OLT is the controller and resource manager of the WOBAN. The ONU is connected directly to the base station of each wireless network to construct a

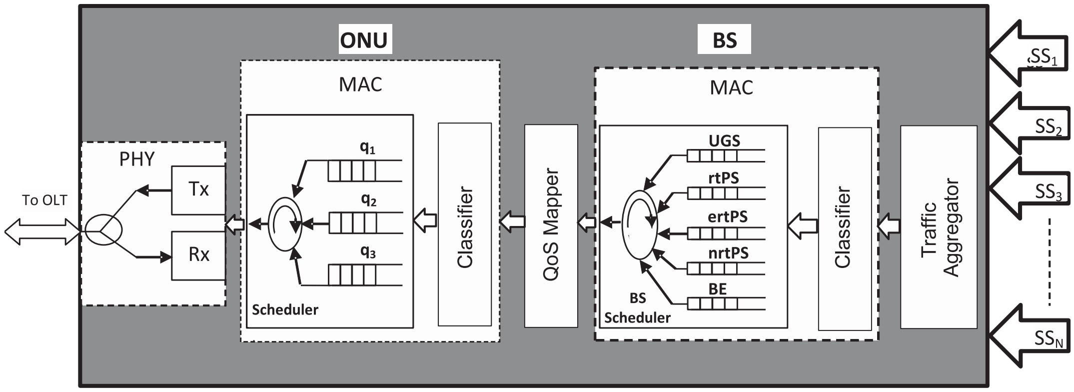

Figure 3.1: Proposed optical network unit (ONU)-base station (BS) architecture.

hybrid unit called ONU-BS/eNB. Therefore, the seamless integration between 10G-EPON and 4G requires a reliable QoS mapper to map the traffic between the two technologies. The BS/eNB controls the transmission of multiple subscriber stations within the cell. Thus, the wireless segment of a WOBAN, in normal operation such as high traffic loads and with no fault in the fibre between ONU and the splitter, is a single hop connection (between SS and BS/eNB). The number of wireless cells is similar to the number of the connected ONUs to the OLT. In WOBAN, the transmission between subscribers within the same cell is controlled by the BS or eNB of the cell, while the transmission between subscribers from different cells is controlled by the OLT. For the upstream, the SS creates a packet and sends it to the BS/eNB of the geographic cell which in turn forwards it to the associated ONU which schedules this packet to the OLT.

The scheduler at the OLT is responsible for scheduling the ONUs by assigning the transmission windows. The ONU is informed of the details of its assigned transmission window by the GATE message (from OLT to ONUs) which tells the intended ONU when and for how long it is allowed to transmit. Because the ONUs should wait until the occurrence of their transmission window, each ONU is constructed with queues to retain the packets until they are transmitted. Moreover, due to the diversity of services that these packets belong to, and the difference in their QoS requirements, the ONUs could use more than one queue to store different packets. Therefore, the scheduling of the packets from different queues needs to be controlled. This control is achieved by the Intra-ONU scheduler, which is located at the ONU. The responsibility of the Intra-ONU scheduler is to nominate the packets and their sequence, which are based on a specific policy, in the next transmission window. In the next sections, the details of the proposed queue management and the Intra-ONU scheduler are described.

3.2 Proposed QoS Management Scheme

ONU and BS adopt dissimilar strategies for handling services that require QoS. For merging the gap in QoS handling, an architecture for the ONU-BS unit is proposed as shown in Fig. 3.1. 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. The queue is also modelled, corresponding to various classes in the ONU unit, and the model shows how to calculate appropriate queue lengths.

3.2.1 QoS Mapping

The queuing system in the proposed ONU-BS architecture is modelled as an M/G/1 system, where M stands for Markovian arrival and G stands for General distribution for the service time and 1 means a single server [66]. The arrival of packets from subscribers is considered as a Poisson random variable with a mean arrival time λ. As different packet lengths require different serving durations, 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. The proposed 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:

-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.

-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.

-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.

This system uses three priority queues at the ONU, because additional queues increase both 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. The arrival of packets in q_1 before or just when a transmission window of ONU has started will qualify those packets to be served since the scheduler did not move to serve q_2. This implementation uses 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. Moreover, the arrival packets in q_1, when the scheduler has moved to serve q_2, will wait until the expiration of this transmission window and the occurence of the next granted transmission window. However, the heavy traffic at q_2 and q_3 will not cause excessive delay for the packets in q_1 because the granted transmission window is bounded by a constraint which is the maximum transmission window.

3.2.2 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 minimised. 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 [67]. A queue with shorter than optimum length reduces delays at the cost of an increased packet dropping rate. The next section addresses 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.

3.2.2.1 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 to, and its 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:

$$E\left(W^{K}\right)=\frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_{k}\times E\left(S_{k}^{2}\right)}{\left(1-\rho_{K-1}^{-}\right)\left(1-\rho_{K}^{-}\right)} \tag{3.1}$$

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 k^{th} 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, $\rho_K^{-}=\sum_{k=1}^{K}\rho_k$; 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:

$$E\left(W^{1}\right)=\frac{\frac{1}{2}\lambda_{1}\times E\left(S_{1}^{2}\right)}{\left(1-\rho_{1}\right)} \tag{3.2}$$

Eq. (3.2) 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 for 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:

$$E\left(W^{2}\right)=\frac{\frac{1}{2}\left(\lambda_{1}\times E\left(S_{1}^{2}\right)+\lambda_{2}\times E\left(S_{2}^{2}\right)\right)}{\left(1-\rho_{1}\right)\left(1-\rho_{1}-\rho_{2}\right)} \tag{3.3}$$

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:

$$E(W^{3})=\frac{\frac{1}{2}\left(\lambda_{1}\times E(S_{1}^{2})+\lambda_{2}\times E(S_{2}^{2})+\lambda_{3}\times E(S_{3}^{2})\right)}{(1-\rho_{1}-\rho_{2})(1-\rho_{1}-\rho_{2}-\rho_{3})} \tag{3.4}$$

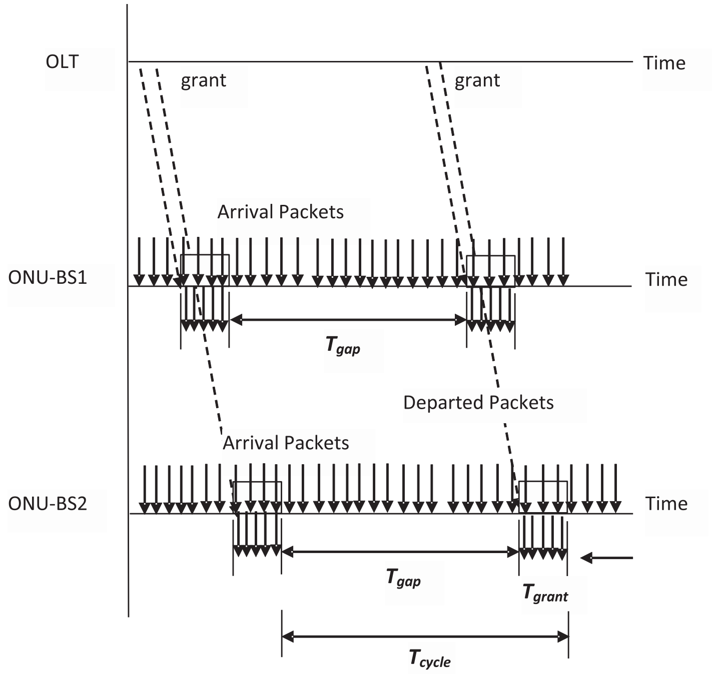

Figure 3.2: The time allocation (grant) of two ONU-BSs in TDM-PON.

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

$$E(S_k^2) = x_k^2 + \sigma^2 \tag{3.5}$$

where x_k denotes the service time for the packets in q_k and is equal to packet size $\times 8$/bandwidth; σ^2denotes the variance of service time and can be adjusted to better reflect the system by using Eq. (3.6)

$$\sigma^2 = \sum_{k=1}^{3} \alpha_k \times (x_k - T_{avg})^2 + \beta \times (T_{gap} - T_{avg})^2 \tag{3.6}$$

In Eq. (3.6), the assumption is that there is a dummy packet with service time equal to T_{gap}, the time period between two successive grant periods (Fig. 3.2) and this T_{gap}, for any ONU in a system consisting of N ONUs, denotes service time of the other (N-1) ONUs and can be calculated as follows:

$$T_{gap} = T_{cycle} - T_{grant} \tag{3.7}$$

$$T_{gap} = (N-1) \times T_{grant} + N \times T_g \tag{3.8}$$

where T_g is the guard time between grant periods of two successive ONUs; T_{grant} and T_{cycle} denote the allocated time for ONU to transmit and polling cycle time, respectively. T_{grant} is fixed in a fixed bandwidth allocation and variable for a dynamic bandwidth allocation. The relationship between T_{grant} and T_{cycle} is as follows:

$$T_{cycle} = \sum_{n=1}^{N} \left[T^n_{grant} + T_g\right] \tag{3.9}$$

Here n denotes the n^{th} ONU.

In Eq. (3.6), T_{avg} denotes the average value of the service time for all types of packets including the dummy packets which has service time equal to T_{gap}. T_{avg} is calculated using Eq. (3.10):

$$T_{avg} = (\sum_{k=1}^{3} x_k + T_{gap})/4 \tag{3.10}$$

In Eq. (3.6), α_k and β are parameters that are linked to the T_{grant} and T_{cycle}, and to the normalized assigned bandwidth (p_k) for queue k (q_k) in the following way:

$$\alpha_k = p_k \times T_{grant}/T_{cycle} \tag{3.11}$$

$$\beta = 1 - T_{grant}/T_{cycle} \tag{3.12}$$

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

From Eq. (3.9), it can be seen that the length of the polling cycle (T_{cycle}) 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 also fixed and can be calculated as follows:

$$T_{cycle} = N \times (T_{grant} + T_g) \tag{3.13}$$

3.2.2.2 Expected Waiting Time Due to the Scheduler Grant

The bandwidth allocation and scheduling mechanisms of the ONUs are centrally controlled by the OLT. Figure 3.2 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:

$$E(T^K) = E(W^K) + c \times T_{gap} \tag{3.14}$$

where $E(T^K)$ denotes the total expected transit time for the packets in k^{th} queue; c denotes a constant $\leq$0.5, which is linked to the differences in waiting time in relation to grant times that are 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).

3.2.2.3 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:

$$L_q^k = \lambda_k \times E(T^k) \tag{3.15}$$

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

By analysing the Eq. (3.14) and adjusting these equations to the proposed system, the queue length can be calculated as:

$$L_q^k = \lambda_k \times \left(\frac{\sum_{i=1}^{k} \lambda_i \times \left(x_i^2 + \sigma^2\right)}{2\left(1 - \rho_{k-1}^-\right)\left(1 - \rho_k^-\right)} + c \times T_{gap} \right) \tag{3.16}$$

where ρ_k denotes the utilization intensity and $\rho_k = \lambda_k \times x_k$. Sections 3.1 and 3.2 show how service class mapping can be achieved in an integrated PON-4G network. The next section extends the service class mapping model for developing a power saving mechanism for the ONU-BS unit.

3.3 Proposed Class-Based Power Saving (CBPS) Model

The CBPS model is based on a DBA scheme which increases the sleep duration of ONUs. This model supports both doze 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 at the ONU wake up, respectively. The following sections

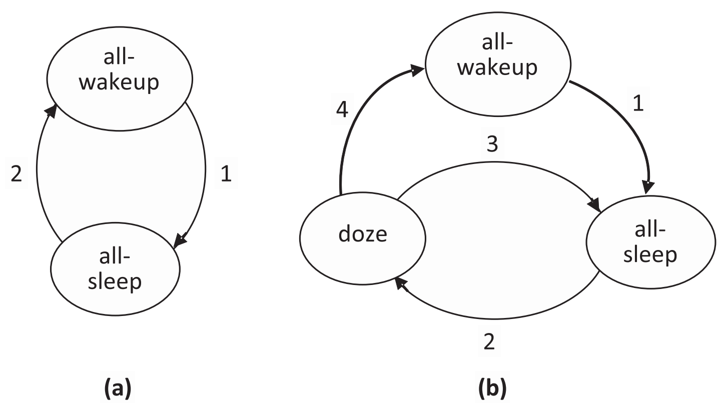

Figure 3.3: Different power states of ONU based on; a- existence and b- absence of a class 1 traffic.

Table 3.1: Maximum Sleep Duration for the ONU's Transmitter for Different Traffic

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

discusses 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 3.2.

Algorithm 3.1 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_cycle − 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 3.2)
            End if
End if
```

Algorithm 3.2 The transition from all-sleep to all-wake-up through doze 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 × T_cycle − T_grant − T_oh;
   a: after s×T_cycle
      ONU_i _Rx is wake_up and read the downstream traffic;
      If ONU_i _Tx_sleep_time > (s×T_cycle)
            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
```

3.3.1 Transition between Different Power States

This proposed scheme uses two different modes for power saving, one for dealing with *class* 1 traffic (Figure 3.3(a)) which is highly time sensitive and requires immediate attention from the transmitter while the other mode (Fig. 3.3(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. 3.3(a), this proposed scheme switches between the all-wake-up and all-sleep states for *class* 1 traffic. The all-wake-up state stays on for $(T_{grant} + T_{oh})$ and all-sleep state stays on for $(T_{cycle} - T_{grant} - T_{oh})$ (Table 3.1). This is because the delay for *class* 1 traffic should not exceed one polling cycle duration. The transition from the all-sleep to all-wake-up occurs when the sleep time expires.

Mode 2 has three power states namely: all-wake-up, all-sleep and doze. These three states arise from the difference between the sleep time duration of the transmitter and the

Table 3.2: Maximum Sleep Duration for Different Sleep Categories for *class* 2 and *class* 3 Traffic

Categories of Sleep	Max. Sleep Time (T_s^{max})
Aggressive	45 ms
Fine	35 ms
Conservative	25 ms

receiver. Because classes 2 and 3 tolerate longer delays, a longer sleep time can be used 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, in order to maintain the QoS of downstream traffic, the receiver needs to wake up in every cycle. Firstly, the decision for sleep time duration in this proposed scheme is based on the upstream traffic condition. Secondly, there is no mechanism to inform the ONU about the priority of downstream classes; therefore, the delay of downstream traffic is restricted within one polling cycle and treated as $class1$ traffic and the reciever has to wake up every polling cycle so the transition between the all-sleep to doze occurs. After receiving the downstream data and GATE, the receiver either turns to sleep if the sleep duration for the transmitter is not expired yet then the transition from doze to all-sleep occurs, or the transition from the doze to all wakeup occurs if the sleep time of the transmitter is expired. The transition between all-wake-up to all-sleep occurs in similar way to mode 1 (Figure 3.3(a)). Because the transmitter is switched to sleep for $(N_{max} \times T_{cycle} - T_{grant} - T_{oh})$ duration (3.1) and the receiver wakes up in every polling cycle, the transition between the all-sleep to doze 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 ($N_{max} = \frac{T_s^{max}\ (Table\,3.2)}{T_{PON}}$). 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 doze 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 3.1 and 3.2 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. (3.17) and (3.18), respectively.

$$T_s = \begin{cases} T_{cycle} - T_{grant} - T_{oh} & class\,1 \\ N_{max} \times T_{cycle} - T_{grant} - T_{oh} & class\,2,\,3 \end{cases} \tag{3.17}$$

$$T_a = T_{grant} + T_{oh} \tag{3.18}$$

In the dynamic bandwidth allocation, the length of polling cycle ($T_{cycle} = \sum_{n=1}^{N}(T_g + T_{grant}^{n})$) varies depending on the load conditions. Because T_{cycle} depicts the polling cycle when all the ONUs are polled and transmitting their upstream traffic (Fig. 3.2), it is necessary to calculate the maximum duration of the polling cycle to estimate the maximum sleep time.

The maximum length of the polling cycle ($T_{cycle}^{max} = N \times (T_g + T_{grant}^{max})$) and ($T_{grant}^{max} = \frac{\sum_{k=1}^{3} q_k^{max} \times P_k \times 8}{R}$), where T_{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 k^{th}; q_k^{max}depicts the maximum queue size for class k^{th}; R depicts the transmission rate.

3.3.2 Sleep Time Duration Policies

The sleep time duration differs in mode 1 & 2 (see Table 3.1) only for the transmitter and the shown sleep durations have been chosen to be less than the maximum gap between two REPORTs, which is 50 ms, to maintain the ONU's registration with the OLT [57] [58]. Moreover, the suggested transmitter sleep time duration depends on the class of traffic. Sleep time was predefined for every traffic class, so that the OLT can calculate the wake up time for each ONU based on firstly, the starting time of the granted transmission window and secondly, on the last REPORT message (as described in Fig. 3.4), which infers the OLT

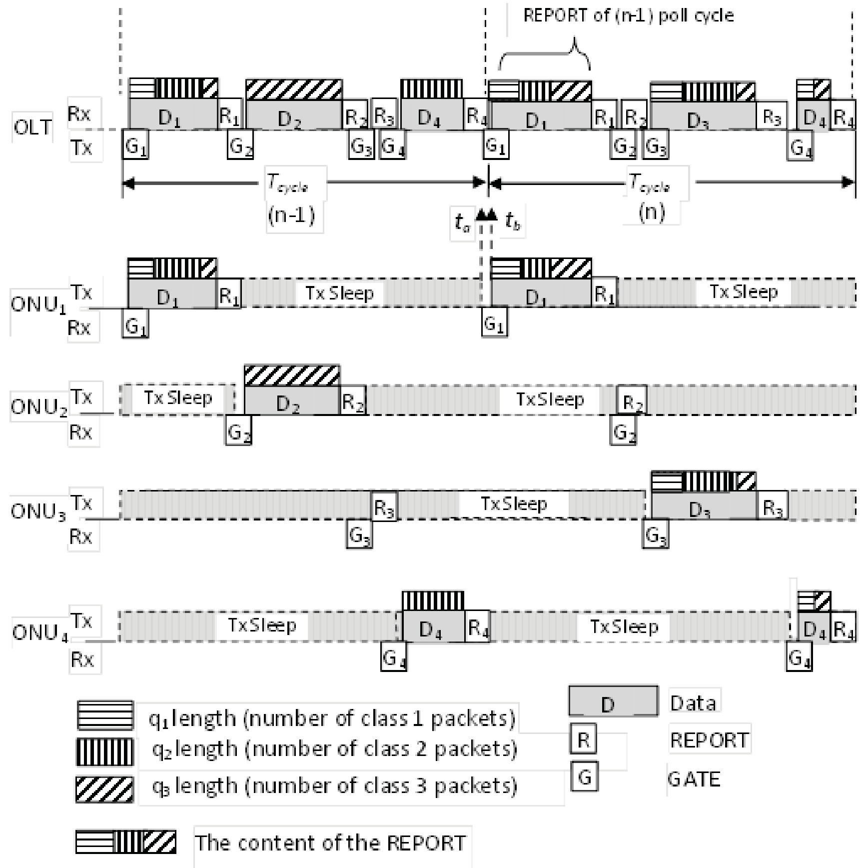

Figure 3.4: Time diagram of CBPS operation for four ONUs.

about the sleep duration of the ONU. 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 ($T_{cycle} - T_{grant} - T_{oh}$). 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 802.3av (the standard for 10G-EPON) as aforementioned.

For *class* 1, the sleep time is set at less than one polling cycle ($T_{cycle} - T_{grant} - T_{oh}$). However, if the polling cycle duration is no longer than the double length of overhead time then it will not activate the sleep state of the transceiver. This work considers the maximum polling cycle as a fixed cycle even with dynamic allocation of the bandwidth, which is a common approach for the dynamic bandwidth allocation, because it facilitates the synchronization. Moreover, in these experiments, only aggressive and conservative sleep times are used with 3 ms and 2 ms overhead times, respectively (see Table 3.2).

Algorithm 3.3 Class-Based Dynamic Bandwidth Allocation (DBA).

```
OLT:
start new polling cycle: Loop ∀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 q1, q2 and q3 occupancies
        If q1 is not empty
            ONU moves into mode 1 (Algorithm3.1)
        Else
            ONU moves into mode 2 (Algorithm 3.2)
        End If
        ONU decides the next wake up time for Rx and Tx
   End If
End If.
```

3.3.3 CBPS based ONU

The essence of the proposed 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 doze mode, the ONU transits to sleep mode only for a short time, and stays in doze state for longer. As previously stated, the transition to sleep and doze states depends on the class priority of the available traffic in the ONUs. Figure 3.4 illustrates the operation of the CBPS mechanism in cooperation with dynamic bandwidth allocation. Only the sleep time of the ONU's transmitter is shown because as stated previously the sleep time of the receiver is bounded to $(T_{cycle} - T_{grant} - T_{oh})$. It can be seen that the OLT periodically polls 4 ONUs and assigns bandwidth for them (only two polling cycles are shown). For instance, in polling cycle (n-1), it can be seen 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 $(T_{cycle} - T_{grant} - T_{oh})$ 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 $(N_{max} \times T_{cycle} - T_{grant} - T_{oh})$ 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_{grant}^{min} of 64 bytes is required to allow the transmission of REPORTS and detection of downstream traffic. Algorithm 3.3 summarizes all the operations of the proposed class-based dynamic bandwidth allocation.

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

$$E = \sum_{n=1}^{N} [a \times E_{cls1}^{n} + (1-a) \times E_{nocls1}^{n}] \tag{3.19}$$

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 (*nonclass1*), Equation (3.20) can be generated as:

Table 3.3: System Parameters for 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

$$E = \sum_{n=1}^{N} (a \times (T_s^n(cls1) \times P_s + T_a^n(cls1) \times P_a) + (1-a) \times (T_s^n(nocls1) \times P_s + T_a^n(nocls1) \times P_a)) \quad (3.20)$$

Eq. (3.20) shows that each ONU's transmitter has both awake and sleep energy consumption.

By specifying the T_s (Eq. (3.17)) and T_a (Eq. (3.18)) of each class, the following can be shown:

$$\begin{aligned} E = \sum_{n=1}^{N} (a \times ((T_{cycle} - T_{grant}^n - T_{oh}) \times P_s \\ + (T_{grant}^n + T_{oh}) \times P_a) \\ + (1-a) \times ((N_{max} \times T_{cycle} - T_{grant}^n - T_{oh}) \times P_s \\ + (T_{grant}^n + T_{oh}) \times P_a)) \end{aligned} \quad (3.21)$$

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. 3.4, $T_{oh} = t_a - t_b$, where t_a denotes wake up time and t_b denotes the transmission start time for the data and REPORT.

Table 3.4: System Parameters for 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

3.4 Performance Evaluation

This section presents the simulation results to validate the proposed queue management and power saving schemes. The Network Simulator 2.29 [68] was chosen and a model was created to simulate the 10G-EPON and WiMAX. WiMAX model was elaborated as a preferable option because of its compatibility with the simulator. The simulation parameters for both the 10G-EPON and WiMAX were configured as shown in Table 3.3 & 3.4, respectively. For traffic modelling, a traffic profile proposed by [45] [69] was used where high priority traffic such as VoIP was modelled by a constant bit rate source, video applications were modelled using a variable bit rate source, and best effort traffic arrived dynamically. Similar traffic profile had been used in other relevant studies [45] [69]. 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 s and 0.2 s for modelling high and low load VoIP applications. For simulating the video streaming (*class* 2) applications, variable bit rate traffic was chosen with packet sizes uniformly distributed between 500 and 1500 bytes. To simulate the BE traffic, a TCP/ Newreno type [68] was used with a packet size of 60 bytes. The arrival rate of applications was changed to investigate the impact of various loads on network performances. It can be noted that because the 10G-EPON with a maximum capacity of 10 Gbps was used, WiMAX system (802.16m), which provides the access network connections in WOBAN, worked as the bottleneck.

The same scenario and traffic profile were maintained when comparing the proposed model against existing models. The comparison was made between the proposed approach, which classified the traffic into three priority queues as presented in Section 3.2, against a model proposed in [46], which classified the traffic into two classes, expedited forwarding (EF) and best effort (BE). The high priority class was expedited forwarding, which was delay sensitive (e.g., VoIP and video streaming). The low priority class was best effort and not delay sensitive (e.g., FTP, email). EF_BE was used to refer to that work (existing work) and Diff_srv was used to refer to the work proposed in this thesis. The EF_BE model assigned a queue of infinite length to each traffic class. If there was no packet in the queue corresponding to EF traffic, the EF_BE model put ONU to sleep mode and the sleep period was calculated based on a state-transition graph model in response to the load in the queue corresponding to EF traffic. The comparison was made against the EF_BE model because this is the only PON model available in the literature that dealt with QoS of broadband services and power savings at the same time.

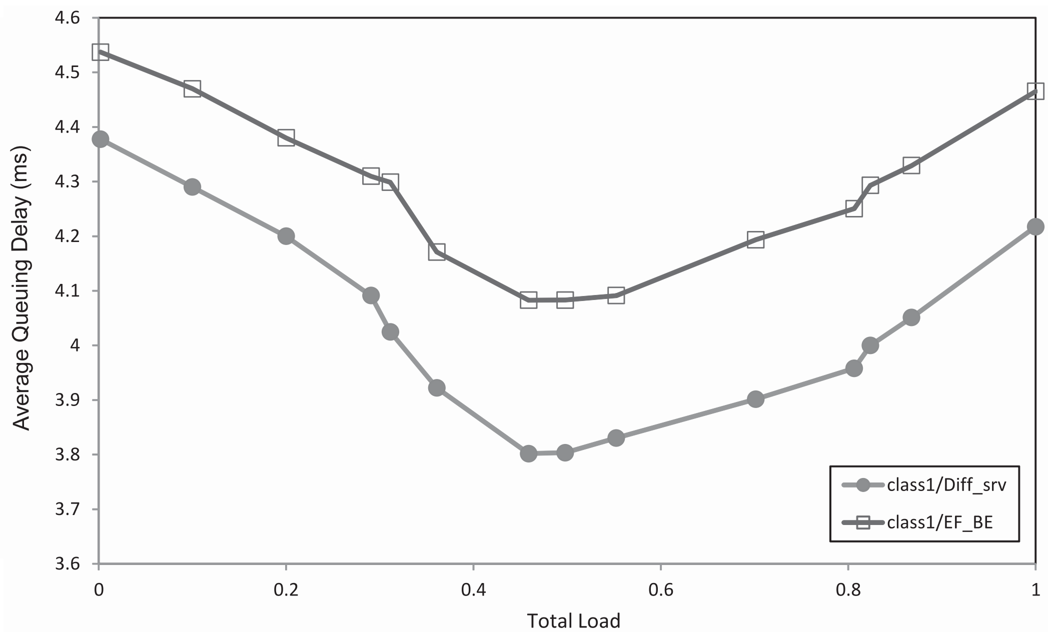

Figure 3.5: Comparison between the proposed Diff_srv and EF_BE model for the average queuing delay with *class* 1 packets.

3.4.1 Queue Management

This section compares the performance of the proposed queue management model against the FE_BE model [46]. In the simulation, the set up of the base station, wireless connections, the ONU and the OLT of the optical part are based on Table 3.3 & 3.4 and the length of each queue in the proposed Diff_srv system was calculated using Eq. (3.16) which is needed before the starting of the simulation. The average queuing delay is monitored in both the Diff_srv and EF_BE systems at various loads as shown in Fig. 3.5 and 3.6. These figures show that the Diff_srv system provides a much lower average queuing delay for *class* 1 flows under all tested loads. Figure 3.5 indicates that the 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 the 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 (Figure 3.6). Moreover, increasing the load causes a difference in the nature of the queuing delay behavior for the *Class* 1 and *Class* 2 (delay in Fig. 3.5 increases after a load of about 0.5 while that in Fig. 3.6 continues to decrease). This is because increasing the load makes more packets

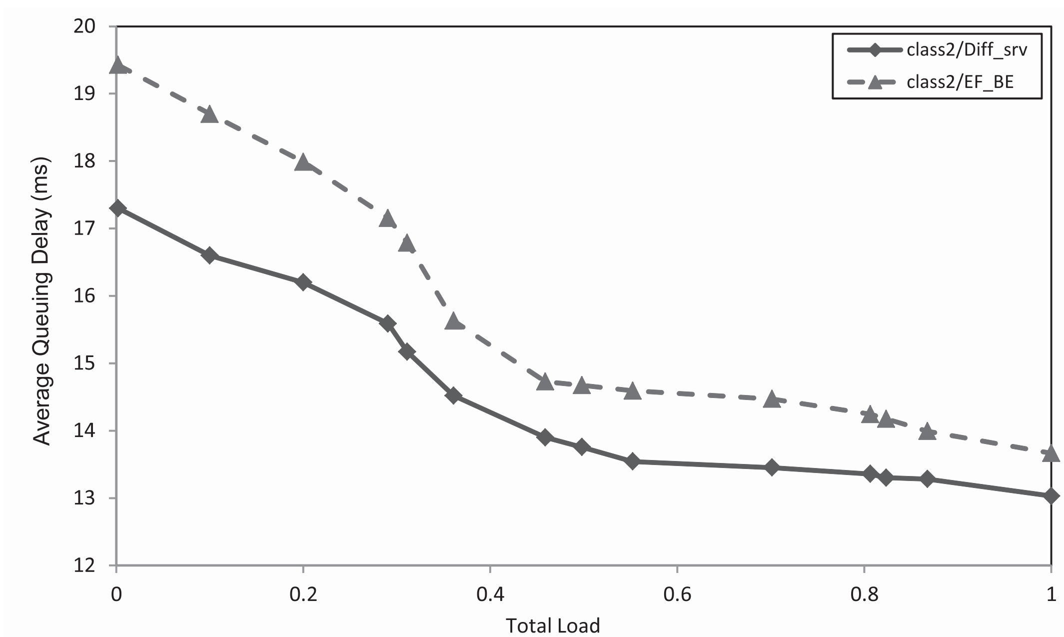

Figure 3.6: Comparison between the proposed Diff_srv and EF_BE model for the average queuing delay with *class* 2 packets.

wait in the queues and lengthens the requested bandwidth. On the one hand, this reduces the queuing delay for *Class* 2 packets as increases the served packets in each transmission window, which then reduces the average queuing delay. On the other hand, causes a longer delay for the *Class* 1 packets as they are affected by longer transmission windows. However, the delay for *Class* 1 traffic is not high as the constraint of a maximum transmission window is applied.

Figure3.7 shows the packet dropping rates for *class* 1 and *class* 2 traffic. It is evident that packet dropping rate is zero for *class* 1 and *class* 2 in the proposed 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.

Figure 3.8 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. 3.9. 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.

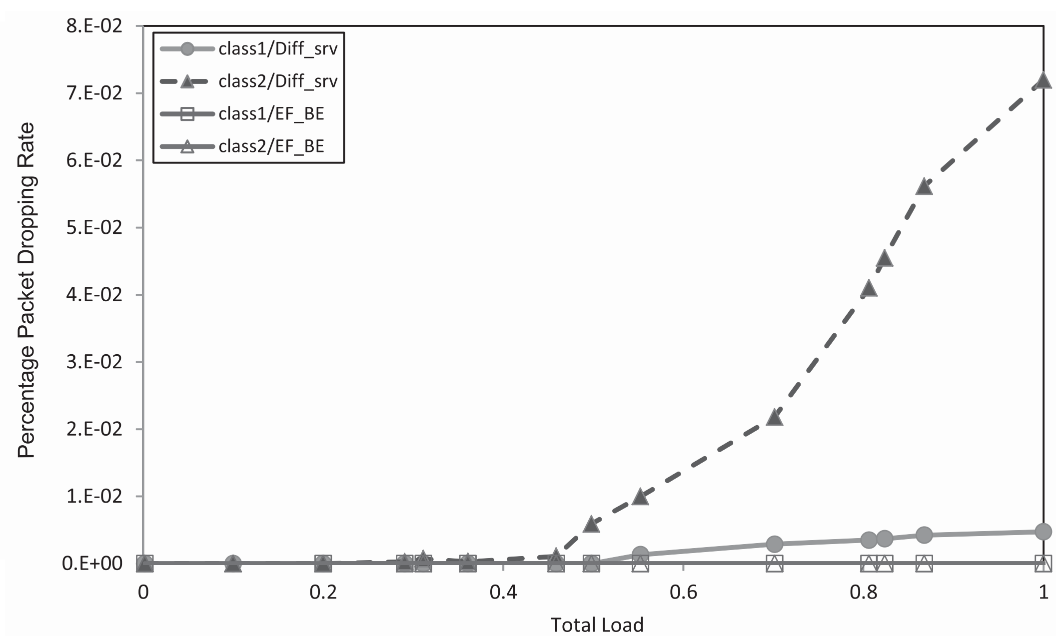

Figure 3.7: Packet dropping rate comparison between the proposed Diff_srv and EF_BE model.

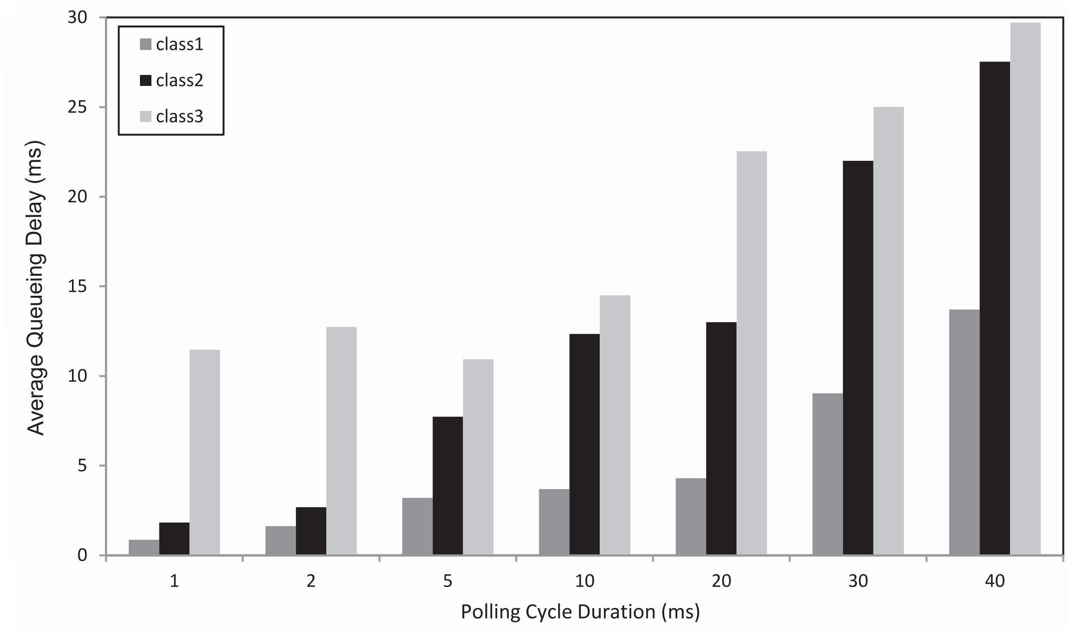

Figure 3.8: Average queuing delay versus different polling cycle duration (T_{cycle}) for Diff_srv model.

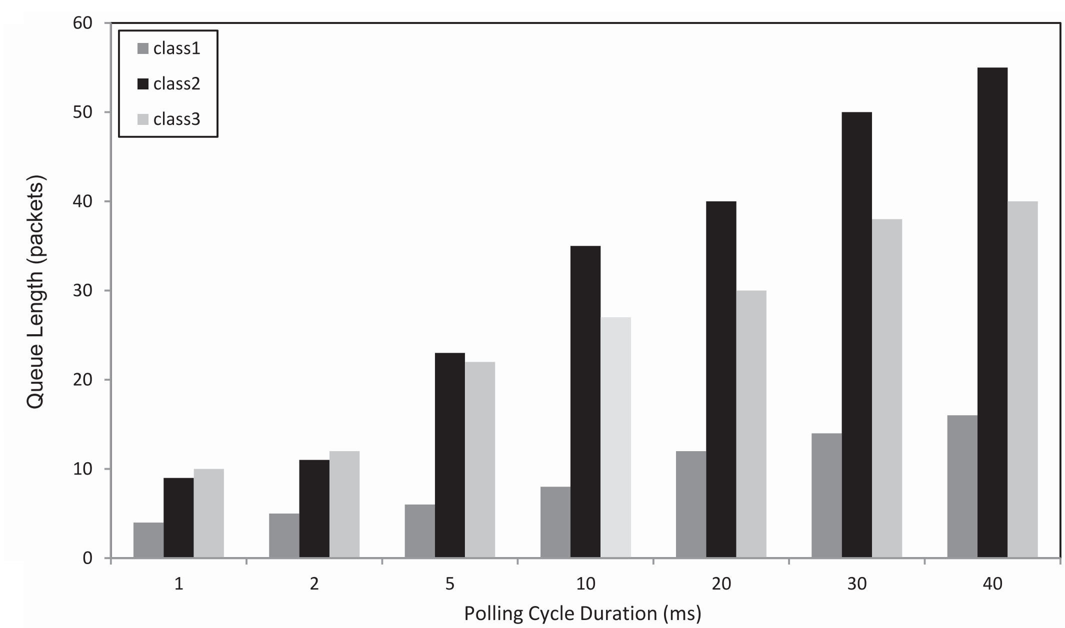

Figure 3.9: Queue length versus different polling cycle duration (T_{cycle}) for Diff_srv model.

Class 3 packets however, are still served at a very low rate because the proposed scheme allows a 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 the 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.

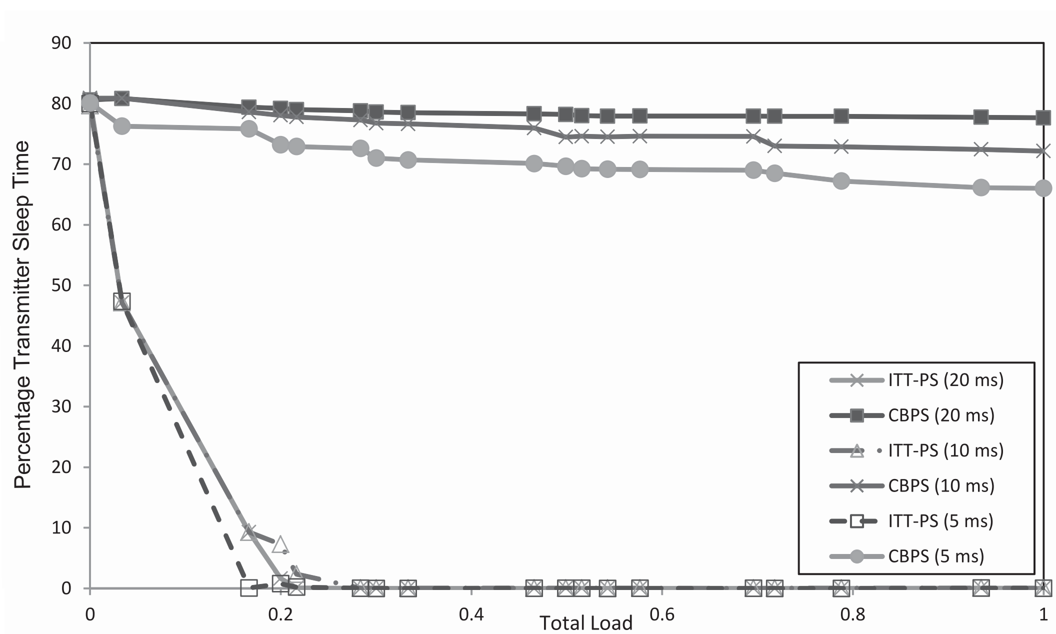

Figure 3.10: Comparison between the CBPS and ITT-PS model for percentage power saving.

3.4.2 Power Saving

This section presents the simulation results of the proposed class-based power saving (CBPS) model against existing models [58] [46] 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. An Idle Transmitter

Time-based Power Saving (ITT-PS) model is used to refer to this concept. The ITT-PS is a framework for calculating the sleep time as a function of the idle transmission period [58]. Initially, the sleep time begins with a specific value. As long as there is no upstream traffic, the sleep time is gradually increased until it reaches the maximum sleep time of 50 ms. The arrival of traffic favours the wake up state for the transmitter.

The simulation parameters for both the 10G-EPON and WiMAX are configured as shown in Table 3.3 & 3.4, respectively. The simulation measures 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 the 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. QoS provisioning for different classes of traffic in the CBPS model is shown.

Figure 3.10 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) [53] 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 percent as high priority traffic increases in the ONU's queues. This occurs because the ITT-PS model makes the ONUs keep awake all the time just when there are any packets flowing in the networks even though those packets can wait at their queues for a relatively long time if they belong to low priority traffic. Moreover, the ITT-PS ignores the idle time of ONUs even with heavy traffic which contributes to about 93 percent for a system consisted of 16 ONUs (the transmission time of each ONU is about 1/16 of the total transmission time). This behavior of the ITT-PS makes the ONUs cannot save energy even they only transmit during their granted transmission windows. 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.

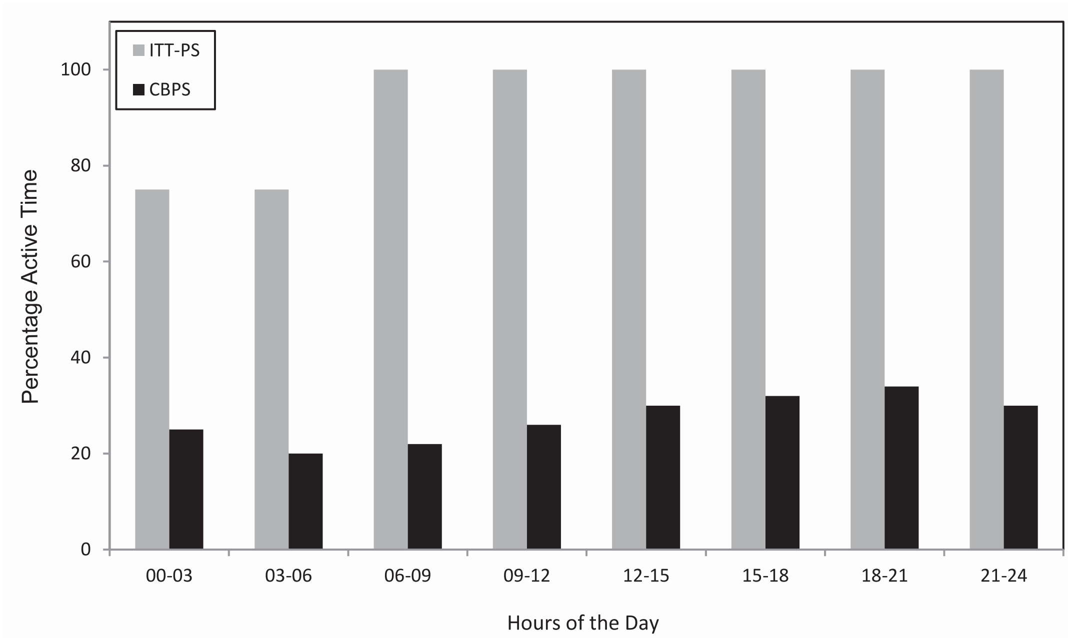

Figure 3.11: Power saving comparison between ITT-PS and CBPS over 24 hours.

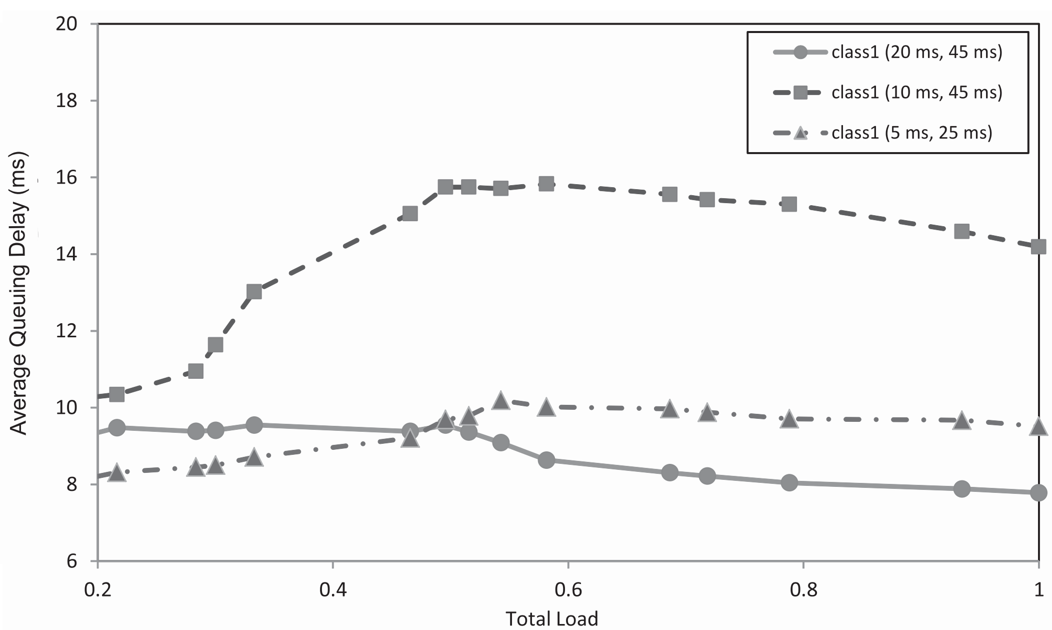

Figure 3.12: Average queuing delay for *class* 1 traffic.

Moreover, there is a difference between the power saving of the 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. Figure 3.11 shows the comparison between the proposed model and the ITT-PS model in terms of ONU's active time (%) for a full 24 hour period following the traffic profile used in [7]. The traffic profile captures the varying traffic loads during a 24 hour period. As indicated in Fig. 3.11, the proposed model consistently outperforms the ITT-PS model by a significant margin.

Figure 3.12 illustrates the influence of CBPS on the behaviour of WOBAN by showing the average queuing delay of *class* 1 traffic for different polling cycle durations. Pairs (T_{cycle}, T_s) are used in the discussion to describe every curve in the figure. As shown, the shortest average delay belongs to the (5 ms, 25 ms) pair when the load is below the half of the full load. Then, with increasing load, (20 ms, 45 ms) exhibits the minimum average delay. The reason is that, with increasing load, the maximum transmission window becomes over-utilized for (5 ms, 25 ms) which causes more unscheduled packets waiting in the queues for the next polling cycle. For the (20 ms, 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 ms, 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.

3.5 Closing Remarks

In this chapter, an efficient QoS mapper for WOBAN, based on differentiating the traffic to three priorities including the priority scheduler at the ONU that satisfies the high priority traffic and does not disadvantage the low priority traffic, has been proposed toward the goal of establishing the basis for the integrated network. It is noted that differentiating the traffic into less than three classes harms the QoS while more than three consumes more power and increases the complexity. This proposed solution for the service class mapping problem, between 4G classes and the 10G-EPON queues, is based on an M/G/1 queuing model. The queue behaviour of these classes is investigated and a model to calculate the required length for different queues is developed which ensures that the QoS for various end applications is not compromised. This is the first work of its kind that addressed service class mapping, scheduling and queue modelling problems at the ONU in WOBAN. In this chapter, the extended concept of service class mapping, a class-based dynamic bandwidth allocation scheme for power savings with a specific sleep period for each class in WOBAN was presented. The proposed scheme significantly increases the sleep period and reduces overheads 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. It was proposed that the differentiated sleep duration power saving exploits the idle periods at the ONUs and the diversity of services that impact quality requirements.

The performance of the proposed mechanism (CBPS) was examined and analysed over WOBAN network for different traffic loads. The simulation results indicated that this proposal significantly reduced power consumption by up to 80 percent, and maintained the packet dropping and delay within the requirements of the service level agreement.

In the next chapter, the mechanism of energy efficiency will include the OLT as it is the major power consumer and at the same time is the controller of the entire WOBAN network.

Chapter 4

Load-Adaptive Power Consumption Model for Green WOBAN

As described in Chapter 2, PON has two types of components that consume power, i.e., ONUs and OLT. A common approach for developing a green WOBAN is to turn off network elements such as ONUs when they are idle. This chapter adopts a similar strategy and presents an efficient resource management scheme that allows putting OLTs and ONUs into sleep for longer periods without sacrificing the QoS of end applications (e.g., voice, video, and data applications). This ultimately reduces power consumption at no additional cost in terms of QoS. The concept of service class differentiation is used to develop a model for calculating the sleep periods for both OLTs and ONUs, in response to the load from various traffic classes. An analytical model is proposed to calculate the sleep and active durations for the OLT and ONUs and the power consumption when the proposed model is applied. There is no other research in the literature that aims to increase sleep periods for the OLT and ONUs, while maintaining the QoS for different traffic classes in WOBAN. Finally, this chapter summarises the significant results of the presented work.

4.1 Proposed Energy-Efficient WOBAN

This work assumes a WOBAN (Fig. 2.9) where N ONUs are connected to an OLT which uses a DBA scheme to allocate resources. In a DBA scheme, bandwidth is allocated dynamically based on ONU demands, and the polling cycle length (T_{cycle}) varies accordingly. The polling cycle length can be expressed as: ($T_{cycle} = \sum_{i=1}^{N}(T_{grant}^{i} + T_g)$), where T_{grant}^{i} depicts the granted bandwidth (time slot) to the i^{th} ONU ; T_g depicts the guard time between two allocated grants for successive ONUs. If the ONUs are heavily loaded, the assigned bandwidth (T_{grant}) is high, thus increasing the polling cycle duration. The long polling cycle affects the QoS (e.g., increases delay and jitter) for delay-sensitive services, whereas, if the ONUs are lightly loaded or have no load, the polling cycle can be made very short. Under no load scenario, a minimum polling cycle is maintained just to allow the transmission of the REPORT message (i.e., 64 bytes). For short polling cycles, the overhead in the form of the guard time and the time required for an ONU to warm up and synchronize following every wake-up, becomes significantly high. This short but frequent polling does not allow the ONU/OLT to go into sleep for longer periods and thereby, contributes to a detrimental performance in terms of power savings and channel utilization. Another aspect of using DBA with sleep enabled ONUs is the variable length of polling cycle which makes the synchronization of the wake up time of the sleeping ONU and the controlling OLT very challenging, often reducing the sleep periods for ONUs/OLTs. One solution to this problem is to use a negotiation process between the ONU and OLT to make an agreement regarding the point of time when ONUs and OLTs go into sleep. Such negotiation process however, introduces an extra level of overhead.

Figure 4.1: Polling cycle duration for different traffic classes.

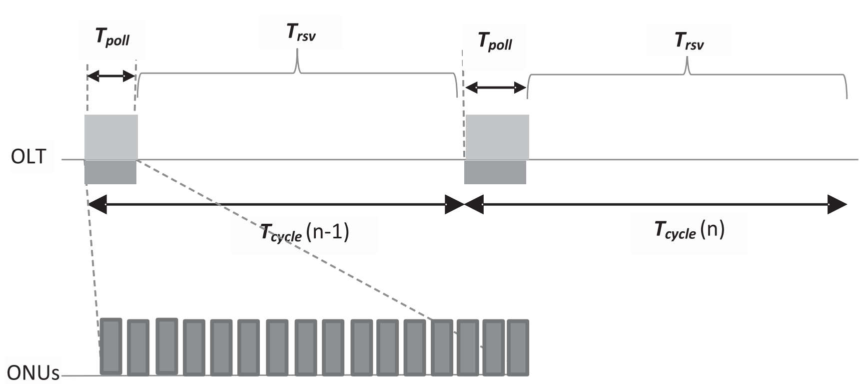

Figure 4.2: Polling cycle for a system with 16 sleep-enabled ONUs.

To mitigate the above mentioned challenges, a DBA mechanism is used to allocate bandwidth for ONUs based on service class differentiation of traffic, which is explained in the following sections.

4.1.1 Polling Cycle Calculation in the Proposed Mechanism

This work deploys the mapping policy that is proposed in Chapter 3 and proposes to use the concept of service class differentiation at ONUs, and presents an adaptive polling cycle solution based on two concepts; the first uses a DBA to allocate bandwidth for the ONUs, but with a fixed polling cycle duration $T_{cycle} \in \{T_{cycle1}, T_{cycle2}, T_{cycle3}\}$ where T_{cycle1}, T_{cycle2} and T_{cycle3} represent the polling cycle corresponding to *class* 1, 2 and 3 traffic, respectively (Fig. 4.1). The second concept is about OLT dynamically controlling the polling cycle duration to adjust to the changes in network traffic (i.e. load from different classes). The proposed mechanism works as follows: the OLT selects the duration of the polling cycle (T_{cycle}) based on the class of the available traffic in the ONUs and the traffic load. This polling cycle is divided into T_{poll} and T_{rsv} (Fig. 4.2), which depict polling sub-cycle and reserved sub-cycle, respectively. Grant allocations for ONUs occur within T_{poll} and T_{rsv} is the duration for ONU and OLT to enter into sleep mode. The duration of the two sub cycles are based on the T_{cycle} which depends on the traffic condition.

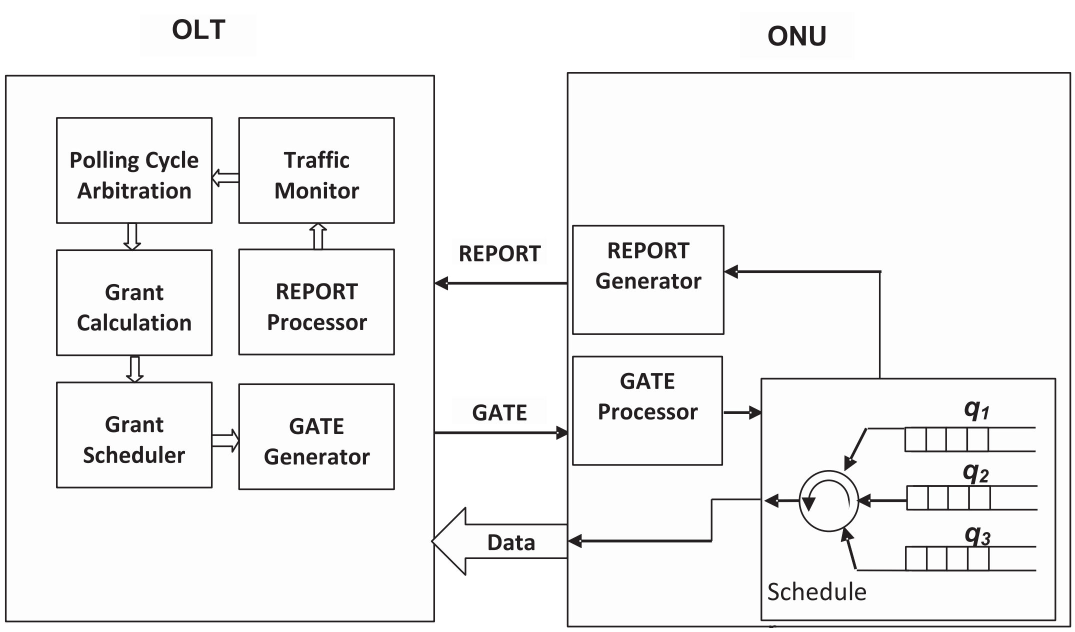

Figure 4.3: System model for the proposed dynamic polling cycle based-power saving mechanism (DPCB-PSM).

The different steps in deciding the cycle duration, in the proposed model, are depicted in Fig. 4.3. All the steps that are performed in OLT are based on the REPORTs received from ONUs. The decision for polling cycle duration arbitration is then made and all ONUs are informed via the GATE messages. Initially, the OLT sets the duration of the polling cycle as the T_{cycle1}, which is the shortest of all polling cycles (Figure 4.1). The OLT then broadcasts the new polling cycle to all ONUs by the GATE messages. In every polling cycle, ONUs request bandwidth from the OLT by sending REPORT messages which contain the instantaneous queue length (i.e., q_1, q_2 and q_3) information at each ONU. The OLT checks the REPORTs to decide the next polling cycle length and allocates bandwidth accordingly. For instance, if any ONU reports q_1 as not equal to zero (i.e., not empty), the OLT keeps the next polling cycle as T_{cycle1}. If none of the ONUs requests bandwidth for q_1, but at least one of them requests bandwidth for q_2, then OLT selects T_{cycle2} as the duration for the next polling cycle. Because the *class* 3 is less affected by the delay, the OLT selects T_{cycle3}, which is the longest duration, as the duration of the polling cycle if no ONU requests bandwidth for q_1 and q_2. If the system in a global sleep mode and packets arrived in q_1, then only the first packets will experience extra delay, because when the system wakes up and the ONU reports the existing of *class*1 traffic the OLT will change to T_{cycle1}. Despite this, the affected packets are only the first arrival ones, but including admission control with the proposed system will solve the problem as the OLT will be aware about the new connection before starting and OLT can choose the proper polling cycle duration.

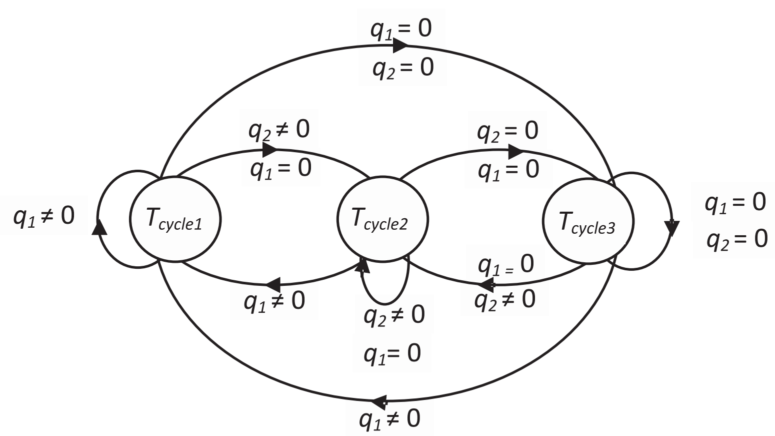

Figure 4.4: Transition between different polling cycle durations at the OLT based on the ONUs' queue occupancies.

Figure 4.4 depicts the transition between different polling cycles at OLT based on queue conditions.

The calculation for T_{cycle1}, T_{cycle2} and T_{cycle3} is shown as follows:

$$\begin{bmatrix} T_{cycle1} \\ T_{cycle2} \\ T_{cycle3} \end{bmatrix} = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & \overline{x_1} \times x_2 & 0 \\ 0 & 0 & \overline{x_1} \times \overline{x_2} \end{bmatrix} \begin{bmatrix} a_1 \times T_{cycle}^{max} \\ a_2 \times T_{cycle}^{max} \\ a_3 \times T_{cycle}^{max} \end{bmatrix} \tag{4.1}$$

where a_1, a_2 and a_3 are three design coefficients ($a_1 < a_2 < a_3$); the first two coefficients (a_1 and a_2) are less than 1 to make the polling cycle lengths suitable for serving real time traffic with acceptable delay. This is because of a relationship between the polling cycle length and the expected waiting time of the packets in the queues as shown in Chapter 3. T_{cycle}^{max} depicts the maximum acceptable duration of the polling cycle; x_1 and $\bar{x}_1$ are binary variables that depict the presence and the absence of *class* 1 traffic, respectively; where x_1 is the logical inverse of $\bar{x}_1$; x_2 and $\bar{x}_2$ are binary variables thar depict the presence and the absence of *class* 2 traffic, respectively; where x_2 is the logical inverse of $\bar{x}_2$. These variables (x_1, $\bar{x}_1$, x_2, $\bar{x}_2$) are updated after each polling process according to the following equation:

$$\forall y \in [1,2]\ x_y = \begin{cases} 0 & if\ class\ y\ does\, not\ exist \\ 1 & if\ class\ y\ exists \end{cases} \tag{4.2}$$

In the proposed mechanism, the polling cycle is refreshed before each polling process and the polling cycle length T_{cycle} can be given as:

$$T_{cycle} = T_{poll} + T_{rsv} \tag{4.3}$$

where

$$T_{poll} = \begin{cases} T_{cycle} & if\, class\, 1\&2\ traffic\ are\ available \\ (1-a) \times T_{cycle} & else \end{cases} \tag{4.4}$$

$$T_{rsrv} \quad = \quad \begin{cases} 0 & if\, class\, 1\&2\ are\ available \\ a \times T_{cycle} & else \end{cases} \tag{4.5}$$

Here, a is a design parameter that controls the ratio between T_{poll} and T_{rsv}, and thereby, influences energy savings.

The OLT calculates the allocated bandwidth for the ONUs according to the T_{poll} where the allocated bandwidth for i^{th} ONU (T^i_{grant}) is given by:

$$T^i_{grant} = \begin{cases} \sum_{j=1}^{3} T_{q_j} & if \ \sum_{j=1}^{3} T_{q_j} \leqslant T_{max} \\ \\ \begin{bmatrix} 1 & \alpha & \delta \end{bmatrix} \begin{bmatrix} T_{q_1} \\ T_{q_2} \\ T_{q_3} \end{bmatrix} & otherwise \end{cases}$$

where T_{max} depicts the maximum transmission window and ($T^i_{grant} \leqslant T_{max}$); α and δ depict coefficients ≤ 1, which means that if the granted transmission window is short and not long enough to send all the packets in all the three queues, the granted transmission window should be at least long enough to send all the *class* 1 packets. T_{q_1}, T_{q_2} and T_{q_3} depict the requested bandwidth by the ONU for q_1, q_2 and q_3, respectively, which can be given as:

$$T_{q_1} = \lambda_1 \times Z_1/R \tag{4.6}$$

$$T_{q_2} = \lambda_2 \times Z_2/R \tag{4.7}$$

$$T_{q_3} = \lambda_2 \times Z_3/R \tag{4.8}$$

where

$$Z_1 = \frac{\lambda_1 \times E\left(S_1^2\right)}{2 \times (1 - \rho_1)} + \gamma \tag{4.9}$$

$$Z_2 = \frac{\lambda_1 \times E\left(S_1^2\right) + \lambda_2 \times E\left(S_2^2\right)}{2 \times (1 - \rho_1)(1 - \rho_1 - \rho_2)} + \gamma \tag{4.10}$$

$$Z_3 = \frac{\lambda_1 \times E(S_1^2) + \lambda_2 \times E(S_2^2) + \lambda_3 \times E(S_3^2)}{2 \times (1 - \rho_1 - \rho_2)(1 - \rho_1 - \rho_2 - \rho_3)} + \gamma \tag{4.11}$$

Here R depicts the data transmission rate of the optical part; Z_1, Z_2 and Z_3 depict the average (expected) waiting time of the packets in the their queues due to their priority and the waiting time for the granted transmission window, which is calculated after an M/G/1 queuing model in Chapter 3; γ depicts a coefficient relates to the effect of the polling cycle length and it's calculation is shown in Chapter 3.

In the proposed mechanism, the OLT centrally controls the scheduling and bandwidth allocation mechanism, to change the duration of the polling cycle to adapt to the class and the volume of the traffic. However, constraints such as MPCP, which is the maximum duration between two successive GATEs or/and two successive REPORTs is 50 ms as constrained by the 802.3ah and 802.3av standards [58], and QoS consideration should be considered by the OLT. Thus, the maximum polling cycle duration (T_{cycle}^{max}) in the proposed model is set as 50 ms.

4.1.2 Modes of the Proposed Dynamic Polling Cycle Based-Power Saving Mechanism (DPCB-PSM)

This section presents the two modes; global sleep and local sleep that are used in the proposed power saving technique. In this proposal, there are two constraints, namely; the delay limit of each class of traffic and the constraints of the MPCP. The following sections describe the effects of the aforementioned constraints, the traffic condition and quantity on the design of the proposed mechanism.

4.1.2.1 Global Sleep Mode (GSM)

In this mode, the maximum energy saving is achieved by making the ONUs and OLT go into controlled sleep. This mode occurs only when no real time traffic is reported by any ONU, i.e., all ONUs send REPORTs of zeros for q_1 and q_2. To achieve this mode, the OLT selects a long polling cycle duration ($T_{cycle} = T_{cycle3} = T_{cycle}^{max}$). Each ONU now is polled within T_{poll}, so $\left(T_{poll} = \sum_{i=1}^{N}(T_{min}^{i} + T_g) \geqslant \sum_{i=1}^{N}\left(\frac{64\times 8}{R} + T_g\right)\right)$ then the duration of T_{rsv} will be ($T_{cycle3} - T_{poll}$). The OLT and all ONUs enter the sleep state during T_{rsv} and they wake up during T_{poll}. In this mode, the duration of the two sub-cycles should be selected to minimize energy wastage. T_{poll} is a short sub-cycle and is divided among the ONUs to send REPORTs with extra bandwidth to send the number of *class* 3 packets. During this sub-cycle, all the ONUs and the OLT are waken, whereas during the second sub-cycle (T_{rsv}) both the ONUs and OLT are in sleep mode. For ONU and OLT, the sleep duration (T_s^{ONU} and T_s^{OLT}, respectively) and the wake up (active) duration ($T_a^{ONU}$and T_a^{OLT}, respectively) can be expressed in Eq. (4.12) and (4.13).

$$T_a^{ONU} = T_a^{OLT} = T_{poll} = \sum_{i=1}^{N}(T_{min}^{i} + T_g) \tag{4.12}$$

$$T_s^{OLT} = T_s^{ONU_i} = T_{rsv} - T_{oh} \tag{4.13}$$

where T_{min}^{i} depicts the minimum granted transmission window for i^{th} ONU; T_{oh} depicts the time required for an ONU and OLT to warm up and synchronize following each wake up as mentioned before.

Algorithm 4.1 Load-Adaptive Polling cycle.

OLT:
Initialization:
    $T_{cycle} = T_{cycle1}$
    OLT polls all ONUs
    OLT collects REPORTs
Decision making: $\forall$ REPORT
if (q_1 is not empty)
    $T_{cycle} = T_{cycle1}$;
    go to LSM;
else if (q_1 is empty)&&(q_2 is not empty)
    $T_{cycle} = T_{cycle2}$;
    go to LSM;
else if (q_1 is empty)&&(q_2 is empty)
    $T_{cycle} = T_{cycle3}$;
    go to GSM;
Advertise:
send GATE with new T_{cycle} to all ONUs
GSM:
Calculate T_{poll} and T_{rsv}
start new polling cycle (T_{poll}): loop $\forall$ ONU
OLT assigns $T_{grant} = (T_{poll} - N \times T_g)/N$
OLT collects DATA+REPORTS
OLT sends downstream DATA
end loop
OLT enters sleep for T_{rsv}
ONU enters sleep for T_{rsv}
after sleep time expiration
ONU and OLT wakes up
go to Decision making;
LSM:
start new polling cycle (T_{cycle}): loop $\forall$ ONU
OLT assigns $T_{grant} = (T_{cycle} - N \times T_g)/N$
OLT collects DATA+REPORTS
OLT sends downstream DATA
ONU enters sleep for $T_s = T_{cycle} - T_{grant} - T_{sync}$
end loop
after sleep time expiration
ONU wakes up
go to Decision making;

4.1.2.2 Local Sleep Mode (LSM)

In this mode, the ONUs can switch to the sleep state, but not the OLT. This mode occurs when the *class* 1 and *class* 2 traffic are present. If *class* 1 packets exist, the constraint is the delay for *class* 1 traffic and this delay is confined within one polling cycle, therefore, a short polling cycle duration is selected. The T_{cycle} is set as T_{cycle1} which is the delay constraint of *class* 1, given ($T_{cycle1} \geq \sum_{i=1}^{N}(T_{min}^{i} + T_g)$ and ($T_{min}^{i} \geq T_{q_1}^{i}$) where T_{q1}^{i} is the requested bandwidth for *class* 1 packets (q_1). In this mode, ($T_s^{OLT} = 0$) and ($T_a^{OLT} = T_{cycle1}$), and the durations of the sleep and wake up time for the ONUs are calculated as follows:

$$T_s^{ONU_i} = T_{cycle1} - T_{grant}^{i} - T_{oh} \tag{4.14}$$

$$T_a^{ONU_i} = T_{grant}^{i} + T_{oh} \tag{4.15}$$

where T_{grant}^{i} denotes the granted transmission window for i^{th} ONU.

However, if there is no *class* 1 packet, but *class* 2 packets are present, the constraint is the delay of *class* 2 traffic ($i.e., T_{cycle} = T_{cycle2}$) where T_{cycle2} is the delay constraint of *class* 2 given $T_{cycle2} \geq \sum_{i=1}^{N}(T_{min}^{i} + T_g)$ where $T_{min} \geq T_{q_2}$. In this case, ($T_s^{OLT} = 0$) and ($T_a^{OLT} = T_{cycle2}$), and the durations of the sleep and wake up time for each ONU, are given as:

$$T_s^{ONU_i} = T_{cycle2} - T_{grant}^{i} - T_{oh} \tag{4.16}$$

$$T_a^{ONU_i} = T_{grant}^{i} + T_{oh} \tag{4.17}$$

Because T_{cycle1} is the shortest polling cycle for enabling sleep within the bandwidth allocation, T_{cycle1} should be long enough and at least double the T_{oh} [58]. The stringent T_{cycle1} may cause a longer delay and a high packet dropping rate for *class* 2 and *class* 3 traffic.

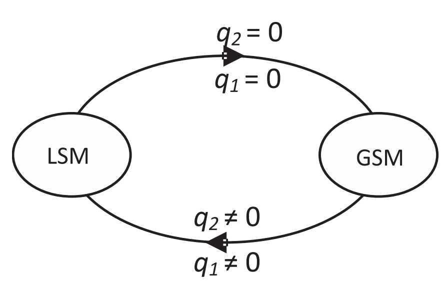

Figure 4.5: Transition between global and local sleep states.

Figure 4.5 depicts the transition between the GSM and LSM based on the queue conditions. All the operations of the proposed dynamic polling cycle based-power saving mechanism are summarized in Algorithm 4.1.

4.1.3 Sleep-Enabled Inter-ONU Scheduler

The durations of the polling cycle T_{cycle} and the two sub cycles T_{poll} and T_{rsv} and the transition between one polling cycle to another are the key design factors for the DPCB-PSM. This scheme supports differentiated sleep time which resolves the problem of energy saving under load variations, i.e., priority and size, at all ONUs. Hence, each ONU enters into sleep mode for a specific duration based on its traffic in order to maintain the QoS. Once this sleep time is specified, the OLT can estimate the wake up time and schedule the grant accordingly.

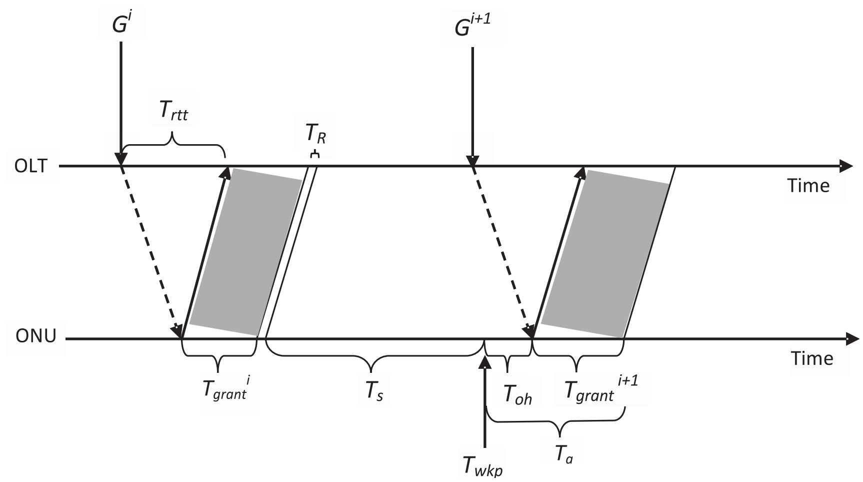

Figure 4.6: Grant scheduling for sleep-enabled ONU.

In the proposed scheme, the OLT schedules an ONU such that its request arrives before its sleep time and at the end of its transmission window (see Fig. 4.6). Based on the grant (G_i), the OLT can calculate the wake up time according to an equation given as:

$$T_{wkp} = T_s + G^i + T_{rtt} + T^i_{grant} \tag{4.18}$$

where T_{wkp} denotes the wake up time for the sleeping ONU; T_{rtt} denotes the round trip time; T_s denotes the sleep duration. From Fig. 4.6, to minimize the waste of energy, the next grant for the same ONU should be:

$$G^{i+1} = G^i + T^i_{grant} + T_s + T_{oh} \tag{4.19}$$

Wake up times for the ONUs change with their traffic condition and the polling cycle durations. Thus the scheduling of the grant, which is based on the wake up time, changes as well in terms of the order of grants and requests. Accordingly, in every polling cycle, the number of scheduled ONUs and the order of grants usually change. The OLT should synchronize the grant with the wake up instant of the ONU to maximize the sleeping time.

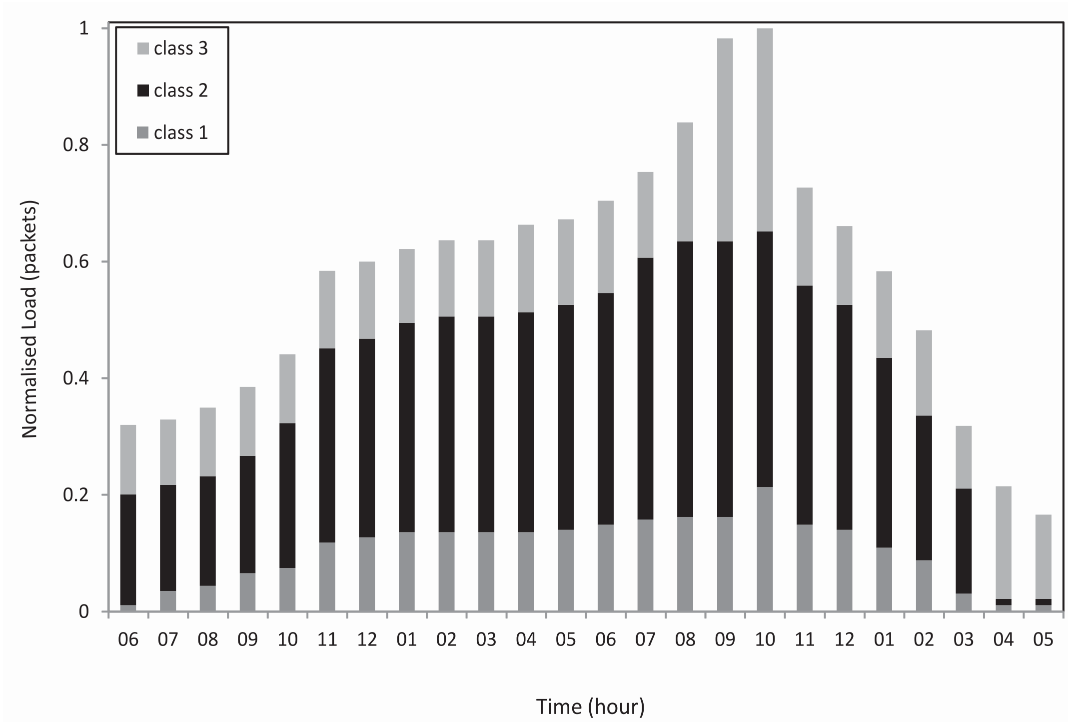

Figure 4.7: Daily traffic profile for different traffic classes at the ONU.

4.1.4 Energy Calculation

Given the sleep and active time for the ONU and OLT in Section 4.1.2, the energy consumption (E) can be calculated for the whole optical part of WOBAN (10G-PON) network using the following equation:

$$E = N \times (T_s^{ONU} \times P_s^{ONU} + T_a^{ONU} \times P_a^{ONU}) + (T_s^{OLT} \times P_s^{OLT} + T_a^{OLT} \times P_a^{OLT}) \quad (4.20)$$

where $P_s^{ONU}$and P_s^{OLT} are the power consumption during the sleep mode for ONU and OLT, respectively. $P_a^{ONU}$and P_a^{OLT} are the power consumption during the sleep and active mode for ONU and OLT, respectively.

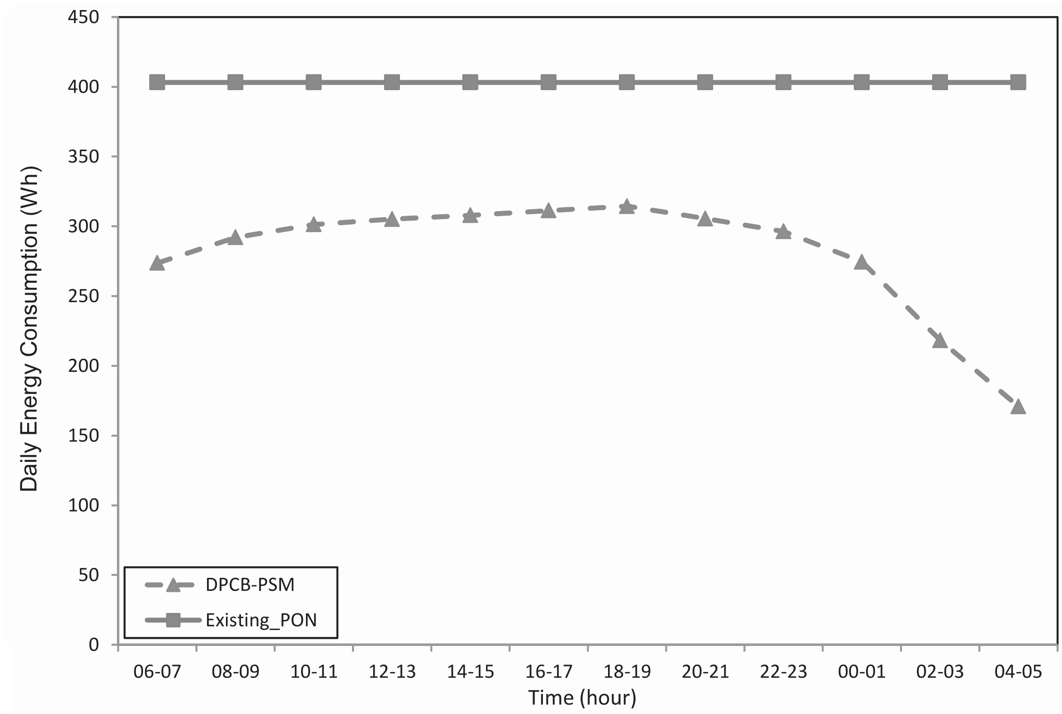

Figure 4.8: Comparison of daily energy consumption between an existing mechanism and the proposed DPCB-PSM.

4.2 Traffic Profile

Internet usage varies considerably during the day in terms of applications, the number of end users and the equipment they use to access the Internet, e.g., smart-phone, tablet, USB and modem. To gain an insight into the traffic classes, the traffic profile over 24 hours periods as shown in Fig. 4.7 is used, which is based on findings reported in [3] [70] [71].

4.3 Performance Evaluation

For performance evaluation, the same scenario was used which was presented in Chapter 3. The proposed Dynamic Polling Cycle Based-Power Saving Mechanism (DPCB-PSM) was applied in a WOBAN, as shown in Fig. (2.9), comprising 16 ONU-BS units. The a is set at 0.98 and the a_1, a_2 and a_3 were set at 0.2, 0.4 and 1, respectively.Within a specified period, the accumulated sleep periods of the ONUs were measured to calculate the percentage sleep time and thus energy saving. All ONUs were considered to have the same load. The

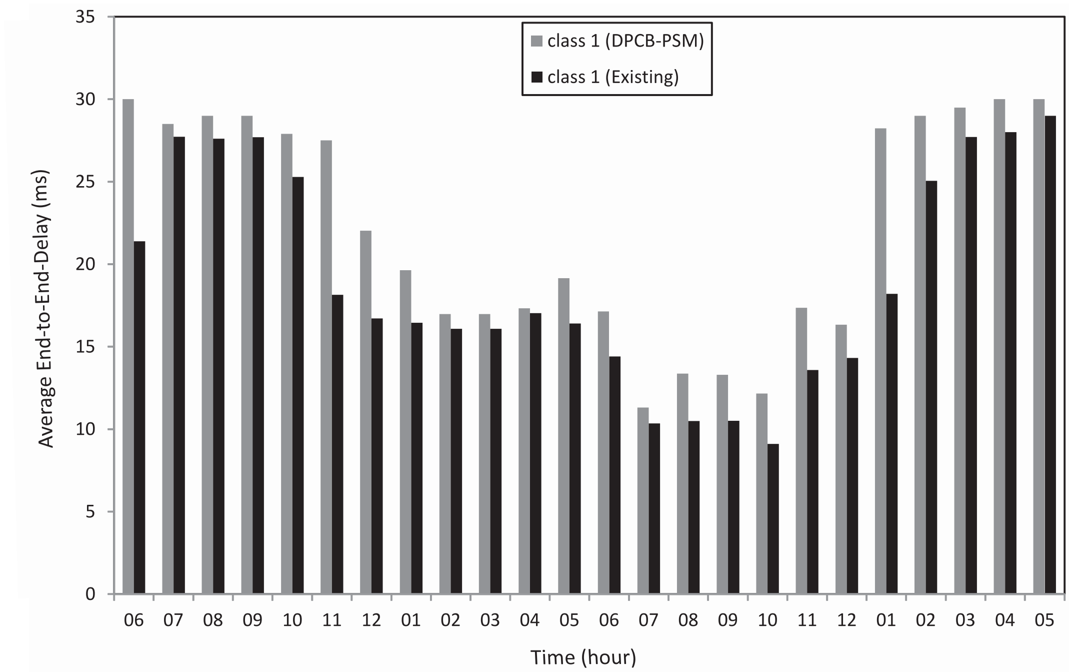

Figure 4.9: End-to-End delay comparisons between an existing mechanism and the proposed DPCB-PSM for *class* 1.

downstream traffic was scheduled synchronous with the scheduling of upstream traffic from a particular ONU, which meant the OLT buffered the traffic for the ONUs and dispatched it when the transmission time of the intended ONU occurred. The power consumption of the ONU was considered as 6.35 W and 0.57 W [54] during the wake up and sleep mode, respectively. Also the power consumption of the OLT was considered as 100 W [72] and 10 W during the wake up and sleep mode, respectively. The accumulated sleep time for the ONUs and OLT were measured within a period. In this work, the proposed mechanism was compared against the existing mechanism presented in [58], which utilized DBA for resource management in a sleep-enabled ONU, but used no service class differentiation. The existing model was simulated with a maximum transmission window restriction that made the maximum polling cycle not exceeding the 10 ms limit which is a preferred duration to suite the high priority traffic. For comparing the proposed mechanism and existing models, the same traffic profile and network scenario were used.

Figure 4.8 shows the energy consumption of PON, observed in the proposed DPCB-PSM and existing DBA mechanism. It is evident that the proposed mechanism significantly improves energy consumption compared to the existing mechanism. On average, the proposed mechanism saves about 44% of energy consumption per day for the simulation scenario. The energy consumption in the proposed model drops to 170 Wh in some cases (e.g., during the 4-5 am period) when the global sleep mode is activated, which saves more than 57% of the total energy consumption compared to the existing scheme. Obviously, the global sleep mode contributes more when only low priority traffic is available in the network. Overall, the proposed mechanism consistently outperforms the existing technique because the existing technique does not consider service class differentiation and puts OLTs and/or ONUs into sleep only when no packets are present in the queue. The proposed mechanism considers service class differentiation and puts OLTs and/or ONUs into sleep even though packets are present in the queue as long as their QoS requirements are met.

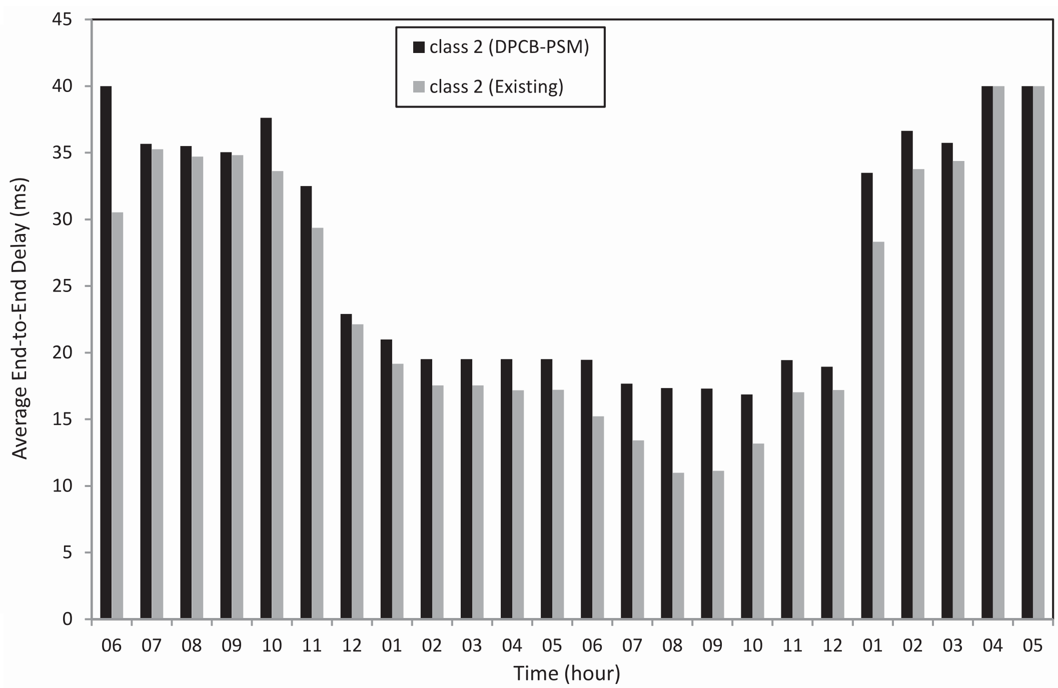

Figure 4.10: End-to-End delay comparisons between an existing mechanism and the proposed DPCB-PSM for *class* 2.

Figures 4.9-4.11 depict the comparison between the DPCB-PSM and existing mechanism in terms of end-to-end delay for different classes of traffic. Figure 4.9 shows the end-to-end delay for *class* 1 traffic in the DPCB-PSM and the existing technique. The average end-to-end delays for delay sensitive services (*class* 1) are comparable for all periods, which is a highly acceptable outcome since the DPCB-PSM saves energy at the same time. The reason for this behavior is the durations of the polling cycles in DPCB-PSM are relatively longer than that for existing work, which eventually increases the sleep duration and causes a slightly longer delay.

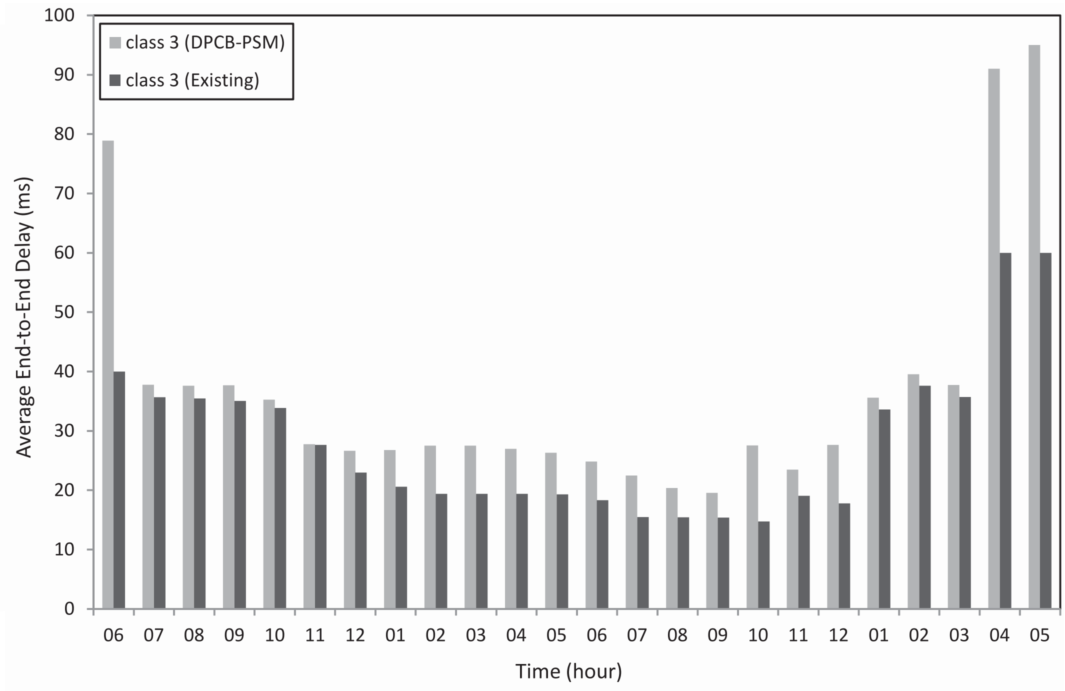

Figure 4.11: End-to-End delay comparisons between an existing mechanism and the proposed DPCB-PSM for *class* 3.

Figure 4.10 shows the end-to-end delay for *class* 2 traffic. Again for the same reason which we mentioned for *class* 1 delay, the end-to-end delays are very close and the same trend is evident for *class* 3 traffic (Fig. 4.11). It can be noted that *class* 1 and 2 traffic can tolerate a maximum delay of 150 ms and 4 s [73], respectively. In this proposed mechanism, *class* 1 and 2 traffic experience an average maximum delay of 30 ms and 40 ms, respectively, which are well within the acceptable limits.

Figure 4.12 depicts the average queuing delay for different traffic types. It is obvious that the Intra-ONU scheduler maintains the queuing delay of the *class* 1 traffic as the minimum among the queuing delays of all the classes, followed by the queuing delay of the *class* 2 traffic.

Figure 4.13 illustrates the average end-to-end delay when the polling cycle is arbitrated dynamically by the OLT. Similar to Fig. 4.12, the average end-to-end delay decreases with the increasing load for the *class* 1 and *class* 2 traffic. This is because the *class* 1 and *class* 2 traffic have less waiting time at queues than *class* 3 and their packets are released with the occurrence of the transmission window. Moreover, because of using 10 ms and 20 ms as polling cycles for *class* 1 and *class* 2, respectively, the length of the transmission window

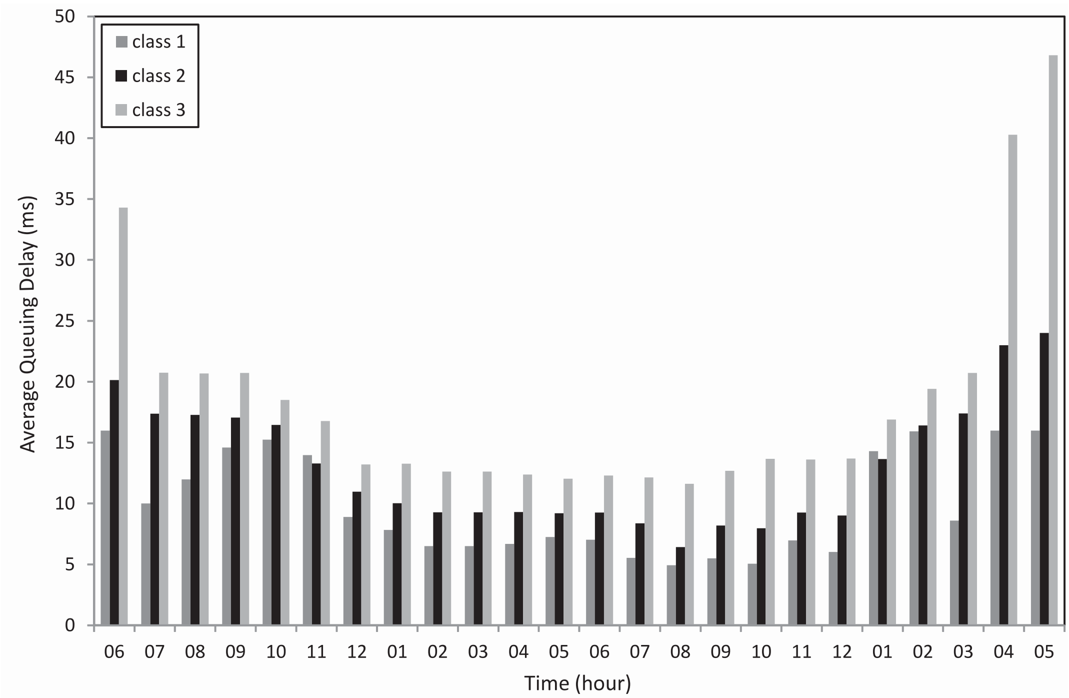

Figure 4.12: Average queuing delay for different traffic classes.

is large enough to transmit all the packets in the queues.

Figure 4.14 depicts the variation of the average throughput for different traffic classes. From this figure, it is evident that the overall throughput for *class* 2 is higher compared to other classes during most of the periods except for the 4-5 am period when the major traffic source is *class* 3.

Figure 4.15 presents the average jitter in WOBAN during the day hours for real time traffic classes. The figure shows that *class* 1 and *class* 2 traffic experiences low jitter even during high load periods.

The impact of the synchronization time interval on the performance of the sleep-aware DBA in WOBAN is also investigated. Figure 4.16 shows that the synchronizing time interval has an impact on the sleep period. It is evident that longer synchronization interval results in decreased accumulated sleep percentage while shorter synchronization interval increases the sleep percentage. This is because longer synchronizing time interval increases the active time, which ultimately decreases the sleep period.

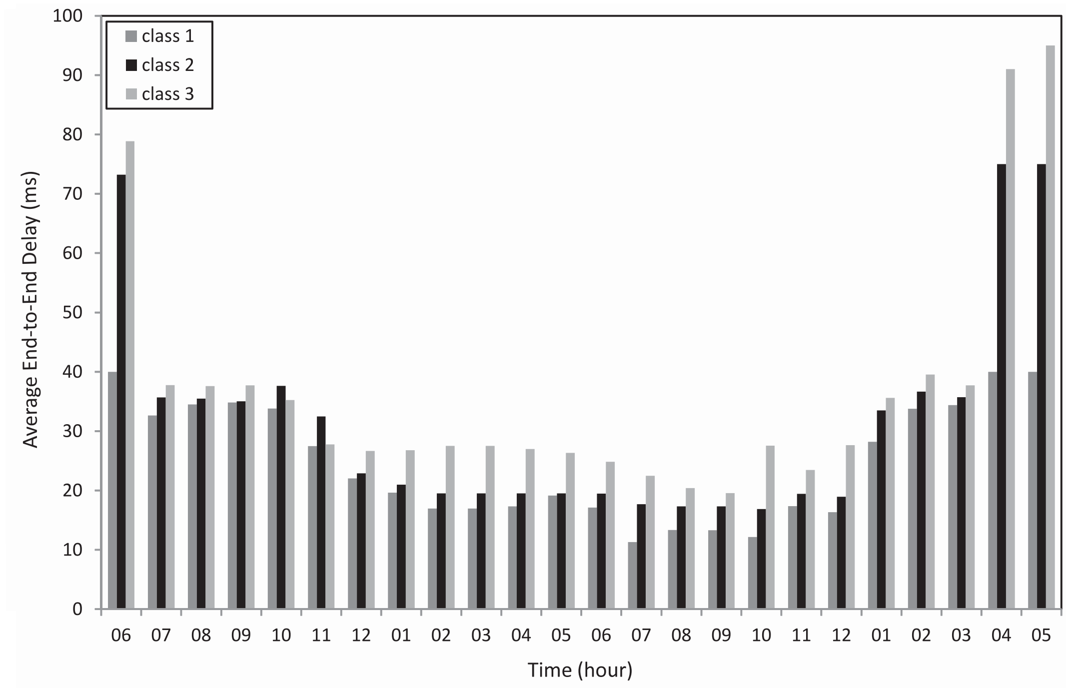

Figure 4.13: Average end-to-end delay for different traffic classes.

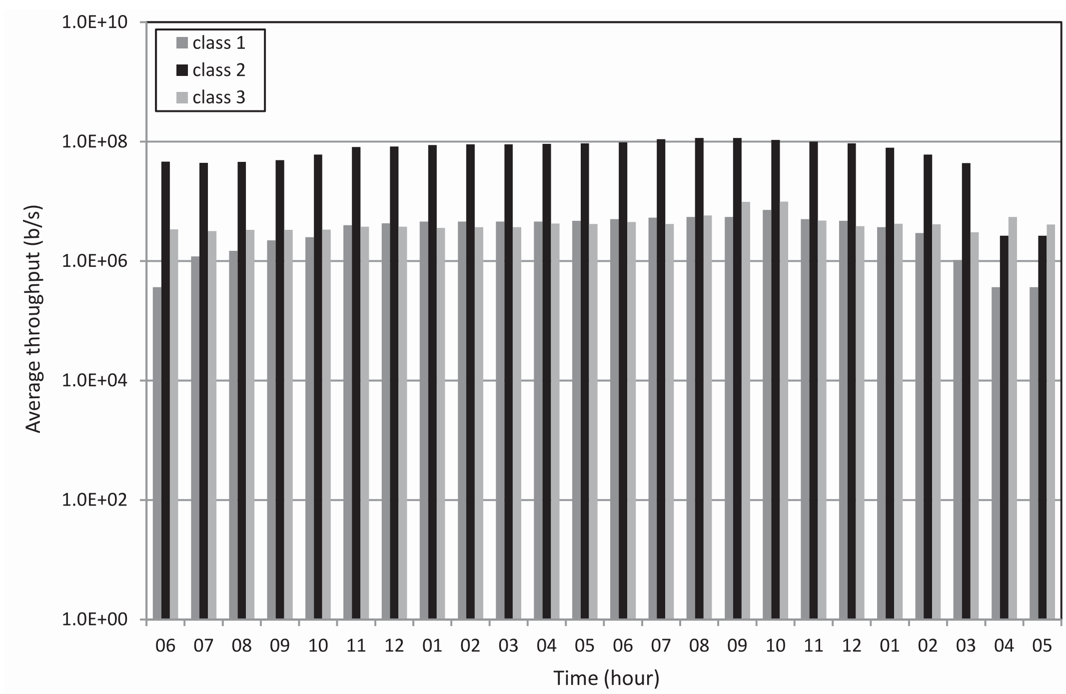

Figure 4.14: Average throughput for different traffic classes.

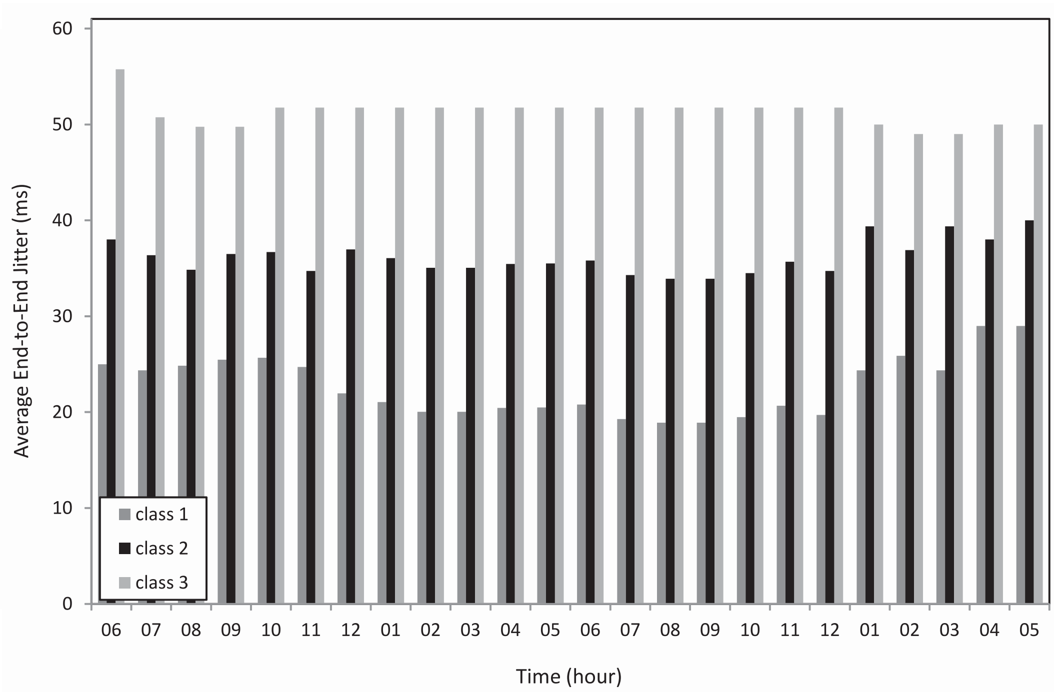

Figure 4.15: Average jitter for different traffic classes.

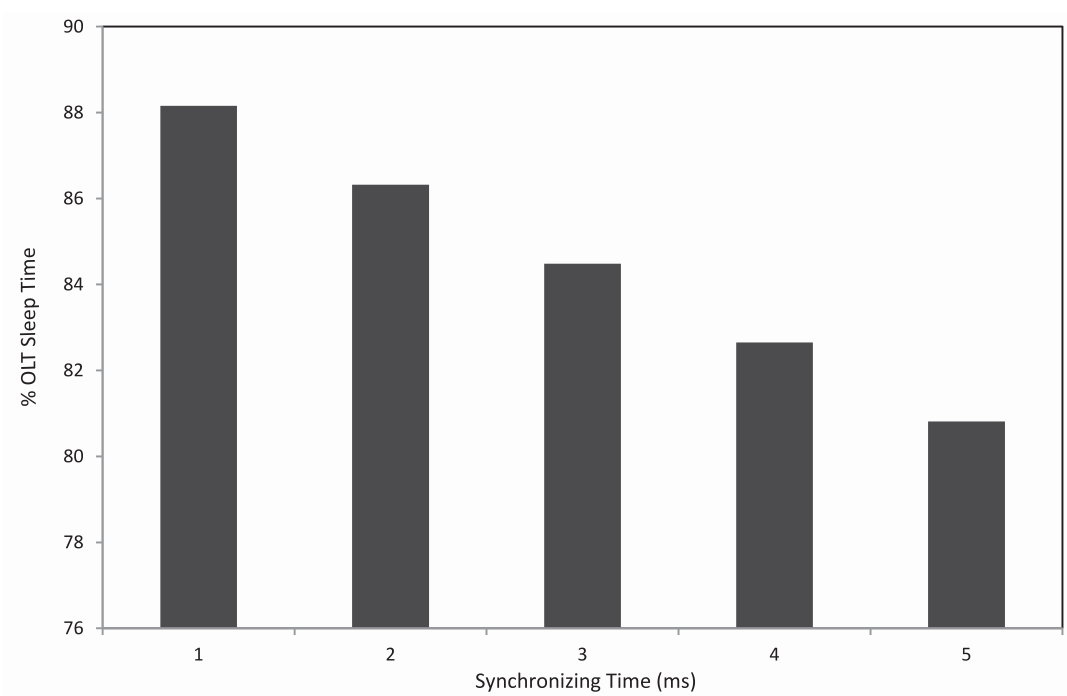

Figure 4.16: Percentage power saving of the OLT with unloaded networks for different synchronization times.

4.4 Closing Remarks

Generally speaking, enabling and controlling the sleep of the ONUs and OLT within the resource management, provide much better QoS support, energy efficiency and channel utilization. Moreover, making resource management dynamics reactive to traffic dynamics is important for supporting QoS. However, without considering sleep mode within resource management, network performance for power consumption and channel utilization will be compromised. Specifically, in light or no traffic load conditions, the dynamic resource management would support QoS but not energy efficiency due to frequent polling associated with assigning short grants for the ONUs. This scenario will lead to the ONUs being almost continuously active. Therefore, to improve energy efficiency, it is important that dynamic resource management operates within predefined durations for polling cycles to sustain reasonable and controlled sleep periods without the need for handshaking between the OLT and ONUs. That synchronisation would increase network overheads.

In this chapter, an efficient resource management scheme was presented that allowed longer sleep periods for OLTs and ONUs without sacrificing the QoS of end applications. This ultimately reduced power consumption with no considerable impact on the QoS. The key innovation was incorporating the concept of service class differentiation, into the power saving mechanism that enabled aggressive use of sleep modes. The proposed approach was underpinned by a model for calculating the sleep periods for both OLTs and ONUs, in response to the varying loads for different traffic classes. While the proposed solution targeted WOBAN in this work, it could be used in any passive optical network irrespective of the last mile access technology. No other research was identified in the literature, that increased sleep mode for the OLT and ONUs, while maintaining the QoS for different traffic classes in WOBAN.

In the next chapter, energy efficiency innovations will be extended to the wireless part of the WOBAN including base stations and subscriber stations.

Chapter 5

Coordinated Resource Management for Green WOBAN

In the previous chapter, the local and global sleep functions were introduced; showing how they improve energy efficiency of the WOBAN by enabling the resource management system to adapt to changes in the existing load. This adaptation is important in order to avoid power wastage by frequent polling the ONUs during periods of light traffic loads. Moreover, channel utilization is enhanced by tailoring the active periods and the sleep periods to be suited to both the requested bandwidth from the ONUs, and the class of the available traffic. This mechanism can also enable power savings to be achieved at the OLT, when a high proportion of best effort traffic is available for transmission during the lightly loaded night time periods. Importantly however, the BSs and SSs are not yet included in this resource management system.

WOBAN is a hierarchical network, and the power efficiency of the wireless part of the network, together with optical part of the network is a challenge, which has not yet been addressed. This chapter deploys similar principles proposed in the previous chapter (Chapter 4), and extends the work to include each of the components of the WOBAN in an energy efficiency mechanism. To this end a coordinated resource management mechanism is proposed; incorporating the resource management of both the wireless part of the network, with the optical part of the network in order to boost the energy efficiency of the WOBAN. The resource management for the wireless and optical parts of the network are investigated, and this chapter provides an understanding of how these two parts can be incorporated together into one coordinated resource management structure. An analytical model of the coordinated resource management for a green WOBAN is presented, showing the calculations for delay, sleep, and active durations of each component of the system, and their power consumption. Finally, this chapter concludes by summarising the important points and significant results.

5.1 WOBAN Network

This work uses the WOBAN architecture which is shown in Fig. 2.9 with the QoS and queue management proposed in Chapter 3. The OLT polls all the ONU-BSs for every new polling cycle (T_{cycle}). The BS polls all the SSs for every new frame size (T_F). Communication with the OLT and BS is based on a poll grant mechanism. The SS and ONU each request bandwidth from the BS and OLT respectively. The BS uses a fixed frame size and a split ratio to assign the resources among the SSs. The split ratio determines the lengths of upstream (T_F^U) and downstream (T_F^D) sub-frames. In upstream communication, the BS slices the upstream sub-frame into a number of time slots and allocates them based on SS requests ($T_{SS}^U = \frac{T_F - T_F^D - T_c - T_R}{n}$) where n depicts the number of SSs; T_c depicts the contention time duration; T_{SS}^U depicts the grant duration (time slot length); and T_R depicts the ranging time duration. The OLT is more flexible than the BS when allocating bandwidth. When utilizing dynamic bandwidth allocation (DBA), the OLT assigns bandwidth based on the amount requested by the ONU, which in turn is an outcome of the queue utilisations. Consequently the polling cycle length changes from cycle to cycle because ($T_{cycle} = \sum_{i=1}^{N} T_{grant}^i + T_g$). The OLT uses the REPORT and GATE messages to implement resource management. The ONU informs the OLT of its bandwidth request by sending a REPORT message, which includes the requests for q_1, q_2, and q_3; and the OLT informs the ONU of the start time and the size of the allocated transmission window, through the GATE message.

In the WOBAN, 10G-EPON and 4G utilize TDMA which means that the ONUs and SSs are in a wait state until the time of their scheduling occurs. As a result, the traffic at the ONUs and the SSs is aggregated at the respective queues, and here the effect of the scheduling duration (which is determined from the length of the polling cycle and the frame size) on the queue length and the QoS is known. As the traffic profile changes during the day, from peak to light load for each of the various classes of data traffic, the required resources change correspondingly. Thus, a fixed scheduling duration is inadequate for unevenly distributed traffic. Furthermore, an inappropriate ratio between the scheduling duration at the BS and OLT creates a bottleneck which causes a degradation of the WOBAN's performance. Therefore, the BS and OLT should change their scheduling durations based on the profile of the existing traffic. The mechanism for changing the scheduling durations for the OLT and BS is described in Section 5.2.

5.1.1 QoS and Energy Efficiency Issues in WOBAN

In the WOBAN, the wireless part of the network represents a bottleneck as the 10G-EPON supports a much higher transmission rate. The resource management of each of the optical and wireless parts of the network affects the QoS of the traffic crossing the interface between these two technologies. In the optical part, during each polling cycle, the OLT polls all the ONUs and assigns a transmission window for each ONU. Within a polling cycle, the ONU only transmits during its grant and is on standby during the rest of the polling cycle; however, the OLT can receive and send data at the same time, as it supports independent downstream and upstream channels.

In the wireless part of the network, the communication between the BS and SSs is via two channels having the same frequency, so the downstream and upstream communication takes place over separate intervals. Consequently the frame is divided into two sub-frames, as described in Section 5.1. The receiver of the SSs and the transmitter of the BS work only during the downstream interval, while the SS's transmitter and the BS's receiver work during the upstream interval where only one SS is allowed to transmit at a time. In this section these idle times at the ONUs, BSs, and SSs within the WOBAN are addressed, and a strategy is presented which aims at reducing them.

As the WOBAN has a hierarchical architecture consisting of heterogeneous networking components, it requires discrete resource management mechanisms operating separately across the optical and the wireless parts of the network. The challenge is to efficiently coordinate the management of these resources across the WOBAN in an effective manner. The design of an appropriate proportion of usage between the two grant cycles (ratio) will strongly influence the WOBAN's performance. A long duration for the polling cycle provides an opportunity for the system to go into sleep mode; but that duration should be continually adjusted in response to the fluctuations and changes within the daily traffic profile. This is an important consideration, as the composition of the network traffic encountered in the so-called last-mile of network access tends to be very unevenly distributed; i.e., different patterns of usage experienced during the day compared with the night time, with correspondingly different classes of data traffic being carried. Accordingly, utilising fixed lengths for the frame size and polling cycle in both the wireless and optical parts of the WOBAN respectively may waste a large amount of energy and use bandwidth inefficiently.

To realise a fully energy efficient WOBAN, the ONUs, OLT, BSs, and SSs should all enter sleep mode during any idle time. The problem arising when implementing such a

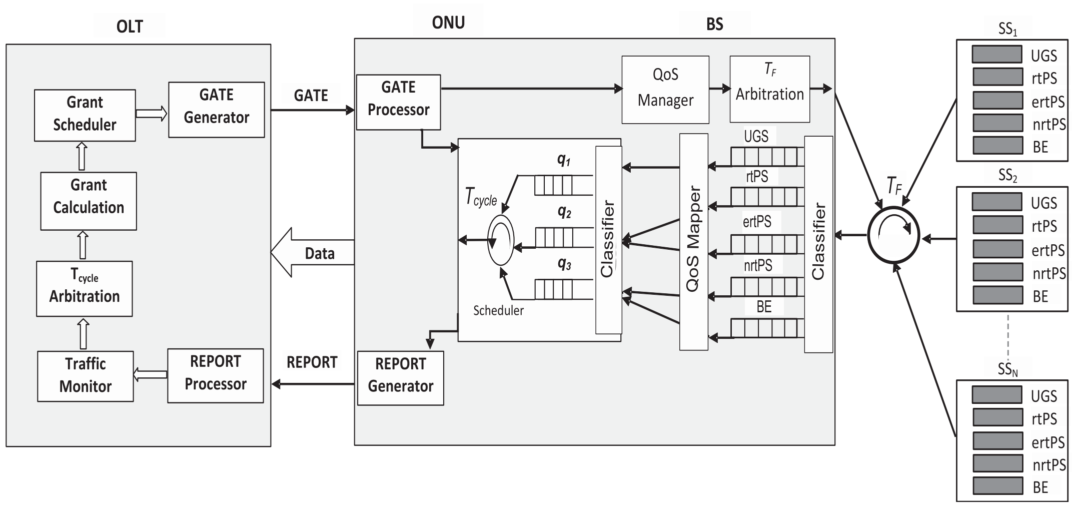

Figure 5.1: Functional model of the proposed Load-adaptive Resource Management strategy in WOBAN.

policy is the need for relatively long polling cycles and long frame sizes at the optical and wireless parts of the network, respectively. These long polling cycles and long frame sizes combine to extend the delays in end-to-end transmission significantly. Furthermore, as the traffic volume and profile fluctuate during the day, the relative usage of different applications also changes. Consequently a mechanism is required to make the distribution of resources and the duration of the grant cycles (polling cycle at the OLT and frame size at the BS) both dynamic and adaptable to the changing traffic conditions. Such an adaptation to the changing traffic conditions means being able to lengthen the grant cycles to the extent that the available class of traffic can accommodate; in order to achieve reasonable sleep durations and to minimise the costs incurred when moving from sleep to wake up state. This adaptation to the changing traffic conditions is able to also facilitate achieving a reduction of power consumption of the OLT and BS. The dynamics of the polling cycle duration of the optical part of the network should be coordinated with the dynamics in the frame size of the wireless part of the network; as both parts contribute to the end-to-end delay, jitter, and other QoS parameters. The interoperability between the resource management strategies at the OLT and BS improves the QoS for different traffic classes.

There is a compelling need for a coordinated approach to resource management. It is proposed that available bandwidth be allocated; based on the dynamic selection of a specific frame size and polling cycle duration at the BS and OLT respectively, in response to the existing load. Such a coordinated approach should result in a dramatic reduction

Algorithm 5.1 Dependent Load-Adaptive Scheduling.

Initialization:
$T_{cycle} = T_{cycle1}$; $T_F = T_{UGS}$
OLT:
Decision making: $\forall$ REPORT
if (q_1 is not empty)
    $T_{cycle} = T_{cycle1}$
    advertise the new T_{cycle}
    go to line BS;
else if (q_1 is empty)&&(q_2 is not empty)
    $T_{cycle} = T_{cycle2}$
    advertise the new T_{cycle}
    go to line BS;
else if (q_1 is empty)&&(q_2 is empty)
    $T_{cycle} = T_{cycle3}$
    advertise the new T_{cycle}
    go to line BS
BS:
if ($T_{cycle} = T_{cycle1}$)
    $T_F = T_{UGS}$
else if ($T_{cycle} = T_{cycle2}$)
    $T_F = T_{rt}$
else if ($T_{cycle} = T_{cycle3}$)
    $T_F = T_{nrt}$

of power consumption across the WOBAN. In the following section, the details of the proposed mechanism are described.

5.2 Proposed Load-Adaptive Resource Management (L-ARM) in Green WOBAN

In this section a load adaptive-resource management (L-ARM) mechanism is proposed; based on coordinating the bandwidth allocation for the wireless and optical parts of the network, in response to load variations being encountered across the WOBAN.

The architecture of the proposed mechanism is presented in Fig. 5.1. The OLT utilises the T_{cycle} arbitration unit in order to select the best polling cycle among a predefined set of polling cycle lengths. Each ONU-BS has a QoS manager unit which monitors the traffic conditions for the BS, and chooses the best available frame size from a set of predefined frame size lengths. The values of T_{cycle} and T_F are continuously being updated in response to the changing traffic conditions. The OLT is at the logical central element of the network, and monitors the traffic from all the ONUs and selects the optimum value of T_{cycle}. Two mechanisms are proposed to enable the QoS manager to change the frame size of the BS, namely: Independent Scheduling (IS) and Dependent Scheduling (DS).

Algorithm 5.2 Independent Load-Adaptive Scheduling.

OLT:
Initialization:
$T_{cycle} = T_{cycle1}$;
Decision making: $\forall$ REPORT
if (q_1 is not empty)
  $T_{cycle} = T_{cycle1}$
  advertise the new T_{cycle}
else if (q_1 is empty)&&(q_2 is not empty)
  $T_{cycle} = T_{cycle2}$
  advertise the new T_{cycle}
else if (q_1 is empty)&&(q_2 is empty)
  $T_{cycle} = T_{cycle3}$
  advertise the new T_{cycle}
BS:
Initialization:
$T_F = T_{UGS}$
Decision making:
if (q_{UGS} is not empty)
  $T_F = T_{UGS}$
else if (q_{UGS} is empty)&&($[q_{rtPS}||q_{ertPS}]$ is not empty)
  $T_F = T_{rt}$
else if (q_{UGS} is empty)&&($[q_{rtPS}||q_{ertPS}]$ is empty) &&($[q_{nrtPS}||q_{BE}]$ is not empty)
  $T_F = T_{nrt}$
where q_{UGS}, q_{rtPS}, q_{ertPS}, q_{nrtPS} and q_{BE} depict the queues for the UGS, rtPS, ertPS, nrtPS and BE, respectively at the BS.

In IS, the QoS manager monitors and evaluates the queue status at the BS in order to make decisions regarding the frame size, and does not need any information from the OLT. In DS, the QoS manager depends on information received from the OLT in order to make decisions regarding the frame size, and utilises the following mechanism. The OLT polls all the ONU-BSs, to which each ONU-BS responds by sending a REPORT message immediately after transmitting the bandwidth requests for each of the queues, q_1, q_2, and q_3. Based on these REPORTs, if at least one ONU requests bandwidth for its q_1, the OLT chooses the shortest polling cycle (T_{cycle1}) and advertises the new polling cycle within the GATE message. Then the QoS manager at each ONU-BS reads the information in the GATE message, and changes the frame size of the BS as needed. The OLT switches to another polling cycle whenever the traffic conditions change. For instance, if *class* 1 traffic is not present in the network, the OLT checks for the existence of *class* 2 traffic; which means that if any ONU sends a REPORT message requesting bandwidth for q_2, the OLT responds by changing to a longer polling cycle (T_{cycle2}). If there is no *class* 1 or *class* 2 traffic, the OLT chooses the longest polling cycle (T_{cycle3}) and advertises the new polling cycle via a GATE message. The QoS manager at each ONU-BS adjusts the frame size of its BS after any change made to the polling cycle. The details of the DS and IS mechanisms are provided in Algorithm 5.1 and Algorithm 5.2, respectively.

The main difference between the DS and IS mechanisms, is that in DS, all ONU-BSs have the same value of T_F because changes in T_F are based on changes to the advertised polling cycle length (T_{cycle}). By way of contrast, in IS every ONU-BS checks its local queues and changes the value of its T_F accordingly. Thus, various ONU-BSs might have different values of T_F when the traffic being encountered at these ONU-BSs does not have the same traffic profile. As a result, DS works better in terms of the end-to-end QoS, as it guarantees the interoperability of the resource management between the BS and the OLT.

As previously noted, the OLT responds to changes in the traffic conditions when updating the polling cycle; in turn affecting the allocated transmission time. The short polling cycle duration reduces the allocated transmission windows and increases the guard time effect. Conversely, a long polling cycle extends the allocated transmission and reduces the guard time effect, but increases the delay. To enable energy saving in the 10G-EPON without requiring negotiation between the ONU and OLT; a dynamic bandwidth allocation in the 10G-EPON is used within a fixed polling cycle length. In 4G, frame size and slot size are fixed. If the frame size is short, the number of the allocated slots will be reduced and eventually the BS will become a bottleneck. However, lengthening the frame size causes an increase in the number of time slots and the length of the delays. Since the traffic profile in the access networks fluctuates during the day, adapting the polling cycle at the OLT and the BS in response to the changing traffic profile, improves network performance and energy conservation. The proposed traffic classification leads to the following cases:

- Case 1 (existing *class* 1 traffic):

The durations of T_{cycle} and T_F should be short as *class* 1 can only tolerates short delays. Then $T_{cycle} = T_{cycle1}$; $T_F = T_{UGS}$

- Case 2 (existing *class* 2 traffic):

The durations of T_{cycle} and T_F can be longer than case 1 as *class* 2 tolerates longer delays. Then $T_{cycle} = T_{cycle2}$; $T_F = T_{rt}$

- Case 3 (existing *class* 3 traffic)

The durations of T_{cycle} and T_F can be relatively long as *class* 3 is not sensitive to delays. Then $T_{cycle} = T_{cycle3}$; $T_F = T_{nrt}$

The relationships between the different durations of polling cycle and frame size are illustrated by Fig. 5.2. The following section explains the details of the proposed mechanism for coordinating resource management at the BS and OLT; which aims to maximize the power saving in Green WOBANs, while maintaining the QoS for the various services.

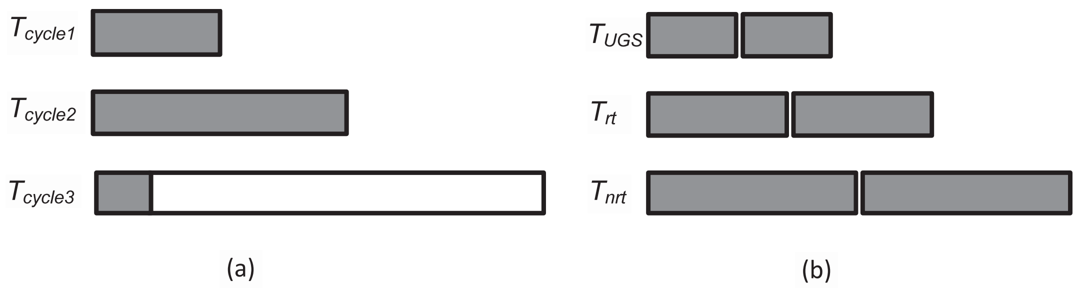

Figure 5.2: Duration of (a) polling cycle and (b) frame size for different traffic classes.

5.2.1 Resource Allocation

In L-ARM, the OLT assigns each ONU the requested bandwidth (T_r) plus extra bandwidth ($\psi \times T_r$), where ψ is a design coefficient, to enable the ONU to send all the packets in its queues, including newly arrived packets (which were not included in the previous REPORT). This flexibility is of particular value when utilising the longer polling cycles. Moreover, the OLT should update the maximum transmission windows ($T_{max} = \frac{T_{cycle} - N \times T_g}{N}$) to enhance the utilization of the channel. Therefore, the transmission window granted for the i^{th}ONU (T^i_{grant}) is:

$$T^i_{grant} = \begin{cases} T_r + \psi \times T_r & if\ (T_r + \psi \times T_r) \leq T_{max} \\ T_{max} & else \end{cases} \quad (5.1)$$

To further enhance channel utilization and fairness among each of the ONUs, it is necessary to take into consideration the lightly loaded ONUs and their excess time; adding the excess time to the loaded ONUs by increasing T_{max} as follows:

$$T_{max} = \frac{T_{cycle} - \sum_{j=1}^{z} T_r^j - N \times T_g}{(N - z)} \quad (5.2)$$

where z depicts the number of the lightly load ONUs.

After describing the mechanism of L-ARM and how a given polling cycle or frame size changes, it now becomes important to show how these dynamics affect the behaviour of the WOBAN. The next section describes and analyses the effect of the proposed L-ARM on the end-to-end delay, jitter, and power consumption of the WOBAN.

5.2.2 Analytical Model for Quantifying the QoS and Energy of L-ARM

The scheduling and propagation delays in the optical and wireless parts of the network, contribute significantly to the overall end-to-end delay (T_{e2e}) of packets when they are moving from the SS to the OLT. The scheduling delay in the optical part of the network can be expressed as:

$$E\left(W_O^K\right) = \frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_k \times E\left(S_k^2\right)}{\left(1-\bar{\rho}_{k-1}\right)\left(1-\bar{\rho}_k\right)} + C \times (T_{cycle} - \frac{T_{cycle} - N \times T_g}{N}) \tag{5.3}$$

The propagation delay of the optical (T_P^O) part of the network can be expressed as:

$$T_P^O = \frac{D^{ONU}}{S_O} \tag{5.4}$$

The total delay in the optical part of the network can be expressed as:

$$T_d^O = E(W_O^K) + T_P^O \tag{5.5}$$

where $E(W_O^K)$ denotes the expected waiting time of the arriving packets in their corresponding queue K at the ONU due to scheduling and queuing delays; T_P^O depicts the propagation delay of the optical network; D^{ONU} denotes the distance between the ONU and the OLT; and S_O denotes the speed of light in the fibre (2×10^8m/s) if the refractive index of the fibre is considered as 1.5. For upstream transmission it is assumed that the delays at the SS and the ONU are approximately equivalent, because both of 4G and 10G-EPON use TDMA, Eq. (5.3) can be adapted to suit the wireless part as:

$$E\left(W_W^K\right) = \frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_k \times E\left(S_k^2\right)}{\left(1-\bar{\rho}_{K-1}\right)\left(1-\bar{\rho}_K\right)} + C \times (T_F - \frac{T_F \times \frac{T_F^U}{T_F^U+T_F^D}}{n})$$

where $E(W_W^K)$ depicts the expected scheduling delay at the wireless part; and n depicts the number of SSs which are connected to the BS. The propagation delay at the wireless part (T_P^W) can be expressed as:

$$T_P^W = \frac{D_{SS}}{S_W} \tag{5.6}$$

where D_{SS} depicts the distance between the SS and the BS; and S_W denotes the speed of light in a vacuum ($3 \times 10^8 \mathrm{m/s}$). The total delay in the wireless part of the network can then be calculated as:

$$T_d^W = E(W_W^K) + T_P^O \tag{5.7}$$

Then the expected end-to-end delay from SS to OLT for *class* K traffic ($E(T_{e2e}^K)$) can be expressed as:

$$E(T_{e2e}^K) = E(T_d^O) + E(T_d^W) \tag{5.8}$$

$$E(T_{e2e}^K) = \frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_k \times E\left(S_k^2\right)}{\left(1-\bar{\rho}_{K-1}\right)\left(1-\bar{\rho}_K\right)} + C \times (T_{cycle} - \frac{T_{cycle} - N \times T_g}{N}) + \frac{D_{ONU}}{S_O} + \frac{\frac{1}{2}\sum_{k=1}^{K}\lambda_k \times E\left(S_k^2\right)}{\left(1-\bar{\rho}_{K-1}\right)\left(1-\bar{\rho}_K\right)} + C_1 \times (T_F - \frac{T_F \times \frac{T_F^U}{T_F^U+T_F^D}}{n}) + \frac{D_{SS}}{S_W} \tag{5.9}$$

$$E(T_{e2e}^{K}) = \frac{\sum_{k=1}^{K} \lambda_k \times E\left(S_k^2\right)}{(1-\bar{\rho}_{K-1})(1-\bar{\rho_K})} + C \times (T_{cycle} - \frac{T_{cycle} - N \times T_g}{N}) + C_1 \times (T_F - \frac{T_F \times \frac{T_F^U}{T_F^U + T_F^D}}{n}) + \frac{D_{SS}}{S_W} + \frac{D_{ONU}}{S_O} \tag{5.10}$$

From Eq. (5.10), it can be seen that the grant cycle (T_{cycle} and T_F) of both the optical and wireless parts of the network affects the end-to-end delay. For this reason, the durations of T_{cycle} and T_F must change in response to the class of traffic in order to control the end-to-end delay.

For example, if *class* 1 is available then the expected end-to-end delay for *class* K traffic is calculated as:

$$E(T_{e2e}^{K}) = \frac{\sum_{k=1}^{K} \lambda_k \times E\left(S_k^2\right)}{(1-\bar{\rho}_{K-1})(1-\bar{\rho_K})} + C \times (T_{cycle1} - \frac{T_{cycle1} - N \times T_g}{N}) + C_1 \times (T_{UGS} - \frac{T_{UGS} \times \frac{T_{UGS}^U}{T_{UGS}^U + T_{UGS}^D}}{n}) + \frac{D_{SS}}{S_W} + \frac{D_{ONU}}{S_O} \tag{5.11}$$

where T_{UGS}^U and T_{UGS}^D depict the upstream and downstream sub-frames for the frame size of T_{UGS}.

If *class* 1 is not available, but *class* 2 is available, then

$$E(T_{e2e}^{K}) = \frac{\sum_{k=2}^{K} \lambda_k \times E\left(S_k^2\right)}{(1-\bar{\rho}_{K-1})(1-\bar{\rho_K})} + C \times (T_{cycle2} - \frac{T_{cycle2} - N \times T_g}{N}) + C_1 \times (T_{rt} - \frac{T_{rt} \times \frac{T_{rt}^U}{T_{rt}^U + T_{rt}^D}}{n}) + \frac{D_{SS}}{S_W} + \frac{D_{ONU}}{S_O} \tag{5.12}$$

where T_{rt}^U and T_{rt}^D depict the upstream and downstream sub-frames for the frame size of T_{rt}.

If *class* 1 and *class* 2 are not available, but *class* 3 is available, then

$$E(T_{e2e}^{K}) = \frac{\lambda_3 \times E\left(S_3^2\right)}{(1-\bar{\rho}_2)(1-\bar{\rho}_3)} + C \times (T_{cycle3} - \frac{T_{cycle3} - N \times T_g}{N}) + C_1 \times (T_{nrt} - \frac{T_{nrt} \times \frac{T_{nrt}^{U}}{T_{nrt}^{U}+T_{nrt}^{D}}}{n}) + \frac{D_{SS}}{S_W} + \frac{D_{ONU}}{S_O} \quad (5.13)$$

where T_{nrt}^{U} and T_{nrt}^{D} depict the upstream and downstream sub-frames for the frame size of T_{nrt}.

Jitter, one of the QoS parameters, captures the delay variation of the received packets. Some services such as CBR traffic are highly sensitive to jitter and need careful management. The difference in the stay time of packets at different queues, due to the scheduling at the wireless and optical parts of the network, contributes to the end-to-end jitter from the SS to the OLT. Again, this can be interpreted as ($E(J_{e2e}) = f(T_{cycle} + T_F)$), where $E(J_{e2e})$ depicts the expected end-to-end jitter.

The total power consumption of the WOBAN (P_{WO}) is the aggregation of the power consumption of the optical components (OLT and ONUs) and the wireless components (BS and SSs) which are (P^{OLT} and P^{ONU}) and (P^{BS} and P^{SS}) respectively. For the WOBAN which is constructed from N ONUs that are connected to the OLT, N BSs, and n SSs that are connected for each BS (Fig. 2.9), the aggregated power consumption can be expressed as following:

$$P^{WO} = P^{OLT} + N \times (P^{ONU} + P^{BS}) + N \times n \times P^{SS} \quad (5.14)$$

To calculate the energy consumption of the green WOBAN it is necessary to analyse the resource management, in order to calculate the sleep and active durations of each component. By using the mechanism presented in Chapter 4, the OLT dynamically moves from one polling cycle to another, according to the type and priority of the traffic being carried within the network at that present time. The ONU is permitted to enter into sleep mode when the duration of the polling cycle is no less than 10 ms. The OLT can enter into sleep mode only with either the existence of *class* 3 traffic, or when there is no traffic at all. The calculations of the sleep time and active time for the ONUs and the OLT are detailed in Chapter 4, and are shown in the following equations:

$$T_S^{ONU} = \begin{cases} T_{cycle1} - T_{grant} - T_{oh} & case\,1 \\ T_{cycle2} - T_{grant} - T_{oh} & case\,2 \\ T_{cycle3} - (\sum T_{grant} + N \times T_g) & case\,3 \end{cases} \tag{5.15}$$

where $(\sum T_{grant} + N \times T_g)$ is a very short time which is about 2% of T_{cycle3}.

$$T_S^{OLT} = \begin{cases} 0 & case\,1 \\ 0 & case\,2 \\ T_{cycle3} - (\sum T_{grant} + N \times T_g) & case\,3 \end{cases} \tag{5.16}$$

$$T_a^{ONU} = \begin{cases} T_{grant} + T_{oh} & case\,1 \\ T_{grant} + T_{oh} & case\,2 \\ \sum T_{grant} + N \times T_g & case\,3 \end{cases} \tag{5.17}$$

$$T_a^{OLT} = \begin{cases} T_{cycle1} & case\,1 \\ T_{cycle2} & case\,2 \\ \sum T_{grant} + N \times T_g & case\,3 \end{cases} \tag{5.18}$$

where $T_S^{ONU}$and T_a^{ONU} depict the sleep and active durations of the ONU respectively; and T_S^{OLT} and T_a^{OLT} depict the sleep and active durations for the OLT, respectively.

The communication between the BS and each of the SSs is half duplex; meanings that only one direction of communication can take place at any time. As shown in Fig. 5.3 in the downstream, the BS sends information and the SSs receive the information; while in the upstream, the SSs send and BS receives. The transceiver system at the BS is the most power demanding component [60], so keeping the transmitter of the BS on during the upstream transmission, or the receiver on during the downstream transmission wastes power. Similarly, keeping the transmitter of each of the SSs on during the downstream

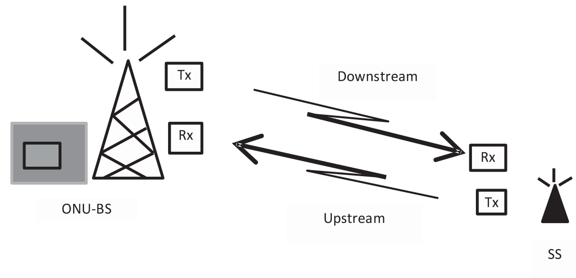

Figure 5.3: The downstream and upstream communication between the BS and one SS where only one direction is activated each time.

transmission, or the receiver on during the upstream transmission also wastes power. In order to reduce energy wastage it is important that the sleep mode for the transmitters and receivers of the BSs and SSs be enabled. Figure 5.4 shows the transition between the

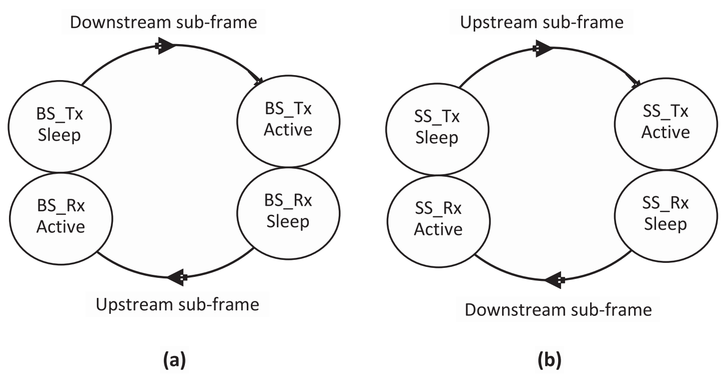

Figure 5.4: Power states of transmitter (Tx) and receiver (Rx) of BS and SS in relation to the upstream and downstream sub-frames.

two power states, sleep and active, for the transmitter and receiver of the BS (Fig. 5.4a) and the SS (Fig. 5.4b). The calculations of the sleep duration for the transmitter $T_S^{Tx(SS)}$ and $T_S^{Tx(BS)}$ and the receiver $T_S^{Rx(SS)}$ and $T_S^{Rx(BS)}$ for the SS and BS, respectively, are:

$$T_S^{Rx(SS)} = T_S^{T_X(BS)} = \sum (T_{nrt} \times \frac{T_{nrt}^U}{T_{nrt}^D + T_{nrt}^U} - T_{ov}) \tag{5.19}$$

$$T_S^{Tx(SS)} = T_S^{R_X(BS)} = \sum (T_{nrt} \times \frac{T_{nrt}^D}{T_{nrt}^D + T_{nrt}^U} - T_{ov}) \tag{5.20}$$

The calculations of the active time duration for the transmitter $T_a^{Tx(SS)}$ and $T_a^{Tx(BS)}$ and the receiver $T_a^{Rx(SS)}$ and $T_a^{Rx(BS)}$ for the SS and BS, respectively, are:

$$T_a^{Tx(SS)} = T_a^{Rx(BS)} = \sum (T_{nrt} \times \frac{T_{nrt}^U}{T_{nrt}^D + T_{nrt}^U} + T_{ov}) \tag{5.21}$$

$$T_a^{Rx(SS)} = T_a^{Tx(BS)} = \sum (T_{nrt} \times \frac{T_{nrt}^D}{T_{nrt}^D + T_{nrt}^U} + T_{ov}) \tag{5.22}$$

Where T_{ov} depicts the overhead time required for the transmitter and receiver to switch from sleep to active state. Enabling the sleep mode requires a longer sub-frame than T_{ov}. Therefore to enable the sleep mode at the BS and SSs where there is a split ratio of 1:1, the T_F should not be less than 20 ms.

In order to calculate the total energy consumption of the green WOBAN within T time, once the sleep and active times have been determined, the power consumption during these times must be calculated for each of the components of the WOBAN.

$$\begin{aligned} T \times P^{WO} = & \sum_T (T_S^{OLT} \times P_S^{OLT} + T_a^{OLT} \times P_a^{OLT} + \\ & N \times (T_S^{ONU} \times P_S^{ONU} + T_a^{ONU} \times P_a^{ONU} + \\ & T_S^{BS} \times P_S^{BS} + T_a^{BS} \times P_a^{BS} + \\ & n \times (T_S^{SS} \times P_S^{SS} + T_a^{SS} \times P_a^{SS}))) \end{aligned} \tag{5.23}$$

Table 1 shows the estimated power consumptions for the OLT, ONU, and BS in

Table 5.1: Power Consumption and Notation for Different Components of WOBAN

Component	Power Consumption (W)		Notation	
	Active	Sleep	Active	Sleep
OLT	100	10	P_a^{OLT}	P_S^{OLT}
ONU	6.35 [54]	0.57	P_a^{ONU}	P_S^{ONU}
BS	50 [60]	8.75	P_a^{BS}	P_S^{BS}
SS	-	-	P_a^{SS}	P_S^{SS}

both sleep and active mode, and their notations. The values for the BS are based on the assumption that a GSM 900 BS is being used, where each TRx contributes about 50 W of the overall power consumption [60]. This work assumes that each BS only has one TRx. The power consumptions for the SSs are not reported, consequently the SSs' consumption are not included in the results.

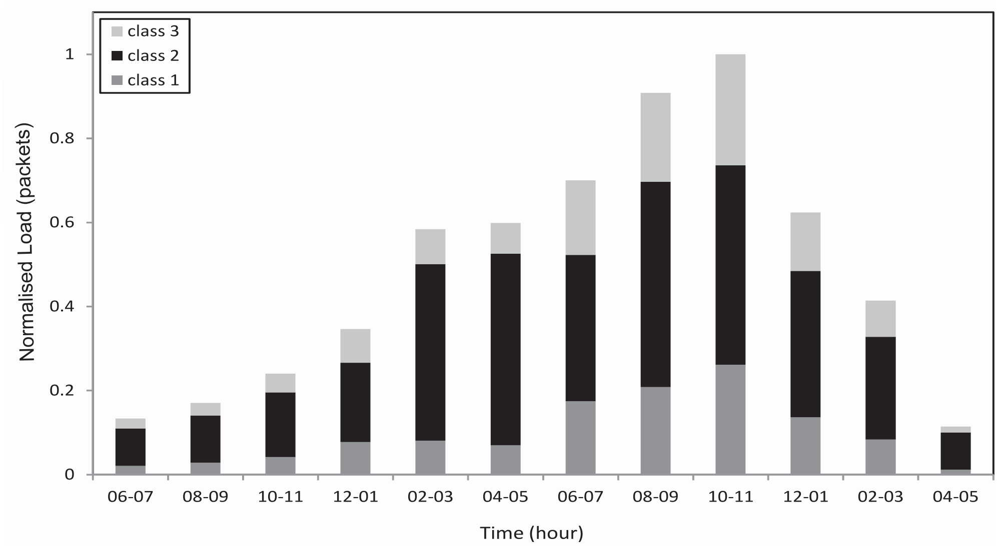

Figure 5.5: Traffic profile taking into consideration traffic classes during the day at the ONU-BS.

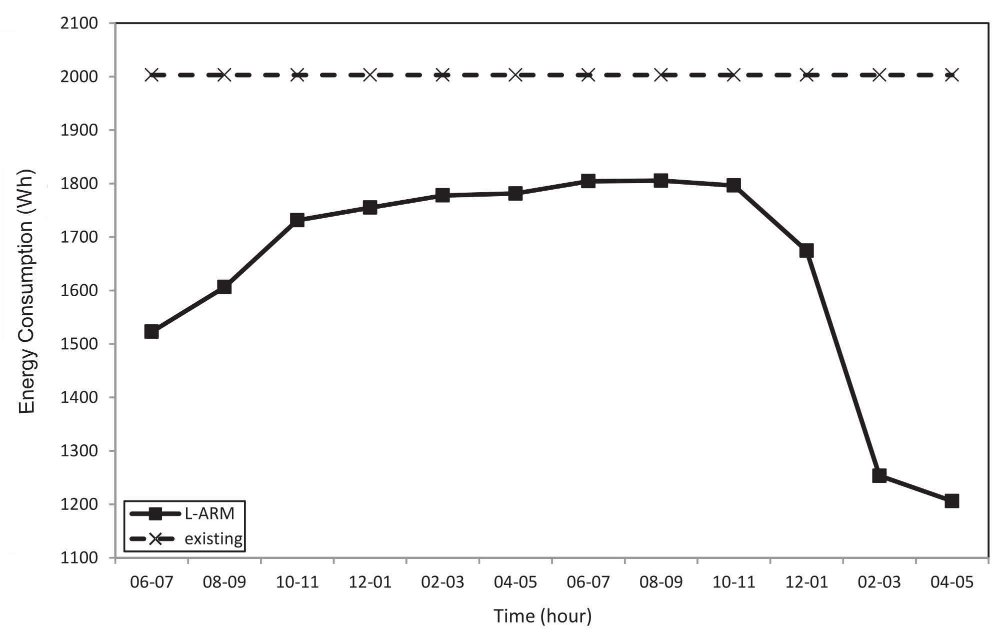

Figure 5.6: Daily power consumption comparison between the L-ARM scheme and existing solution.

5.3 Performance Evaluation

In this section the performance evaluation of the proposed coordinated resource management mechanism for the green WOBAN is presented. The profile of the generated traffic is shown in Fig. 5.5. The traffic rate is altered by changing the packet generation interval and the number of connections. A comparison is made between the performance of the proposed solution against the existing solution presented in [8] [58]. In their dynamic bandwidth allocation schemes, the authors used a fixed polling cycle duration to mitigate the effect of the guard time. This concept was implemented by distributing the excess bandwidth, which is the residual bandwidth of the lightly loaded ONUs, evenly across the heavily loaded ONUs, and by using a pair (T_{cycle}, T_F) to refer to each curve of the existing work under discussion.

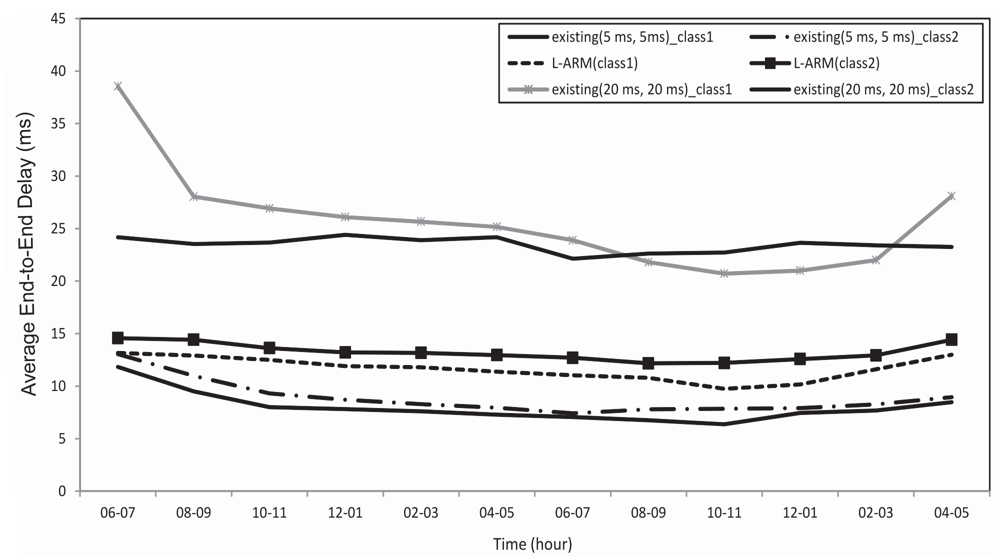

Figure 5.7: End-to-end delay comparison for *class* 1 and *class* 2 between the L-ARM scheme and existing solutions (with 5 ms and 20 ms durations of T_{cycle} and T_F).

Figure 5.6 shows a comparison between the performance of the L-ARM scheme and the existing solutions, which permit sleep when there is no traffic. The traffic profile in Fig. 5.5 reveals the existence of traffic during all periods, but the traffic is shown to be comprised of a variety of volumes and classes. When this load is applied to the WOBAN network, the proposed mechanism performs efficiently in terms of power saving, as illustrated in Fig. 5.6 where most of the energy saving is achieved over the light load periods. This figure takes into consideration the power consumption of the BSs, the ONUs, and the OLT. This proposed mechanism outperforms the existing solutions, as it exploits the idle durations in order to enable devices to enter sleep mode across both the wireless and optical parts of the network.Power savings of up to 40 % are achieved during off-peak hours,

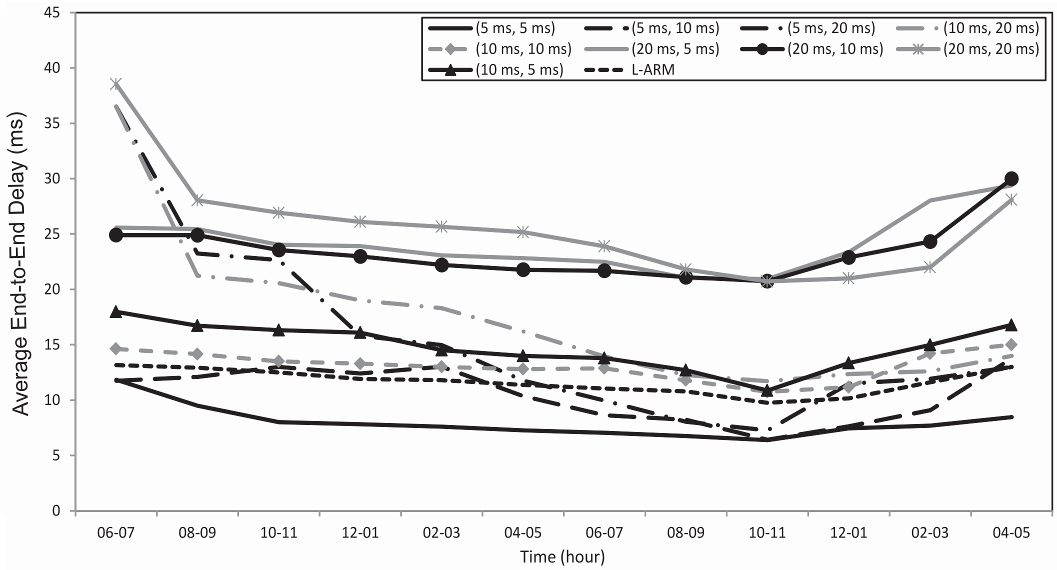

Figure 5.8: Average end-to-end delay for *class* 1 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

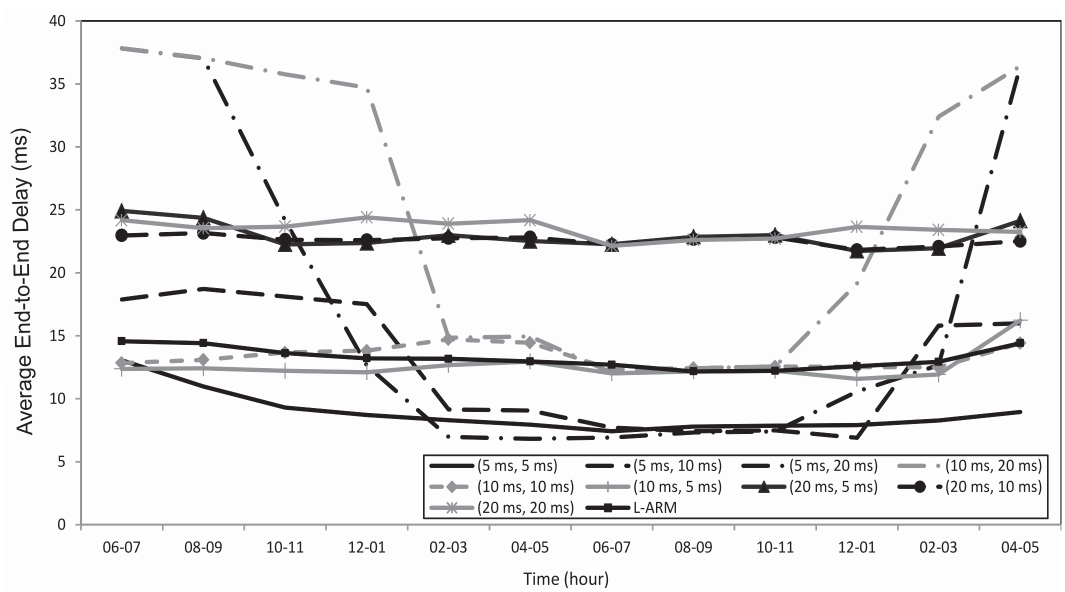

Figure 5.9: Average end-to-end delay for *class* 2 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

Furthermore, the proposed mechanism consistently delivers an acceptable end-to-end delay for *class* 1 and *class* 2 traffic when compared to the existing work for (5 ms, 5 ms) and (20 ms, 20 ms) as shown in Fig. 5.7 because the proposed mechanism adjusts the setting of the network to the available load, which means adjusting the sleep durations and the resources according to the avaiable load. This figure shows that the L-ARM scheme produces significantly lower end-to-end delays over existing solutions for (20 ms, 20 ms); but slightly longer delays for(5 ms, 5 ms). Importantly however, in the proposed mechanism a polling cycle of less than 10 ms is not used.

The comparison also includes the impact of various T_{cycle} and T_F combinations (existing solutions) and the proposed L-ARM scheme on the performance of the WOBAN. Figures 5.8 - 5.11 describe average end-to-end and queuing delays for *class* 1 and *class* 2 traffic. In these figures L-ARM performs better than the other combinations for existing solutions, except where T_{cycle} and T_F are equal to 5 ms. It can be seen that the behaviour of the average queuing delays is similar to that for end-to-end delays. For the same T_{cycle}, the delay increases with longer T_F. The maximum delay occurs on the existing solution at (20 ms, 20 ms), reaching a value of 40 ms. When T_{cycle} is 20 ms the delay is considerably higher, whatever the value of T_F. This figure confirms that a T_{cycle} duration of more than 10 ms is unsuitable for *class* 1 if an end-to-end delay of less than 30 ms is required.

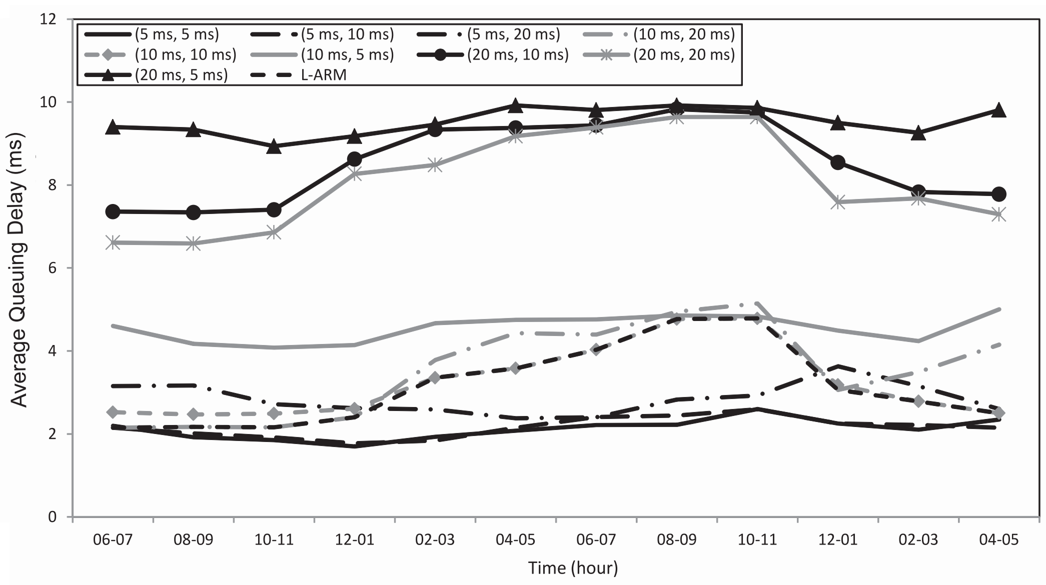

Figure 5.10: Average queuing delay for *class* 1 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

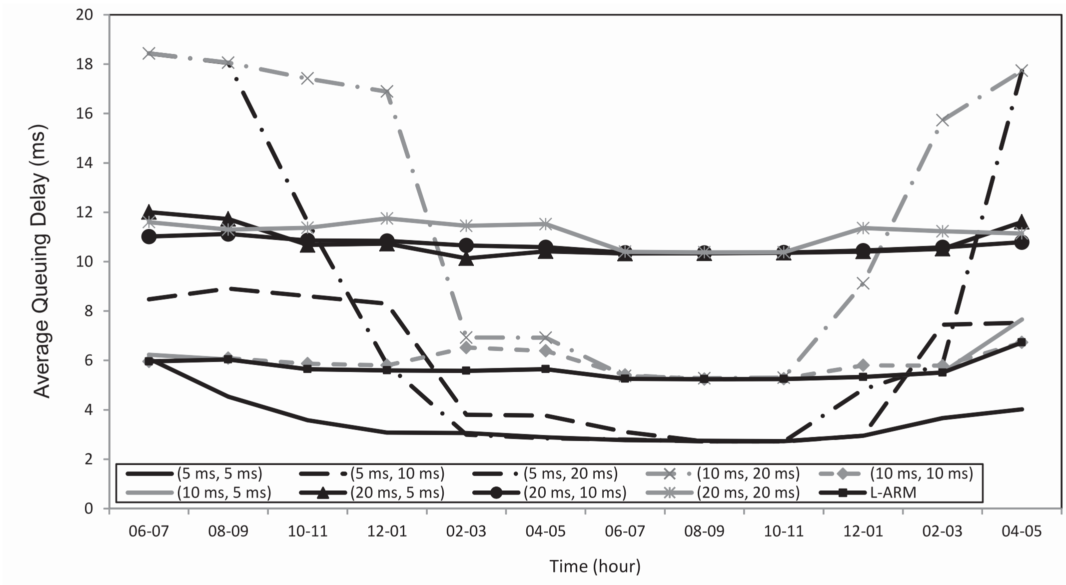

Figure 5.11: Average queuing delay for *class* 2 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

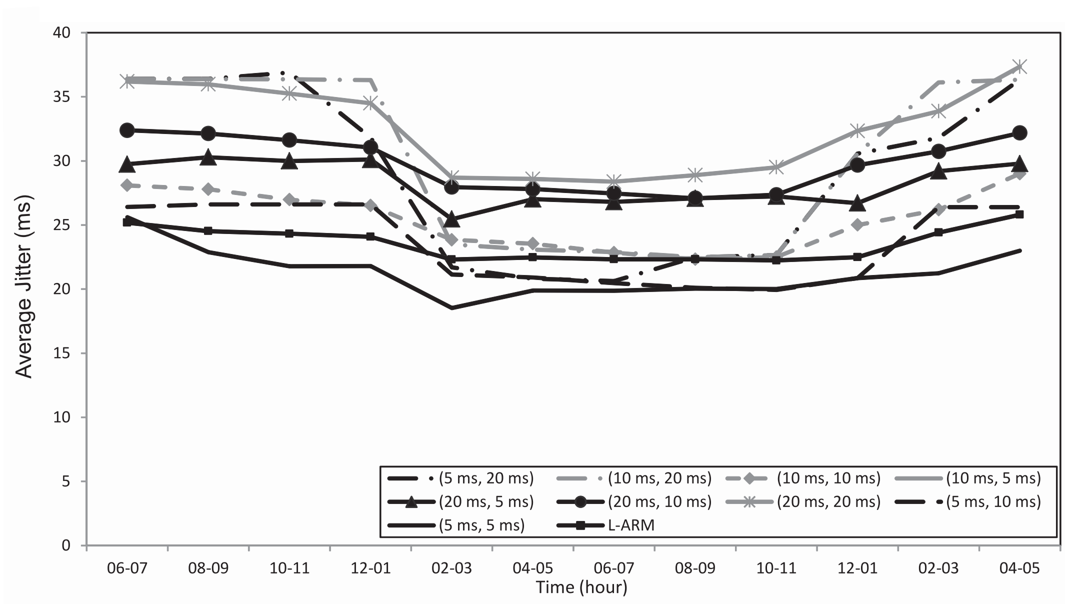

Figure 5.13: Average jitter for *class* 2 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

The reason being that the value of T_{cycle} affects the queuing delay at the ONU more than the value of T_F. The shortest T_{cycle} produces the shortest end-to-end delay, because it causes a shorter queuing delay. Moreover, the delay experienced when using the proposed

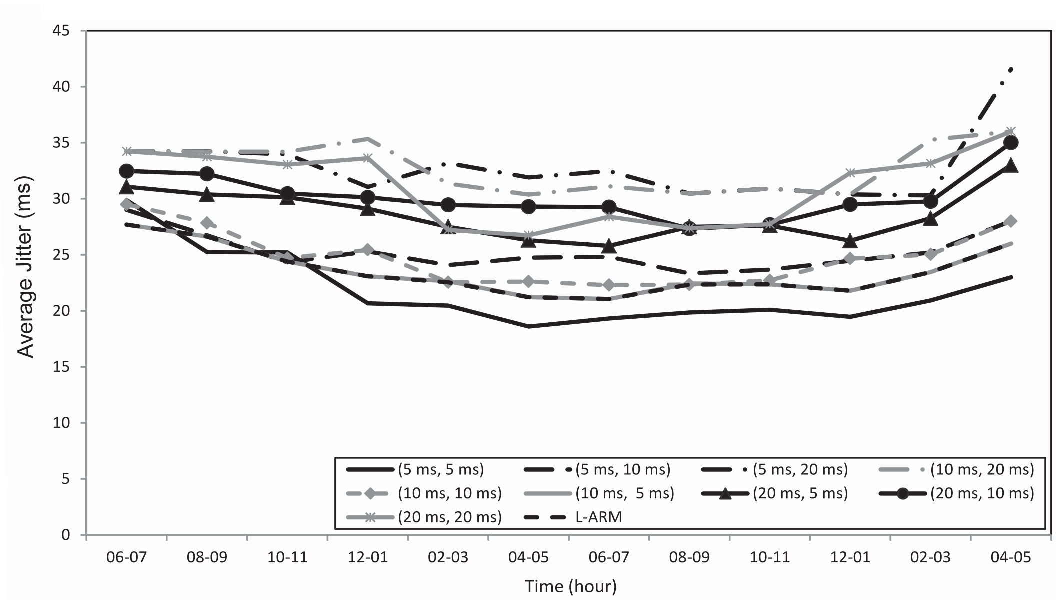

Figure 5.12: Average jitter for *class* 1 traffic for L-ARM and existing solutions (with different durations of T_{cycle} and T_F).

mechanism is less than that achievable by the existing solutions for all the other combinations, except as mentioned above for (5 ms, 5 ms). The shorter delays are achieved in the proposed mechanism because the shortest polling cycle is 10 ms when any ONU reports the existence of *class* 1 traffic. For *class* 2 traffic, it can be seen again that the end-to-end delay is relatively high at a T_{cycle} of 20 ms (Fig. 5.9). However, the delay is more stable and is acceptable for *class* 2 traffic (about 25 ms). This means that 20 ms is the most acceptable polling cycle duration for *class* 2 traffic.

Figure 5.12 and 5.13 describe the average jitter for *class* 1 and *class* 2 traffic with various combinations of T_{cycle} and T_F . For *class* 1 traffic (Fig. 5.12), L-ARM is the equal second best performer with (10 ms, 5 ms), and with (5 ms, 5 ms) it is the best. The maximum jitter for L-ARM is less than 30 ms which is acceptable [74]. For *class* 2 traffic, L-ARM is consistently the best performer for all the combinations except (5 ms, 5 ms). Again, it should be noted thata T_{cycle} of less than 10 ms would not be used when deploying the L-ARM mechanism. In Fig. 5.12, it is observed that the worst combinations are those with either T_{cycle} or T_F equal to 5 ms and the other value equal to 20 ms. For instance, the existing solutions at either (5 ms, 5 m) or (10 ms, 5 ms) achieve the least jitter, which means that the best performing ratios of T_{cycle} to T_F are 1:1, and 2:1 provided that T_{cycle} and T_F do not exceed 10 ms. The worst jitter results from the use of the combinations (5 ms, 20 ms) and (20 ms, 5 ms) which means the ratios of 4:1 and 1:4 between T_{cycle}

and T_F are not suitable for *class* 1 traffic. Furthermore, these results confirm that the combinations of a polling cycle and a frame size where either is 20 ms are not suitable for *class* 1 traffic.

Clearly in the majority of cases, the proposed scheme achieves the shortest end-to-end delay for *class* 1 and *class* 2 traffic, regardless of the volume of traffic in the network. Without an appropriate cooperative resource management mechanism, the wireless part of the network is the bottleneck because of its low data rates compared to the optical part of the network and degrades the network performance by affecting the QoS. Efficient resource management is achieved by changing both the frame size carried by the wireless part of the network, and the polling cycle duration. Conversely, switching to the appropriate frame size improves channel utilization and saves energy.

In terms of throughput, there is no significant difference between the proposed mechanism (L-ARM) and the existing solutions.

5.4 Closing Remarks

In this chapter, a sleep mode was adopted for all WOBAN components by using a dynamic and cooperative resource management within both the optical and wireless parts of the network. The aim was to enhance the energy efficiency and the QoS of the WOBAN, and to exploit the changing traffic profiles and their dynamically changing class compositions typically being experienced over the course of the day. To this end, the means to achieve interoperability of resource management between 4G and 10G-EPON systems in order to realize a green WOBAN was addressed; and a load-adaptive resource management mechanism for green WOBAN was proposed. In the proposed mechanism, the frame size of the wireless (4G) part of the network, directly affects the resource management for the wireless subscribers. The choice of scheduling mechanism to be utilized was between having either a dependent scheduling (DS) or an independent scheduling (IS) mechanism. With DS, the frame selection by the BS is based on the information sent from the OLT, while in the case of IS, the frame size selection depends on traffic monitoring by the BSs. However, whether DS or IS was used, the selection of frame size and polling size followed the same policies to prevent any conflict between resource management between the wireless and optical parts of the network. The performance of L-ARM on the WOBAN network was examined for a range of traffic loads. Simulation results showed that the proposed mechanism maximizes energy savings; while maintaining the end-to-end delay and jitter within acceptable boundaries for heterogeneous traffic conditions. The improvement in energy efficiency was achieved in part by placing the BSs along with the OLT and ONUs under resource management control. Further improvements were achieved by making the resource allocation in the optical and wireless parts of the network dynamically responsive to changes in the available traffic load. No other previously published studies considered cooperation between the resource management of the optical and wireless parts of the network to increase sleep duration and support the QoS.

Chapter 6

Conclusion

This thesis addressed the challenge of realising a power efficient, environmentally friendly, or green wireless-optical broadband access network (WOBAN). The aim was to develop a future-proofed access network, having improved QoS, using less power, and where the resulting service price can be significantly reduced for end users. A complete design for a green WOBAN having these attributes has been proposed. This design encompasses the establishment of an integrated network; and the development of resource management functionality which considers the class of traffic as another dimension for reducing power consumption, while maintaining the QoS for various types of traffic. In this design it was proposed firstly that a QoS mapper and queue management functionality be implemented at the interface between the optical and wireless systems, in order to achieve the fusion between these two technologies. Secondly, mechanisms to enable cyclic sleep within the resource management were proposed to enhance the channel utilization and QoS. In relation to cyclic sleep, both a global and local sleep mechanisms have been proposed to take advantage of having the OLT and ONUs in cyclic sleep to save additional power. Consequently, with the fluctuating daily traffic for access network usage, the green WOBAN can exploit opportunities as they arise to further optimise its resource management. The idle durations are increased to the extent that the given class of traffic can tolerate, with subsequent reduction in the net power consumption during those durations. In addition, the proposed design provides for improved channel utilization, by avoiding the frequent allocation of resources when the traffic volume is low. The inter-operability between the resource management at the wireless and optical interface is important, and should consider all the components, both optical and wireless, when attempting to maximise the energy efficiency of the system.

This chapter summarises the findings and contributions of this thesis and suggests several research directions that can be considered for future work.

6.1 Contribution of the Thesis

The main contributions of this thesis are summarized as follows:

- The synthesis of a robust architecture for an integrated wireless-optical broadband access network (WOBAN). The primary requirement for realising the integration between the optical and wireless systems, is finding a solution for the gap between the two different standards that serve the traffic according to their individual QoS requirements. The wireless standard is capable of dealing with the QoS requirements for different data traffic, while the optical standard uses queues to differentiate between the various types of traffic. A QoS mapper and queue management system, based on M/G/1, supported by a special design for a priority scheduler were proposed; to ensure the maintenance of the QoS for different types of traffic.

- The development of a new class-based power saving model (CBPS) for the ONUs in the WOBAN architecture, which reduces the overall power consumption and improves the quality of service across the WOBAN. In CBPS, the sleep durations are enabled within the resource management. The proposed model significantly improves power saving by efficiently exploiting the idle periods within the tolerated delay of various services, thereby increasing the total sleep time of each ONU. In addition to the resource management, a mechanism for calculating the energy consumption was formulated. The results reveal that by using the CBPS model, power savings of up to 80 percent are obtained. The proposed class-based queue model also provides a better quality of service compared to existing models.

- The design of a new resource management system that adjusts the power consumption in response to the changing volume and composition of the daily traffic, by monitoring additional attributes such as the class of the traffic. This thesis proposes a novel solution to this problem through the introduction of a dynamic polling cycle based-power saving mechanism (DPCB-PSM). The identification of each class of traffic is critical, as some classes of traffic do not need urgent service while others are very sensitive to delays. For those classes of traffic which are not delay-sensitive, such as the best effort traffic, the longest polling cycle is deployed. This long polling cycle is then divided into two sub cycles; one for use in polling the ONUs, and the other for use in putting the ONUs and the OLT into a sleep state. By monitoring the class of traffic currently being transmitted, the OLT can adjust the polling cycle so that delays and the allocated transmission window are able to respond to the instantaneous traffic conditions.

- The design of a new mechanism for coordinating the resource management in the wireless part of the network, able to adjust to the changes of the resource management in the optical part of the network. The aim is to enhance the behaviour of the WOBAN and to boost its energy efficiency. A load-adaptive resource management (L-ARM) mechanism is proposed, where the changes in the frame size of the wireless part of the network are either dependent on the advertised polling cycle of the optical part of the network, or are an outcome of monitoring the traffic in the queues at the base station (BS).

6.2 Future Work

This section suggests some extensions to the work presented in this thesis.

6.2.1 Using a Mixed Line Rate Strategy

An adaptive resource management mechanism was proposed in this thesis to improve channel utilization, QoS, and energy efficiency. The aim was to enable the network to identify the optimum polling cycle duration. Future work may include a mixed line rate strategy that offers a range of data rates. Any extended design should consider the i) existing work, ii) the volume and the classes of traffic, and iii) the system's ability to make decisions to either increase or decrease the data rate; in order to further control the power consumption, as lower data rates consume less power.

6.2.2 Load Balancing for the WOBAN

The load for each cell differs, depending on the number of currently active users. Further developments to this thesis could aim at maintaining the QoS by performing load balancing within the wireless components of the network by using router, power control, or handover strategies. This is an important consideration during periods of high load, in order to avoid blocking or dropping of traffic when users want to connect with an overloaded BS, which has neighbouring BSs that are not overloaded. WOBAN has the advantage of being centralized, thus enabling the OLT to: monitor the entire network and control the turning on and off of routers; or change the coverage of the base station by changing the transmission power. A future load balancing mechanism should consider the existing work and define new messages to extend control to the BSs and routers.

6.2.3 Dynamic Split Ratio

The proposed mechanism to coordinate the resource management at the BS and OLT is based on changing the durations of the polling cycle of the optical part and the frame size of the wireless part of the WOBAN; in order to adjust the allocated bandwidth and sleep durations for the available load. The concept of using dynamic durations for the polling cycle and frame size is proposed in this thesis, introducing the capability of responding to changes in the required bandwidth and the available load. As has been clearly shown, this model has an advantage in terms of improving energy efficiency, especially during the periods when the dominant traffic is of the best effort class. This mechanism is based on a ratio of 1:1 between the upstream and downstream sub-frames in the wireless part of the network. Future work aimed at improving the coordination of resource management in the WOBAN, may investigate the effect of using a dynamically split ratio between the upstream and downstream sub-frames. The advantage of such a solution would be a significant increase in sleep durations and thus energy efficiency.

Bibliography

[1] Radio-Electronics.com, "WiMAX QoS quality of Service, last access Feb. 2015, [Online]. Available," http://www.radio-electronics.com/info/wireless/wimax/qos-quality-of-service.php/.

[2] L. Korowajczuk, "Free webinar: How to dimension user traffic in 4G networks, last access feb. 2015, [online]. available," http://www.celplan.com/webinars/webinar-20140507.asp/.

[3] Amsix, last access April 2015, [Online]. Available. http://www.ams-ix.net/technical/statistics/.

[4] Y. Li, J. Wang, C. Qiao, A. Gumaste, Y. Xu, and Y. Xu, "Integrated fiber-wireless (FiWi) access networks supporting inter-onu communications," *Journal of Lightwave Technology*, vol. 28, no. 11, pp. 714–724, 2010.

[5] W.-T. Shaw, S.-W. Wong, N. Cheng, K. Balasubramanian, X. Zhu, M. Maier, and L. G. Kazovsky, "Hybrid architecture and integrated routing in a scalable optical–wireless access network," *Journal of Lightwave Technology*, vol. 25, no. 11, pp. 3443–3451, 2007.

[6] S. Sarkar, S. Dixit, and B. Mukherjee, "Hybrid wireless-optical broadband-access network (WOBAN): A review of relevant challenges," *Lightwave Technology, Journal of*, vol. 25, no. 11, pp. 3329–3340, 2007.

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

[8] K. Yang, S. Ou, K. Guild, and C. Hsiao-Hwa, "Convergence of ethernet PON and IEEE 802.16 broadband access networks and its QoS-aware dynamic bandwidth allocation scheme," *Selected Areas in Communications, IEEE Journal on*, vol. 27, no. 2, pp. 101–116, 2009.

[9] M. Lopes. Video may make up 84 percent of Internet traffic by 2018: Cisco, last access Mar. 2015, [Online]. Available . http://www.reuters.com/article/2014/06/10/us-internet-consumers-cisco-systems-idUSKBN0EL15E20140610/.

[10] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2014-2019, cisco, last access feb. 2015, [online]. available," http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-/.

[11] mathscinotes, "High definition television bandwidth and compression math, last access april 2015, [online]. available," http://mathscinotes.com/2012/05/high-definition-television-bandwidth-and-compression-math/, 2012.

[12] M. J. O'Mahony, C. Politi, D. Klonidis, R. Nejabati, and D. Simeonidou, "Future optical networks," *Journal of Lightwave Technology*, vol. 24, no. 12, pp. 4684–4696, 2006.

[13] J. M. Pereira, "Fourth generation: now, it is personal!" in *Personal, Indoor and Mobile Radio Communications, 2000. PIMRC 2000. The 11th IEEE International Symposium on*, vol. 2. IEEE, 2000, pp. 1009–1016.

[14] R. B. Marks, J. Costa, and B. Kiernan, "The evolution of WirelessMAN," *Microwave Magazine, IEEE*, vol. 9, no. 4, pp. 72–79, 2008.

[15] N. Bhushan, J. Li, D. Malladi, R. Gilmore, D. Brenner, A. Damnjanovic, R. Sukhavasi, C. Patel, and S. Geirhofer, "Network densification: the dominant theme for wireless evolution into 5G," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 82–89, 2014.

[16] M. Maier, N. Ghazisaidi, and M. Reisslein, "The audacity of fiber-wireless (FiWi) networks," *AccessNets*, pp. 16–35, 2009.

[17] M. Mills, "The cloud begins with coal big data, big networks, big infrastructure, and big power an overview of the electricity used by the global digital ecosystem, last access april 2015, [online]. available," http://www.tech-pundit.com/wp-content/uploads/2013/07/Cloud_Begins_With_Coal./, 2013.

[18] W. N. Association, "Greenhouse gas emmissions avoided through use of nuclear energy last access april 2015, [online]. available," http://www.world-nuclear.org/Nuclear-Basics/Greenhouse-gas-emissions-avoided/, 2013.

[19] C. Lange, D. Kosiankowski, R. Weidmann, and A. Gladisch, "Energy consumption of telecommunication networks and related improvement options," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 17, no. 2, pp. 285–295, 2011.

[20] Q. Li, G. Li, W. Lee, M. Lee, D. Mazzarese, B. Clerckx, and Z. Li, "MIMO techniques in WiMAX and LTE: a feature overview," *Communications Magazine, IEEE*, vol. 48, no. 5, pp. 86–92, 2010.

[21] "IEEE standard for local and metropolitan area networks part 16: Air interface for broadband wireless access systems amendment 3: Advanced air interface," *IEEE Std 802.16m-2011(Amendment to IEEE Std 802.16-2009)*, pp. 1–1112, May 2011.

[22] "LTE-Advanced (3GPP Release 10 and beyond), Dec. 2009, last access April 2015." [Online]. Available: ftp://www.3gpp.org/workshop/2009-12.../REV-090006.pdf/

[23] H. Wei, S. Ganguly, R. Izmailov, and Z. J. Haas, "Interference-aware IEEE 802.16 WiMAX mesh networks," in *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, vol. 5. IEEE, 2005, pp. 3102–3106.

[24] WiMAX FORUM, "Mobile WiMAX-part I: A technical overview and performance evalution," June, 2006.

[25] A. Golaup, M. Mustapha, and L. Patanapongpibul, "Femtocell access control strategy in UMTS and LTE," *Communications Magazine, IEEE*, vol. 47, no. 9, pp. 117–123, September 2009.

[26] J. Gan, Z. Guo, F. Rui, L. Weihong, W. Hai, K. Sandlund, L. Jianjun, S. Xiaodong, and L. Guangyi, "LTE in-band relay prototype and field measurement," in *Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th*. IEEE, 2012, pp. 1–5.

[27] S. Parkvall and D. Astely, "The evolution of LTE towards IMT-advanced," *Journal of communications*, vol. 4, no. 3, pp. 146–154, 2009.

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

[29] N. Madamopoulos, S. Peiris, N. Antoniades, D. Richards, B. Pathak, G. Ellinas, R. Dorsinville, and M. A. Ali, "A fully distributed 10G-EPON-based converged fixed–mobile networking transport infrastructure for next generation broadband access," *Journal of Optical Communications and Networking*, vol. 4, no. 5, pp. 366–377, 2012.

[30] A. Dixit, B. Lannoo, G. Das, D. Colle, M. Pickavet, and P. Demeester, "Dynamic bandwidth allocation with SLA awareness for QoS in ethernet passive optical networks," *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 5, no. 3, pp. 240–253, 2013.

[31] J. Zheng and H. Mouftah, "A survey of dynamic bandwidth allocation algorithms for ethernet passive optical networks," *Optical Switching and Networking*, vol. 6, no. 3, pp. 151–162, 2009.

[32] M. Maier, M. Herzog, and M. Reisslein, "STARGATE: the next evolutionary step toward unleashing the potential of WDM EPONs [Topics in Optical Communications]," *Communications Magazine, IEEE*, vol. 45, no. 5, pp. 50–56, 2007.

[33] N. Madamopoulos, S. Peiris, N. Antoniades, D. Richards, B. Pathak, G. Ellinas, R. Dorsinville, and M. Ali, "A fully distributed 10G-EPON-based converged fixed–mobile networking transport infrastructure for next generation broadband access," *Journal of Optical Communications and Networking*, vol. 4, no. 5, pp. 366–377, 2012.

[34] J.-i. Kani, F. Bourgart, A. Cui, A. Rafel, M. Campbell, R. Davey, and S. Rodrigues, "Next-generation PON-part I: Technology roadmap and general requirements," *Communications Magazine, IEEE*, vol. 47, no. 11, pp. 43 –49, november 2009.

[35] J. Zhang and N. Ansari, "Minimizing the arrayed waveguide grating cost and the optical cable cost in deploying WDM passive optical networks," *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 1, no. 5, pp. 352 –365, oct. 2009.

[36] F. Effenberger, D. Clearly, O. Haran, G. Kramer, L. Ruo Ding, M. Oron, and T. Pfeiffer, "An introduction to PON technologies [topics in optical communications]," *Communications Magazine, IEEE*, vol. 45, no. 3, pp. S17–S25, 2007.

[37] R. Davey, J. Kani, F. Bourgart, and K. McCammon, "Options for future optical access networks," *Communications Magazine, IEEE*, vol. 44, no. 10, pp. 50 –56, oct. 2006.

[38] E. Wong, "Next-generation broadband access networks and technologies," *Lightwave Technology, Journal of*, vol. 30, no. 4, pp. 597–608, 2012.

[39] N. Ghazisaidi, M. Scheutzow, and M. Maier, "Survivability analysis of next-generation passive optical networks and fiber-wireless access networks," *Reliability, IEEE Transactions on*, vol. 60, no. 2, pp. 479 –492, june 2011.

[40] N. Cheng, Z. Liao, and F. Effenberger, "Large splitting and long reach passive optical networks with mode coupling receivers," in *Optical Communication (ECOC), 2010 36th European Conference and Exhibition on*. IEEE, 2010, pp. 1–3.

[41] S. Sarkar, Y. Hong-Hsu, S. Dixit, and B. Mukherjee, "DARA: Delay-aware routing algorithm in a hybrid wireless-optical broadband access network (WOBAN)," in *Communications, 2007. ICC '07. IEEE International Conference on*, pp. 2480–2484.

[42] ——, "A mixed integer programming model for optimum placement of base stations and optical network units in a hybrid wireless-optical broadband access network (WOBAN)," in *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, pp. 3907–3911.

[43] 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]," *Communications Magazine, IEEE*, vol. 45, no. 8, pp. 44–50, 2007.

[44] B. Jung, J. Choi, Y. Han, M. Kim, and M. Kang, "Centralized scheduling mechanism for enhanced end-to-end delay and QoS support in integrated architecture of EPON and WiMAX," *Lightwave Technology, Journal of*, vol. 28, no. 16, pp. 2277–2288, 2010.

[45] C. M. Assi, Y. Ye, S. Dixit, and M. A. Ali, "Dynamic bandwidth allocation for quality-of-service over ethernet PONs," *Selected Areas in Communications, IEEE Journal on*, vol. 21, no. 9, pp. 1467–1477, 2003.

[46] L. Shi, B. Mukherjee, and S. S. Lee, "Energy-efficient PON with sleep-mode ONU: progress, challenges, and solutions," *Network, IEEE*, vol. 26, no. 2, pp. 36–41, 2012.

[47] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON: design and analysis of an optical access network," *Phonotic Network Commun.*, vol. 3, no. 3, pp. 307–319, 2001.

[48] 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," *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 1, no. 4, pp. 259–273, 2009.

[49] C. Gunaratne, K. Christensen, B. Nordman, and S. Suen, "Reducing the energy consumption of ethernet with adaptive link rate (ALR)," *Computers, IEEE Transactions on*, vol. 57, no. 4, pp. 448–461, 2008.

[50] P. Chowdhury, M. Tornatore, A. Nag, E. Ip, T. Wang, and B. Mukherjee, "On the design of energy-efficient mixed-line-rate (MLR) optical networks," *Lightwave Technology, Journal of*, vol. 30, no. 1, pp. 130–139, 2012.

[51] R. Kubo, J. I. Kani, H. Ujikawa, T. Sakamoto, Y. Fujimoto, N. Yoshimoto, and H. Hadama, "Study and demonstration of sleep and adaptive link rate control mechanisms for energy efficient 10G-EPON," *Optical Communications and Networking, IEEE/OSA Journal of*, vol. 2, no. 9, pp. 716–729, September 2010.

[52] J. Zhang, T. Ting Wang, and N. Ansari, "Designing energy-efficient optical line terminal for TDM passive optical networks," in *Sarnoff Symposium, 2011 34th IEEE*, May 2011, pp. 1–5.

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

[54] J. Mandin. EPON power saving via sleep mode," presented at the IEEE P802.3av 10GEPON Task Force Meeting, Sep. 2008 [online]. available:. http://www.ieee802.org/3/av/public/2008_09/3av_0809_mandin_1.pdf.

[55] Y. Yan, S. W. Wong, L. Valcarenghi, S. H. Yen, D. R. Campelo, S. Yamashita, L. Kazovsky, and L. Dittmann, "Energy management mechanism for ethernet passive optical networks (EPONs)," in *Communications (ICC), 2010 IEEE International Conference on*, May 2010, pp. 1–5.

[56] S. S. Lee and A. Chen, "Design and analysis of a novel energy efficient ethernet passive optical network," in *Networks (ICN), 2010 Ninth International Conference on.* IEEE, 2010, pp. 6–9.

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

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

[59] 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, 2010.

[60] J. Lorincz, T. Garma, and G. Petrovic, “Measurements and modelling of base station power consumption under real traffic loads,” *Sensors*, vol. 12, no. 4, pp. 4281–4310, 2012.

[61] E. Oh, B. Krishnamachari, X. Liu, and Z. Niu, “Toward dynamic energy-efficient operation of cellular network infrastructure,” *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 56–61, 2011.

[62] A. Conte, A. Feki, L. Chiaraviglio, D. Ciullo, M. Meo, and M. A. Marsan, “Cell wilting and blossoming for energy efficiency,” *Wireless Communications, IEEE*, vol. 18, no. 5, pp. 50–57, 2011.

[63] L. Saker, S.-E. Elayoubi, and T. Chahed, “Minimizing energy consumption via sleep mode in green base station,” in *Wireless Communications and Networking Conference (WCNC), 2010 IEEE*, April 2010, pp. 1–6.

[64] Y. Liu, L. Guo, and J. Yang, “Energy-efficient topology reconfiguration in green fiber-wireless (fiwi) access network,” in *Wireless and Optical Communication Conference (WOCC), 2013 22nd.* IEEE, 2013, pp. 555–559.

[65] K. Togashi, H. Nishiyama, N. Kato, H. Ujikawa, K. Suzuki, and N. Yoshimoto, “On the effect of cooperation between power saving mechanisms in wlans and pons,” in *Communications (ICC), 2013 IEEE International Conference on.* IEEE, 2013, pp. 6225–6229.

[66] D. P. Heyman, “Optimal operating policies for m/g/1 queuing systems,” *Operations Research*, vol. 16, no. 2, pp. 362–382, 1968.

[67] 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,” *Communications Magazine, IEEE*, vol. 49, no. 6, pp. 62–69, June 2011.

[68] “The network simulator -ns-2, last access april 2015, [online]. available,” http://www.isi.edu/nsnam/ns/.

[69] A. R. Dhaini, C. M. Assi, M. Maier, and A. Shami, “Per-stream QoS and admission control in ethernet passive optical networks (EPONs),” *Lightwave Technology, Journal of*, vol. 25, no. 7, pp. 1659–1669, July 2007.

[70] "DE.CIX. last access Feb. 2015 [Online]. Available," http://de-cix.net/content/network.html.

[71] Cisco, "Cisco visual networking index: Global mobile data traffic forecast update, 2013-2018, last access feb. 2015, [online]. available," http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/.

[72] J. Baliga, R. Ayre, W. V. Sorin, K. Hinton, and R. S. Tucker, "Energy consumption in access networks," in *Optical Fiber communication/National Fiber Optic Engineers Conference, 2008. OFC/NFOEC 2008. Conference on*, Feb 2008, pp. 1–3.

[73] T. Szigeti and C. Hattingh, "Quality of service design overview, last access Feb. 2015, [Online]. Available ," http://www.ciscopress.com/articles/article.asp?p=357102i/.

[74] Cisco, "Quality of service for voice over IP, last access (Mar. 2015) ." [Online]. Available: http://www.cisco.com/c/en/us/td/docs/ios/solutions_docs/qos_solutions/QoSVoIP/QoSVoIP.html/