Dynamic Bandwidth Allocation Algorithms for Differentiated Services enabled Ethernet Passive Optical Networks with Centralized Admission Control

Dawid Nowak

under the supervision of

Dr. John Murphy

A thesis submitted in March 2005 for the degree of
Doctor of Philosophy in
the School of Electronic Engineering,
Dublin City University

DCU

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: David Nel

ID No.: 50161938
15th of July, 2005

Rodzinie i przyjaciołom — za ich wsparcie.

To my family and friends — for their support.

Abstract

Fiber based access networks can deliver performance that can support the increasing demands for high speed connections. One of the new technologies that has emerged in recent years is Ethernet Passive Optical Networks. The key features of this approach are the simplicity of the architecture and compatibility with existing Ethernet based local area networks.

To make Ethernet Passive Optical Networks (EPONs) a fully functional part of the telecommunication system, support for classes of traffic with different Quality of Service (QoS) requirements is mandatory. Much research has been done on the optimal bandwidth allocation algorithms that would have the capability of supporting Differentiated Services (DiffServ) in EPONs.

This thesis proposes that the access control mechanism should be centralized and performed by the Optical Line Terminal (OLT). It is shown that this approach can give greater flexibility to adjust to changing traffic conditions, can simplify the structure of the Optical Network Units, and can allow the easy adoption of Service Level Agreements.

This thesis introduces a novel EPON simulator that allows testing of various types of bandwidth allocation algorithms. It is possible to evaluate the allocation mechanism under different traffic conditions and with network configurations that closely resemble real systems.

New algorithms are presented based on a paradigm of centralized access control. Simulation results showed that they offer good performance and support for the DiffServ architecture.

List of Publications

- **Dawid Nowak** and John Murphy, "Fiber to the Home: Comparison of Current System Architectures." in *Proceedings of OptoIreland*, Apr. 2005.
- **Dawid Nowak**, John Murphy, and Philip Perry, "Bandwidth Allocation for Service Level Agreement Aware Ethernet Passive Optical Networks," in *Proceedings of GLOBECOM – Global Telecommunications Conference*, Dec. 2004, vol. 3, pp. 1953-1957
- **Dawid Nowak**, John Murphy, and Philip Perry, "Adaptive Dynamic Bandwidth Allocation Algorithm for Ethernet PONs," in *Proceedings of 30th European Conference on Optical Communication*, Sept. 2004, vol. 3, pp. 740–741.
- **Dawid Nowak**, John Murphy, and Philip Perry, "A novel Service Level Agreement based algorithm for Differentiated Services enabled Ethernet PONs," in *Proceedings of OptoElectronics and Communications Conference/International Conference on Optical Internet*, 2004, vol. 1, pp. 598–599.
- **Dawid Nowak**, John Murphy, and Philip Perry, "Bandwidth Scheduling Techniques for Differentiated Services Support in Ethernet Passive Opticacl Networks," in *Proceedings of 4th International Symposium on Communication Systems, Networks and Digital Signal Processing*, 2004, vol. 1, pp. 67–70.
- **Dawid Nowak** and Bartłomiej Kłusek, "Analysis of Waveband Connecting Probability in Hybrid Optical Networks," in *Proceedings of 2nd IFIP-TC6 International Conference on Optical Communications and Networks*, Oct. 2003.

- **Dawid Nowak**, Bartłomiej Kłusek, and Tommy Curran, "New Method for Calculating Probability of Waveband Connection in All Optical Networks," in *Proceedings of IEI/IEE Telecommunications System Research Symposium*, May 2003.

- Bartłomiej Kłusek, **Dawid Nowak**, and Tommy Curran, "Wavelength Assignment Algorithms in an OBS Node," in *Proceedings of IEI/IEE Telecommunications System Research Symposium*, May 2003.

- **Dawid Nowak** and Tommy Curran, "Simulation of a Signalling Channel on Top of DiffServ Networks," in *Proceedings of IEI/IEE Telecommunications System Research Symposium*, Nov. 2001.

Contents

Chapter 1

Introduction

The idea of Passive Optical Network (PON) was first conceived nearly 20 years ago [1–3] as a response to the problem of limited bandwidth available in the local loop between a customer and the local exchange. At the time, the quality of optical fibers was improving. Efficient transmitters and receivers appeared and it seemed possible to build an access network that would be based on optical technology. Due to the lack of any active units in the light path such as amplifiers, signal regenerators or O-E converters, the architecture of the system was simple and offered bandwidth was not and still is not possible to achieve by other access methods.

However, the initial progress in the development of optical networks was slowed to a halt by economic and technological factors. The Internet was not as widespread then as it is now and customers were reluctant to pay higher prices for their broadband access. At the time it was envisaged that narrow band Integrated Services Digital Network (ISDN) access would be sufficient for most of the users until 2010 [4] while a handful of business customers would need a truly broadband access. Rapid progress in development of different types of the Digital Subscriber Loop (DSL) technology and its widespread adoption in the nineties was another factor impeding the deployment of PONs. As the average transfer rate of 6 Mbps could be achieved over DSL, the cost of replacing existing copper infrastructure with optical cables was not justified from an economic point of view.

In recent years there has been new interest in the PON technology. From the social point of view the Internet has become phenomenally popular and the number of customers requiring broadband access and willing to pay for it, is increasing steadily despite the slumps in the global economy [5]. New services have been developed such as video on demand or high definition television, requiring more bandwidth than can be provided by DSL or Cable Modems. Furthermore, a growth in the number of people requiring broadband access is forecasted. Much of the current development of broadband networks based on the optical technology has been achieved as a result of an active and stimulating role of governmental agencies. These agencies are responsible for implementing policies that encourage the development of broadband access networks as a way to increase computer literacy rates amongst citizens. This is especially the case in Asian countries, where the penetration of broadband access is the greatest. The Japanese Ministry of Public Management estimates that by March 2006 there will be 7.73 million households connected to the Internet by means of optical fiber in Japan. This number will be higher than the number of customers connected via DSL or CATV technologies [5–8].

In Korea, 74% of population already has a broadband connection to the Internet. The main telecommunications companies are heavily investing in their optical infrastructure in order to maintain their competitive edge [5,8–12].

In recent years China has been catching up with the global leaders. The recent forecasts indicate that in the near future, China will take first place in terms of the number of households with a broadband connection [13].

Also in Europe, broadband access in general is getting more and more attention. "eEurope 2005" [14–16] program was launched by the European Council in 2002. Its agenda was primarily concerned with promoting broadband Internet in member countries. A plan of action endorsed by the Council included:

- most of the public services would be accessed online (e-government, e-learning, e-health);
- an environment for e-business would be created;
- the broadband access would be widespread and available at competitive prices.

Substantial resources were allocated under Framework Programme 5 and 6 to spur research and development of broadband access and many projects received considerable levels of funding:

GigaPON Access Network (GIANT)[1] The primary objective of GIANT project is to study and simulate networks based on the GPON protocol. Partners from five universities are participating in this venture. The total cost of the project was estimated to be 7.07 million euro with 3.85 million euro funded by the European Council.

Next Generation Optical Networks for Broadband European Leadership[2] The goal of NOBEL is to create intelligent and flexible optical networks that will enable broadband services for all. The total project budget is 24.5 million euro with 13.7 million euro of contribution from the European Council .

Multi Service Access Everywhere (MUSE)[3] The scope of MUSE project is to test different types of a broadband access for residential subscribers. With 27 organizations involved, the cost of the project is 34 million euro with 18.6 million euro provided by the European Council.

Fiber based access networks are also implemented on national and regional scales in Europe without the support of the European Council. This policy was successfully introduced in Sweden which became a regional leader with broadband access based on Fiber-to-the-Home (FTTH) technology available to the largest number of people per 100 inhabitants[4]. The other Scandinavian countries are closely following the leader. Interesting examples of the local deployment are city owned FTTH networks in Rotterdam and similar one planned in Amsterdam [5, 8, 16, 17]. Also, in Italy FastWeb[5] [17] offers *"triple-play"* services over fiber to its customers.

[1] http://www.alcatel.be/giant
[2] http://www.ist-nobel.org
[3] http://www.ist-muse.org
[4] http://www.oecd.org/sti/telecom
[5] http://www.fastweb.ie

In USA, the market is not as strongly regulated as in Europe and there is no clear government policy on the development of broadband access networks. Despite this, PONs were deployed by the local communities in around 40 different locations between 2001-2002[6]. Moreover, it is envisaged that PON access will achieve dominant position in the nearest future due to the population being more distributed than in European or Asian countries [18,19].

Another reason for a growing interest in a broadband access is the global crisis after September 11th. The burst of the "economy bubble" hit big telecommunications companies especially hard and they were forcing them to revise their business strategies. In [20], Paltridge shows that the main cause of the crisis was inflated expectations about constant demands for new equipment. Paltridge points out that long haul core networks were no longer the biggest source of revenue. Instead the attention of industry is shifting towards mobile and broadband access networks, as they recorded the fastest rate of growth and are seen as a new primary source of revenue.

The last factor is technological in nature. In last twenty years huge progress has been made in the electronics as well as in the optical signal processing domain. High quality optical cables are generally available, efficient light sources and receivers can be bought "off the shelf" at reasonable prices. Also electronic circuits are available that can match the speed of optical devices. The availability of components makes building a PON system not as costly as it used to be. New business models [21,22] showed that PON systems could make better revenue than DSL or Cable systems.

1.1 Motivation

In recent years a noticeable change occurred in the networking paradigm. Plain data oriented services are no longer sufficient as modern applications such as Voice over IP (VoIP) and Video on Demand (VoD) [23,24] have stronger traffic requirements

[6]http://www.ponforum.org/market/trials.asp

and demand better QoS than any other service based on the Internet Protocol (IP) architecture.

Many see networks with a variable level of support offered to different classes of traffic as a way forward and much research has been done on the optimal architecture of future IP networks. Specification of the Differentiated Services (DiffServ) network framework [25,26] has unified and further spurred these activities.

Enabling PONs to be a part of the DiffServ framework is essential. As an access point to the Internet, EPON has to be responsible for making sure that Service Level Agreements (SLAs) are obeyed, hence there is a need for efficient admission control and bandwidth allocation algorithms.

1.2 Thesis Contribution

In this thesis the idea of a centralized approach to bandwidth allocation problem in EPON is advocated. The architecture of the PON clearly makes the Optical Line Terminal (OLT) the most important part of the domain. Moreover, it is the only entity that has knowledge about the status of all other units. Despite this fact most of algorithms published in literature are based on a distributed approach. The primary objective of this thesis was to design a framework that would facilitate deployment of DiffServ-aware bandwidth allocation algorithms in EPONs.

- The first main contribution of this thesis is to demonstrate that a centralized approach allows for efficient creation, maintenance and enforcement of SLAs and at the same time offers performance comparable to current solutions. Highlighted are necessary extensions to the Multi-Point Control Protocol (MPCP) used to facilitate the adoption of centralized bandwidth allocation.

- A Cyclic Polling with Adaptive Polling Cycle scheme is shown in an EPON with a centralized bandwidth allocation architecture. This new scheme allows for more flexible approach to calculating the length of the polling cycle which results in smaller average delays and better bandwidth utilization. To show the flexibility of the new approach, a new mechanism of Grant Multiplexing is introduced. The adoption of the Grant Multiplexing scheme considerably reduces the jitter experienced by high priority packets.

- The next main contribution of this dissertation is two dynamic bandwidth allocation algorithms based on a centralized approach. Their behavior and performance were tested under various traffic conditions in systems with different configurations.
- Another contribution is a new customized EPON simulator. This tool is optimized to speed up the process of development of bandwidth allocation algorithms and to test their performance in different situations. A fully functional simulator is implemented in C++, as high performance is crucial. It is possible to port the application's code to different operating systems as only widely available libraries were used.

1.3 Thesis Overview

In this thesis some new contributions are proposed that could improve the level of support for classes of traffic with different QoS requirements in EPONs. The outline of the thesis is as follows:

Chapter 2 The architecture of PONs is summarized and different flavors of PONs are presented. The basics of the bandwidth allocation process are described. Also some physical properties of the PON system are mentioned that put constraints on its performance.

Chapter 3 In this chapter the architecture of EPON is presented in greater detail, along with in-depth review of current research.

Chapter 4 The contents of this chapter are dedicated to introducing in greater detail the concept of centralized bandwidth allocation. The modifications to the

MPCP are presented, as well as new Cyclic Polling with Adaptive Polling Cycle scheme. Furthermore, the Gant Multiplexing scheme is introduced to show the flexibility of a centralized approach.

Chapter 5 In this chapter the bandwidth allocation algorithms are described and their behavior is explained in detail.

Chapter 6 The performance of EPONs with algorithms introduced in the previous chapter was tested on the custom built simulator. In Chapter 6 the architecture of this program is explained, the analytical model used to validate the output is described and the results of some experiments are compared with the model.

Chapter 7 This chapter contains the results of extensive simulation experiments designed to show the performance of proposed algorithms in different situations. The detailed discussion and drawn conclusions are included.

Chapter 8 In the last chapter the final conclusions are drawn.

Chapter 2

Passive Optical Networks

A short introduction to the PON architecture is presented in this chapter. Different types of PONs based on different underlying protocols are outlined. Also some physical properties of the system are mentioned which have impact on the bandwidth allocation process and which must be taken into account during the design and testing of the new bandwidth allocation algorithms.

2.1 Introduction

The PON is a technology that gives a customer truly broadband access to the Internet. Contrary to existing access methods which utilize the properties of copper links or radiowaves, PONs try to exploit the properties of optical fiber. The unique properties of the fiber provide PONs with numerous advantages over other means of access in the "last-mile" area.

- Much higher transfer rates over longer distances are achievable. In the current specifications, rates over 60 Mbps can be offered to 32 customers over a distance of 20 kilometers.
- The core of the network is transparent to the signal, which makes deploying new services or protocols easier.
- The cost of network maintenance is considerably lower in comparison to the existing copper infrastructure. Complex active elements placed in cabinets near customer's premises and requiring an independent source of power and

laborious supervision can be replaced with small and reliable passive splitters located at the splice tray.

- In comparison with the Point-to-Point access, the PON architecture minimizes the number of resources needed in the local exchange and in the local loop, and the total cost of the equipment is shared amongst more customers.

Given those features, the PON was first envisaged as a technology that would provide access to the cabinet and from there the signal would be distributed using copper cables [3]. The Time Division Multiplexing Access (TDMA) approach was used to multiplex streams from different customers and such an approach was more than sufficient for Plain Old Telephony Services (POTS).

With the advent of broadband, data oriented services and Asynchronous Transfer Mode (ATM) protocol, the PON was envisaged as a technology that could enable a truly high speed access to the backbone network [27,28]. However, it was not until late nineties that the first standard describing PONs was released. In the last three years new interest in PON architecture has been observed as a result of growing popularity and accessibility of the Internet.

2.2 The PON Standards Overview

The existing PON standards are the results of efforts of two different groups of network providers and equipment vendors. The standards represent the different views and attitudes of these groups towards the problem and the possible future of the telecommunication market.

2.2.1 Broadband PON Standard

The Broadband Passive Optical Network (BPON) [29] standard was introduced first. In 1999, it was accepted by the International Telecommunications Union (ITU). The standard was endorsed by a number of network providers and equipment vendors which cooperated together in the Full Service Network Access (FSAN) [1] group. The FSAN group proposed that the ATM protocol should be used to carry user data, hence sometimes access networks based on this standard are referred to as APONs [30,31].

The architecture of the BPON is very flexible and adapts well to different scenarios. The underlying ATM protocol provides support for different types of service by means of adaptation layers. The small size of ATM cells and the use of virtual channels and links allow the allocation of available bandwidth to end users with a fine granularity. Moreover, the deployment of ATM in a backbone of metropolitan networks and easy mapping into SDH/SONET containers allow the use of only one protocol from one end user to another.

Yet, the advantages of ATM proved to be the main obstacle in deployment of BPON and despite many field trails [32–34] BPON did not gain much popularity. The complexity of the ATM protocol made it difficult to implement and in many cases superfluous. The much simpler, data only oriented Ethernet protocol found a widespread use in local area networks and started to replace ATM in many metropolitan area and backbone networks.

2.2.2 Gigabit PON Standard

Progress in the technology, need for larger bandwidths and the unquestionable complexity of ATM forced the FSAN group to revise their approach. As a result a new standard called Gigabit Passive Optical Network (GPON) [35] was released and adopted by the ITU in 2003.

The GPON's functionality was heavily based on its predecessor, although it is no longer reliant on ATM as an underlying protocol. Instead a much simpler Generic Framing Procedure (GFP) is used to provide support for both voice and data oriented services. A big advantage of the GPON over other schemes is that interfaces to all main services are provided. Employing GFP guaranteed that packets belong-

[1] www.fsanweb.org

ing to different protocols could be transmitted in their native formats. Functionality was also provided which allowed seamless interoperability with other GPONs or BPONs. As in modern networks the security of transmitted data is a key issue. A sophisticated mechanism based on the Advanced Encryption Standard (AES) and a complex exchange of unique keys was built into the GPON architecture.

In comparison with the BPON standard much higher transmission rates are specified, the GPON being capable of supporting transfer rates of up to 2.48 Gbps in the downstream as well as the upstream direction.

2.2.3 Ethernet PON Standard

The EPON standard is the result of work in the vendor driven cooperation, the Ethernet in the First Mile Alliance (EFMA)[2]. Noticing that the majority of traffic in the network is data oriented and that efficient mechanisms enabling support for real time services were in place, the group decided that the sophisticated functionality of the BPON and GPON protocols was no longer needed. Instead, as the Ethernet protocol had became widespread and had a dominant role in the data oriented local and metropolitan networks [36], the EFMA decided to promote its functionality in PONs. The final version of the new protocol and necessary amendments to the existing ones were accepted by the standards body and released as IEEE 802.3ah in September 2004 [37]. The main goal was to achieve a full compatibility with other Ethernet based networks. Hence, the functionality of Ethernet's Media Access Control (MAC) layer is maintained and extensions are provided to encompass the features of PONs. The achieved solution is simple and straightforward, and allows legacy equipment and technologies to be reused.

[2]http://efmaalliance.org

2.3 The PON Architecture

The most important aspect of PON's architecture is its simplicity. The Optical Line Terminal (OLT) is the main element of the network and is usually placed in the Local Exchange. Optical Network Units (ONUs) serve as an interface to the network and are deployed on a the customer's side. ONUs are connected to the OLT by means of an optical fiber and no active elements are present in the link. A single ONU can serve as point of access for one (Fiber to the Home) or multiple (Fiber to the Block or Curb) customers and can be deployed either at the customer's premises (Fiber to the Home or Block) or on the street in a cabinet (Fiber to the Curb). The PON can exist in three basic configurations [38]:

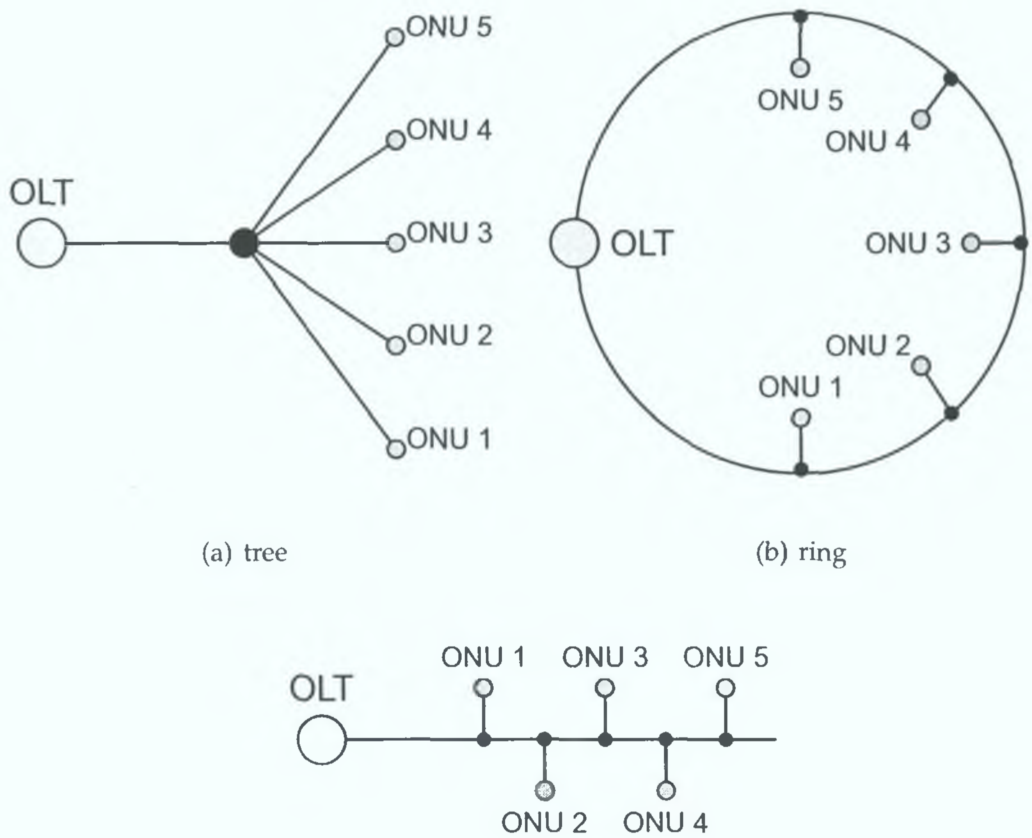
Tree topology Tree topology is shown in Figure 2.1(a). The optical splitter is used to divide the optical signal to every ONU. The typical split ratios are 1:16 and 1:32.

Ring topology In this topology ONUs are connected to the main ring using tap couplers with 2 input and 2 output ports. A schematic structure of the ring is shown in Figure 2.1(b).

Bus topology In a bus topology the stations are connected to the main link using tap couplers with 1 input and 2 output ports as shown in Figure 2.1(c)

These topologies can be further modified and redundant fiber can be added to arrange necessary protection of the network.

Although different topologies perform best in different situations, the physical properties of modern burst mode receivers make the implementation of bus and ring PONs difficult to achieve as the distances between the OLT and ONUs vary significantly. Conversely, in the tree configuration the distances between OLT and ONUs are similar. Hence, signal power levels at the receiver in the OLT vary less. This feature has a crucial effect on the performance of the PON. Firstly, large variations in the signal power can be a source of errors at the receiver. Secondly, the



(a) tree

(b) ring

(c) bus

Figure 2.1: Example PON Topologies.

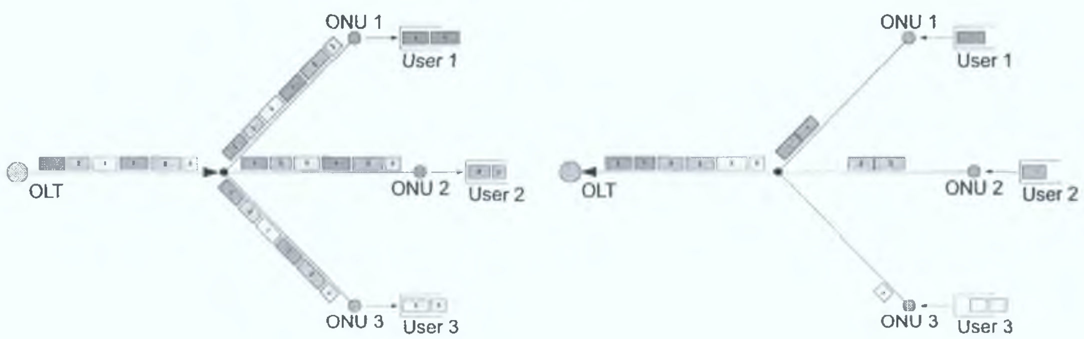
mechanisms that compensate for the imperfections of the receiver are easier to implement.

The main element that distinguishes PONs from other wired networks is a splitter, which is the source of asymmetric behavior in EPONs and it makes them similar to wireless or satellite systems.

The downstream direction signals sent by the OLT arrive at the splitter's input and then the same signal reaches every ONU. The signal is attenuated but otherwise its structure and properties are the same. Although the point-to-multipoint nature of the PON in the downstream direction may create issues concerning security, complicated traffic management and bandwidth allocation algorithms are not necessary. In principle this situation in PONs could be compared with a situation where the same signal sent by a satellite is received by different stations on the ground.

In the other direction, from ONUs to the OLT, the situation is different. The signals from different ONUs arrive at the inputs of the splitter. Although the signals can not reach different ONUs, as they traverse through the splitter they get mixed with each other and the superposition of all signals is received at the OLT. Similarly, in mobile and satellite systems only one station can transmit at any time to the satellite or the base station in order to avoid the interference of signals in the atmosphere. In both cases the outcome of multiple stations sending at the same time is unacceptable and every station is given an opportunity to transmit during a unique and non-overlapping time slot. Hence, in the upstream direction, TDMA is used and methods similar to those devised for satellite communications can be used.

A schematic version of the principle of operation of PONs is depicted in Figure 2.2. While waiting for its opportunity, an ONU buffers all incoming data. The contents of the queues are transmitted in a single burst using the whole available bandwidth of the channel upon the start of an allocated transmission window.



(a) downstream transmission

(b) upstream transmission

Figure 2.2: PON Principle of Operation.

2.4 Bandwidth Allocation in PON

The optimal allocation of available resources is crucial from performance point of view. Being a part of the global communication system there are some requirements imposed that every network must meet.

In the BPON standard every upstream frame consists of 53 cells of 56 bytes. The first three bytes create an overhead which is used for maintenance purposes such as collision avoidance, bit synchronization, amplitude recovery and cell start delineation. The remaining 53 bytes creates a standard ATM cell. At the rate specified by the OLT, the ONU inserts Physical Layer Operations and Maintenance (OAM) cells which convey the information about the current status of its physical layer. In the G.983.1 version of the BPON standard no mechanism was provided for ONUs to explicitly report their buffer statuses to the OLT. Hence, the OLT made the decision about bandwidth allocation based on information from upper layers.

This situation was rectified in the G.983.4 [39] version of the standard which was published in 2001. The functionality of the OLT was improved and the OLT was enabled to support ONUs with and without reporting capabilities. This change made the dynamic bandwidth allocation scheme more efficient but at the same time even further complicated the complex structure of the BPON protocol.

In G.983.4, Transmission Containers (T-CONs) were introduced as a way to provide support for traffic with different QoS requirements. T-CON, in principle, was a buffer in which incoming ATM cells were stored. The decision about bandwidth assignment was based on the buffer occupancy of reported T-CONs [39, 40]. The standard specified five different types of T-CONs and their importance depended on the type of service that they were assigned to. Such an approach gave the OLT better opportunity to allocate bandwidth efficiently and avoid congestion. The standard did not specify any dynamic allocation methods but it outlined requirements that such an algorithm would have to meet. A similar approach based on the use of T-CONs was adopted in the GPON standard.

In comparison with the well defined and functionally complex Dynamic Bandwidth Allocation (DBA) process in the BPON and GPON standards, a simpler and less sophisticated approach was adopted in the EPON standard. Here only one type of ONU is supported and messages reporting the buffer occupancy have to be sent in every cycle. The OLT was responsible for allocating transmission windows to the ONUs. The OLT made the decision about the size of the allocated transmission based on some internal policy. The implemented policy could take into account the report messages received from ONUs, but it was not deemed obligatory. On the other hand the ONU was free to utilize assigned bandwidth according to some local policy. In Table 2.1 a comparison of DBA features in different standards is presented.

2.5 PON Design Issues

PONs are envisaged to provide high-speed broadband access over much longer distances than can be supported with current copper based access technologies. The physical properties of available equipment have a strong impact on the possible performance, available bandwidth and network reach. These parameters determine the network architecture and must be considered in the process of designing and testing any efficient bandwidth allocation algorithm.

Table 2.1: Bandwidth Allocation Overview in PONs.

	BPON/GPON	EPON
Supported services	voice and data	data
Number of supported services	5 TCONs, up to 49 TCONs per ONU	up to 8 queues
Bandwidth allocation	centralized	distributed
Reporting mechanism	implicit and explicit supported	explicit

2.5.1 Power Budget

The main factors influencing the performance of PONs are the devices and materials used to build the network. The transmitter's power and receiver's sensitivity are two parameters that define the possible reach of the access network. In Table 2.2 typical parameters of commercially available burst mode transceivers capable of supporting 1.25 Gbps are shown. To calculate the worst case scenario power budget the minimum receiver sensitivity is subtracted from the minimum transmitter power. For devices presented in Table 2.2 the available power budget can be estimated to be around 23 dB and 22 dB respectively. Based on these values the total loss in the network is known and the maximum reach of the network can be calculated from (2.1) , where P is power budget, FCA is Fiber Cable Attenuation in dB/m, L is a distance and SL is a splitter loss.

$$P = FCA \cdot L + SL + \text{Penalties} \tag{2.1}$$

Penalties stand for additional costs such as losses at splices and connectors. The typical attenuation of a single mode fiber is about 0.4 dB/m for a wavelength of 1310 nm and 0.35 dB/m for 1550 nm.

Table 2.2: Transceivers Comparison.

Manufacturer/Model	Transmitter power		Receiver sensitivity
	min	max	max
Infineon Technologies AG[3] V23870-G5131-x610	-1 dBm	4 dBm	-24 dBm
LuminentOIC[4] SFFB-34-1250-BP-SCA-T	-2 dBm	3 dBm	-24 dBm
Fiberxon[5] FTM-9712P-F20	2 dBm	7 dBm	-30 dBm

[3]www.inifneon.com, [4]www.luminentoic.com , [5]www.fiberxon.com

As an example, assuming that the power budget is equal to 23 dB, a single mode fiber operating at the wavelength of 1550 nm is used, SL is 14 dB, and there are two mechanical splices and two connectors, the maximum reach of the network can be calculated from (2.1) :

$$\frac{23\,[\mathrm{dB}] - \mathrm{SL} - 2\cdot 0.5\,[\mathrm{dB}] - 2\cdot 0.5\,[\mathrm{dB}]}{\mathrm{FCA}\,[\frac{\mathrm{dB}}{\mathrm{km}}]} \approx 20\,\mathrm{km}$$

The minimum power budgets for typical configurations of PONs are shown in Table 2.3.

2.5.2 Receiver

The receiver in a PON must be capable of detecting signals with different amplitudes from sources located at different distances. Due to the attenuation in the optical fiber the received power levels might vary considerable between time slots. This very undesirable situation is known as a "near-far problem" [41]. In a situation where a strong signal was used in the receiver adjustment process the ones of a weak signal might be interpreted as zeros. On the contrary strong signal zeros can be interpreted as ones if the receiver is adjusted to low power input.

The best solution to this problem is to use a burst receiver at the OLT. At the beginning of every slot a training sequence is sent that allows the burst receiver to adjust itself to a new power level. This requires that for every burst of data there

Table 2.3: The Minimum Power Budget for Different PON Configurations.

ONUs	L	Wavelength	FCA	SL	Penalties	Required power budget
16	10 km	1310 nm	0.4 dB/m	14.5 dB	2.5 dB	21 dB
16	20 km	1550 nm	0.35 dB/m	14.5 dB	2.5 dB	24 dB
32	10 km	1310 nm	0.4 dB/m	17 dB	2.5 dB	23.5 dB
32	20 km	1550 nm	0.35 dB/m	17 dB	2.5 dB	26.5 dB

must be a preamble attached that allows the receiver to adjust its threshold values. This method is proposed in the IEEE 802.3ah standard.

A different approach was proposed in the ITU G.931 and G.941 recommendations. To ensure that the signal power from different ONUs have similar strength at the OLT a feedback was created between the OLT and ONUs and every ONU has to adjust the power of its transmitter according to the directions received from the OLT. Implementation of such a mechanism significantly increased the complexity of the protocol and required more complex and expensive light sources in ONUs to be used.

2.5.3 Transmitter

Transmitter performance must also be taken into account during the design phase of the network. Especially important is the so called "capture effect" [41,42], where the source emits a noise signal even though data is absent at the input. In order to avoid the interference of noise with data streams from other ONUs at the splitter, an ONU must switch off its laser completely during the idle time. As a result some bandwidth must be sacrificed to guarding gaps between time slots from different ONUs to compensate for time needed to warm up and cool down the laser. The length of the guarding gap is dependent on five parameters:

"on time" Time needed to turn the laser on.

"off time" Time needed to switch the laser off.

timing drift tolerance Time needed to synchronize clocks in the OLT and an ONU.

level recovery Time needed to adjust the gain of the receiver to an acceptable level.

clock recovery Time needed to recover the bit timing from the signal.

In EPON, the maximum guarding time cannot be longer than $1.3\,\mu s$[6], whereas in GPON, the minimal guarding time is equal to 96 bits at a current link rate. In

[6] [37] Sections 60.7.13.1 and 65.3.2.1; Tables 60-3, 60-6, 60-6,60-8

the case of 1244.6 Mbps link rate[7], this gives a guarding gap length of 77 ns. These considerable differences in the length of the guarding gaps are the result of using different types of receivers and with contrasting approaches to the power levelling mechanism. In the BPON and GPON ONUs have to adjust the power of their signal so the differences between them are minimized and so shorter time is needed to perform level and clock recovery. Conversely, ONUs do not have to perform any power adjustments in EPONs. Burst mode receivers are used in the OLT and they need more time to adjust to different power levels.

2.6 Conclusion

PONs are clearly a powerful alternative to other access technologies:

- Optical fiber deployed in the access loop is a "future proof" solution. Not only are very large transfer rates achievable with the existing technology, but also as the technology matures even higher rates can be achieved on the same medium.
- PONs help reduce the cost of the network. Active elements which require maintenance and external sources of power, are replaced with inexpensive, robust and passive elements.
- The third advantage of PONs over other solutions is the possibility of sharing the cost of the network infrastructure amongst a number of customers. The possible number of customers can be easily increased when terminals operating on faster data rates are implemented or a mixture of wavelength and time division multiple access is used [44,45].

EPON is a solution based on the PON architecture. Although slightly less efficient than its counterpart GPON, its key advantages are simplicity and widespread popularity of other Ethernet based networks.

[7] [43] Appendix I; Table I.2

Chapter 3

EPON Overview

In this chapter the architecture of EPONs is outlined. The format of control packets and the signalling process between the OLT and ONUs are explained in greater detail. A proper understanding of these factors is necessary in designing successful SLA-aware bandwidth allocation methods.

A detailed overview of the main trends in the development of EPONs is also presented and different approaches to the problem of optimal bandwidth allocation are discussed.

3.1 Introduction

The Ethernet PON standard is a result of work in the vendor driven EFMA. The main goal of the group was to promote the development of Ethernet as a means of access to backbone networks.

The EPON standard is based on providing an extension to Ethernet's MAC layer frame, thus obtaining a simple and effective solution where legacy equipment and technologies can be reused. From this point of view, ensuring the compatibility with existing Ethernet network standards was the key issue. Achieving a seamless cooperation between EPON and existing systems required EPONs to be compliant with the generic architecture of the Ethernet protocol as specified in IEEE 802 standard [46] .

3.2 EPON Compliance with Existing Protocols

In the IEEE 802 [46] suite two different types of media access are described: shared media and full-duplex. In a shared medium access mode stations are connected to a single link. In this situation only one station can transmit at a time and all the remaining stations must be in the listening mode. The Full-duplex configuration allows simultaneous data transmission by two stations over a point-to-point link.

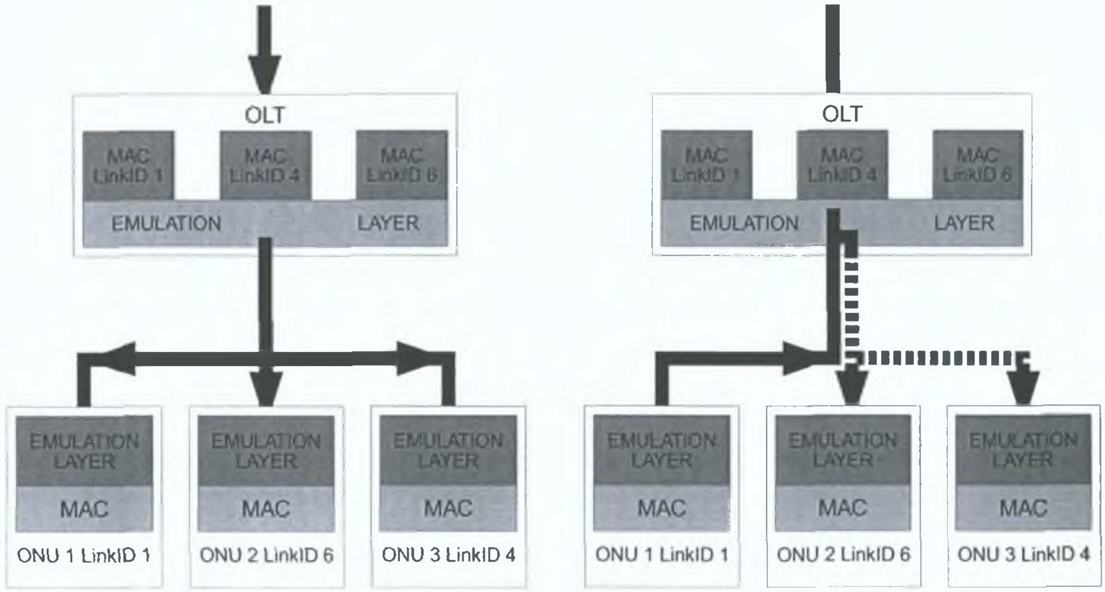
The IEEE Ethernet LAN standard relies heavily on a concept of *bridging*. A bridge is a network station responsible for interconnecting multiple access domains [46–48]. In such a network, end stations can communicate with end stations in other access domains as if they were sharing a common access medium. Assuming that in an access domain stations can communicate with each other, the bridge does not relay traffic to its ingress ports.

Such an approach is a source of incompatibility between EPON and the Ethernet architecture. In EPON in the upstream direction ONUs can not communicate with each other due to properties of the splitter. Providing inter-ONU communication involves using at least the network layer of the Open Systems Interconnection OSI stack in the OLT. This situation is clearly in contrast with the existing standards.

In [49] Kramer et al. introduced a special sublayer that helps to resolve these compatibility problems and ensures smooth integration of EPON with other Ethernet networks. The operation of this sublayer is based on marking Ethernet frames with *link ID* tags. Every ONU is assigned a unique tag which is transmitted in the preamble before every frame.

3.2.1 Emulation of Shared Medium Access

As specified in the standard [46], in Shared Medium Access mode packets sent by any station should be received by every other station. In emulating this behavior the OLT inserts a broadcast address in the preamble so every ONU passes such a packet into the MAC layer. In the downstream direction, upon receiving such a broadcast tag from a particular ONU, the OLT is required to replicate the frame and forward it to every other ONU in the domain. The schematic of Shared Medium Access emulation is shown in Figure 3.1.



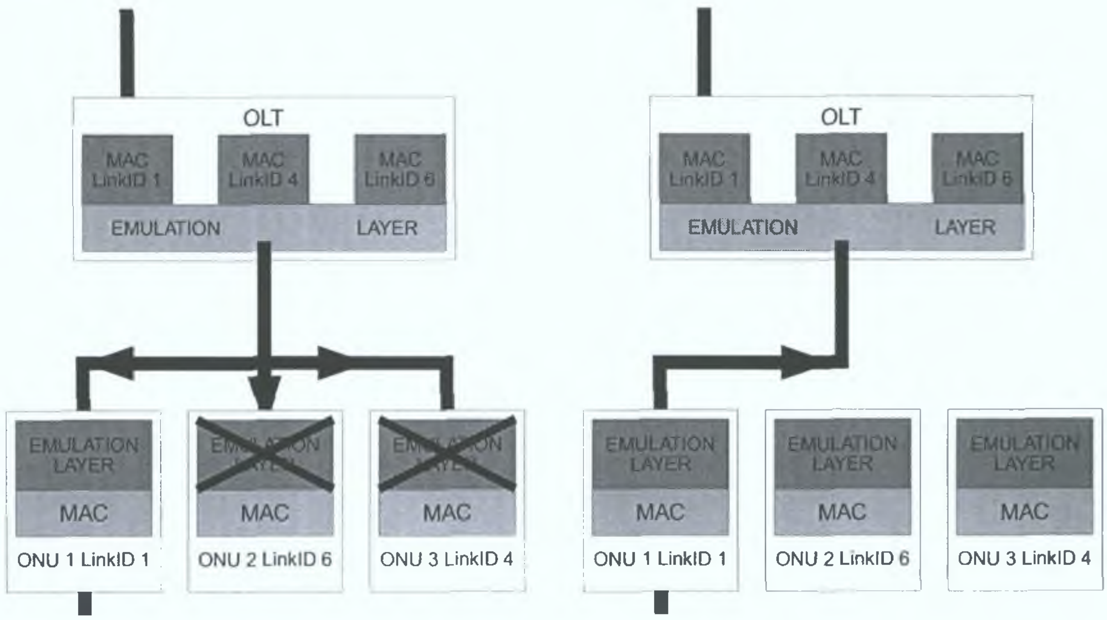
(a) downstream direction (b) upstream direction

Figure 3.1: Shared Medium Access Emulation.

3.2.2 Emulation of Point-to-Point Access

Providing successful emulation of point-to-point transmission requires the OLT to assign a unique *link ID* to every ONU. Even though all frames sent by the OLT are broadcast to every ONU in the domain the emulation sublayer filters out and passes to the MAC layer only frames with *link ID* associated with a particular ONU. In the upstream direction the ONU inserts its *link ID* in the preamble and based on its value the OLT demultiplexes the frame to the proper port. This is shown in Figure 3.2.

Introduction of an extra sublayer to resolve the problem of the EPON protocol non-conformance with existing standards creates some problems with providing fair access to the network resources. In the Shared Medium Access mode the frames received from an ONU must compete for bandwidth with traffic coming from the



(a) downstream direction (b) upstream direction

Figure 3.2: Point-to-Point Medium Access Emulation.

Internet and this can lead to a situation where the local traffic suffers from worse than expected QoS.

In the emulation of Point-to-Point Access the OLT is required to perform extra processing which in turn makes its functionality more complicated. In cases where more than one *link ID* is assigned to an ONU the OLT must be capable of handling multiple requests simultaneously. This puts additional strain on the performance of the OLT and faster processing speeds are required.

3.3 Multi-Point Control Protocol

In EPON, a new control protocol called Multi-Point Control Protocol (MPCP) is implemented to carry control data between the OLT and ONUs. As the protocol packets are built on the structure of the Ethernet frame a compatibility with legacy solutions is provided. The standardization process of the new scheme was carried out by the IEEE 802.3ah [37] group and the final version of the draft was approved by the Standards Board in June 2004.

3.3.1 General Structure of the MPCP Data Unit

In Figure 3.3 the structure of the MPCP frame is presented. The total length of the packet is 64 octets. This length has its roots in limitations imposed by the functionality of the MAC layer. The header field consists of a 6 octet long Destination Address (DA), a 6 octet long Source Address (SA) and a 2 octet long Length/Type field which indicates the type of the Ethernet frame. For the MPCP the value of Length/Type field is set to "88 – 08" in hexadecimal notation. The MPCP specific information is carried in the Logical Link Control (LLC) part. The LLC consists of mandatory OPCODE and TIMESTAMP and additional optional fields whose presence depends on the type of message. The field OPCODE is 2 octets long and is used to distinguish the type of the MPCP message in the MAC Control sublayer. Each packet must also contain a TIMESTAMP field of 4 octets. Information about the value of the EPON counter is carried in this message and its value is used to calculate the Round Trip Time (RTT) of a control packet between the OLT and an ONU. This value is later used in bandwidth allocation mechanism and it limits the minimum cycle length. The value of the counter is also used to synchronize ONUs to the common clock.

The total length of optional fields must be 40 octets long. Padding consisting of octets with all bits equal to zero is used in cases where the total length of optional fields is shorter. The frame is always closed by the FCS field which contains the value of the Cyclic Redundancy Check (CRC) calculated for DA, SA, Length/Type field, LLC and Padding.

OCTETS	
6	DESTINATION ADDRESS
6	SOURCE ADDRESS
2	LENGTH/TYPE (88-08)
2	OPCODE (00-0X)
4	TIMESTAMP
40	DATA/RESERVED/PADDING
4	FRAME CHECK SEQUENCE

Figure 3.3: MPCP Message Structure.

3.3.2 Overview of MPCP Messages

In the MPCP specification the format of five messages is defined. Two of these messages are used primarily during the bandwidth allocation process. The remaining three are used to register new sources with the OLT.

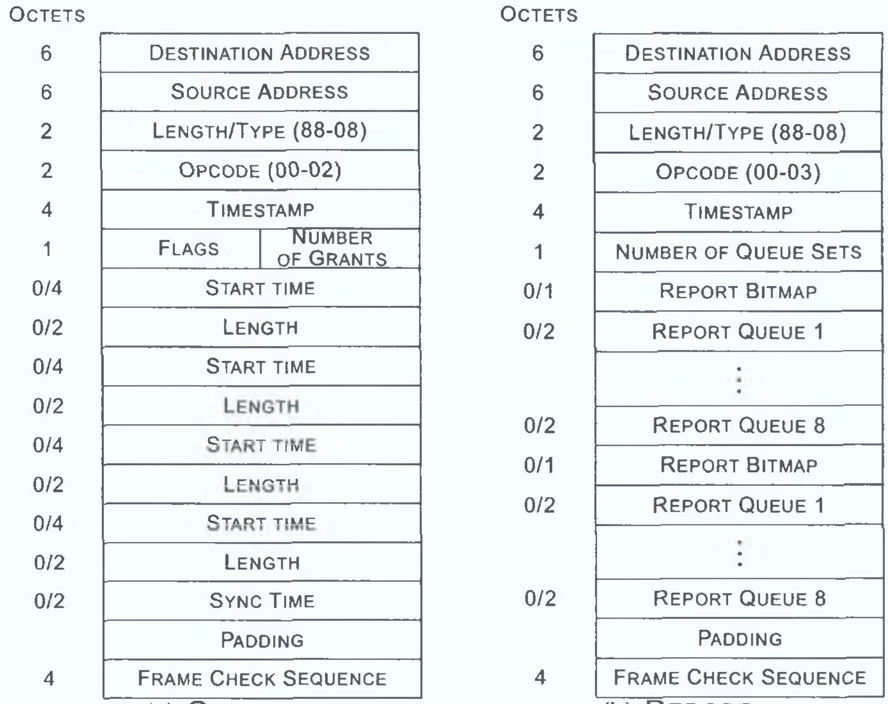
The GATE message is sent by the OLT and is used to inform an ONU about the beginning and the length of a time slot allocated to it. Up to four different slots can be allocated to a single ONU in one cycle. An ONU is free to utilize a given opportunity to the best of its ability. The existing standard does not define the default bandwidth allocation method. This leaves room for research in this area and many different approaches were developed and published in literature.

The GATE message is also used to give idle ONUs a chance to register with the OLT. During a normal network operation, special GATE messages are issued periodically indicating the start and duration of a slot in which idle ONUs can start the discovery process. The GATE message structure as defined in the IEEE 802.3ah standard is presented in Figure 3.4(a).

The REPORT message is sent by ONUs to the OLT. It carries information about the current state of all queues hosted by a given ONU. The OLT is free to use this knowledge during the bandwidth allocation process or ignore it. An ONU should periodically sent REPORT messages to inform the OLT of its status. The format of the REPORT message is presented in Figure 3.4(b). The REPORT BITMAP field indicates the queues for which buffer status is sent in the current message.

Messages REGISTER REQUEST, REGISTER and REGISTER ACK, presented in Figure 3.5, are used during a three way handshake process to register new ONUs with the OLT. The REGISTER REQUEST message has one extra mandatory field, the P2PE REQUEST field. REGISTER REQUEST is issued by the ONU and its main purpose is to announce to the OLT the number of logical links that this ONU is ready to handle. Apart from the mandatory field the REGISTER REQUEST message can enclose optional fields with a list of additional ONU's capabilities.

Upon receipt of REGISTER REQUEST, the OLT answers with a REGISTER message.



(a) GATE (b) REPORT

Figure 3.4: Format of MPCP GATE and REPORT Messages.

The number of assigned ports is transferred in the P2PE PORTS field. For each assigned port its physical MAC address is put in 2 octet long P2PE PHYSICAL ID field. The OLT can inform an ONU about its additional capabilities by attaching optional fields to the REGISTER message. A copy of optional fields received from an ONU can also be attached. In the last phase of the registration process an ONU replies with the REGISTER ACK message which in turn might include a copy of optional fields received from the OLT in the REGISTER message.

3.4 Bandwidth Allocation Algorithms in EPON

As shown in the previous sections EPON is a data oriented service. As its simplicity and compatibility with existing Ethernet standards was a key issue, no provisions were made to support services with strong traffic requirements such as voice or video on demand and mechanisms native to the third layer of the Open Systems Interconnection (OSI) stack must be used.

Octets	
6	Destination Address
6	Source Address
2	Length/Type (88-08)
2	Opcode (00-04)
4	Timestamp
1	Flags
1	Pending Grants
	Padding
4	Frame Check Sequence

(a) REGISTER REQUEST

Octets	
6	Destination Address
6	Source Address
2	Length/Type (88-08)
2	Opcode (00-05)
4	Timestamp
2	Assigned Port
1	Flags
2	Sync Time
1	Echoed Pending Grants
	Padding
4	Frame Check Sequence

(b) REGISTER

Octets	
6	Destination Address
6	Source Address
2	Length/Type (88-08)
2	Opcode (00-06)
4	Timestamp
1	Flags
2	Echoed Assigned Port
1	Echoed Sync Time
	Padding
4	Frame Check Sequence

(c) REGISTER ACK

Figure 3.5: MPCP Messages.

If EPON is to become a part of the global network, support for DiffServ must be provided and traffic parameters negotiated during the setup phase must be guaranteed during the lifetime of a connection.

The bandwidth allocation algorithm is a key component in EPON's architecture and the only one that is in charge of all decisions concerned with assignment of transmission windows to ONUs. From this point of view finding an algorithm that is able to cope with classes of traffic requiring different QoS is necessary.

As mentioned in Section 3.3.2 the whole bandwidth allocation process is based on the exchange of the GATE and REPORT messages. The GATE message notifies a particular ONU about the start and duration of its time slot. The REPORT message is periodically sent by an ONU to inform the OLT about lengths of all queues hosted by an ONU. The distance between the OLT and ONUs and the size of assigned time slots are of paramount importance as they affect the length of the polling cycle.

> *The polling cycle, which is sometimes called the granting cycle, could be defined as the time that elapsed between two transmission windows assigned to the same ONU or queue.*

Extending the length of the polling cycle leads to an increase in the average delays but at the same time the system is more efficient since:

- the control messages are sent less often, hence more bandwidth can be dedicated to the customer's payload;

- when the size of the transmission windows increases, the bandwidth dedicated to the guarding gaps is proportionally smaller.

3.4.1 Polling Cycle

Two polling schemes have been analyzed in the EPON system. Kramer et al. introduced "Interleaved Polling with Adaptive Cycle Time (IPACT)" in [41,50,51]. The functionality of their algorithm is based on following assumptions:

- The OLT knows exactly how many bytes are buffered by ONUs and the RTT to every ONU. The OLT allows an ONU to send some number of bytes and notifies an ONU about the allocated transmission window with the GATE message.
- Based on the allocated grant and the RTT, the OLT can precisely calculate at what time the window allocated to the next ONU should start and when the GATE message should be sent.
- At the end of the time slot allocated to a given ONU it issues the REPORT message with the number of bytes that were left in the buffer.
- In the next cycle the OLT allocates the number of bytes reported in the previous cycle.

The schematic figure showing the flow of messages is depicted in Figure 3.6. The obvious advantage of the IPACT algorithm is its simplicity. On the other hand, its main disadvantage lies in the fact that packets that arrived outside of the time slot are not taken into account during the bandwidth allocation process and must wait until the next cycle, thus experiencing much larger delays.

Also some questions must be asked about the amount of the downstream bandwidth used by GATE messages. In a scenario where the offered load is small or

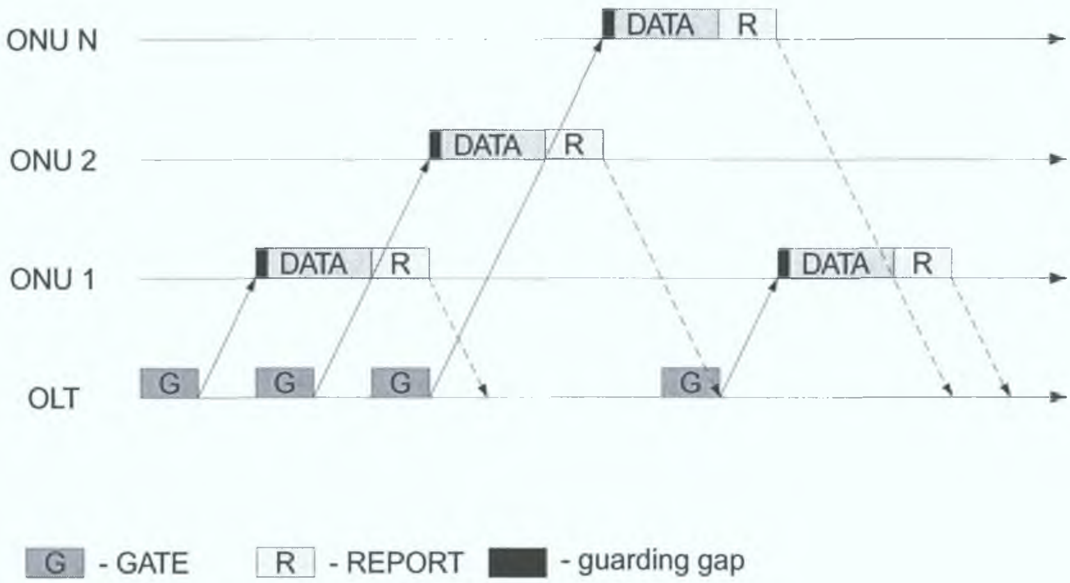


Figure 3.6: IPACT Polling Scheme.

the distance between an ONU and the OLTis small, the control messages are sent more often than in situation where ONUs are further apart from the OLT or traffic in the network is heavier. This means that a variable amount of bandwidth must be allocated to the control channel depending on the length of the polling cycle. In Table 3.1 a comparison is made of the proportions of bandwidth dedicated to the control channel for different lengths of the polling cycle. The applicability of IPACT

Table 3.1: IPACT Performance.

RTT	Approximate minimum distance	Minimum cycle time	Control channel bandwidth in %
50 μs	5 km	50 μs	21.5
100 μs	15 km	100 μs	10.7
200 μs	20 km	200 μs	5.376

link capacity = 1 Gbps, number of ONUs = 16, packet size = 84 bytes (Inter Frame Gap + Preamble + GATE)

in Wavelength Division Multiplexing (WDM) PONs was shown in [52].

In [53], Su-il Choi presented a different approach to this problem and cyclic polling cycling mechanism was proposed instead. A diagram of this mechanism is presented in Figure 3.7. The author compared the performance of these two polling schemes in a situation where there was one queue per ONU. The results presented showed that IPACT achieved better performance in terms of the average delays. However, the main advantage of the cyclic polling scheme over its counterpart was that the size of the control channel was considerably smaller and constant. This meant that not only more bandwidth can be dedicated to the customers' payload but also the necessary bandwidth is easier to provide. Choi et al. showed that for a constant 2 ms long polling cycle only 0.5376 % of available bandwidth was used to transfer the control messages.

3.4.2 Inter-ONU Bandwidth Allocation

An inter-ONU scheduler is responsible for allocating a single or many non--overlapping transmission windows to all ONUs attached to the OLT. During the

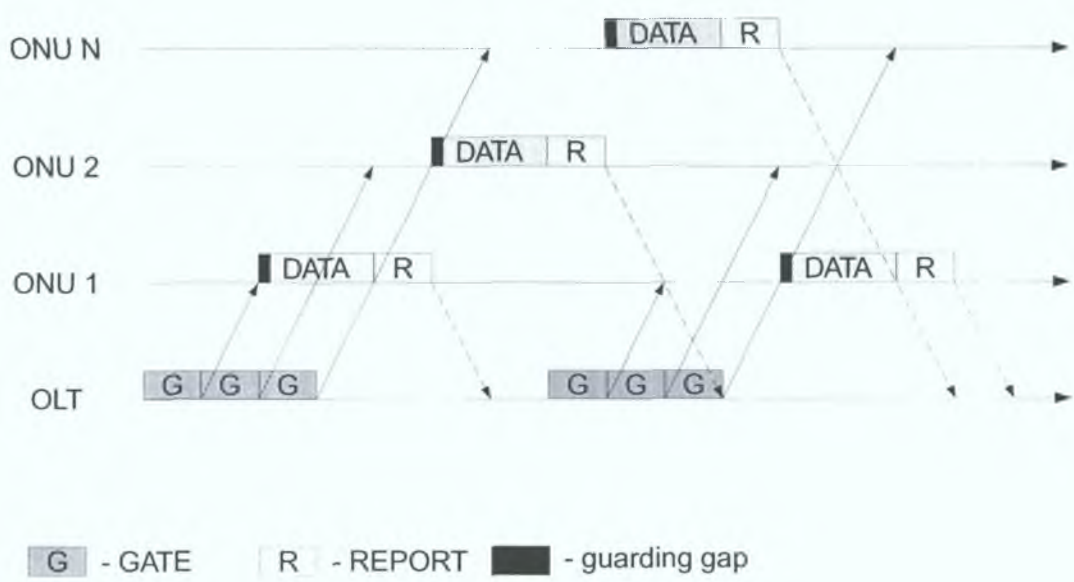


Figure 3.7: Cyclic Polling Scheme.

allocation process a control unit might take into consideration the state which ONUs are in, although scheduling policies could be implemented that do not rely on such information. Based on REPORT messages the inter-ONU scheduler knows the exact status of the system and is able to make the optimal decision.

Kramer et al. presented various types of distributed bandwidth allocation algorithms in [49]. To show the impact of the bandwidth allocation policy implemented in the OLT on the average delays experienced by packets results were presented where no intra-ONU scheduling was present and every ONU hosted only one queue. The performance of the following six different bandwidth assignment methods was compared:

Fixed Every ONU receives the same amount of bandwidth regardless of the reported number of bytes. This scheme is best characterized by the constant cycle time.

Limited In this method one tries to allocate the requested number of bytes as long as some maximum threshold value is not exceeded.

Gated In this approach the OLT always grants ONU the requested number of bytes. The limiting factor is dependent on the implementation of ONUs and it is assumed that an ONU can not request more bytes than the maximum length of its buffers.

Constant credit In this scheme the algorithm tries to compensate for the delay in processing of REPORT messages and the window time is increased by a constant value to compensate for the arrival of new packets after the REPORT message has been sent. Again the time slot can not be longer than the maximum allowed.

Linear credit This scheme is similar in concept to Constant credit. The difference is that the window time is increased not by a constant value but proportionally accordingly to the requested bandwidth.

Elastic In this scheme the only limiting factor is the length of the granting cycle. All ONUs are assigned time slots proportionally according to the requested bandwidth.

The results presented showed that there was no clear winner. For small loads of up to 40 percent of the total link capacity the average delays given by different algorithms were very close to each other. At loads between 40 and 60 percent all algorithms showed a sudden increase in the recorded average delays. Further increase in the offered load did not change the overall trend and as the total load approached 100 percent of the link capacity the difference between algorithms became negligible again. Generally the performance of all algorithms did not exceed the boundaries set by the results recorded for Fixed and Gated approaches. The Fixed algorithm gave the worst results as it suffered from an initial large delay due to the fixed size of a time slot. On the other hand the results achieved by Gated algorithm were slightly better than those of other algorithms. This was due to the fact that there was no maximum limit to allocated bandwidth and the length of the polling cycle could be extended to accommodate for extra burst of packets.

An interesting insight into the behavior of all of the algorithms was gained when the mean cycle times was plotted as a function of the offered load. In contrast to the previous results, the Gated approach was significantly worse than the others. For large loads its mean cycle time was almost 1000 times longer. In the Fixed approach the mean cycle length was constant and did not depend on the load.

As none of the presented algorithms were optimal, authors recommended choosing the Limited approach as the most suitable bandwidth allocation technique due to its conservative behavior.

In addition to cyclic polling Su-il Choi also proposed a cyclic polling algorithm with support for DiffServ [53]. He suggested that traffic should be classified into three categories. High priority traffic would be allocated a guaranteed amount of bandwidth regardless of the number of packets waiting in the queue. Medium priority classes would be allocated bandwidth based on a following policy:

$$\min(R_i^M, (B_{total} - \sum_i^n G_i^H)\frac{R_i^M}{\sum_i^n R_i^M}),$$

where R_i^M is the requested rate and G_i^H is the bandwidth allocated to high priority traffic. The remaining bandwidth was allocated to low priority traffic proportionally according to the total amount of traffic requested by an ONU.

Assi et al. [54, 55] proposed an improved version of the Limited approach to bandwidth allocation. The authors pointed out that in the original version some bandwidth might not have been allocated after the scheduler cycle. The size of this excess bandwidth could be considerable when a small load was offered to a number of ONUs. The improved algorithm calculated the amount of spare bandwidth and distributed it evenly amongst all ONUs whose requests were not fulfilled.

Assi et al. extended this concept even further and proposed a bandwidth allocation algorithm which would provide support for DiffServ. The main drawback of Kramer's approach was the fact that the packets that arrived after the REPORT message was sent were not taken into account during bandwidth allocation process. This means that the calculated transmission window was not long enough to transmit all data. Assi et al., in their approach, tried to anticipate the number of high priority packets that arrived during the waiting time and readjust the calculated length of the transmission window. This would allow an ONU to send packets that arrived during the waiting time during in the next transmission window. Otherwise these packets would have to wait in the buffer for another full cycle suffering much larger delays.

In [56], Nikolova et al. presented a dynamic bandwidth allocation algorithm with a new version of the Limited approach. The novelty of her approach was the fact that not only the maximum, but also the minimum amount of bandwidth allocated to ONUs was limited. This meant that the minimum and maximum lengths of the polling cycle were the function of different threshold values set for every ONU in the domain. In this dissertation a different approach is advocated, where the minimum and maximum cycle length is independent from the bandwidth allocation algorithm.

A bandwidth allocation scheme based on cyclic polling was also mentioned by Yang et al. in [57]. They proposed Delta DBA (DDBA) in which the difference is calculated between the reports received in two consecutive cycles. Later this value is used to calculate the new length of the transmission window.

Lee et al. in [58] introduced a two step scheduling algorithm that supported two different bandwidth allocation policies SBA and DBA. The functionality of the algorithm was based on separating the process of generating GATE messages from calculating the start time of the transmission window. In their approach the OLT hosted a Grant Scheduler and in the first phase GATE messages were queued in four separate buffers. SBA-GATE messages were generated periodically and had the highest priority. Middle priority queues were dedicated to different priority GATE messages generated by the DBA algorithm. The lowest priority queue was used for GATE messages in the discovery mode. In the second phase the decision was made about the start and duration of the transmission window and a method similar to IPACT [41,50] was proposed.

A novel approach to the problem of optimal bandwidth allocation was proposed by Zhu et al. in [59]. The authors introduced the idea of dividing ONUs into two disjoint groups namely, Bandwidth Guaranteed (BG) and Bandwidth Not Guaranteed (not-BG). All ONUs were assigned to one of these groups based on parameters agreed in the SLAs such as the maximum peak rate an ONU was committed to and maximum waiting time the ONU could tolerate. After the classification process was finished the algorithm entered phase II, where the Evenly Distributed Algorithm was executed and an Entry Table for all BG was created. The Entry Table held the number of time slots assigned to ONUs. At this stage not-BG ONUs did not have any time slots assigned. Upon completion the algorithm entered phase III, where all ONUs were polled according to the Entry Table and the list of not-BG ONUs sorted by their IDs. Bandwidth unallocated to BG ONUs was allocated dynamically to non-BG ONUs.

A similar approach was proposed by An et al. in [60]. In the Hybrid Slot-Size/Rate DBA (HSSR) algorithm, the authors suggested dividing the available bandwidth into two parts. The first part would be further split into a number of fixed size slots that could be allocated to services with strong traffic requirements. The other part would be treated as one big slot and its bandwidth would be dedicated to low priority traffic and shared among all ONUs. The strong point of this approach is that the beginning of every slot dedicated to the high priority traffic is aligned with the beginning of the frame so the packet delay variation is minimized. The authors generally assumed that only Poisson distributed voice traffic was carried in fixed size time slots and hence the variation in packet inter-arrival was small. It is hard to agree with this postulate as video on demand and voice over IP services are getting more and more important.

An approach similar to HSSR was published in [61], where Xie et al. proposed a Two-Layer Bandwidth Allocation (TLBA). The method is based on hierarchical allocation scheme and bandwidth is allocated dynamically according to current requests from ONUs. The main difference between the HSSR and TLBA algorithms is that instead of fixed size slots dedicated to high priority traffic a dynamic allocation is used.

The TLBA algorithm performs the task of allocating the bandwidth in two steps. In the first step the all available bandwidth is divided into a number of classes representing different types of QoS. This division is not fixed and the assignment of an ONU to a particular group depends on the buffer occupancy reported by ONUs for different queues. As a safety precaution to avoid dedicating the whole bandwidth to one class under heavy load a set of weights is assigned to each class. After the OLT allocates bandwidth to a class the second part of the algorithm is executed. At this stage the OLT distributes bandwidth allocated to a given class among ONUs based on their requests using a max-min policy, which adhere to following principles:

- ONUs are allocated bandwidth in an order which is based on the amount of requested bandwidth. ONUs requesting less bandwidth are allocated their share first.

- ONU can not be allocated more than the maximum value.
- Unused bandwidth is shared evenly among ONUs with unfulfilled demands.

The major weakness of the HSSR and TLBA algorithms lies in the fact that an ONU must know precisely the sequence in which the queues are serviced. Otherwise packets from the queue assigned to the best effort traffic might be transmitted in a slot dedicated to high priority data. From this point of view these algorithms must rely on some proprietary solutions as a mechanism providing such a feedback has not been implemented in the EPON protocol. Later, in Chapter 4 a new format of the GATE message is shown that could rectify this omission.

Another approach based on a max-min policy was presented by Son et al. in [62]. The authors proposed two DBA schemes that tried to improve bandwidth utilization under non-uniform traffic.

3.4.3 Intra-ONU Bandwidth Allocation

Contrary to inter-ONU scheduling the scope of the intra-ONU allocation is restricted to the admission control methods implemented in an ONU. According to the EPON standard an ONU is free to use the allocated transmission window in the best way. This means that different policies might be deployed and, what is even more important, different ONUs might operate under the control of different internal schedulers.

Achieving efficient DiffServ support has always been a primary concern in designing bandwidth allocation algorithms in EPON. Kramer et al. proposed the incorporation of a strict priority scheduling scheme in EPONs [49,50]. They showed that assigning traffic streams with different QoS requirements to queues with different priorities greatly improved the performance of the system and average delays experienced by the high priority traffic were approximately a thousand times shorter than delays recorded for the Best Effort (BE) traffic. Although strict priority queueing is easy to implement, its application in networks supporting the DiffServ paradigm can have a negative impact on the system performance and make enforcement of existing SLAs very difficult. As packets with a higher priority are always sent before packets with lower priorities, the presence of large numbers of unanticipated high priority packets will result in smaller bandwidth offered to all traffic streams with lower priorities. This is could lead to a breach of the SLA contract between the service provider and the customers.

It was also pointed out that implementing strict priority scheduling with the IPACT algorithm led to a *"light load"* phenomenon, where average delay times measured in a network with a small load were larger than in the network with the medium load level. The authors provided an explanation of that paradox in [50] and they noted that the source of the anomaly could be explained by the fact that the OLT assigned only the requested amount of bandwidth. Therefore there was no extra bandwidth allocated for packets that arrived during time when the ONU was waiting for its transmission window. As the calculated window size was too small, newly arrived packets with lower priorities had to be deferred until packets with higher priorities were transmitted. In Section 7.4.1 a different explanation to the *"light load"* phenomenon was provided

To alleviate this situation the authors proposed a two stage buffering scheme. The first stage system is built from multiple priority queues. The second stage consists of a single First-In-First-Out buffer. The ONU only reports packets that are in the second stage so the allocated time window is always fully utilized. The drawback of this architecture, however, is that high priority traffic experiences longer delays due to the additional time that high priority packets must spent in buffers of the first stage.

When the ONU tries to send the packet at the head of the queue it might happen that the time needed for its transmission is longer than the remaining part of a time slot. If the algorithm decides to stop transmitting a fraction of bandwidth will be left unused even though there might be some shorter packets still waiting in the queue. As mentioned earlier the amount of this unused bandwidth could be reduced by employing two-stage buffering, but this caused extra delays. In [63] Kramer et al. outlined a method based on a partial packet reordering that improved this situation. In this method the authors proposed that the ONU should keep track of all incoming traffic streams based on their source and destination addresses. If a situation occurs where a packet from a particular stream was too long to transmit, the algorithm tries sending the first packet from a different stream. Hence, a situation is avoided where packets from the same stream arrive at the destination in a wrong order. The efficiency of the partial reordering algorithm depends on the number of tracked streams. More streams give a better chance to find a packet that fits the gap but on the other hand it increases the amount of memory needed to store all packets and additional computational resources are needed to sort and divide all incoming packets into appropriate queues.

An improved version of the priority scheduler was proposed by Assi et al. in [54, 55] in which the intra-ONU allocation algorithm operating in two modes is proposed. Firstly only packets that arrived in the previous cycle before the REPORT message was sent, are transmitted. The queues are emptied based on their priority and all packets that arrive during this phase are buffered. When all reported packets have been transmitted the algorithm enters phase two, where packets received during idle time are sent according to a strict priority rule. A similar approach was adopted by Nikolova et al. in [56].

A structure of an ONU with a separate queue dedicated to each class of traffic but with common buffering space was proposed in [61]. As the queues share limited buffer space, a method of discarding some packets must be used to enable DiffServ support. The authors propose the use of a weighted Random Early Detection (RED) [64,65] mechanism to drop packets in order to monitor admission control and avoid congestion. The algorithms based on the RED approach have been investigated in the literature and adopted by Internet Engineering Task Force (IETF) to use with

DiffServ networks.

Ghani et al. in [66] proposed to adapt algorithms used for Connection Admission Control in ATM and IP networks [67–69] to EPON needs. They proposed an intra-ONU scheduler in which a virtual time approximation of the ideal generalized processor discipline is employed [67,68]. The authors claimed that their approach offers two main benefits:

- fine bandwidth control can be achieved and
- the functionality of intra- and inter-ONU bandwidth allocation is decoupled and different inter-ONU algorithms can be used.

The authors also introduced a modified version of Start-Time Fair Queueing [68,69] which is less complex than other virtual queueing systems where all packets are time stamped.

Yet another approach to the problem of bandwidth allocation was presented by Sherif et al. in [70]. The novelty of the proposed scheme is in the fact that the upstream control channel is redirected in the splitter and sent back to other ONUs in the domain. In principle this gives an ONU detailed knowledge about the buffer occupancy in other ONUs and, based on this information, an ONU can make a better decision about its bandwidth allocation. Although this is an interesting idea, introducing a feedback loop and making ONU responsible for bandwidth allocation has number of disadvantages:

- An ONU's architecture becomes very complex as extra functionality is added not only to perform the bandwidth allocation but also to extract the control signal from additional wavelength.
- With such a decentralized approach support for DiffServ is questionable as a control mechanism is needed to update the bandwidth assignment policies in every ONU in the network.

- The architecture of the EPON system becomes more complex and less functional. As the simplicity of the system is the main advantage of EPONs in comparison with other PON protocols and the proposed algorithm did not offer substantial improvements in bandwidth utilization this significantly reduces its competitiveness.
- The applicability of the proposed solution is also questionable from the optical budget point of view. As creating a feedback loop requires redirecting the output signal back through the splitter to ONUs, the signal power is attenuated twice. Assuming that the optical budget of EPON is 23 dB and a typical insertion loss of a 1:16 coupler is 13 dB, it is seen clearly that the signal after traversing the splitter for the second time will be too weak to decode properly in the receiver.

3.5 Conclusion

In this chapter an overview of EPONs was presented. The MPCP that all bandwidth allocation must rely on was highlighted and its most important features were described in detail.

The exchange of control messages leads to create a polling cycle. Two polling schemes: IPACT and cyclic polling were introduced in this section and compared with each other. The comparison shows that the IPACT tried to minimize the polling cycle. That gave shorter delays but also resulted in a variable control channel. The cycle length was fixed in cyclic polling and that resulted in longer delays for small loads. The advantage of such an approach is that longer cycles result in less bandwidth dedicated to guarding gaps.

Another implication arising from the structure of MPCP is a necessity of distributed inter- and intra-ONU bandwidth allocation. In inter-ONU allocation bandwidth is allocated amongst all ONUs served by the OLT. During intra-ONU allocation, an ONU makes the decision about allocating bandwidth to different classes of traffic. In this chapter a review of existing algorithms responsible for inter- and intra-ONU bandwidth allocation was presented.

Chapter 4

Centralized Bandwidth Allocation in EPON

New mechanisms enabling EPONs to support the DiffServ paradigm are introduced in this chapter. The format of a new control message which ensures that both distributed and centralized bandwidth allocation algorithms can be used in one EPON system is described. Another contribution is a new polling scheme which is deemed to be more versatile than other existing methods. The third contribution presented in this chapter is a novel Grant Multiplexing scheme which can significantly improve the QoS offered to jitter-sensitive classes of traffic.

4.1 Centralized Bandwidth Allocation

A centralized approach to the problem of bandwidth allocation is presented in this chapter. In this scheme both inter- and intra-ONU scheduling is done by the OLT. The centralized approach has numerous benefits and provides better DiffServ support. Relieving ONUs from scheduling responsibilities allows for the use of generic, simple and easy to manufacture equipment which gives the network provider a crucial edge in a highly competitive market.

From the overview of bandwidth assignment algorithms presented in Section 3.4 of Chapter 3, it is clearly seen that a majority of cases rely on some level of cooperation between the OLT and ONUs. In this thesis, it is claimed that this approach has two main disadvantages

- Cooperation between an ONU and the OLT might cause problems with compatibility when equipment with different bandwidth allocation algorithms is used. Currently the standard does not provide guidelines regarding the implementation of bandwidth allocation algorithms and it is more than likely that different vendors will release their own proprietary solutions where a tight co-operation between the OLT and ONUs is assumed. From a customer or service provider perspective, such a situation is dangerous. It would significantly decrease competition in the market as it is doubtful that interoperability could be provided between equipment from different vendors.

- In an EPON with separate inter- and intra-ONU, scheduling, support for SLAs is more cumbersome. The addition, maintenance and deletion of customers' contracts result in updating scheduling policies not only in the OLT but also in ONUs. This requires some control channels to be set up and results in a further increase in the ONU's complexity and hence the cost of design and manufacturing. The resulting increase in the cost of the equipment could be a deciding factor between EPON or another access method, especially in a situation where the initial investments in the fiber infrastructure are similar.

The centralized bandwidth assignment proposed here is based on the principle that both inter- and intra-ONU scheduling is done by the OLT. In this sense an ONU has no intelligence of its own and must fully rely on the decisions made by the OLT. The main advantage of such an approach is that the OLT is the only entity in the domain that has precise knowledge about the current state of all other units. This knowledge is obtained through the exchange of GATE and REPORT messages. Hence, the OLT is able to assign bandwidth in an efficient manner to an arbitrary queue in any ONU connected in the network. Such a "master-slave" configuration has the following benefits:

- Various algorithms can be used at different times depending on the network conditions. Moreover a transition between different methods is smooth, since due to placing intra-ONU scheduling in the OLT, ONUs are unaware of the whole process.

- SLAs are created and stored in the same logical unit as and are readily available to an algorithm that might want to use them. They are easier to access and modify, and the whole process does not influence normal network operation.
- There is no need to create a new signalling channel between the OLT and ONUs. Hence, ONUs are less complex and expensive to make.

The scalability and performance of the centralized approach is as good as the performance of the distributed algorithms. This is because the complexity of the bandwidth allocation algorithm grows with the number of active ONUs in the network and this feature is common for both centralized and distributed approaches. Obviously, the centralized algorithm must perform the extra tasks of allocating bandwidth to queues but as effect ONUs are simpler, more generic and robust.

4.2 The Format of the New GATE Message

The centralized approach, which is the subject of this thesis, cannot be implemented in EPON based on the current standard. As bandwidth allocation algorithms are not in the scope of the standard, it was assumed at an early stage that the OLT was only responsible for inter-ONU admission control. The intra-ONU scheduling was solely the duty of an ONU. Hence, as mentioned in Section 3.3.2, the GATE message conveys information about the bandwidth allocated to a particular ONU rather than to a particular queue. Solving this problem requires either introducing the new GATE message or making changes to the structure of the existing one. Although adding a new message to the protocol suite increases the overall complexity of the system, the undisputed advantage of such an approach is that backward compatibility is ensured and two types of ONUs can exist in the same domain. Here, a new GATE message is proposed whose format gives the OLT more flexibility in controlling the process of bandwidth allocation. At the same time, it provides a clear and logical way to separate the functionality of the bandwidth allocation algorithm from the underlying infrastructure and ONUs from different vendors can exist and cooperate in the same domain under some arbitrary bandwidth assignment scheme. The format of the proposed GATE message is shown in Figure 4.1. Although the differences in the structure of the old and the proposed GATE messages are not great, the new format gives the OLT far greater flexibility in controlling the bandwidth allocation process.

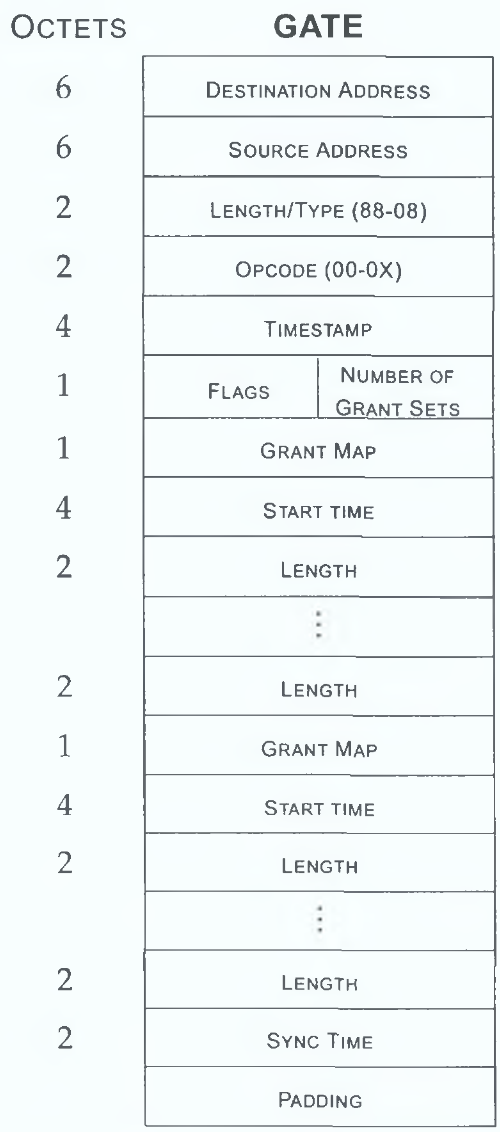


Figure 4.1: The Structure of New GATE Message.

4.2.1 The Encoding of the FLAGS/NUMBER OF GRANT SETS Field

In the IEEE 802.3ah standard [37], it is specified that the FLAGS/NUMBER OF GRANT SETS field is used to indicate the purpose of the GATE message and the number of grants that are included. In the proposed new format of the GATE message, the general structure of this field remains the same. Bits 0-3 are used to encode the number of Grant Sets, instead of indicating the number of grants. Bits 4-7 are dedicated to the FLAGS field. In the old format of the FLAGS field, particular bits were set or cleared to indicate the purpose of the GATE message. It is proposed here, that an integer number representation should be used instead. This greatly increases the flexibility of the protocol as 32 different operation codes can be used. Based on the decoded value, an ONU is able to identify the type of the GATE message and to extract information about allocated transmission windows from included Grant Sets. This also allows to maintain the backward compatibility. ONUs that do not support centralized bandwidth allocation will discard the whole message upon the receipt of the message with unknown operation code.

The format of the new message has the advantage over the old GATE message as it allows assignment of transmission windows to all eight queues supported by the ONU. The different configurations allowed by the proposed format of GATE are shown in Table 4.1.

4.2.2 Grant Sets

The main difference between the two GATE messages is the way in which information about allocated bandwidth is represented. In the proposed format of the GATE message the Grant Sets are introduced to convey information about allocated time slots. Each Grant Set consists of a GRANT MAP field, followed by a START TIME field which indicates the start of the allocated transmission window. The START TIME field is followed by number of LENGTH fields which carry information about the duration of consecutive transmission windows allocated to particular queues. The

Table 4.1: Different Configurations of Possible Grants in the GATE Message.

Number of Grant Sets	Number of Grants	Encoding	
		Number of Octets	Field
1	8	$1 \cdot 1$	FLAGS/NUMBER OF GRANT SETS
		$1 \cdot 1$	GRANT MAP
		$1 \cdot 4$	START TIME
		$8 \cdot 2$	LENGTH
		$1 \cdot 2$	SYNC TIME
		$16 \cdot 1$	PADDING
1	5	$1 \cdot 1$	FLAGS/NUMBER OF GRANT SETS
		$5 \cdot 1$	GRANT MAP
		$5 \cdot 4$	START TIME
		$5 \cdot 2$	LENGTH
		$1 \cdot 2$	SYNC TIME
		$2 \cdot 1$	PADDING

GRANT MAP is a register in which bits indicate queues to which those transmission windows are assigned to. An ONU should switch its laser off at the end of the last transmission window in any Grant Set. If an ONU operates under a centralized bandwidth allocation scheme the OLT must indicate in the GRANT MAP at least one queue for which a transmission window is allocated. Conversely, if distributed bandwidth allocation is implemented and Grant Sets are not used, all bits in the GRANT MAP must be set to one and this field is followed by a single tuple of START TIME and LENGTH fields delineating a transmission window allocated to all queues in an ONU. During the discovery process the encoding of the allocated transmission window is the same as for distributed bandwidth allocation as in this mode, ONUs sends only short control packets.

4.3 Cyclic Polling with Adaptive Cycle Length

In *TDMA* based networks there are three main factors that affect packet delays [71].

Cycle length This is the time in which all active stations may transmit the data. Cycle length is measured as the time elapsed between two GATE messages sent to the same ONU. Intuitively, the longer the cycle, the longer stations have to wait for their turn and the longer packets have to be buffered. On average, if the order in which ONUs transmit their data is random, the waiting time is equal to half of the cycle length.

Transmission time This term refers to the time needed to send a packet and propagate it between two end stations. In modern systems operating at very high data rates over much longer distances the propagation time cannot be neglected and it becomes significant part of the transmission time. As an example it takes only 8 μs to transmit a 1000 byte long packet on 1 Gbit/s link and almost 100 μs to propagate the signal over a 20 km long link. For a link operating at 10 Mb/s rate the time needed to send 1000 bytes is 800 μs which is significantly longer in comparison with the propagation time.

Queueing time The length of time that a packet is buffered at the station.

It is clearly seen that the first component is the most important, especially for small and medium loads where time spent in a buffer is usually short. In EPON the length of the cycle depends on the implemented policy and different approaches have been presented in Section 3.4.

In this dissertation a new approach Cyclic Polling with Adaptive Cycle Length based on the modified cyclic polling scheme is advocated. Similar to the method presented by Choi in [53], new grants are allocated to all stations at the same time. Contrary to his approach, the length of the polling cycle is adaptive and the minimum and maximum length of the cycle are not dependent on the bandwidth allocation algorithm deployed in the network. The functionality of this approach is briefly outlined by the following steps:

1. The total number of bytes Q^{total} in all queues is calculated based on latest reports received.

2. The cycle time τ is calculated from (4.1) where C_L is the link rate.

$$\tau(n) = \frac{Q^{total}}{C_L} \tag{4.1}$$

3. It must be ensured that time $\tau(n)$ fulfills (4.2) where τ^{min} and τ^{max} are the minimum and maximum length of the cycle.

$$\tau^{min} \leq \tau(n) \leq \tau^{max} \tag{4.2}$$

   The maximum cycle time does not depend on the physical properties of the system and can be set to any value. The minimum cycle time must be such that enough time is provided to process all REPORT messages that arrived during the last polling cycle and that GATE messages have enough time to arrive at all ONUs.

4. The new time slots are calculated for the next cycle based on reports received in the current one. The GATE messages are issued for all ONUs at the same instance.

The schematic exchange of control messages in Cyclic Polling Algorithm with Adaptive Cycle Length is shown in Figure 4.2. This approach is an interesting alternative to IPACT and cyclic polling with constant polling cycle and it is a compromise between these two algorithms.

Similar to the IPACT, in approach presented here, the polling cycle is variable. In a situation where low volume traffic is present in the network this results in a shorter length of the polling cycle and thus smaller average delays. Shortening of the polling cycle requires that GATE and REPORT messages be sent more often and the bandwidth available to customers' data is reduced. This is especially a problem

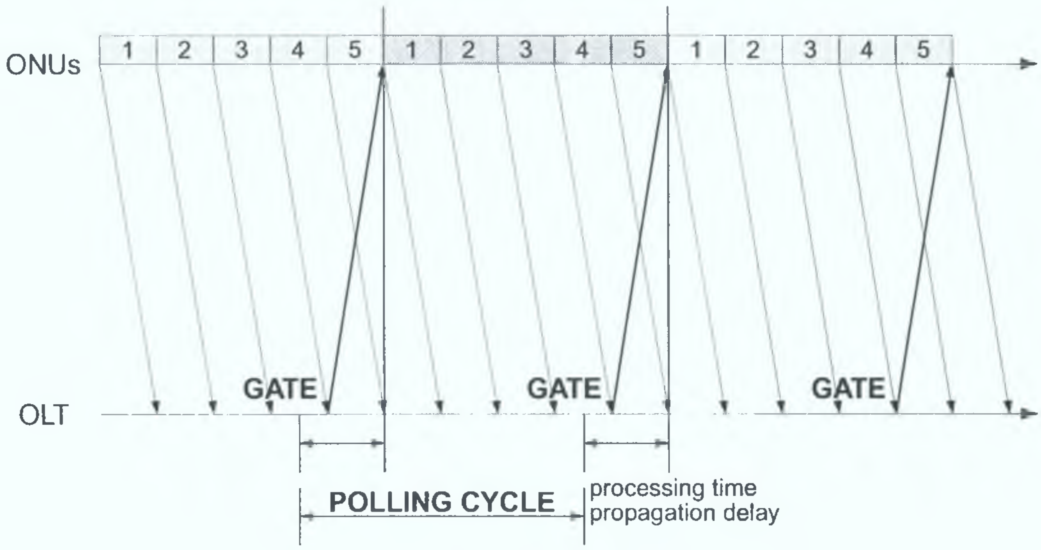


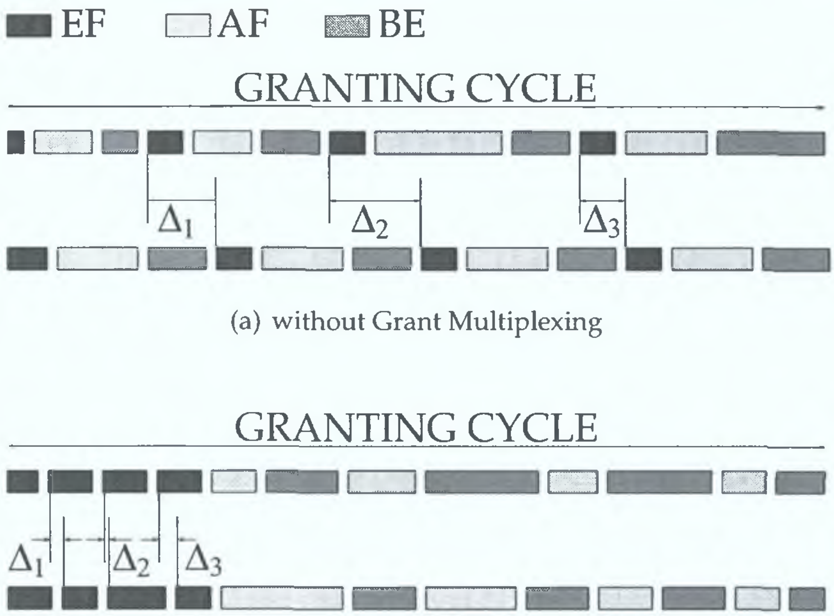
Figure 4.2: Cyclic Polling with Adaptive Cycle Length.

when traffic in the downstream direction has a much larger volume than in the opposite direction. In such a case the calculated cycle time is short, and in order to provide enough bandwidth to send GATE messages, some customers' packets must be dropped to avoid congestion. Increasing the minimal length of the cycle will improve this situation significantly. As GATE messages must be sent at larger intervals, less bandwidth is needed for the control channel and the need to drop customers' packets is less likely to arise. By changing the minimum length of the polling cycle, the OLT can regulate the size of the control channel and improve the QoS delivered in either the upstream or downstream direction.

4.4 Grant Multiplexing Scheme

> *Jitter can be defined as the difference between the minimal and maximal interarrival times of packets at different times. [72]*

Voice and real-time video services are delay sensitive and require low values of jitter. Employment of a DBA scheme may result in considerable variation in the



(a) without Grant Multiplexing

(b) Grant Multiplexing

Figure 4.3: Grant Multiplexing Scheme.

assigned bandwidth between consecutive polling cycles. Hence, it is difficult to maintain stable and good traffic parameters for the high priority classes. The mechanism of Grant Multiplexing guarantees considerable improvements of QoS offered to jitter sensitive classes.

The mechanism proposed in this thesis is based on the assumption that the bandwidth requested by high priority traffic is small in comparison with the bandwidth needed by medium and low priority classes. It is also assumed that high priority traffic is mainly voice and video, which is not bursty in nature. On the other hand, medium and low priority classes carry busty Internet traffic and so the amount of requested bandwidth can fluctuate considerably. This results in large deviations in the offset time measured between the start of the polling cycle and start of the transmission window as shown in Figure 4.3(a). The values of Δ_1,Δ_2 and Δ_3 emphasize the differences between starting points of transmission windows assigned to the same ONUs in different cycles.

However, with the new GATE message, the OLT can assign bandwidth to an arbitrary queue rather than to an ONU and it is possible to grant time slots to high priority queues at the beginning of the cycle. The new situation, where high priority classes are allocated at the beginning of the cycle, is shown in Figure 4.3(b). It is expected that the size of transmission windows allocated to high priority classes will vary less in this way, hence the variance in the offset time and jitter could be considerably reduced.

The main disadvantage of this approach is that the available bandwidth is utilized less efficiently in comparison to networks without Grant Multiplexing Scheme. In EPON, transmissions from different ONUs must be separated by a guard time and hence twice as much bandwidth must be dedicated to provide necessary isolation when the Grant Multiplexing scheme is employed.

4.5 Conclusion

In this chapter the idea of a centralized bandwidth allocation scheme was outlined and a new GATE message was introduced giving the OLT much greater flexibility in the control of the bandwidth allocation process. The main advantage of a new GATE message is that a clear way is provided to separate bandwidth allocation algorithms from the underlying MPCP. This allows for deployment of sophisticated algorithms utilizing both the centralized and distributed inter- and intra-ONU bandwidth allocation, as envisaged in the EPON standard [37]. Also the new format provides backward compatibility so ONUs of both types can be supported in the same network.

The main advantage of the centralized inter- and intra-ONU can offer is the possible savings in the cost of the equipment. As both inter- and intra-ONU scheduling can be performed by the OLT, ONUs with only rudimentary access control mechanisms are needed.

Another strong point of centralized approach is its flexibility. Different bandwidth allocation methods can be deployed at virtually any time. New methods can be tailored to the existing traffic conditions. In the distributed scenario such an approach has only limited applicability, as changing policies would require updating parameters in every ONU in the network.

The new mechanism also offers substantial flexibility in the allocation of transmission windows. This was described in Section 4.4 at the example of new Grant Multiplexing mechanism. This mechanism is anticipated to improve the QoS received by the high priority classes. The mechanism is based on the assumption that high priority traffic is not bursty and that all transmission windows allocated to privileged classes are allocated at the beginning of the polling cycle. It is envisaged that such an approach could reduce jitter measured for packets belonging to these classes.

Chapter 5

Bandwidth Assignment Algorithms

Bandwidth assignment algorithms are necessary in providing optimal utilization of available resources. The implementation of the algorithms is not in the scope of the existing standard but their functionality must rely on the MPCP protocol in exchanging information between ONUs and the OLT. In this chapter algorithms based on a centralized approach are presented. It is assumed that the OLT is responsible for both inter and intra-ONU bandwidth allocation as it has knowledge about the current status of the network gathered from REPORT messages received from all ONUs. The main focus of the algorithms presented here is on providing efficient support for the DiffServ architecture.

The performance of the proposed algorithms in the presence of heavy loads and the impact of policies set in the SLAs on the general performance were evaluated. In this chapter the adaptation to centralized scheme of two well known Static Bandwidth Allocation (SBA) and Proportional Dynamic Bandwidth Allocation (P-DBA) algorithms is presented. These algorithms were used as a base in development of new methods. In this chapter three new algorithms are proposed and their functionality is described in detail:

- Strict Priority Dynamic Bandwidth Allocation (SP-DBA),
- Service Level Agreement aware Dynamic Bandwidth Allocation (SLA-DBA),
- Adaptive Dynamic Bandwidth Allocation (A-DBA).

5.1 Static Bandwidth Assignment

In SBA, the decision about bandwidth allocation is not influenced by information received from ONUs. Let $\gamma_i(j)$ be the fraction of bandwidth which queue i in ONU j is guaranteed in the SLA. Then the size of the transmission window $\omega_i(j)$ is obtained from (5.1) , where $\tau(n)$ is the length of a cycle.

$$\omega_i^{[n]}(j) = \tau(n) \cdot \gamma_i(j) \tag{5.1}$$

In SBA, more than in the other algorithms, the average packet delays are directly dependent on the bandwidth assigned to a particular class of traffic. If all the expected volume of traffic can be sent in the allocated transmission window, a small penalty is incured. Conversely, delays will grow if the allocate transmission window is too small, as no adjustments are made to the allocated bandwidth.

The SBA is only suitable in a situation where the amount of traffic can be accurately forecasted in advance and changes in volume follow some pattern known to the OLT. Traffic present in modern computer networks has been shown to be very bursty in nature with large variance from the mean volume [73–75] and in such conditions the functionality of SBA can lead to increased delays and lower throughput.

5.2 Proportional Bandwidth Assignment

The P-DBA algorithm is based on an scheme that fully utilizes updates on the status of connected ONUs to calculate the size of transmission windows. The algorithm works on the principle that bandwidth is divided amongst all queues proportionally to the reported buffer occupancy. Let $Q^i(j)$ be number of bytes reported in queue i by ONU j. A fraction of bandwidth allocated to a particular queue is obtained from (5.2) .

$$\beta_i(j) = \frac{Q_i(j)}{\sum_{i,j} Q_i(j)} \tag{5.2}$$

Assuming that the cycle length is known and equal to $\tau(n)$ the length of of transmission window $\omega_i(j)$ allocated to queue j in ONU i is given by (5.3) . Moreover this approach ensures that all available bandwidth is utilized efficiently.

$$\omega_i(j) = \tau(n) \cdot \frac{Q_i(j)}{\sum_{i,j} Q_i(j)} \tag{5.3}$$

The main advantages of P-DBA are its responsiveness and fairness. As the allocation is based on latest reports, the possible congestion can be quickly resolved and backlogged packets transmitted. The proportional assignment is very unbiased and all types of traffic are treated equally. This is great as long as all users comply with their SLAs. Otherwise, traffic from non-compliant sources will affect QoS delivered to other sources. In this sense, P-DBA is not suitable for DiffServ applications.

5.3 Priority Bandwidth Assignment

In SP-DBA, support for classes of traffic demanding various QoS is implemented by introducing strict priority scheduling. The bandwidth allocation process consists of three steps:

1. Based on received reports the total bandwidth requested by different types of service is calculated.

2. The requested bandwidth is allocated to classes of service based on their priority. The top priority classes are looked after first. A situation is possible where lower priority classes will be allocated less than requested or no bandwidth at all.

3. The bandwidth allocated to a given class of service is divided among all queues as in the P-DBA scheme. This ensures that if less bandwidth was allocated than requested all queues will be treated equally.

Two goals were envisaged before process of implementing SP-DBA in EPON with centralized bandwidth allocation is started:

- better support for the DiffServ paradigm,

- showing that with the centralized approach the same functionality could be achieved as in EPON where distributed inter and intra-ONU scheduling was implemented [49,50,54,55].

It is be anticipated that strict priority scheduling will be able to provide requested QoS for the classes of traffic with high priority. On the other hand, this might lead to worse performance of medium and low traffic, since for heavy loads the bandwidth allocated to these classes can be severely limited. The pseudo code of SP-DBA scheduler is shown in Figure 5.1.

Let P indicate the number of queues with different priorities in the system and let 0 be the highest and P the lowest priority level. Also, let $\beta^{total}(k)$ be total number of bytes requested by queues with priority k and $P_i(j)$ be the priority of queue j in ONU i. Finally let β^{max} be the total number of bytes that can be transmit in current cycle.

for $k = 0$ to P **do**
  for $i = 0$ to Onu **do**
    for $j = 0$ to $Queue$ **do**
      $\beta^{total}(k) \Leftarrow \beta_i^k(j) + Q_i^k(j)$
    end for
  end for
end for

for $k = 0$ to P **do**
  if $\beta^{total}(k) > (\beta^{max} - \sum_{l=0}^{l<k} \beta^{total}(k))$ **then**
    for $i = 0$ to $i < Onu$ **do**
      for $j = 0$ to $j < Queue$ **do**
        if $k < P_i(j)$ **then**
          $\beta_i(j) \Leftarrow \beta_i(j)$
        end if
        if $k = P_i(j)$ **then**
          $\beta_i(j) \Leftarrow \beta_i(j) \cdot \frac{\beta^{total}(k)}{\beta^{max} - \sum_{l=0}^{l<k} \beta^{total}(k)}$
        end if
        if $k > P_i(j)$ **then**
          $\beta_i(j) \Leftarrow 0$
        end if
      end for
    end for
  end if
end for

Figure 5.1: Pseudocode of SP-DBA Scheduler.

5.4 SLA aware P-DBA

In this section the SLA-DBA algorithm is outlined. The core functionality of this algorithm is based on the P-DBA approach to manage the optimal allocation of available resources and good responsiveness to changing network conditions. In the SLA-DBA algorithm achieving different QoS for different types of traffic is the primary goal.

From the functionality point of view this approach consists of three steps. In the first phase the algorithm allocates bandwidth proportionally to the reported queue length. This is identical to the functionality of P-DBA. As mentioned in Section 5.2 this approach gives no protection of traffic parameters. In order to secure a right amount of bandwidth for every queue, in the second stage the constraints agreed in the SLA are taken into account. A pseudo code for SLA-DBA is presented in Figure 5.2.

Let $Q_i(j)$ be the number of bytes reported in queue j by ONU i and $\beta_i^k(j)$ is the number of bytes allocated to this queue in step k of the algorithm. The number of bytes that can be sent in a particular granting cycle is given by $\beta^{[n]}$ which must honor the constraints:

$$\tau^{min} \leq \frac{\beta^{[n]} \cdot 8}{C_L} \leq \tau^{max}, \tag{5.4}$$

where τ^{min} and τ^{max} are the minimum and maximum time of a granting cycle and C_L is the link capacity in bits per second.

In Phase I, the amount of assigned bandwidth to a particular queue can be expressed as:

$$\beta_i^I(j) = \frac{Q_i(j)}{\sum_{i,j} Q_i(j)} \cdot \beta^{[n]} \tag{5.5}$$

Also in this step, the excess bandwidth β^{ex} is calculated as the sum of bandwidth allocated to all low priority queues. At later stage this excess bandwidth is divided amongst the queues with higher priorities if the bandwidth allocated in the first phase was smaller than the expected minimum.

Let $\gamma_i^{min}(j)$ and $\gamma_i^{max}(j)$ be the minimum and maximum number of bytes guar-

```
:STEP I
Require: ∃Q^i_j > 0
  for i = 0 to Onu do
    for j = 0 to Queue do
      Req ⇐ Req + Q_i(j)
    end for
  end for
  for i = 0 to Onu do
    for j = 0 to Queue do
      β^I_i(j) ⇐ 1/Req · Q_i(j) · β^max
    end for
  end for
  :STEP II
  for i = 0 to Onu do
    for j = 0 to Queue do
      if β^I_i(j) − γ^max_i(j) ≥ 0 then
        β^II_i(j) ⇐ γ^max_i(j)
        β^ex ⇐ β^ex + (β^I_i(j) − γ^max_i(j))
      end if
      if β^I_i(j) ≥ γ^min_i(j)(j) ∧ β^I_i(j) < γ^min_i(j) then
        β^II_i(j) ⇐ β^I_i(j)
      end if
      if β^I_i(j) < γ^min_i(j) then
        if β^ex − γ^min_i(j) > 0 then
          β^II_i(j) ⇐ γ^min_i(j)
          β^ex ⇐ β^ex − γ^min_i(j)
        else
          β^II_i(j) ⇐ β^I_i(j)
        end if
      end if
    end for
  end for
  :STEP III
  if β^ex > 0 then
    for i = 0 to Onu do
      for j = 0 to Queue do
        β^III_i(j) ⇐ β^II_i(j) + β^ex · Q_i(j) / Σ_{i,j} Q_i(j)
      end for
    end for
  end if
```

Figure 5.2: Pseudocode of SLA-DBA Scheduler.

anteed to a particular queue. In the Phase II the constraints given in the SLAs are applied to all high and medium priority classes. Three distinct situations have to be considered at this stage:

1. $\beta_i^I(j) \geq \gamma_i^{max}(j)$ – Assigned bandwidth has exceeded the amount promised in the SLA. The bandwidth allocated to a particular queue is thus reduced to $\beta_i^{II}(j) = \gamma_i^{max}(j)$.

2. $\beta_i^I(j) \geq \gamma_i^{min}(j)$ and $\beta_i^I(j) < \gamma_i^{max}(j)$ – Requested bandwidth is within the limits of the SLA. No changes are made and $\beta_i^{II}(j) = \beta_i^I(j)$

3. $\beta_i^I(j) < \gamma_i^{min}(j)$ – In a situation where there is enough of excess bandwidth β^{ex} the allocated bandwidth is equal to $\gamma_i^{min}(j)$. Otherwise the amount of the allocated bandwidth does not change.

Bandwidth that is not allocated during the second phase is shared amongst all queues in Phase III. The amount of bandwidth assigned to a queue is expressed as:

$$\beta_i^{III}(j) = \beta_i^{II}(j) + \frac{\beta^{ex}}{\beta_i^I(j)} \tag{5.6}$$

After obtaining the new bandwidth allocations for every queue the new transmission windows are assigned. The size of a new transmission window is calculated from (5.7) where $\tau(n)$ is the length of the cycle as calculated from (4.1) .

$$\omega_i^{[n]}(j) = \tau(n) \cdot \frac{\beta_i(j)}{\sum_{i,j} \beta_i(j)} \tag{5.7}$$

5.5 SLA aware Adaptive DBA

Analysis of the structure of SBA and P-DBA algorithms leads to a different approach to the problem. As shown earlier in the SBA algorithm a fixed amount of bandwidth is allocated to a class of traffic. Conversely, P-DBA is quick to react to the changing conditions as the bandwidth allocation is based solely on reports received from

ONUs. In an effort to incorporate positive aspects of both SBA and P-DBA in one approach the A-DBA algorithm was designed.

In order to achieve the best performance it is assumed that the amount of assigned bandwidth is dependent on the reported length of a queue. To emulate the level of QoS support offered by the SBA, the maximum allowable bandwidth that can be assigned to the queue was introduced. The amount of allowable bandwidth is subject to off-line negotiations between the customer and network provider and is set in the SLA. The parameters in the SLA are chosen in a way such that as long as a particular source transmits packets at a rate lower than the maximum, they are guaranteed to be forwarded without an additional delay. If the source exceeds the agreed maximum rate then its packets will be sent at the maximum allowed rate and surplus data will be buffered until the source decreases its rate below the maximum rate or other sources have no data to send.

As in previous sections, let $Q_{i,j}^{[n]}$ be the queue length in bytes reported for queue j by ONU i in granting cycle n. Let $\beta_{i,j}^{[n]}$ be a number of bytes assigned in granting cycle n and $\beta_{i,j}^{max}$ the maximum number of bytes that can be sent by a particular queue in one cycle. The granting cycle time $\tau(n)$ is dependent on the total amount of bandwidth allocated to the queues and is given by (5.8) . Time $\tau(n)$ can not be longer than τ^{max} which is calculated from (5.9) .

$$\tau(n) = \sum_{i,j} \beta_i^{[n]}(j) \frac{8}{C_L} \tag{5.8}$$

$$\tau^{max} = \sum_{i,j} \beta_i^{max}(j) \frac{8}{C_L} \tag{5.9}$$

The block diagram of the A-DBA algorithm is shown in Figure 5.3. Based on values obtained from the scheduler the OLT calculates new transmission windows from (5.10) , where $\tau(n)$, as in the previous section, is the length of the cycle as calculated from (4.1) .

$$\omega_i^{[n]}(j) = \tau(n) \cdot \frac{\beta_i^{[n]}(j)}{\sum_{i,j} \beta_i^{[n]}(j)} \tag{5.10}$$

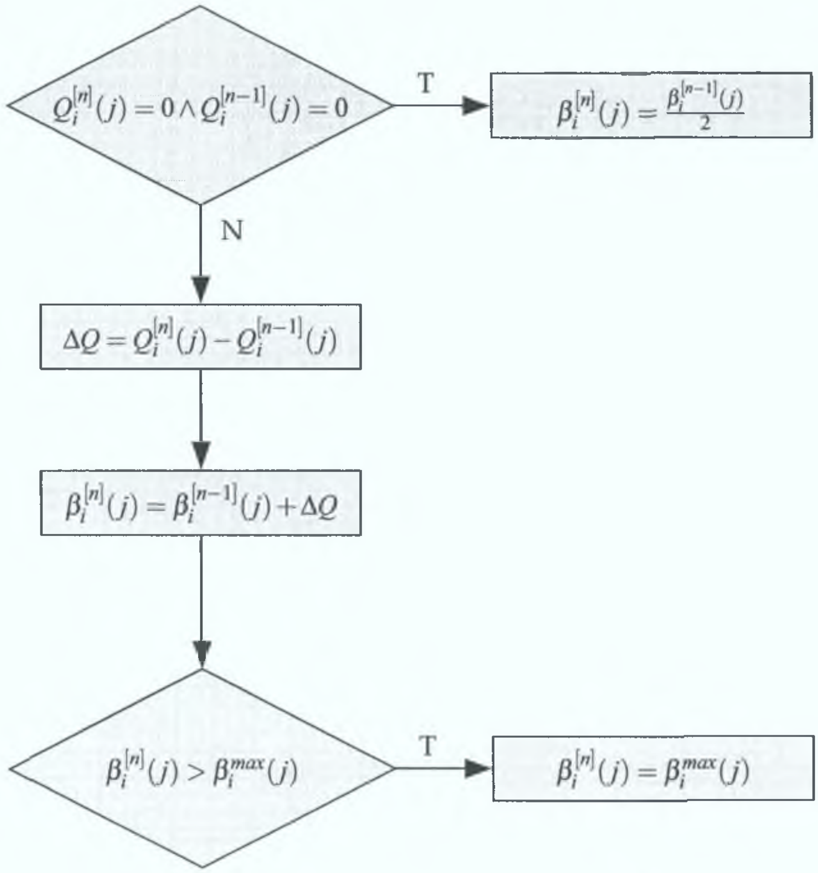


Figure 5.3: Block Diagram of the A-DBA Algorithm.

5.6 Conclusion

In this chapter the architecture of different algorithms based on centralized inter and intra-ONU bandwidth allocation were presented:

- Well known SBA and P-DBA algorithms were adapted to the centralized environment. As the functionality of these algorithms is well known, the results recorded for them can be used to benchmark the newly proposed algorithms.
- The centralized version of SP-DBA algorithm was also presented in this section. This well known solution is often employed as a default per hop behav-

ior in EPONs. It is important to show that the centralized approach can be at least as flexible as the distributed inter- and intra-ONU allocation scheme. It is thought that the comparison between the results recorded for both versions of SP-DBA could show that the performance of centralized version would be similar.

- A new SLA-DBA algorithm was proposed in this section. This three step algorithm is based on the P-DBA scheme. The existing approach was enhanced and extra functionality was added to enable support for SLAs.
- The new A-DBA algorithm was also outlined in this section. This algorithm is based on the SBA scheme. The algorithm tries to maintain the same level of support for SLA in different traffic conditions by assigning some maximum transmission window size to every class of traffic.

All algorithms presented in this section are independent from the cycle length and they operate on relative values. The calculated sizes of transmission windows must be scaled to the absolute length of the polling cycle. Such an approach guarantees that the change in the cycle length will not affect the functionality of the algorithms.

Chapter 6

Simulator

The architecture of PONs is simple and straightforward, yet it eludes mathematical analysis and no models which produce accurate results have been developed. This is primarily a result of complex dependencies between packet delays, the polling cycle length and the choice of bandwidth allocation method. Thus, the analysis of EPON's behavior presented in the literature is based on simulation techniques and it was decided to follow the same path in this thesis.

6.1 Method of Choice

Simulating high speed networks is not a trivial problem and some questions regarding the performance and time scale of the experiments must be considered. At the early stages of the research, it was decided to build a custom EPON simulator and not to use currently available tools like NS-2[1] or OPNET[2]. The main reason was performance. Both simulation packages offer a wide functionality and can be used to model a variety of problems. However, this functionality puts a strain on the performance of the system, particularly where high speed connections are under investigation. This is because even a short simulation period requires generation of an enormous number of events that must be processed. As packets flows over 1 Gbps links have to be simulated the estimated number of events, even in the simplest experiment, can exceed 400,000 per one second of simulated real time. In such a situation, with a streamlined tool where only bare functionality is implemented,

[1]http://www.isi.edu.pl/nsnam

[2]http://www.opnet.com

more complex and demanding experiments can be created and executed in a shorter period of time.

Another factor that influenced the decision to implement a custom simulator was objective problems with NS-2 and OPNET. As no PON modules were present at the time, their development would demand an in-depth knowledge of the architecture of the simulation tools. This is especially true in the case of NS-2, which is distributed as Open Source software. Different people have contributed sections of code, and therefore to gain a sufficient working knowledge of the software package would require substantial amount of time. In case of OPNET, which is distributed by OPNET Technologies, Inc. prohibitive pricing and complex licensing issues were the limiting factors and it was decided to implement the simulator in C++.

6.2 Simulator Architecture

The best definition of simulation is offered by Banks in [76]:

> *Simulation is the imitation of the operation of a real-world process or system over time.*

Simulation as a problem solving method is an especially important and indispensable part of the analysis of modern communication systems, where dependencies between different parts of the system are hard to grasp and elude other forms of analysis [76–80]. Building a simulation experiment relies on a detailed model of the system under investigation. The model must describe the real entity with enough accuracy to include all important relationships. On the other side, excessive complexity of the model might obscure the analyzed problem and a large number of events might lead to slow performance.

6.2.1 Model

During the design phase it was decided that the adoption of a discrete event-scheduling method of simulation was the best way to implement the EPON's functionality. The decision was based on following factors:

- Event-driven simulation allows the behavior of the investigated system to be shown over a long period of time. This is extremely important as capturing the flow of packets in the network is essential.
- The life cycle of a packet in the system can be described in terms of discrete events, e.g. packet arrival at the queue, packet departure from the queue, packet arrival at the OLT, etc.
- The time in the system is advanced from one event to another. New events are scheduled as a result of the processing of other events that have occurred and been dispatched earlier. This implies that no change in the state of the system is allowed in between the scheduled events.

6.2.2 Structure

One of the main steps in designing a simulator is to analyze the performance of the system that is under investigation and to represent it as a network of independent entities that cooperate with each other and communicate by means of discrete messages. Such a representation simplifies the process of modeling and allows for better understanding of the system [76, 77]. At this stage of building the EPON simulator three type of entities were distinguished as being crucially important:

OLT From the simulation point of view two aspects of the OLT functionality, bandwidth allocation and time slot scheduling, have to be modelled in detail.

ONU The model describing an ONU must provide mechanisms for efficient handling of packet arrival, queueing and departures from the queues.

USER The user's functionality is limited to periodical generation of data packets that enter the system with some arbitrary distribution of inter-arrival times.

In a real network the communication between all units is carried out by means of packets with their arrival and departure triggering some specific actions. Thus, it is convenient to model the cooperation of modules in the simulator as a set of discrete events. Such a representation simplifies the modeling of different types of delays that might affect the time a packet spends in the system. A schematic of the relationship between the main entities and flow of events is depicted in Figure 6.1.

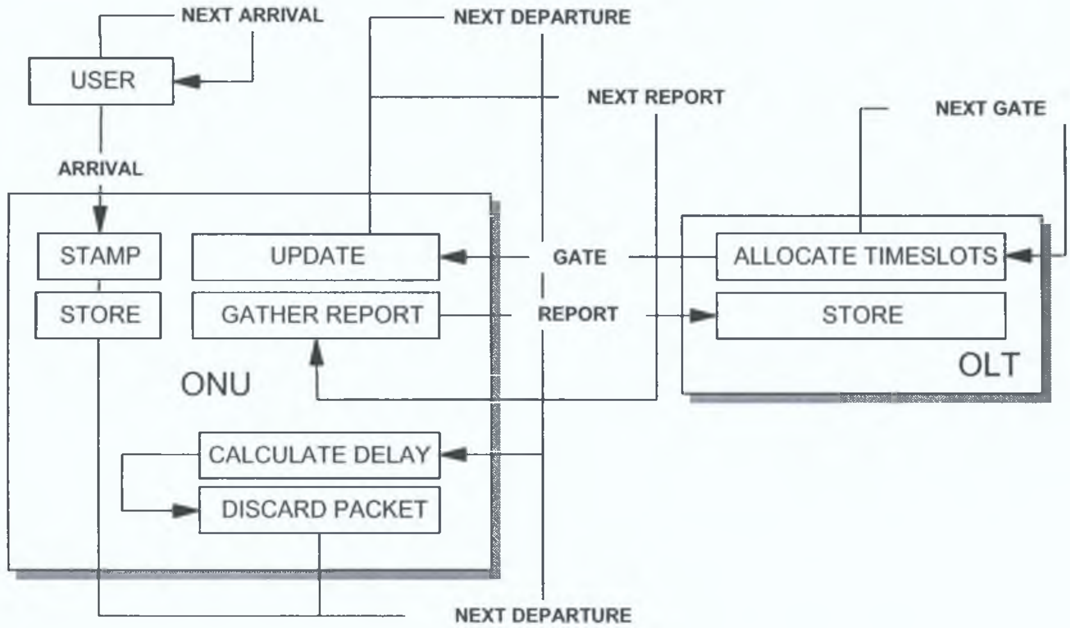


Figure 6.1: Main Entities and Flow of Events in the Simulator Model.

6.3 Flow of Events

Implementing a discrete event simulator involves the periodical processing of events generated by different entities in the system. When a new event is created, the time of its execution is set and it is added to a queue which maintains a chronological order of all unprocessed events. The event at the head of this queue is always processed next.

6.3.1 Description of Events

Analysis of Figure 6.1 reveals that only four main events are used in the EPON simulator:

- GATE
- REPORT
- PACKET ARRIVAL
- PACKET DEPARTURE

In processing of GATE events the functionality of the OLT must be replicated as closely as possible. Based on information received during processing of REPORT events, new time slots are calculated and notification is sent to respective ONUs. Adoption of object oriented objective programming techniques such as polymorphism and inheritance allows the implementation of different scheduling algorithms without introducing changes to the core functionality of the simulator. An a result the functionality of the simulation module and the OLT module are independent from each other. When all activities connected with processing of the current GATE event are completed the next GATE event is scheduled and added to the event queue. REPORT and PACKET DEPARTURE events are also scheduled at this time. The purpose of the REPORT event is to mark the time when an ONU should notify the OLT about its queue lengths.

The lifetime of a packet in the system is limited by PACKET ARRIVAL and PACKET DEPARTURE events. On the PACKET ARRIVAL event, a new packet arrives from a source to an ONU and enters one of the queues depending on its class. In the case where the packet arrives to an empty queue and the time slot assigned to that queue has not yet ended, an instance of the PACKET DEPARTURE event for this packet is created. Otherwise the packet is buffered. After the PACKET ARRIVAL event is processed, the arrival of the next packet from the same source is scheduled.

The packet is removed from the system on the PACKET DEPARTURE event. The total time it has spent in the system is recorded for further statistical analysis purposes. In the simulated EPON system a packet can only leave the system in a situation when the following conditions are met:

- The packet is at the head of the queue.
- The time slot for a given queue has started and it will not end before the transmission of the whole packet currently at the head of the queue is completed.

The next PACKET DEPARTURE event is only scheduled if the previous packet was sent. Otherwise packets must wait for the next time slot. The example of flow of events in the simulated EPON system is presented in Figure 6.2.

Time	Event Type	Onu:Queue	Description/Action	Scheduled Time
0000	PACKET ARRIVAL	0:0		0005
0000	PACKET ARRIVAL	0:1		0008
0000	PACKET ARRIVAL	1:0		0016
0000	PACKET ARRIVAL	1:1		0014
0000	GATE	0:0	Time slot duration 10	
		0:1	Time slot duration 10	
		1:0	Time slot duration 10	
		1:1	Time slot duration 10	
		0:0	REPORT	0010
		0:1	REPORT	0020
		1:0	REPORT	0030
		1:1	REPORT	0040
		0:0	PACKET DEPARTURE	0000
		0:1	PACKET DEPARTURE	0010
		1:0	PACKET DEPARTURE	0020
		1:1	PACKET DEPARTURE	0030
		0:0	GATE	0040
0000	PACKET DEPARTURE	0:0		
0005	PACKET ARRIVAL	0:0	Packet size 2	
		0:0	Queue	
		0:0	PACKET DEPARTURE	0005
		0:0	PACKET ARRIVAL	0009
0005	PACKET DEPARTURE	0:0	Calculate delay	
		0:0	PACKET DEPARTURE 0:0	0007
0007	PACKET DEPARTURE	0:0		
0008	PACKET ARRIVAL	0:1	Packet size 4	
		0:1	Queue	
		0:1	PACKET ARRIVAL	0011
0009	PACKET ARRIVAL	0:0	Packet size 4	
		0:0	Queue	
		0:0	PACKET ARRIVAL	0013
0010	REPORT	0:0	1 packet	
0010	PACKET DEPARTURE	0:1	Calculate delay	
		0:1	PACKET DEPARTURE	0014
0011	PACKET ARRIVAL	0:1	Packet size 3	
		0:1	Queue	
		0:1	PACKET ARRIVAL	0017

Figure 6.2: Sample Flow of Events.

6.3.2 Processing of Events

The discrete event method of simulation requires that events be created in logical and chronological order. Events arriving out of sequence can affect the generated output. The simulator used to test EPONs has been designed to minimize the possibility of such an out of sequence event occurring. Only when one event has been processed can new events be added to the priority queue. The simulation progresses as system time advances from one event to another and it is run for as long as there are events in the queue or the simulation time has not expired. The schematic diagram of the event processing loop is presented in Figure 6.3.

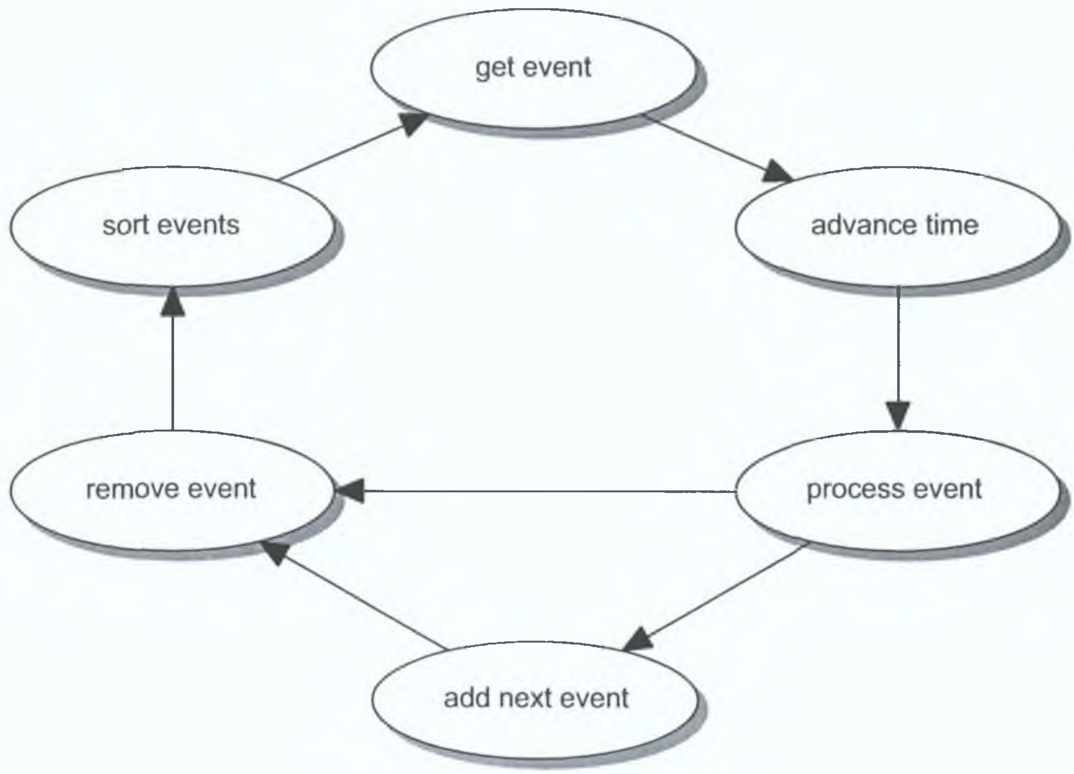


Figure 6.3: Events Processing Loop.

6.4 Traffic Sources

Probably the most important part of every simulator is the pseudo random generator. It is used to generate the numbers used to indicate the arrival of new events and their duration. The most common method to obtain a pseudo random sequence of numbers with an arbitrary distribution is based on finding a solution to (6.1) , where F^{-1} is the inverse of a distribution function and u is a random variable such that $0 < u < 1$ [81–83].

$$x = F^{-1}(u) \tag{6.1}$$

On the other hand a problem remains of finding a good generator producing "random" numbers which are uniformly distributed between 0 and 1. The most common method to generate pseudo-random sequences is the Linear Congruential Method (LCG) method, where the computation of the successive random number is based on the previous numbers [82]. The negative effect of such an approach is that generator is periodic. After sufficiently long sequence the generator will get back to its starting point and the numbers will start repeating. Generators with too short cycles are unacceptable in demanding simulation experiments as the correlation between generated numbers can affect the outcome of the experiment.

In the EPON simulator a family of generators used by Kramer[3] et al. was incorporated. As many ideas presented in this dissertation are alternative to the approach advocated by Kramer et al. it seemed reasonable to build the simulator on similar foundations. This ensures that the results are comparable with each other in terms of the structure and properties of generated traffic.

The adopted generators use the "Mersenne Twister" algorithm described by Makoto Matsumoto and Takuji Nishimura [84], and are implemented by Richard J. Wagner to obtain uniformly distributed numbers between 0 and 1. As shown in [84,85] the "Mersenne Twister" algorithm has extremely good "randomness" and a short execution time. As its cycle length can be as high as $2^19937 - 1$ it is highly

[3] http://wwwcsif.cs.ucdavis.edu/~kramer/code/trf_gen2.html

improbable that the generated numbers will start repeating during a simulation.

The adopted generators allow the generation of packet streams with three different distributions of interarrival times:

Exponential The distribution of interarrival times is given in (6.2) , where λ is the mean interarrival time and t is time between adjacent arrivals [86]traffic sources.

$$D(x) = 1 - e^{-\lambda t} \tag{6.2}$$

Exponential distribution of interarrival times is probably most widely adopted in queuing theory and interarrival times in the Poisson arrival process are exponentially distributed. The exponential distribution has a property of being *memoryless*, which means that the future events are not dependent on the past. The inverse function $F^{-1}(u)$ is given by (6.3) .

$$F^{-1}(u) = -\frac{1}{\lambda} \ln 1 - u \tag{6.3}$$

Pareto It was shown in [73–75, 87, 88] that the traffic in modern data networks can be described by the terms *"self-similarity"* and *"long-range dependent"*. Self-similarity of traffic means that traffic patterns are similar to each other irrespective of the time scale. A process that is long-range dependent can be described by a stochastic process with a *heavy-tail* distribution. One of the key features of such processes, is that their variance is infinite. In the long-range dependent process rare events contribute significantly to the final outcome and their presence can not be neglected in the analysis [89]. The probability function of a random variable that has a heavy-tailed distribution must obey (6.4) .

$$\Pr\{Z > x\} \sim cx^{-\alpha}, \qquad x \to \infty \tag{6.4}$$

where $0 < \alpha < 2$ and c is a positive constant. It was shown in [73] that a set of sources with Pareto distribution (6.5) can be used to model self-similar and long-range dependent traffic.

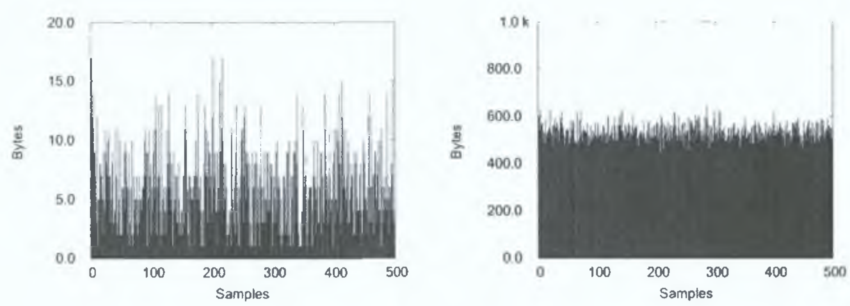
$$D(x) = 1 - \left(\frac{b}{x}\right)^{\alpha}, \qquad b \leq x \tag{6.5}$$

The inverse function $F^{-1}(u)$ used for generating random numbers with Pareto distribution is given in (6.6) .

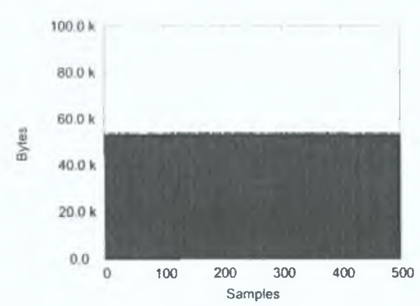
$$F^{-1}(u) = (1-u)^{-1/c} \tag{6.6}$$

Constant In this distribution the interarrival times are constant. This type of generator is used to model services that demand a Constant Bit Rate (CBR) such as some video and voice streaming applications [90,91].

The results of generator tests and sample traffic characteristics are presented in Figures 6.4, 6.5 and 6.6. It can be seen that Pareto source is the only one that shows self-similarity. Traffic from other sources, although very bursty for small timescales, becomes very even for larger timescales.

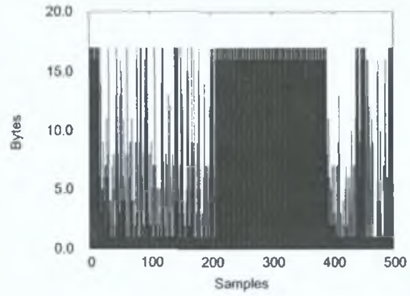


(a) sampling period 0.0001 s (b) sampling period 0.01 s

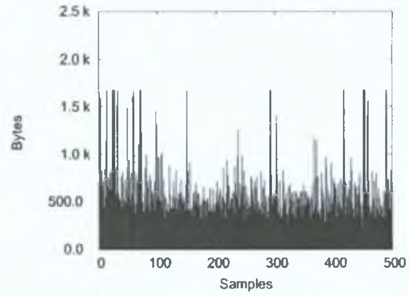


(c) sampling period 1 s

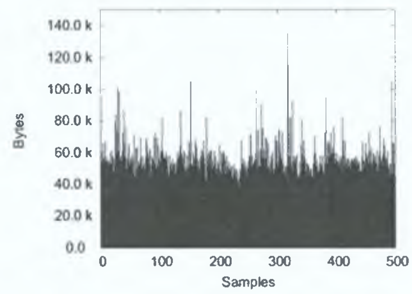
Figure 6.4: Generated Traffic with Exponential Interarrival Times.



(a) sampling period 0.0001 s

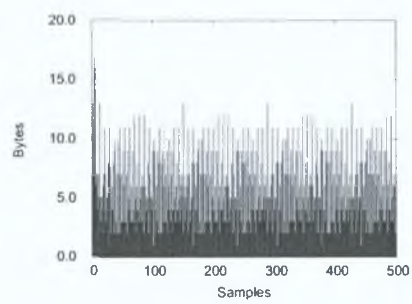


(b) sampling period 0.01 s

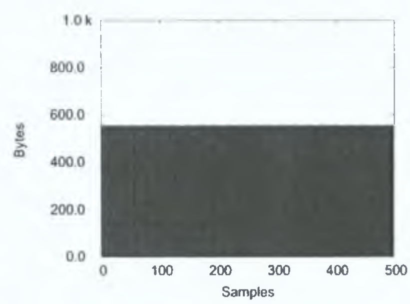


(c) sampling period 1 s

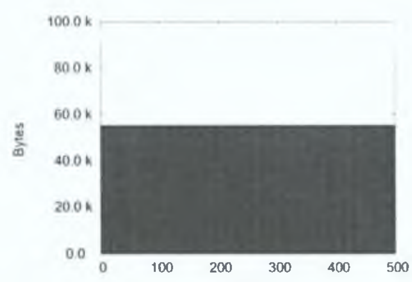
Figure 6.5: Generated Traffic with Pareto Interarrival Times.



(a) sampling period 0.0001 s



(b) sampling period 0.01 s



(c) sampling period 1 s

Figure 6.6: Generated Traffic with Constant Interarrival Times.

6.5 Simulator Validation

The validation of the simulator is an important part of the design [76, 77]. In this section the approach is described that was adopted to ensure that the general architecture of the EPON simulator is correct and gives accurate results.

As bandwidth assignment in a downstream direction in EPON is based on the TDMA method, it was decided to compare the output from the generator with a mathematical model of such a network given in [71]. Let M be the number of ONUs that share the same channel with rate R bits per second and let all ONUs send packets of length X. The times between packets arrival to the same ONU are independent of each other and have exponential distribution with mean λ in packets per second. Hence, they create a Poisson process. To simplify the analysis let assume that the length of a single transmission window is equal to the length of the packet. Thus, in any cycle an ONU can send only one packet. Based on this assumption, the length of the cycle is calculated from (6.7) .

$$C = M\frac{X}{R} \tag{6.7}$$

From analysis of the TDMA scheme it can be noted that there are three main factors contributing to the total delay of a packet in the system:

- The packet transmission time, which is equal to $\frac{X}{R}$.
- The waiting delay, which is the time that ONU spends waiting for its turn to send data. On average the waiting delay is equal to $\frac{C}{2}$.
- The queuing delay equal to the time a packet spends in a buffer.

To calculate the queueing delay the model of the TDMA network presented in Figure 6.7 was considered where every ONU was modeled as an independent and separate $M/D/1$ queue where λ is the mean arrival rate.

The average queueing delay in the TDMA system is given by (6.8) [71,86] where $\bar{x}$ is the average service time calculated from (6.9) .

$$W = \frac{\rho}{2(1-\rho)}\bar{x} \tag{6.8}$$

$$\bar{x} = \frac{MX}{R} \tag{6.9}$$

Hence, the average delay in the system can be presented as a sum of the transmission, waiting and queueing delay and calculated from (6.10) .

$$T = \frac{X}{R} + \frac{MX}{2R} + \frac{XM}{R}\frac{\rho}{2(1-\rho)} \tag{6.10}$$

To simulate such the TDMA network a following experiment was created:

- A number of ONUs were connected to the OLT in the tree topology.
- All ONUs operated at the same rate and are equally important.
- The packet size was fixed and equal to $\frac{C}{M}$, hence only one packet was sent per cycle.
- Packets inter arrival times were exponentially distributed.
- All other parameters native to the Ethernet architecture that could affect the validation process such as: guarding times, Interframe Gap (IFG) and link propagation time were set to zero.

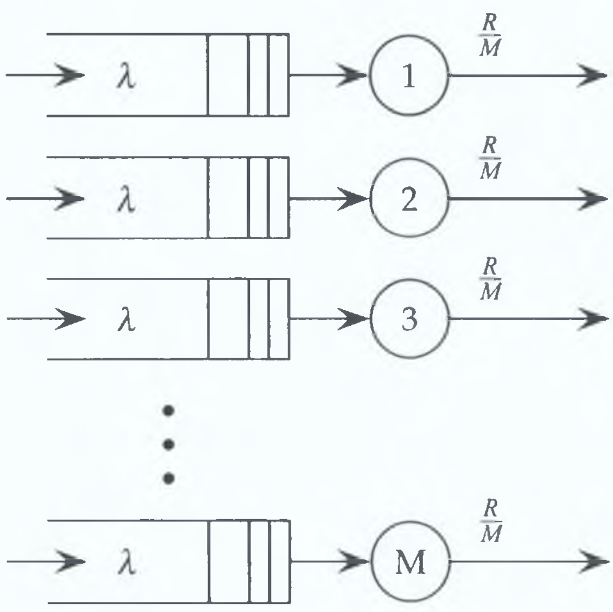
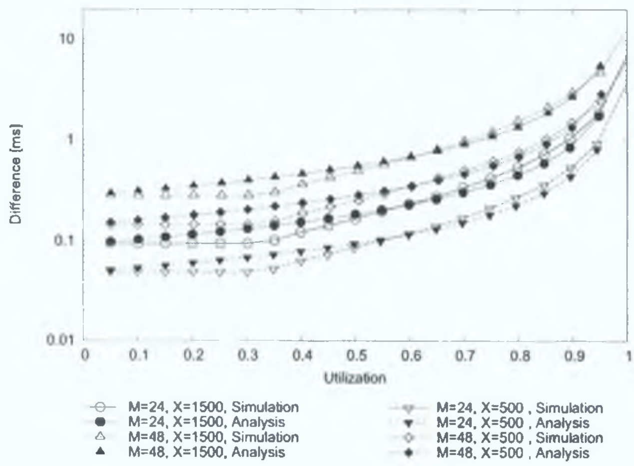
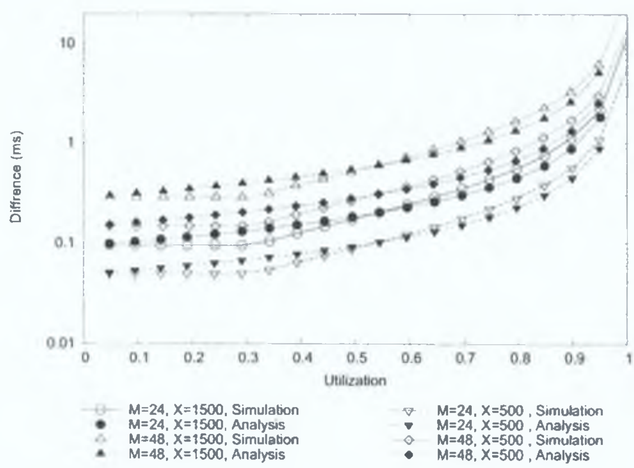


Figure 6.7: The TDMA model.

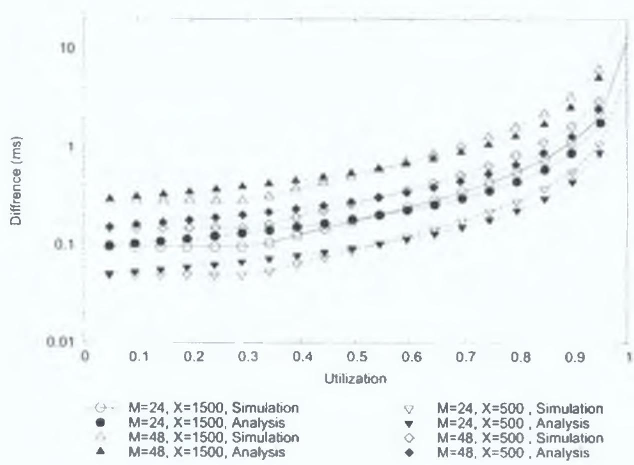
Comparison of simulation experiments and analytical results for various parameters of M and X are shown in Figure 6.8. From the diagrams it can be concluded that there is a good correlation between the model and simulation experiments. Small inaccuracies exist but this is to be expected as the generators used are only an approximation of the ideal system. Moreover, due to their functionality, the total load can only be estimated from the number of generated packets. As this number is highly random only coarse values can be obtained. Despite those problems, the biggest difference between theoretical and simulated values did not exceed 0.5 ms. In Figure 6.9, the differences in results for the simulation and model are shown. It can be observed that the error values follow a similar pattern for different experiments. Moreover, similar values are recorded. This allows an estimate of the error that is introduced by the simulator. In Figure 6.9, the relative error values are compared. It can be seen that, as in the previous comparison, the values do not show dependence on the setup of the experiments.



(a) simulation time 2 ms

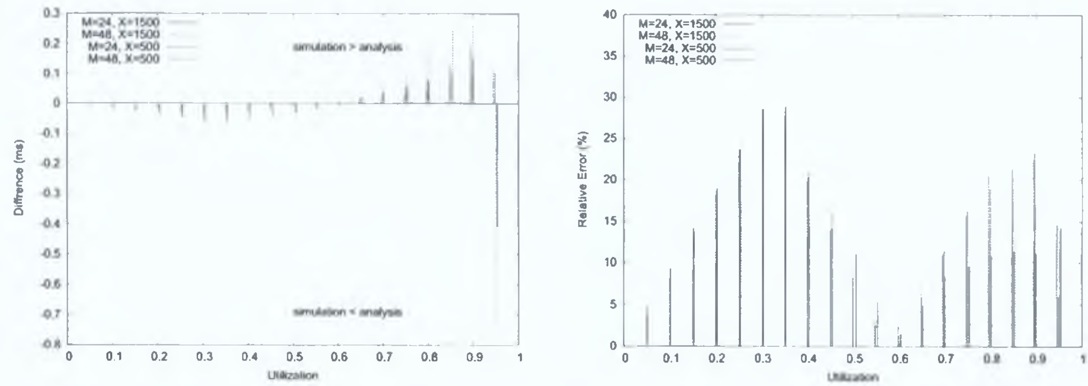


(b) simulation time 60 ms

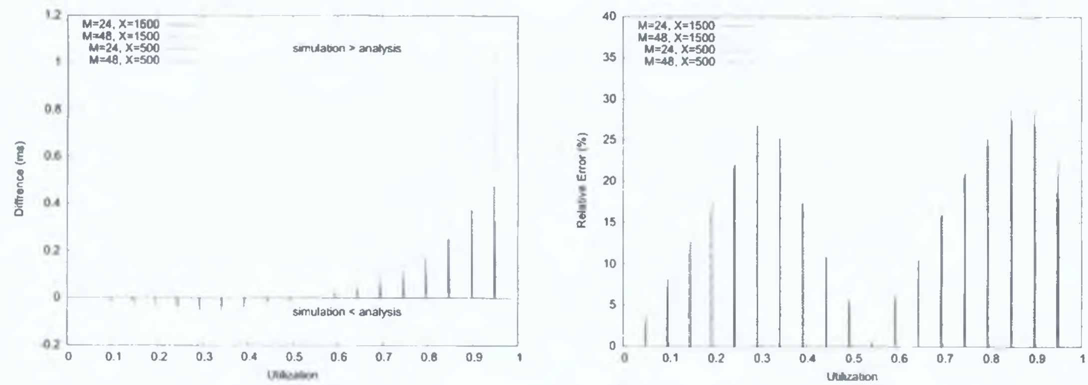


(c) simulation time 600 ms

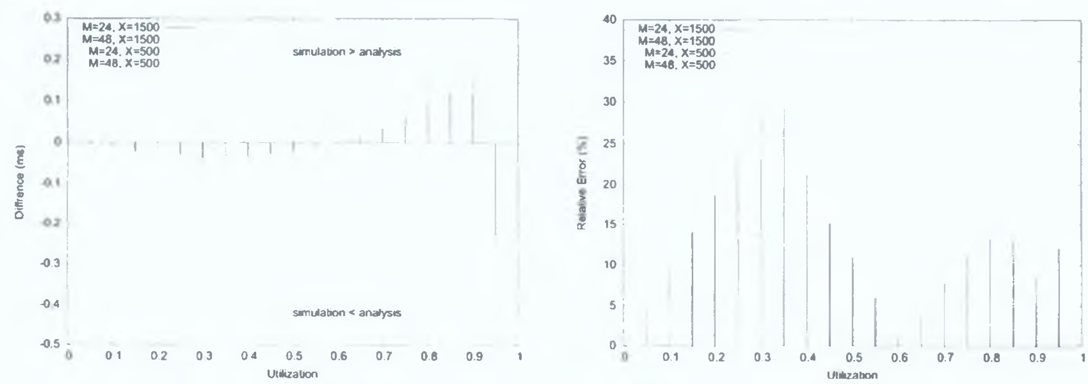
Figure 6.8: Results Validation for Different Simulation Times.



(a) simulation length 2 s



(b) simulation length 60 s



(c) simulation length 600 s

Figure 6.9: Difference and Relative Error Comparison.

6.6 Conclusion

The simulation method is best suited to measure the performance of the proposed centralized framework. Modelling the complex behavior of EPONs requires that the flow of all packets and timing between modules is depicted with great accuracy. In a situation like this discrete event simulation is the only choice.

The architecture of the simulator that was presented in this chapter was designed to achieve the best performance, hence the lack of a complicated user interface. As it is based on Standard Template Library the code can be recompiled and run on different platforms.

The random number generators are the heart of the simulation. In the approach presented in this thesis all possible precautions were taken to ensure that errors introduced by psudo-random sources are minimalized. This was achieved by incorporating a state of the art random "Mersenne Twister" generator designed by Makoto Matsumoto and Takuji Nishimura [84]. This generator offers superior performance to other well known approaches.

In every experiment it is crucial to validate the achieved results. The accuracy of results generated by the simulator was validated by comparison with the results obtained from the mathematical model of TDMA system. The figures presented in this chapter show that there is a good correlation between the simulator and the model. They indicate that the differences between recorded results do not depend on the experiments setup, neither they depend on the length of the simulation. This allows to calibrate the simulator to compensate for its imperfections.

Chapter 7

Simulation Experiments and Results

Evaluation of the ideas presented in this dissertation requires a number of carefully planned experiments. In this chapter these experiments are described in detail and the measured results are shown. The outcome of all experiments is discussed in detail and general conclusions are drawn.

7.1 Simulator Setup

7.1.1 General Architecture

During the simulation experiments, the network is connected in a tree topology, as this kind of setup is the most common and also the only one accepted in the standard [37]. The maximum link capacity was 1 Gbit/s, even though 1.25 Gbit/s is endorsed in the EPON specification. The difference is a result of adopting the 8b/10b channel encoding scheme, which decreases bandwidth utilization by 20 %.

The experiments were performed for a network of 16 ONUs. Every ONU had three separate queues with independent buffering space assigned to different classes of traffic, each with an independent buffering space. The maximum queue length was equal to 10 Mbits. The guarding interval between time slots allocated to different ONUs was equal to 1 μs in accord with the existing standard [37]. Interframe Gap (IFG) between Ethernet frames was used to model the flow of frames as closely as possible to the reality. The standard value of IFG for different architectures of Ethernet networks is equal to 96 bits [37, 92, 93]. A summary of experimental setup is shown in Table 7.1.

All algorithms were tested in normal conditions and in congested networks. In normal conditions the offered load was smaller than the maximum link capacity. During simulations of congested systems the offered load was significantly larger than the maximum link capacity.

Table 7.1: Simulation Experiments Parameters.

Parameter Name	Value
Link speed	1 Gbit/s
Network topology	tree
Number of ONUs	16
Number of queues per ONU	3
Inter-Frame Gap (IFG)	96 bits
Guarding time	1 μs

7.1.2 Traffic Models

In the simulation experiments it is important to model traffic accurately to guarantee that the obtained results can be applied to a real system. As the main goal was to measure the level of support of the centralized approach for classes of traffic with different QoS requirements, tests of the EPON system were performed for three different traffic mixtures. Traffic assigned to the high priority class was created from packets generated by sources with exponential or constant interarrival time distribution and its volume was generally smaller than the volume of medium or low priority traffic. Packets belonging to medium and low priority traffic were mainly generated by sources with Pareto distributed interarrival times, although other types of sources were also used. The proportions of high, medium and low priority traffic in the different Traffic Mixtures (TMs) are shown in Table 7.2.

To measure the influence of types of traffic sources on the performance of the proposed algorithms the experiments were carried out for three different scenarios, where the mixture of traffic in the network consisted of packets coming from a number of sources with different packet interarrival time distributions. The details regarding the type and number of sources in different Source Sets can be found in Table 7.3.

7.1.3 Simulation Time

Due to the complexity of the simulated system and the speed at which the link was operating the duration of simulations was fairly limited. In all experiments the simulated period of time was equal to 2 seconds. Over this interval, for loads equal to the maximum link capacity, the estimated number of generated packets reached 350,000. This made the simulation long enough to capture the behavior of tested algorithms and at the same a simulation could finish in an acceptable period. The network was initialized during first the 5 % of the simulation time to make sure that the impact of initial conditions on the final outcome could be neglected. In the last 5% of the simulation no new packets were generated to make sure that packets queued could leave the system.

Table 7.2: Traffic Mixtures.

Traffic Mixture	Class of Traffic		
	EF	AF	BE
TM 1	20%	40%	40%
TM 2	10%	60%	30%
TM 3	33%	33%	33%

Table 7.3: Traffic Sources Setup in Different Scenarios.

(a) Source Set I

Class of Service	Source Number	Source Type	Packet Length(bytes)
EF	10	Exponential	70
AF	5	Pareto	100
	5		500
	5		1000
	5		1500
BE	5	Pareto	100
	5		500
	5		1000
	5		1500

(b) Source Set II

Class of Service	Source Number	Source Type	Packet Length(bytes)
EF	10	Exponential	70
AF	5	CBR	100
	5		500
	5		1000
	5		1500
BE	5	CBR	100
	5		500
	5		1000
	5		1500

(c) Source Set III

Class of Service	Source Number	Source Type	Packet Length(bytes)
EF	10	Exponential	70
AF	5	Exponential	100
	5		500
	5		1000
	5		1500
BE	5	Exponential	100
	5		500
	5		1000
	5		1500

7.1.4 Service Level Agreements

Simulating the performance of SLA-DBA and A-DBA algorithms requires the use of SLAs which set the parameters describing the expected grade of service. The minimum and maximum values used here indicate the lower and upper limit of bandwidth that can be allocated to a given class. During the experiment these limits are calculated as a percent of the nominal rate set for a class of traffic in the TM used in the particular scenario. This ensured that TMs and SLAs were independent from each other.

Different SLAs were used to test the behavior of the SLA-DBA and A-DBA algorithms and to measure the impact of parameters agreed in the contract on the algorithm's performance. In Table 7.4, the SLAs used during the simulation experiments are presented.

Table 7.4: SLA Parameters for SLA-DBA and A-DBA Algorithms.

	SLA-DBA								
	SLA 1			SLA 2			SLA 3		
	EF	AF	BE	EF	AF	BE	EF	AF	BE
min	100%	95%	–	80%	80%	–	50%	50%	–
max	110%	100%	–	100%	100%	–	100%	100%	–
	A-DBA								
max	130%	130%	40%	110%	110%	40%	100%	100%	100%

7.2 The Effects of Adaptive and Constant Polling Cycle Time

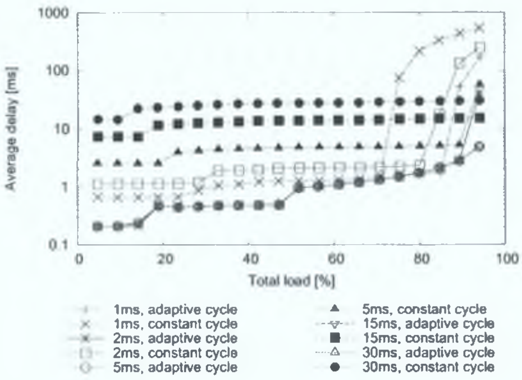
The length of the polling cycle is one of the most important factors influencing the overall performance of the EPON system. In this thesis Cyclic Polling with Adaptive Cycle Length scheme is proposed. This scheme offers functionality which is a compromise between the constant polling cycle proposed in [53] and the IPACT algorithm proposed in [41]. In this section the effects that the polling time has on the recorded average delays and queue lengths are compared. The values used during this experiment are shown in Table 7.5. The results of the simulation experiments

Table 7.5: The Setup for the Cycle Comparison Experiment.

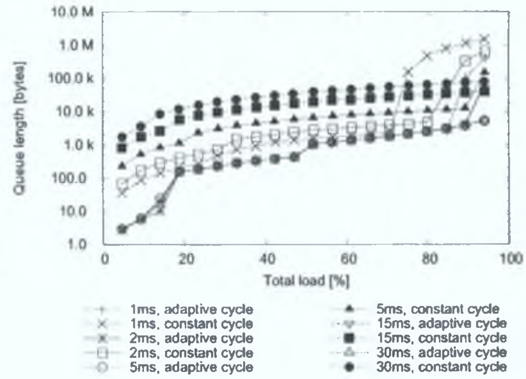
Parameter Description	Value
Algorithm	P-DBA
The maximum cycle length	2 ms, 5 ms, 15 ms, 30 ms
Traffic Mix	TM3
Traffic Type 1	all classes assigned 5 sources with exponentially distributed interarrival times; the packet size is 500 bytes
Traffic Type 2	all classes assigned 5 sources with Pareto distributed interarrival times; the packet size is 500 bytes
Traffic Type 3	all classes assigned 5 sources with constant interarrival times; the packet size is 500 bytes

for Traffic Type 1 are shown in Figure 7.1. Results for other traffic types are shown in Appendix A. It can be observed from the results presented that the average delays recorded for the adaptive polling cycle scheme are generally lower in comparison to the results for the constant cycle. The difference becomes more obvious when the value of the maximum length of the polling cycle grows. An interesting effect can be observed where as the offered loads became heavy there was a sudden and significant increase in the average delays is observed for all cases. It can also be noted that the recorded maximum delay for loads reaching the maximum link capacity decreased as the maximum length of the polling cycle was extended.

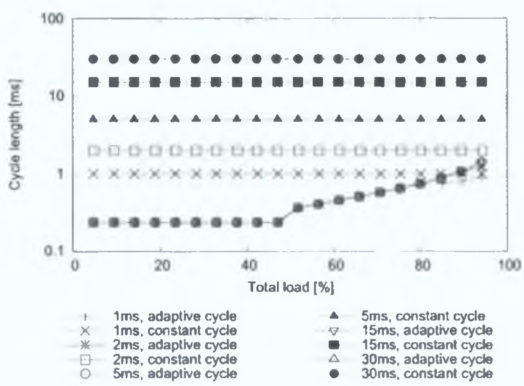
The explanation for this behavior lies in the bursty arrival of packets to queues. In the adaptive approach, the length of the cycle is calculated based on the information received in the REPORT messages from ONUs. It is possible that in a given period of time the calculated length would be longer than the maximum. When the length of the cycle is restricted, the amount of allocated bandwidth is smaller



(a) Average delay



(b) Average queue length



(c) Average cycle length

Figure 7.1: The Comparison of Cyclic Polling with Adaptive and Constant Cycle Time. Traffic Type 1.

than requested and superfluous packets have to be buffered until the next cycle. For small values of maximum cycle length it is more likely that the total amount of data is larger than the available bandwidth. Hence, its effect is stronger and manifests itself even for medium loads. As the cycle length lengthens there is more room for the temporary expansion of the polling cycle and more packets can be sent in a single cycle. Since the delay introduced by the polling cycle contributes significantly to the total delay it is better to extend the cycle and accommodate the burst than to have a short cycle length and defer packets until the next cycle.

In the scheme with constant cycle a similar observations can be made. The situation here is even worse as the length of the cycle is fixed and there is no room for expansion. Hence, for small cycle lengths traffic is penalized more severely. As the length of the cycle increases, the effect was less visible. The packets are arriving over a longer period of time and the temporary variations do not have as big an impact on the decision about bandwidth allocation as for shorter cycle lengths.

Another issue that has to be considered in EPON systems is the dependency between the length of the cycle and the size of the control channel. This problem is especially visible in networks where the traffic patterns are asymmetrical (i.e. the downstream traffic is much larger than in the upstream direction.) In such networks the average delays will be low because the volume of traffic is small. As mentioned in Section 3.4.1, the results show that the execution of the scheduler happens more frequently and proportionally more bandwidth must be allocated to GATE messages. At the same time the volume of the downstream traffic is large and provisioning bandwidth to the control channel restricts the customers' traffic. A number of conclusions can be drawn from this scenario:

- It is hard to find an optimal solution as there is a tradeoff between average delay and the size of the control channel. Since it is difficult to find one perfect value that fits in all scenarios the length of the polling cycle should be dependent on current traffic conditions.
- The OLT should be able to control the minimum and maximum length of the polling cycle. In the situation described above it would be able for the OLT to increase the minimum cycle length and less bandwidth would be needed for the control channel. When the volume of traffic in the upstream direction is small, the decrease in QoS will be small.
- The mechanism responsible for adjustments to the cycle length should be independent from the bandwidth allocation algorithm.

The Cyclic Polling with Adaptive Cycle Length scheme proposed in this thesis not only gives lower average delays in comparison with the fixed polling cycle scheme but also permits adjustments to the maximum and minimum cycle length. Also, the functionality of all algorithms introduced in this thesis is independent from the cycle length and this gives freedom to the OLT to adapt the cycle length to current conditions.

Based on results shown in Figure 7.1 and drawn conclusions it was decided that for the clarity of the presentation only results of experiments for the Cyclic Polling with Adaptive Cycle Time scheme with the maximum cycle length equal to 5 ms will be presented in this thesis. As seen in Figure 7.1 for this value the maximum delays are two orders of magnitude smaller than for 1 ms long polling cycle. Further extending the length of the polling cycle did not improve the overall performance.

7.3 Grant Multiplexing Scheme

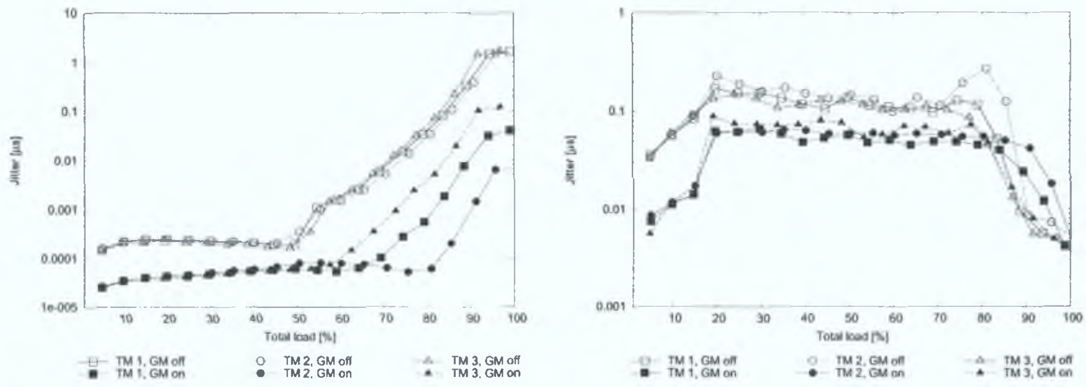
The Grant Multiplexing scheme was introduced in this thesis as a way to improve QoS offered to jitter sensitive services, such as real time voice or video. The general mechanism of this scheme was outlined in Section 4.4. Here, the results of simulation experiments are presented.

The jitter was measured calculated according to (7.1), where t_0 is the beginning of cycle n and $t_{\beta_j(i)}$ is the beginning of the transmission window allocated to ONU i and queue j.

$$S_N^2 = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2 \qquad \text{where} \qquad x_n = t_0 - t_{\beta_j(i)} \tag{7.1}$$

The results obtained for various proportions of high, medium and low priority traffic are presented in Figure 7.2 and also in Appendix B.

The recorded results show a strong correlation with previous expectations and it can be seen that the Grant Multiplexing scheme significantly reduces the jitter and some significant gains are achieved where schemes with both the adaptive or constant cycle length are used. The biggest improvement in a scenario where adaptive polling cycle was used (Figure 7.2(a)), was recorded for Traffic Type 2, where medium and low priority traffic was dominant. This again was as expected. As



(a) adaptive cycle length (b) constant cycle length

Figure 7.2: Jitter Comparison. Source Set I.

outlined in Section 4.4, the scheme was meant to reduce the impact of other bursty traffic on the QoS received by privileged classes by moving the slots assigned to them to the beginning of the cycle. Hence, as the most bandwidth is dedicated to classes carrying bursty traffic, the largest gains were recorded for Traffic Type 2.

In a situation where the length of the polling cycle is constant (Figure 7.2(b)) the gains are generally smaller than in the previous case. It can be seen than the biggest gains were recorded for the small and medium loads. For large loads, all classes are requesting their maximum and the variation in the size of the assigned transmission windows is small, resulting in smaller jitter.

The negative side of the Grant Multiplexing scheme is that more bandwidth is dedicated to the guarding intervals between time slots allocated to different ONUs. The amount of bandwidth that has to be sacrificed is dependent on the cycle length; the longer the cycle the smaller the proportion of bandwidth is lost. The comparison between bandwidth dedicated to the guarding bands and cycle length is shown in Table 7.6. It is clear that the longer the cycle length the proportionally less bandwidth has to be dedicated to the guarding gaps.

The application of the Grant Multiplexing scheme in real situations should of course be based on a thorough analysis of the existing traffic patterns. In some situations this scheme will not be necessary as the gains will not justify worsened

Table 7.6: Size of Guarding Gaps in Grant Multiplexing Scheme.

Cycle length	Bandwidth Dedicated to Guarding Gaps	
	GM on	GM off
2 ms	1.60 %	0.80 %
5 ms	0.64 %	0.32 %
15 ms	0.21 %	0.11 %

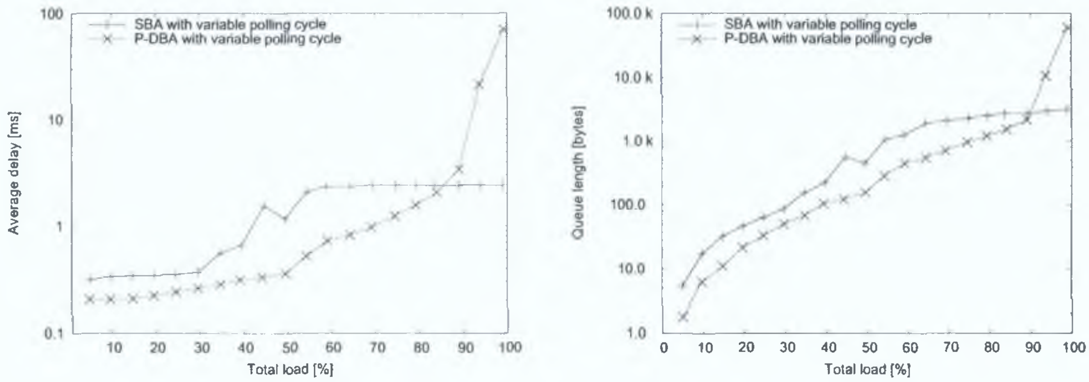
efficiency. But the decision about the application of Grant Multiplexing should be made by the network provider. In EPONs supporting the proposed GATE message, introducing this mechanism is easy and it does not involve any cooperation from ONUs, leaving them unaware of the whole process.

7.4 The Performance of SBA and P-DBA Algorithms

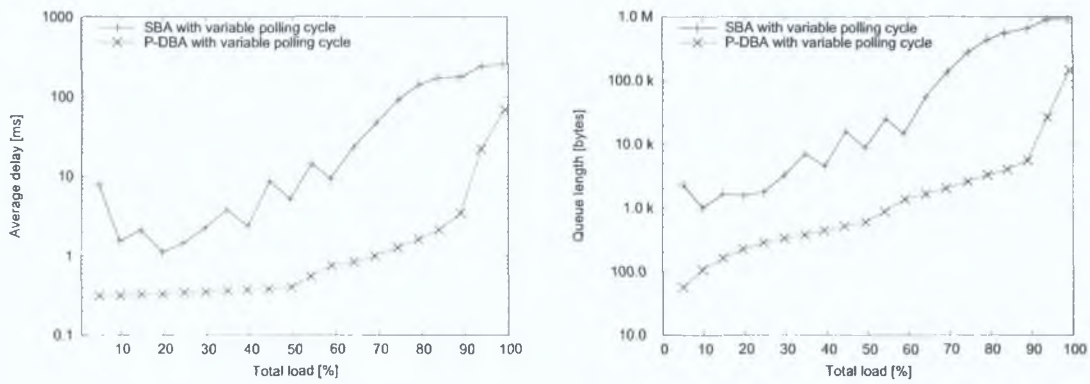
The SBA and P-DBA algorithms are two extremes of bandwidth allocation. In SBA, the bandwidth is allocated based on some "off-line" agreement and the process does not take into account REPORT messages received from ONUs. Conversely, in P-DBA, the decision about allocated bandwidth is based on information about the current queue length thus the size of the assigned transmission window can radically change from one cycle to another. Good understanding of SBA and P-DBA algorithms behavior and their performance in EPONs lays the foundation on top of which the improved bandwidth assignment methods can be built. Simulation experiments were designed to show the behavior of the SBA and P-DBA under variable traffic conditions and for different traffic mixtures.

7.4.1 Normal Traffic Conditions

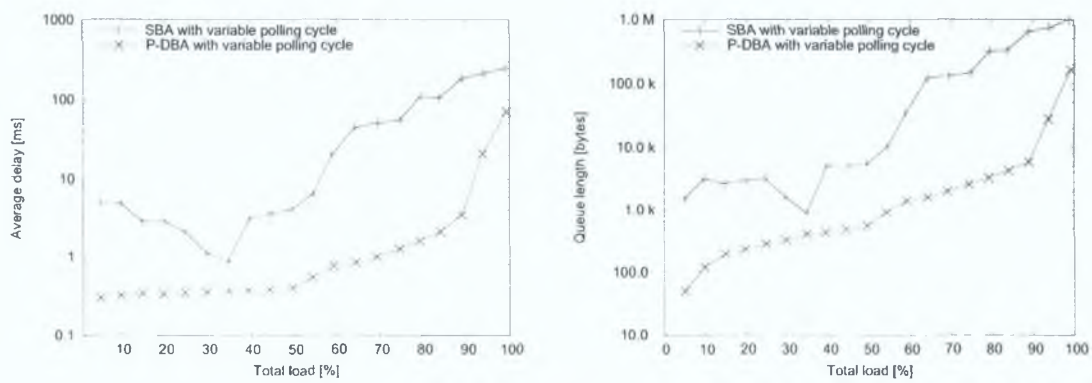
In a situation where the EPON is not congested, SBA and P-DBA show very different performance in terms of average delay and queue length. The comparison of the results recorded for these two algorithms is shown in Figures 7.3 and 7.4.



(a) Expedited Forwarding (EF)

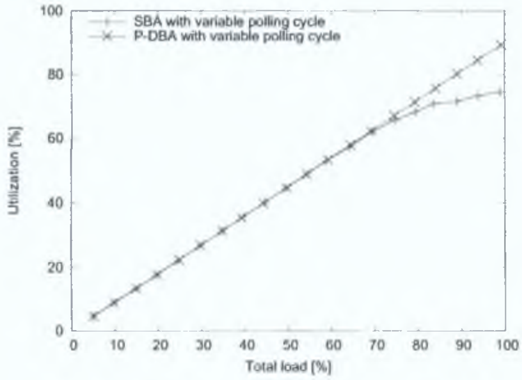


(b) AF

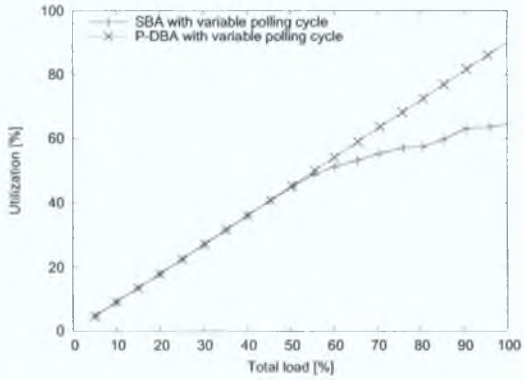


(c) BE

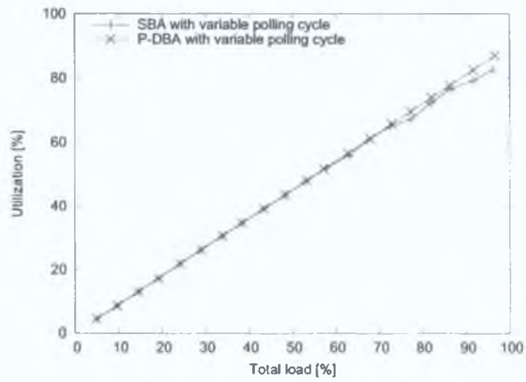
Figure 7.3: Algorithm Performance Comparison. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) TM 1



(b) TM 2



(c) TM 3

Figure 7.4: Throughput Comparison. Maximum Cycle Length 5 ms.

The SBA Algorithm

The general performance of the SBA algorithm depends on the amount of bandwidth assigned to a given class of traffic or ONU. As long this value is larger than the requested amount of bandwidth the average delays will be small. On the other hand, if at some point the requested bandwidth is larger than the allocated bandwidth the average packet delay will start growing very fast. This mechanism can be understood better when the results from different experiments are compared. From Figure 7.3(a), it is clearly seen that the EF class was allocated enough bandwidth to transmit all the buffered packets, hence the recorded average delays were small and constant for increasing loads. But if this situation is compared with the results obtained for TM 3, it is obvious that in this case the EF class of traffic experiences much worse QoS. Here, the allocated time slots were too short to empty the queues. As more EF traffic entered the system, the size of the allocated transmission window did not change and resulting increasing average delays. The experiments performed for other traffic sources showed that, although the average delays recorded in particular cases vary considerably (compare results for Source Set I and II), the general characteristics of the SBA algorithm remains the same.

In Figures 7.3(b) and 7.3(c), the so called *"light load"* phenomenon can be noticed for the SBA algorithm, where average delays recorded for the small loads are much larger than delays for the medium loads. The recorded results show that for 10% of link utilization the average delays can be 10 ms longer than for 30% of link utilization. This paradox was first described by Kramer et al. in [50]. In the SBA algorithm with a variable polling cycle, the calculation of the total length of the cycle is based on the length of the queues reported by ONUs. When the load in the network is small, the majority of the queues are empty and the cycle length is short. But, with the static allocation, the queues are assigned some bandwidth even though they did not request any and the calculated length of the time slot is divided between all queues according to some arbitrary policy. In the case where a large packet is at the head of a queue, the assigned time slot might not be long enough to send it in one fragment and a particular queue is blocked until the size of the allocated transmission window is large enough to accommodate the obstructing packet. As the load increases, so does the length of the cycle; the allocated time slots are proportionally longer and obstruction is less likely.

The P-DBA Algorithm

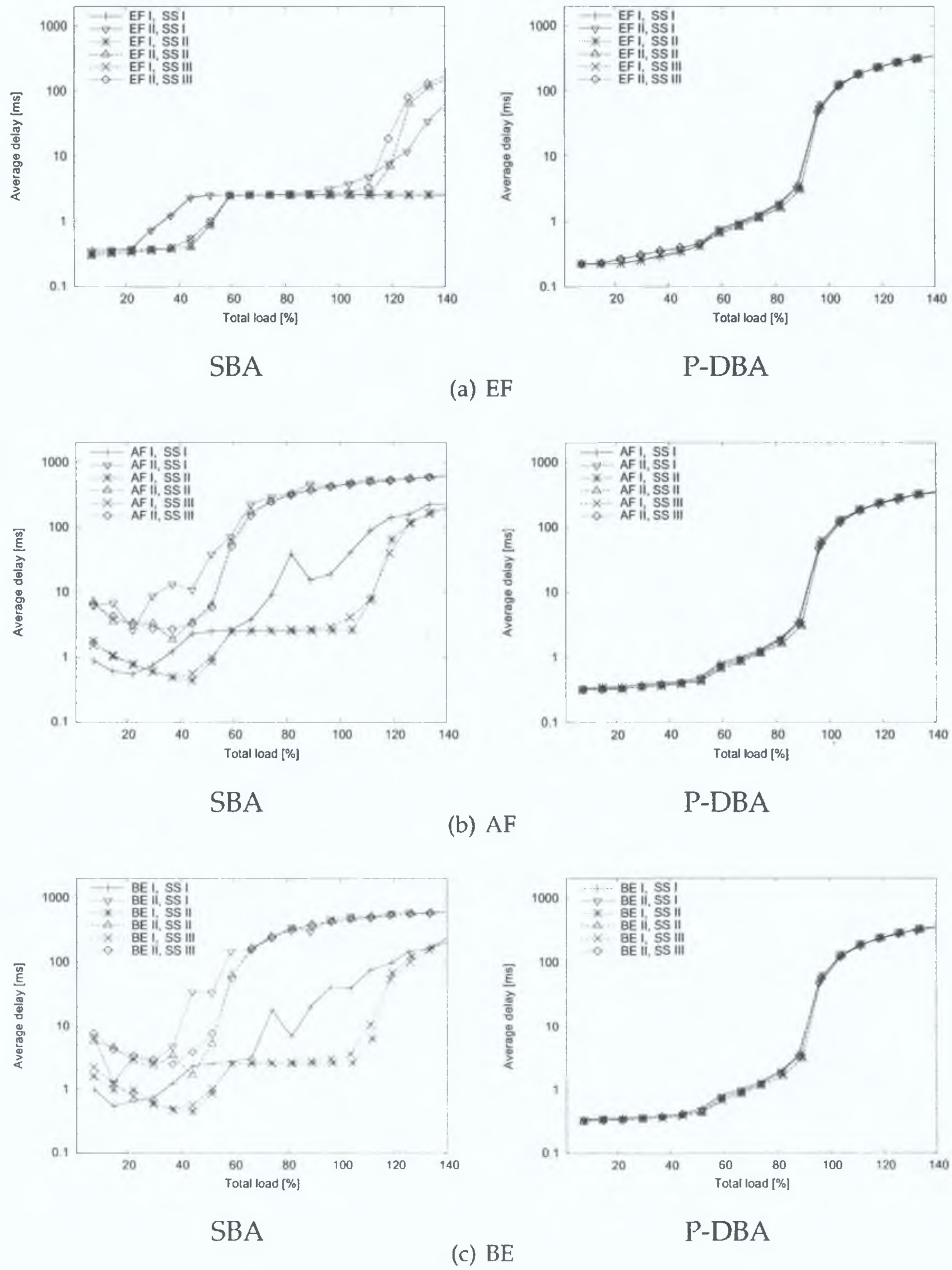
The P-DBA algorithm allocates bandwidth in a dynamic manner and the size of a time slot assigned to a particular queue can vary considerably between cycles. This approach is especially suitable for medium and low priority data-oriented services, as they are generally bursty in nature and the volume of data traffic is much bigger. The results concur with this viewpoint. P-DBA with a constant polling cycle gives larger values of average delays in the EF class of traffic when compared to the SBA algorithm in the same configuration. Conversely, for Assured Forwarding (AF) and BE classes of traffic the recorded delays are much lower. This is because the bandwidth is assigned proportionally to the reported queue length, which means that bandwidth is utilized more efficiently in comparison with the SBA algorithm and if one class requests less bandwidth, the other can be allocated more to compensate for a sudden burst. The negative side of this approach is that all classes are treated equally and worse QoS is delivered to high priority traffic in comparison with the SBA algorithm. Conversely, the functionality of the P-DBA algorithm does not permit the *"light load"* phenomenon to happen, which results in much smaller delays recorded for the medium and large loads.

7.4.2 Algorithms' Performance in Congested Network

To measure the performance of the algorithms in congested networks an experiment was designed in which half of the sources were transmitting their data two times faster than the nominal rate. This resulted in the network being heavily congested. The performance of the SBA and P-DBA algorithms was measured in these conditions. The sample results depicting the properties of both algorithms are shown in Figure 7.5 and 7.6. More results can be found in Appendix C.

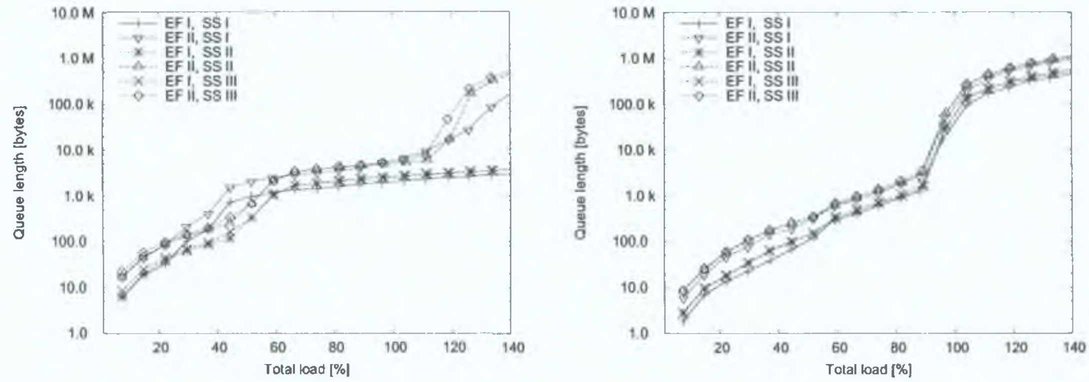
Comparison of these two algorithms reveals that indeed, the P-DBA algorithm does not offer any guarantees. It is clearly seen that all classes of traffic receive the same QoS despite their transmission rate.

In contrast the SBA algorithm is able to offer guarantees. As the amount of bandwidth assigned to a queue is fixed, traffic from sources transmitting at higher rates does not have an impact on the QoS received by classes of traffic from other sources. In comparison with the P-DBA scheme, the SBA algorithm, although less efficient from a bandwidth utilization point of view, guarantees that classes will receive the QoS that was agreed in the SLAs.



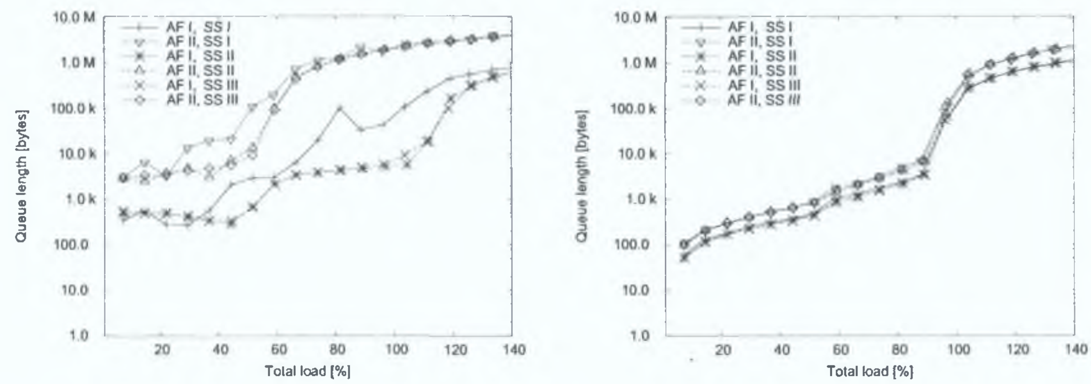
I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure 7.5: Average Delays Comparison. Congested Network. Traffic Mix 1.



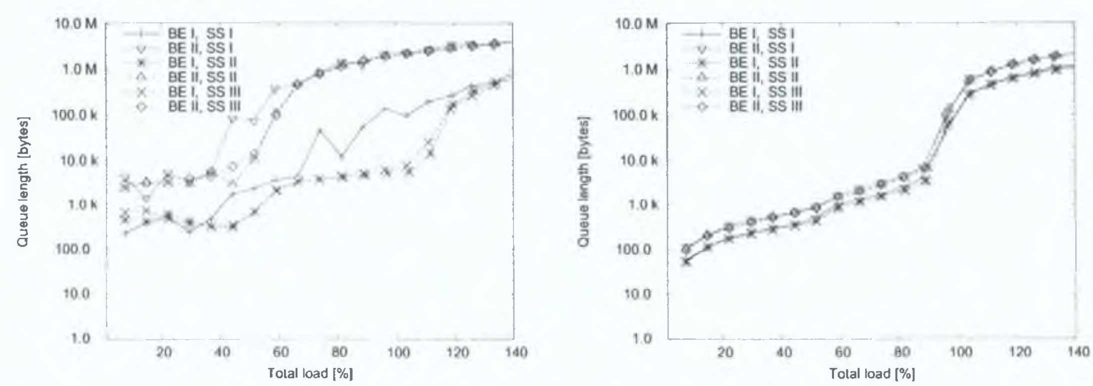
SBA P-DBA

(a) EF



SBA P-DBA

(b) AF



SBA P-DBA

(c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure 7.6: Average Queue Length Comparison. Congested Network. Traffic Mix 1.

7.4.3 Summary

In the summary of this section it must be noted that neither the P-DBA nor SBA algorithms are perfect but both of them have features that are desirable in EPONs capable of supporting the DiffServ architecture. The strong point of the P-DBA algorithm is very efficient bandwidth utilization and it clearly outperforms its counterpart. The proportional bandwidth assignment also prevents the *"light load"* phenomenon from happening.

On the negative side, it turns out to be the main obstacle in implementing this algorithm in DiffServ networks where QoS agreed in SLAs must be guaranteed. As shown by the results of the experiments with the heavily congested system, the P-DBA algorithm does not offer any assurances as no distinction is made between the good sources which transmitted at their nominal rate and the bad ones which were violating the agreement.

The situation is different in the SBA scheme. Due to a fixed bandwidth assignment, the QoS delivered to a source of traffic is not affected by the volume of traffic sent by other sources. This guarantees that the classes will always receive their promised service and from this viewpoint the SBA algorithm is much better suited for DiffServ-aware EPONs. The fixed bandwidth assignment also has its negative sides. In comparison with the P-DBA algorithm the utilization of available resources is worse and the *"light load"* phenomenon has a big impact on delays recorded for small loads.

7.5 The Performance of the SP-DBA Algorithm

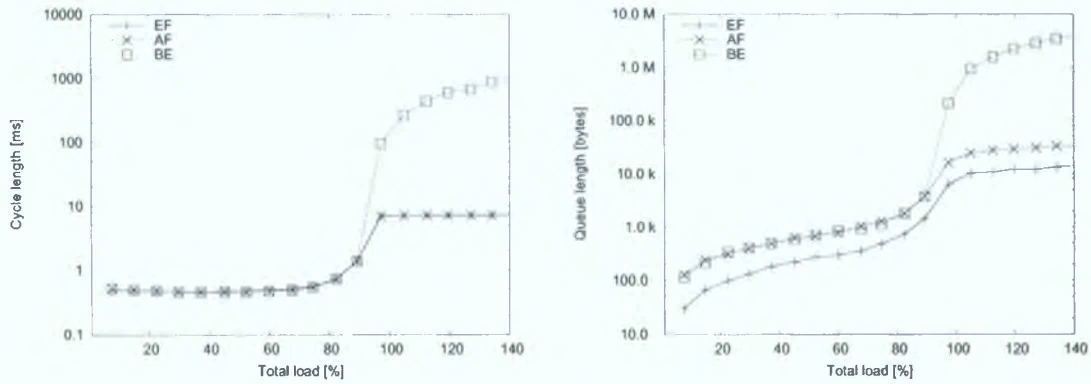
Priority scheduling is a very popular concept and different variations of this approach have been implemented in IP networks. Priority scheduling was also introduced as an admission control mechanism in EPONs [49, 50, 54–56]. To show that both inter- and intra-ONU scheduling can be done in the OLT it was decided to implement the centralized version of the strict priority scheduler, described in Section 5.3. In this section, the behavior of the SP-DBA algorithm is analyzed.

To show the positive and negative aspects of the SP-DBA algorithm an experiment was designed in which the transmission rate of all ONUs was gradually increased up to 140 % of their nominal rate. This resulted in a situation where the network was severely congested. It was anticipated that the low priority classes would be serviced with a worsening QoS as the load increased.

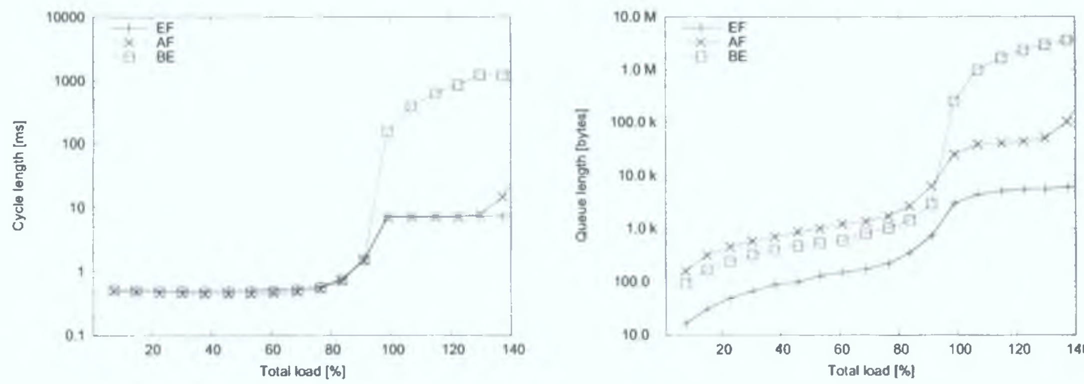
The results for Source Set I with a maximum cycle length of 5 ms are presented in Figure 7.7. The results for other types of sources are presented in Appendix D. It can be seen that the general performance of SP-DBA is inline with expectations. As the total load increases, the average delays get progressively longer. When the volume of the load is equal to the link capacity, the QoS offered to the BE class of traffic is sacrificed and the average delays recorded for this class grow rapidly. On the other hand, the EF and AF classes of traffic do not experience worse QoS. This situation lasts for as long as the volume of traffic in the EF and AF classes is smaller than the link capacity. When the combined traffic in the EF and AF classes becomes larger than the link capacity, the AF class starts getting less bandwidth and the delays recorded for medium priority packets are longer. This situation occurs when the volume of the total load is around 130 % for TM 2 and 3. For TM 1, where the BE class of traffic constituted 40 % of the total load, this effect is only visible when the volume of traffic is larger.

It has to be mentioned that only two seconds of real time were simulated in this experiment. If the simulation time were longer, the recorded average delays for the BE and later for the AF class of traffic would also be larger as these queues would not be allowed any bandwidth and packets belonging to these classes would be buffered for the whole duration of the simulation.

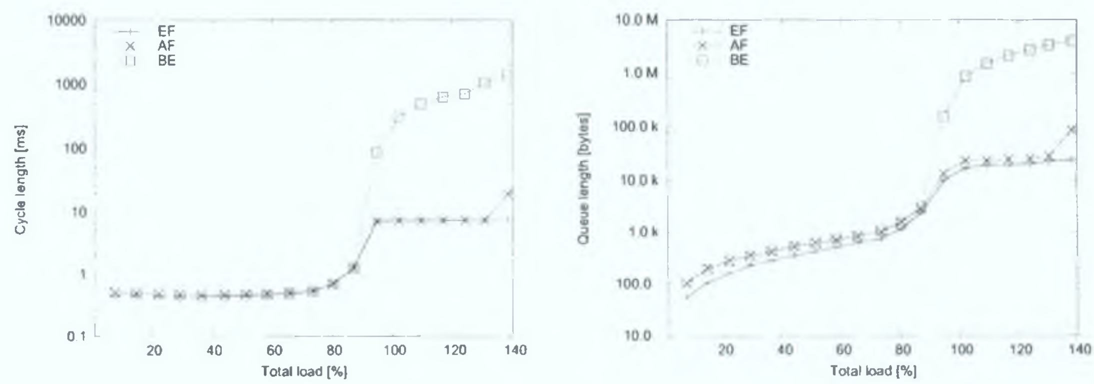
Two things are obvious in the performance of the SP-DBA algorithm. Firstly, the traffic with the highest priority will always receive the best QoS. From this point of view such an approach is best for classes with very strong requirements. Secondly, it is hard to provide any guarantees regarding traffic properties, as the only



(a) TM 1



(b) TM 2



(c) TM 3

Figure 7.7: SP-DBA Performance Comparison. Source Set I. Maximum Cycle Length 5 ms.

thing that is known precisely is the order in which the classes are served and it is hard to predict the amount of bandwidth allocated to every queue. This property of the SP-DBA algorithm makes it hard to create sensible SLAs, as the minimum and the maximum amount of bandwidth cannot be determined with sufficient accuracy. Moreover, in a situation where one class of traffic is sending more data than it should, the QoS of all classes with lower priorities will be affected and no mechanism can be provided to penalize packets from the non-compliant queue.

Although the performance of the SP-DBA algorithm in the centralized and distributed mode is virtually the same, the former has a certain advantage as updating the parameters affecting the behavior of the algorithm can be done at one point. Hence, there is no need for a specialized protocol to control the ONUs which is the case in the distributed approach. In the long term, ONUs with a simpler, more generic and robust architecture can be designed.

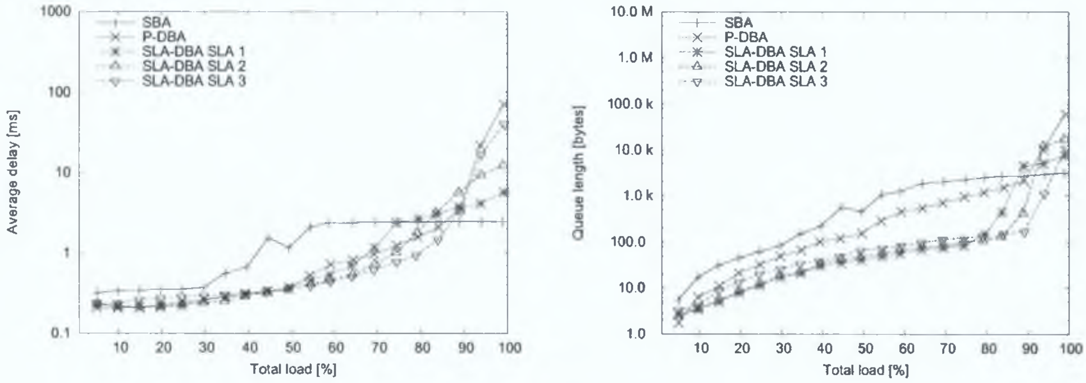
7.6 The SLA-DBA Algorithm

The functionality of the SLA-DBA algorithm has been already presented in Section 5.4. The algorithm is based on the P-DBA approach because it achieved the best bandwidth utilization as shown in Section 7.4.1. Here, the results of simulation experiments are shown. The experiments were performed for different types of traffic. As earlier, only algorithms with a variable polling cycle were tested.

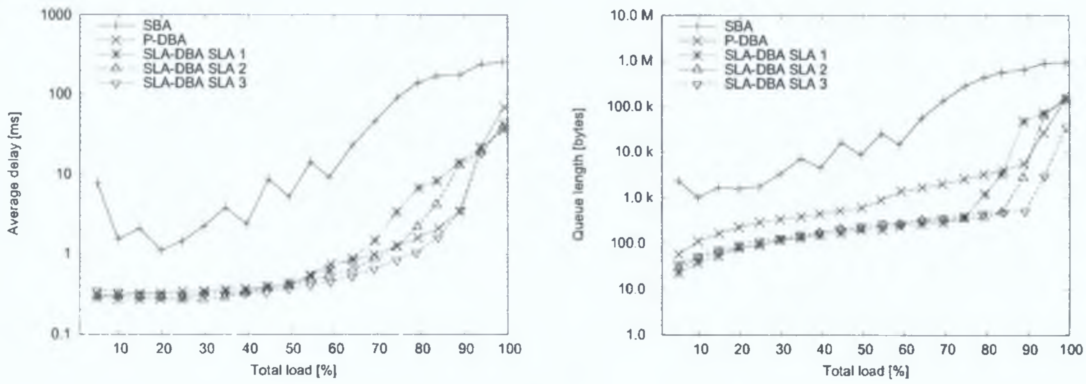
7.6.1 Performance in Normal Conditions

In this section the results of the experiments are shown in a situation where the total load seldom exceeds the total link capacity. An example of the performance of the SLA-DBA for Source Set I is shown in Figures 7.8, 7.9 and 7.10.

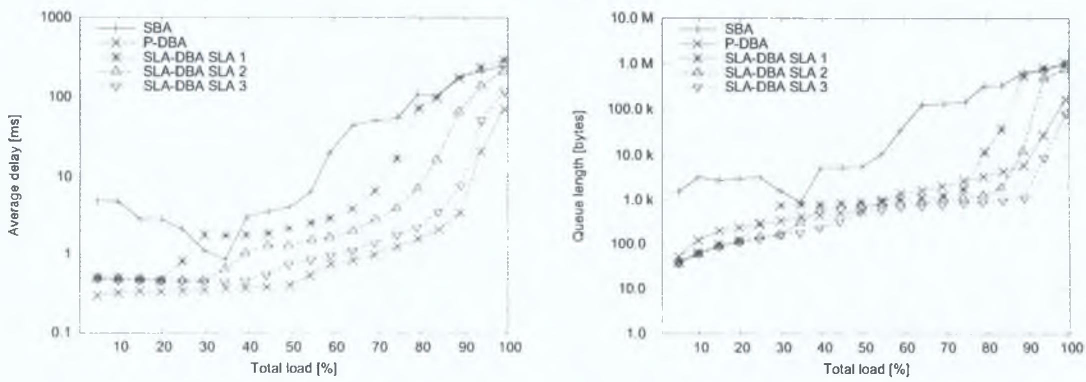
The results of experiments for different types of sources and traffic mixes are included in Appendix E. It is clearly seen that the general trend in the performance of the SLA-DBA algorithm does not differ much from the performance of the P-DBA



(a) EF

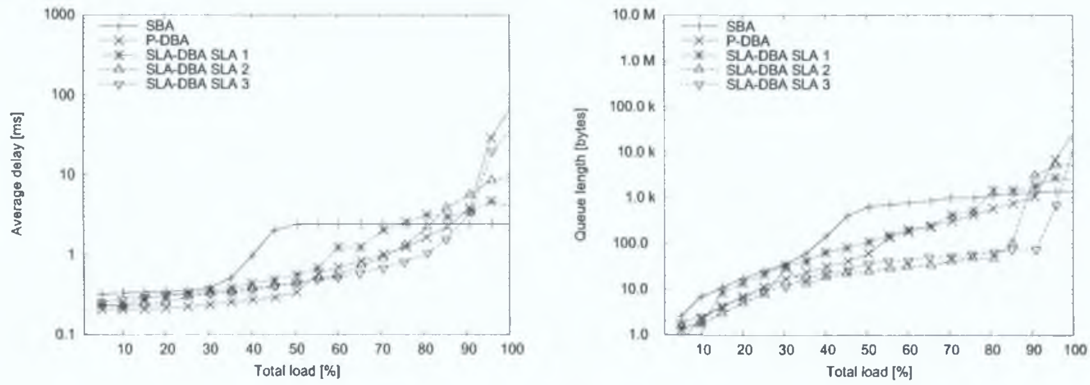


(b) AF

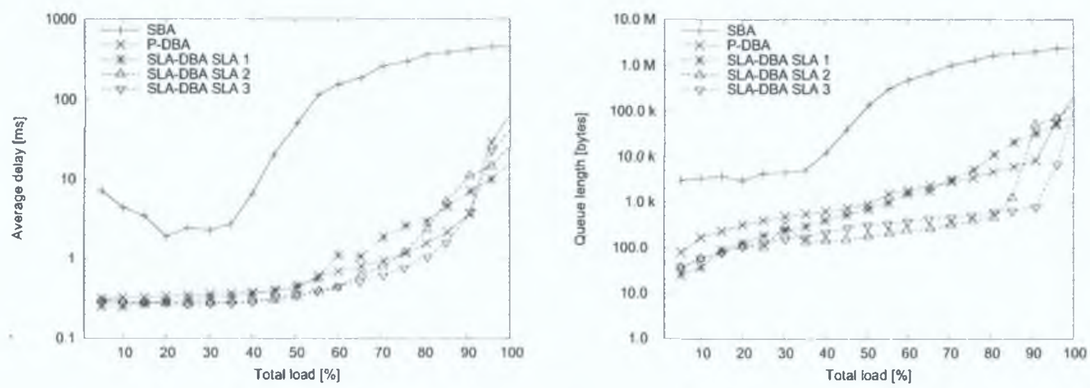


(c) BE

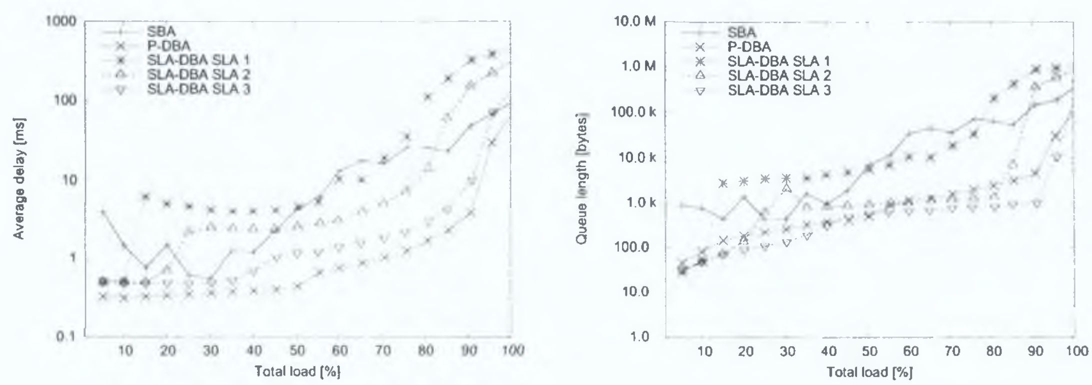
Figure 7.8: Algorithm Performance Comparison. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) EF

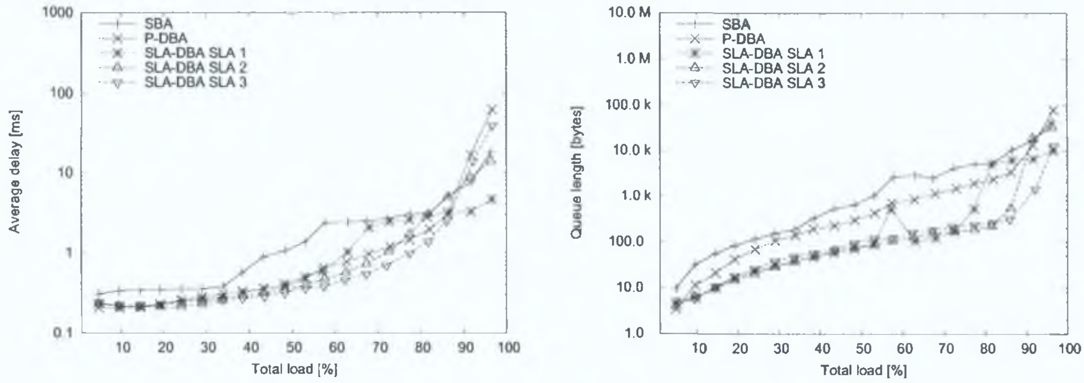


(b) AF

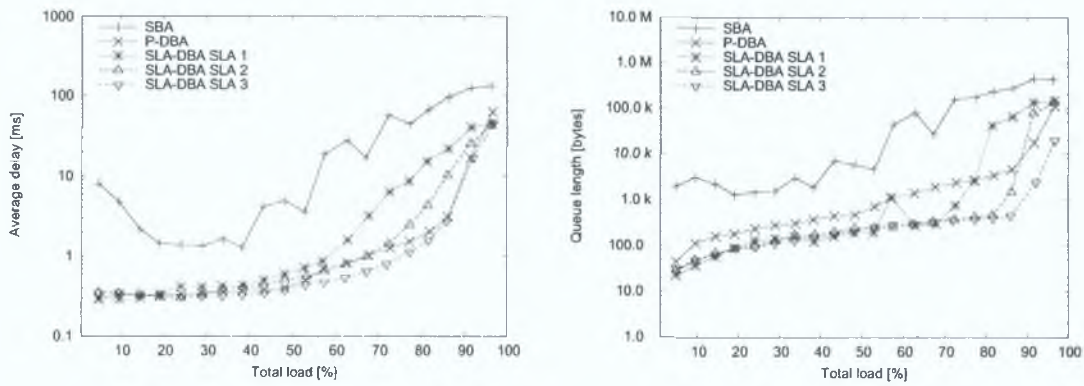


(c) BE

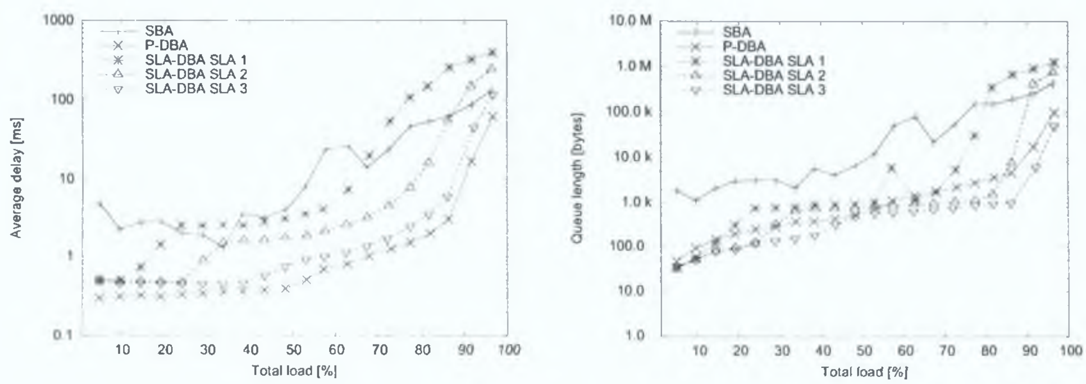
Figure 7.9: Algorithm Performance Comparison. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



(a) EF



(b) AF



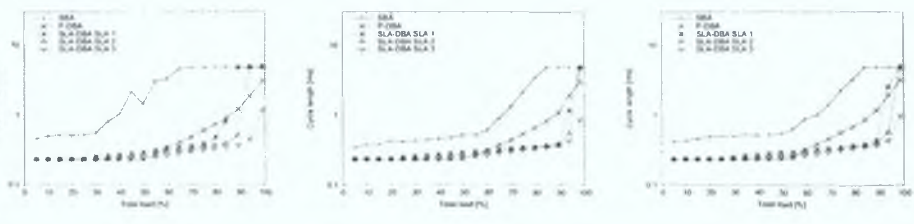
(c) BE

Figure 7.10: Algorithm Performance Comparison. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.

algorithm. The EF and AF classes of traffic receive better QoS for high loads at the expense of worse QoS offered to the BE traffic. If the effects of the different the SLAs are compared, it can be said that for the small and medium loads, there is only a small difference. SLA-DBA under SLA 1 has worse performance for loads between 70 % and 85 % of the total link capacity, but for larger loads the recorded delays are smaller than for SLA-DBA with SLA 2 or SLA 3. The opposite can be said about the results recorded for SLA-DBA and SLA 3. It can be concluded from presented diagrams that numbers measured for SLA-DBA with SLA 2 can be placed between the results recorded for SLA-DBA with SLA 1 and SLA 3. This is not a coincidence and it reflects the way the parameters were set in the SLAs in Table 7.4.

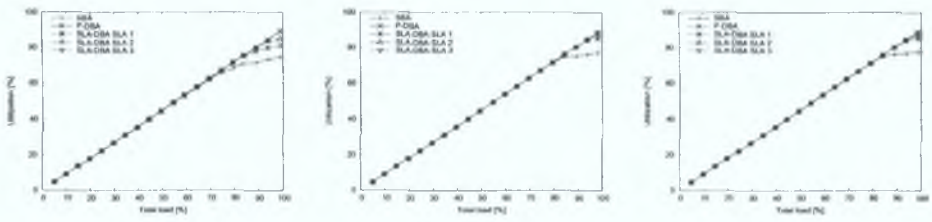
If the performance for the BE traffic is compared, it can be seen that the average delays are the largest for SLA-DBA with SLA 1 and the smallest for SLA-DBA with SLA 3. Again, the performance for SLA-DBA and SLA 2 lies between these two.

In Figures 7.11 and 7.12, the recorded throughput and average cycle length are compared. It is clearly seen that P-DBA and SLA-DBA with SLA 3 achieved better bandwidth utilization than their counterparts. Although, here the cycle length does not affect the recorded average delays, it is worth pointing out that the SLA-DBA with SLA 3 has the shortest average cycle time, especially for very large loads. When the behavior of the SLA-DBA algorithms is compared with SBA, similar comments can be made as in the situation where P-DBA was compared to SBA.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure 7.11: Cycle Length Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure 7.12: Throughput Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.

7.6.2 Performance in Congested Network

If SLA-DBA is to be implemented in DiffServ-aware EPONs, it should allocate guaranteed bandwidth even during heavy congestion. The results of experiments show that in a situation where the total traffic is much bigger than the link capacity, the performance of the algorithm depends on the SLA assigned to a particular class or traffic source. In Figure 7.13, the average delays and queue lengths for different SLAs are compared. It can be seen that the best separation of traffic from sources transmitting at the nominal and double rate was achieved by SLA-DBA and SLA 1, followed by SLA-DBA with SLA 2. Similar to the normal traffic conditions, SLA-DBA with SLA 3 was worst. It is interesting to point out that although the average delays recorded for the sources sending at the nominal rate were small, they started to grow when the total load increased above a certain level. As mentioned in Section 5.4, the value of this level is dependent on the values set in the SLAs and the mix of traffic used. In particular, it was sensitive to the proportion of the BE traffic in the total load. For SLA 1 and TM 1, the volume of BE traffic is large enough in proportion to the EF and AF traffic and no negative effects are visible. For the same SLA and TM 2, where the volume of the BE traffic is proportionally smaller, very undesirable instabilities in the recorded results appear as shown in Figure E.20.

Another negative effect present for very high loads is the consequence of the limited size of the cycle length. As in congested networks where the reported number of bytes is larger than the maximum number of bytes that can be sent in one cycle

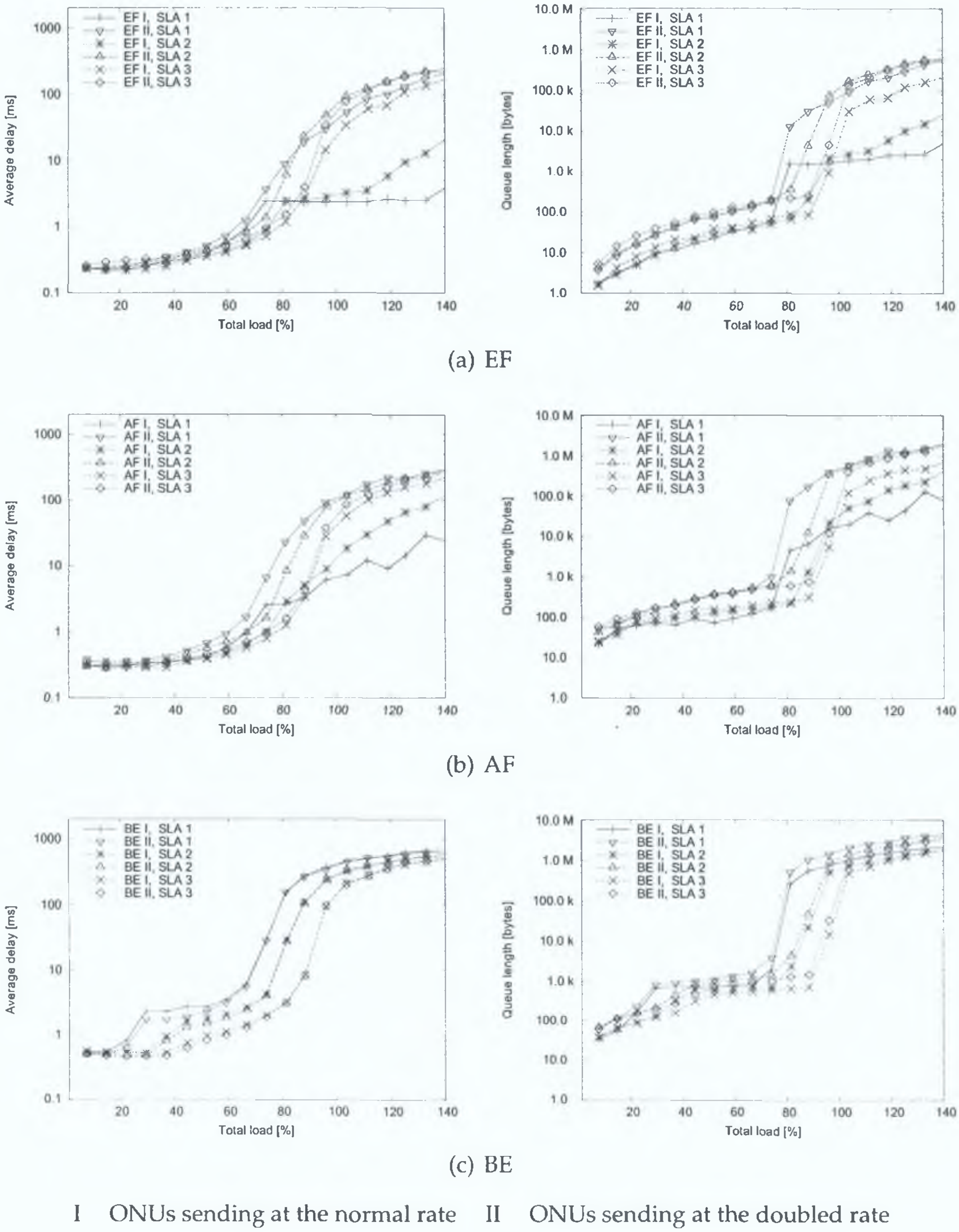


Figure 7.13: Algorithm Performance Comparison. Congested network. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.

and all classes are assigned less bandwidth than they requested. This leads to an increase in the recorded delays even for the high priority sources transmitting at the normal rate.

These two negative effects can occur at the same time and lead to a situation where the behavior the SLA-DBA algorithm is hard to predict for very large loads. It is very unlikely that this will be a concern to the network provider. It is highly improbable that the situation where the offered loads will reach 150% of the link capacity can persist.

The results of simulation experiments carried out for other traffic proportions are included in Appendix E as they do not show a major departure from the general trend.

7.6.3 Summary

The SLA-DBA algorithm proposed in this thesis is an interesting choice as it enables EPONs to support the DiffServ architecture. It has been shown that by controlling the SLAs, the grade of service received by different sources can be regulated. The results presented highlight its particularly good performance for the loads not exceeding the maximum link capacity. In congested networks the algorithm shows good properties and sources violating their SLAs receive much worse QoS. The adoption of proportional bandwidth assignment as a part of the SLA-DBA has a strong impact on the general behavior.

- The sources that breach their contracts are not penalized enough. As their queues get longer, they request more bandwidth. As the SLA-DBA algorithm is based on the proportional bandwidth assignment it reacts by assigning proportionally more resources to non-compliant queues. Obviously, this comes at the expense of sources staying within the limits. Although at later stages the algorithm tries to rectify this situation, the non-compliant sources has an advantage over other sources.

- The proportional allocation in the SLA-DBA can be a source of fluctuations in QoS for some particular combinations of SLAs and traffic mix. The parameters $\gamma_i^{min}(j)$ and $\gamma_i^{max}(j)$ describing the minimum and maximum bandwidth allocated to a queue must be chosen with care as in some cases wrong values can lead to unexpected behavior for very large loads. It must be ensured that there is enough excess bandwidth in the system that can be allocated to the more demanding classes. Otherwise, the high priority queues which happened to be served towards the the end of the allocation process receive substantially worse QoS.

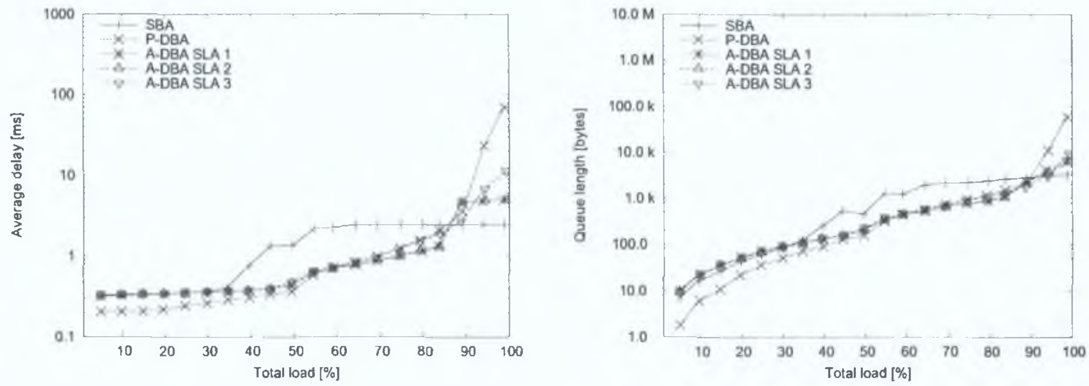
7.7 The Analysis of the A-DBA Algorithm

The architecture of the A-DBA algorithm was fully described in in Section 5.5. Its functionality differs significantly from P-DBA and SLA-DBA algorithms which are based on proportional bandwidth assignment and in its essence is more similar to the SBA algorithm. During the simulation experiments, various traffic conditions and mixtures were used and again only algorithms with a variable polling cycle were tested.

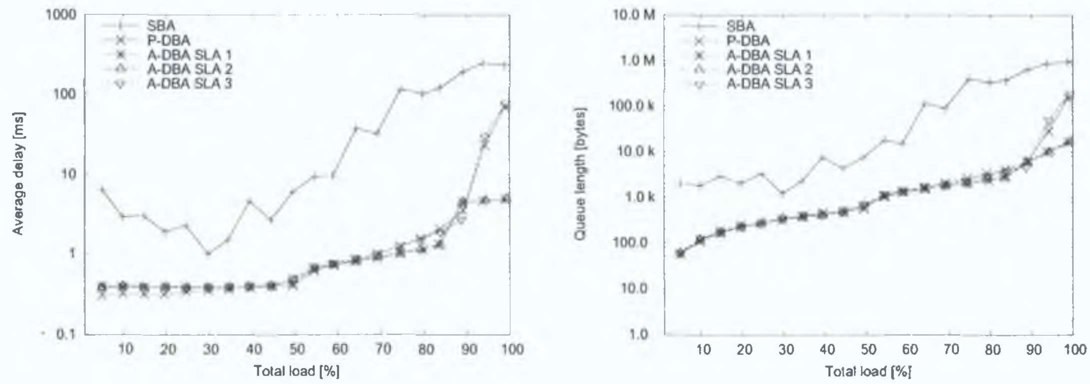
7.7.1 Performance in Normal Conditions

In this section the results of the experiments are shown for normal conditions where the total load does not exceed the total link capacity. The recorded average delays and queue lengths for Traffic Mix 1 are shown in Figure 7.14. More results can be found in Appendix F.

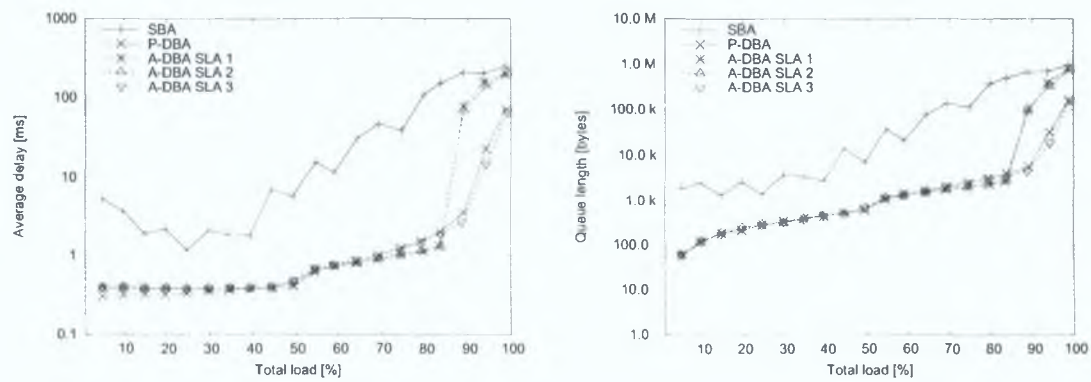
It can be seen that the performance of the A-DBA algorithm for small and medium loads and EF and AF traffic is comparable to the performance of the P-DBA algorithm. The important difference is in the behavior of both algorithms for loads approaching the maximum link capacity. The average delays recorded for the A-DBA with different SLAs were significantly lower than the values obtained for



(a) EF



(b) AF



(c) BE

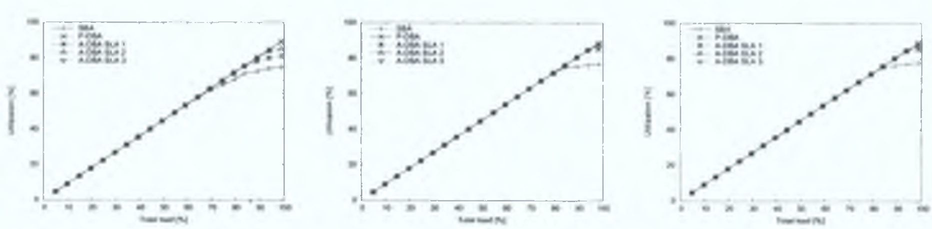
Figure 7.14: Algorithm Performance Comparison. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms

the P-DBA algorithm. Where, for A-DBA the maximum delay did not exceed 5 ms, it was above 60 ms for the P-DBA algorithm.

The improved performance for the EF and AF classes of traffic did not result in worse level of QoS for the BE traffic. This can be seen in Figure 7.14(c), where the results recorded for A-DBA with SLA 3 are comparable with the results measured for P-DBA.

If the performance of the A-DBA algorithms is measured against the SBA algorithm similar patterns are observed and generally the same conclusions can be drawn as in the case of the P-DBA algorithm.

The throughput comparison shown in Figure 7.15 reveals that the achieved bandwidth utilization is not as good as that achieved by the P-DBA algorithm, but at the same time, it is better than that of the SBA algorithm.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure 7.15: Throughput Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.

The comparison of the average cycle length is presented in Figure 7.16. It can be pointed out that for small and medium loads the average cycle length of the A-DBA algorithm is shorter than the average cycle length of the P-DBA algorithm. The opposite is true when the total load is close to the total link capacity. In this situation the cycle length approaches the allowable maximum.

7.7.2 Performance in Congested Network

The rationale for implementing the A-DBA in DiffServ-aware EPONs requires that it show good performance and enforce agreed contracts when the traffic is heavy.

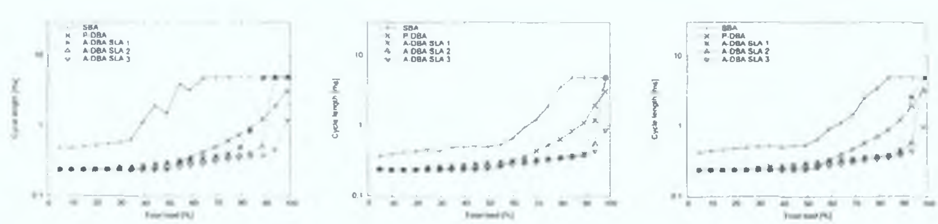
The results of experiments show that, in a situation where the total traffic is much bigger than the link capacity, the performance of the algorithm depends on the SLA assigned to the particular class or traffic source. In Figure 7.17 the average delays and queue lengths for A-DBA with different SLAs from Table 7.4 are compared. It can be seen that the EF class of traffic sources sending at their nominal rate have low average delays for all SLAs. The differences between particular SLAs in this case are very small. Conversely the average delays measured for sources that sent at double rate grow steadily as the load increases.

The performance of the A-DBA algorithm for the AF class of traffic is similar to its behavior for the EF class. Again, traffic from sources sending at double rate is heavily penalized.

Extremely good performance for high and medium priority traffic again comes at the expense of BE traffic, which suffers from large delays. Also the recorded average queue lengths are much longer than for classes with higher priorities.

The uniqueness of the A-DBA algorithm lies in the fact that good protection of the traffic parameters is not achieved at the expense of the BE class of traffic. Although in the case of A-DBA with SLA 1 or SLA 2 the average delays recorded for the BE are large, this is because of the way the limits were set in the SLAs. At the example of A-DBA with SLA 3, it can be shown that, even for the BE class of traffic, non-compliant sources were severely penalized.

The results recorded for other experiments are shown in Appendix F, but they



(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure 7.16: Cycle Length Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.

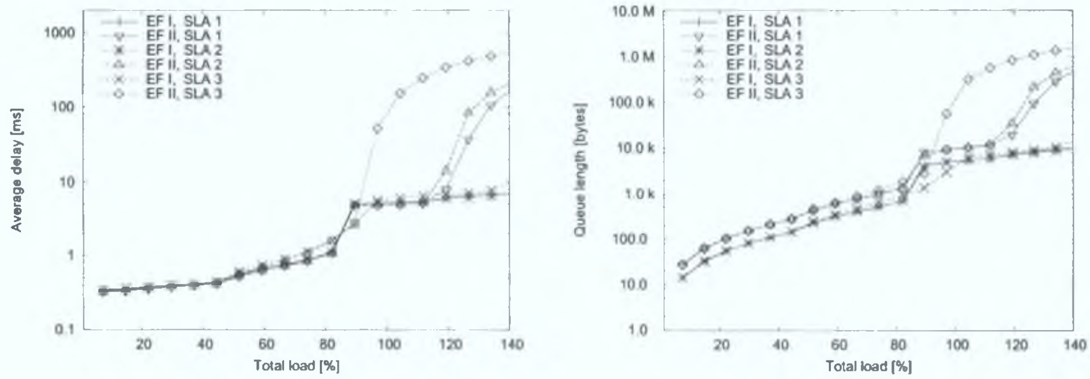
do not reveal any major departures from the general trend.

7.7.3 Summary

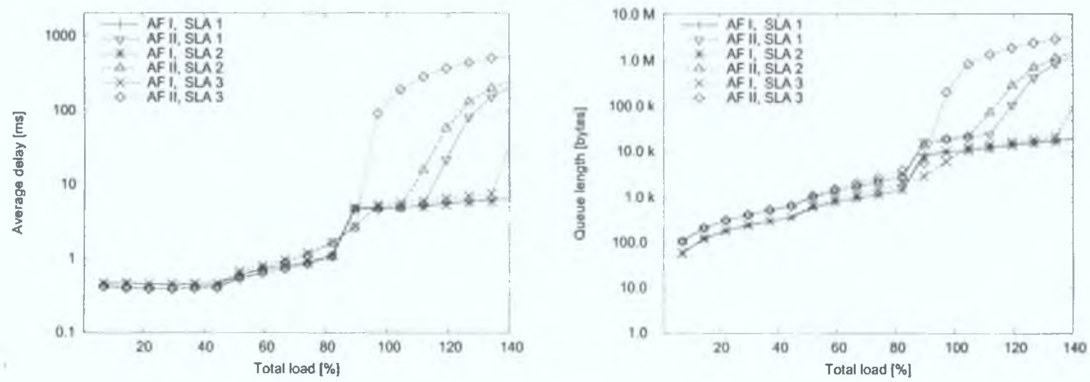
The design of the A-DBA algorithm was based on the properties of the SBA algorithm. As shown earlier, the SBA algorithm provides good protection of traffic parameters agreed in the SLAs. In the A-DBA algorithm, the fixed allocation is removed as a source of poor bandwidth utilization. Instead an adaptive approach is used to dynamically allocate bandwidth. As a result the A-DBA algorithm shows good behavior under normal as well as heavy traffic conditions and sources breaching their SLAs do not affect the QoS received by classes staying within the limits. The results show that the average delays recorded for sources that conform to their SLAs are low and not longer than the maximum cycle time. Although this could be a disadvantage when the cycle time is long, it is shown in Section 7.2 that long cycles do not improve the overall performance.

7.8 SLA-DBA and A-DBA Algorithms Comparison

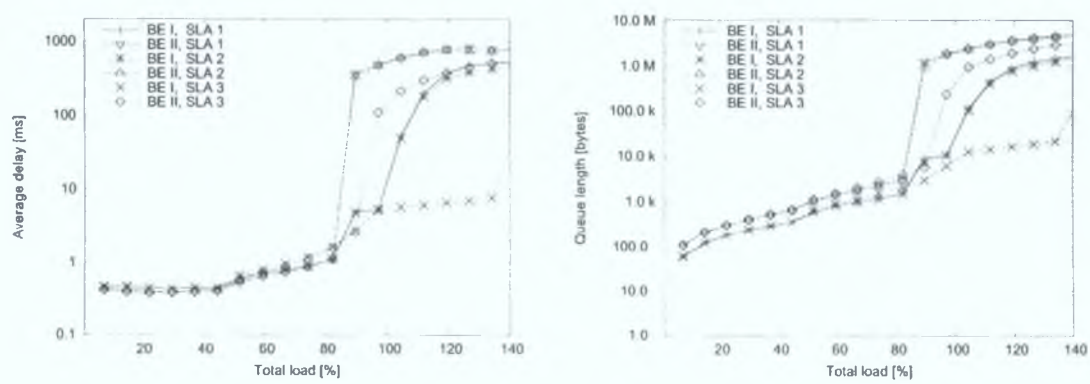
In this section a comparison of the SLA-DBA and A-DBA algorithms is presented. In Figure 7.18, the performance of both algorithms is drawn side by side under the normal traffic loads. The throughput and average cycle length are compared for both algorithms in Figures 7.19 and 7.20 respectively. The analysis of the presented graphs reveals that although the differences between the algorithms are narrow, the A-DBA algorithm gives smaller average delays for large loads. This difference can be as much as 60 ms when the link is fully utilized. This trend is observed for all classes of service. The recorded queue lengths are generally longer for the A-DBA algorithm. On the other hand it must be pointed out that the growth in queue length is more gradual and for the large loads the A-DBA algorithm outperforms the SLA-DBA scheme. Also if the recorded throughput is compared it can be seen that the A-DBA algorithms achieves slightly better results. The comparison of the performance



(a) EF



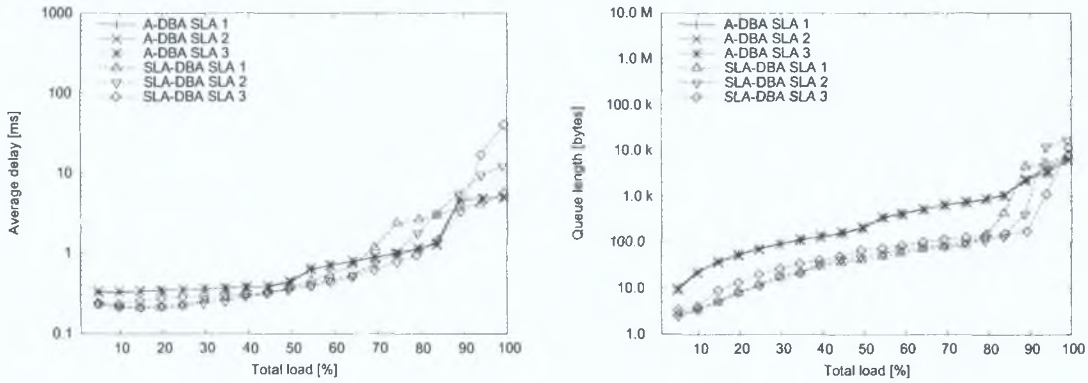
(b) AF



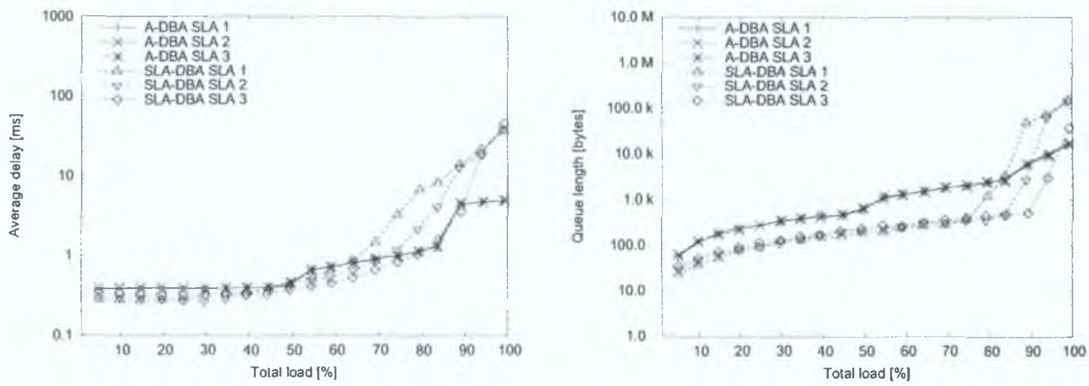
(c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

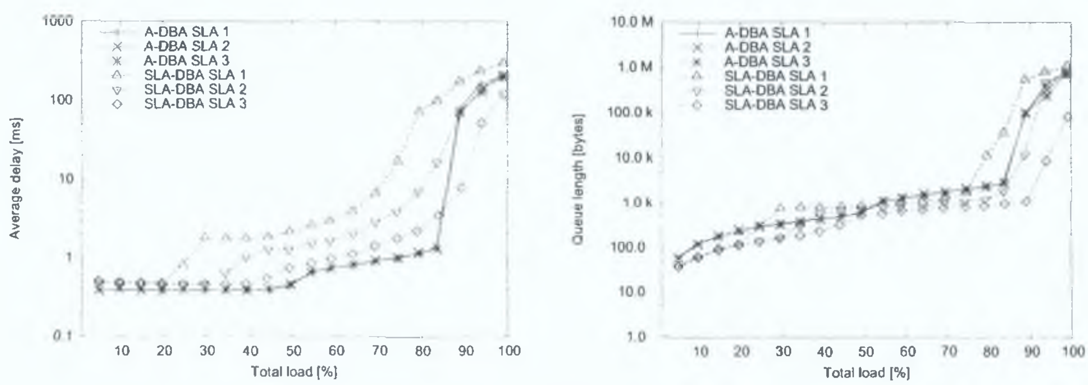
Figure 7.17: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) EF

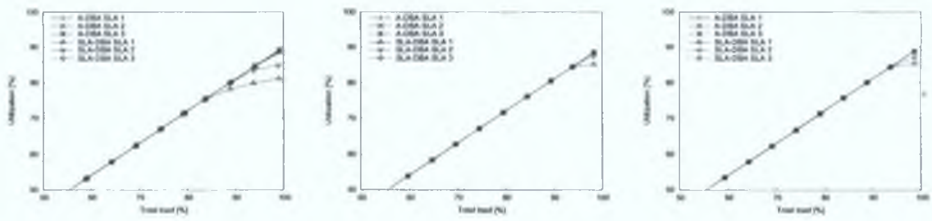


(b) AF



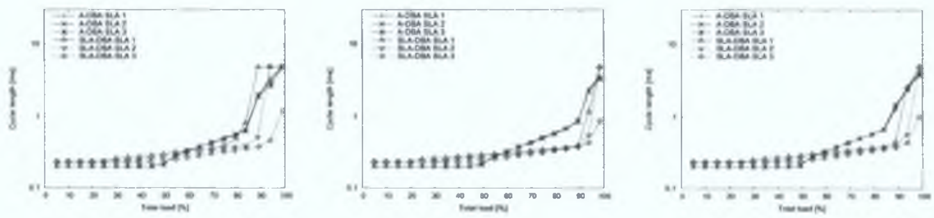
(c) BE

Figure 7.18: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure 7.19: Throughput Comparison. Normal Conditions. Traffic Mix 1. Maximum Cycle Length 5 ms.

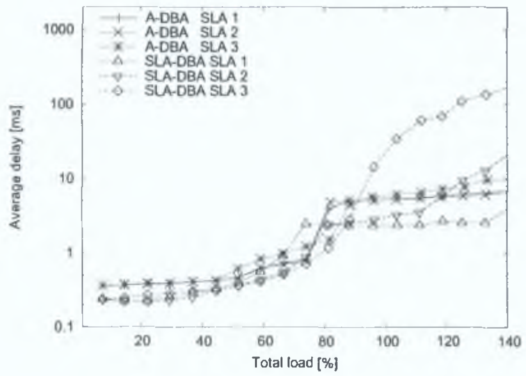


(a) Sources Set I (b) Sources Set II (c) Sources Set III

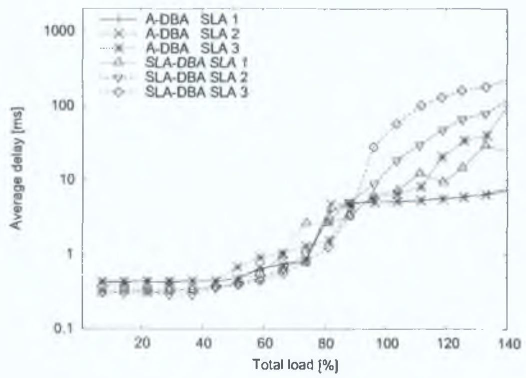
Figure 7.20: Cycle Length Comparison. Normal Conditions. Traffic Mix 1. Maximum Cycle Length 5 ms.

of these two algorithms in congested networks is shown in Figure 7.21.

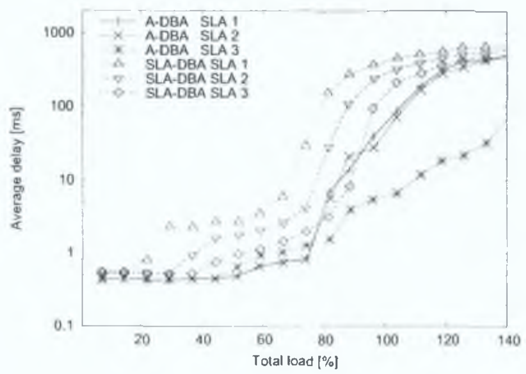
It is hard to conclude which algorithm performs better in congested networks. For SLA-DBA the difference between average delays recorded for compliant and non-compliant sources is strongly dependent on traffic mix, used SLA, class of traffic. For example for SLA-DBA, TM 1, SS I and SLA 1 the difference recorded for compliant and non-compliant sources can be between 200 and 250 ms for EF class of traffic. At the same time for SLA 2 and AF class of traffic the difference between compliant and non-compliant sources is between 50 and 100 ms. From this point of view it can be said that the A-DBA algorithm has better performance as the traffic parameters and SLAs have smaller impact on QoS received by compliant and non-compliant sources. At the same time the A-DBA algorithm can be outperformed by the SLA-DBA algorithm for some SLAs. For example in Figure 7.21(a) it is seen that SLA-DBA for SLA 1 gives average delays 5 ms shorter than given by A-DBA.



(a) EF



(b) AF



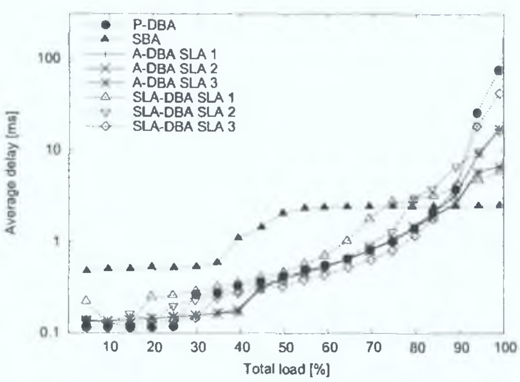
(c) BE

left ONUs sending at the normal rate right ONUs sending at the doubled rate

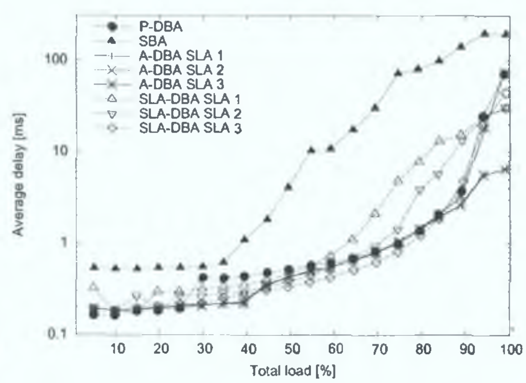
Figure 7.21: Average Delay Comparison. Congested Network. Traffic Mix 1. Source Set I. Maximum Cycle Length5 ms.

7.9 Scalability of the Algorithms

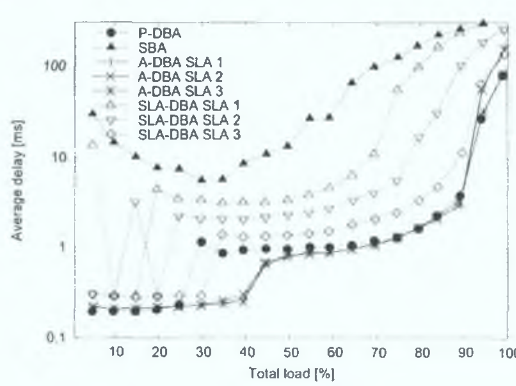
Although the current standard specifies that sixteen ONUs must be supported by the OLT, more endpoints could be supported as the quality of equipment improves and higher data rates are achieved. In this section the results of simulations for an EPON with 32 ONUs are shown. From Figure 7.22 it can be concluded that all algorithms based on the centralized approach scale well as the number of ONUs increases. If the recorded results are compared with average delays measured for a network with 16 endpoints it can be observed that in both cases the general performance is similar. Obviously, as the number of ONUs increases more bandwidth



(a) EF



(b) AF



(c) BE

Figure 7.22: Algorithms performance comparison for network with 32 ONUs

must be dedicated to the guarding gaps and hence packets experience larger delays. This effect is especially visible for heavy loads where the differences between recorded results can be of as much as 10 ms for high priority traffic.

7.10 Chapter Summary

In this chapter the analysis of the performance of different schemes based on the centralized approach to bandwidth allocation is presented.

- The effects of a new polling method with the variable length of the polling cycle, which was introduced in Section 4.3, are outlined. It is shown that with Cyclic Polling with Adaptive Cycle Length significant improvements in QoS can be achieved and the recorded average delays are much lower in comparison to other polling methods. In the new scheme the OLT can easily adjust the size of the control channel. It is shown that this results in better performance of the system in situations where traffic in the downstream and upstream directions is asymmetric.

  It was observed that the short polling cycle resulted in large delays for the loads close to the link capacity. By extending the cycle length the maximum delays can be reduced without worsening the average delays recorded for small and medium loads. It is shown that by making the cycle too long, no further improvements are obtained. Based on these conclusions it was decided to perform other experiments only for the maximum cycle length.

- A Grant Multiplexing scheme based on centralized bandwidth allocation was analyzed. It is shown that it offers significant improvement in the service given to classes with stringent traffic requirements. This was achieved by placing the time slots allocated to high priority traffic at the beginning of the polling cycle. Presented results show that at least order of magnitude decrease in experienced jitter can be achieved.

- The performance of the P-DBA and SBA methods with the proposed polling scheme is analyzed for normal and congested networks. The results show that although the P-DBA algorithm gives best bandwidth utilization it does not offer any guarantees regarding the QoS delivered to the queues. On the contrary, the SBA method showed excellent support for DiffServ but lacked good bandwidth utilization.

- Simulation experiments for the SP-DBA algorithm in its centralized version were carried out. They prove that the centralized scheme is flexible and tasks which are usually performed by ONUs could be successfully implemented in the OLT without degrading the overall performance of the system.

- Two new algorithms were introduced to enable efficient support of the DiffServ architecture in EPONs. The SLA-DBA method is based on the P-DBA approach and the A-DBA scheme tries to exploit the behavior of SBA. The performance of new algorithms proposed was compared for different traffic mixes and different characteristics of traffic sources. The results of extensive simulation experiments show that both algorithms have good performance and that it is difficult to point out a clear winner.

  Both algorithms show similar performance in terms of average delays for normal traffic conditions, where the total load was smaller than the maximum link capacity. The A-DBA algorithm showed slightly better performance for very large loads. In this situation recorded average delays were around 60 ms shorter than for SLA-DBA.

  In more adverse conditions where the network was heavily congested, the algorithms differ substantially in their performance. The performance of the SLA-DBA algorithm is strongly dependent on the chosen SLA parameters. For SLA 1 the performance of SLA-DBA is significantly better than for SLA-DBA and SLA 3. This strong dependence on SLAs can lead sometimes to unpredictable behavior. For specific combinations of traffic mix and SLAs, the behavior of SLA-DBA algorithm can become unstable in very congested networks. In such a situation the agreed levels of QoS could not be sustained.

The A-DBA algorithm proved to give small average delays, although the separation in terms of recorded average delays between compliant and non-compliant sources in congested networks was sometimes not sufficient. It also proved to be more stable than SLA-DBA for very heavy loads.

The results of extensive simulation experiments clearly showed that, with centralized bandwidth allocation and the new GATE message, a flexible architecture can be built where different algorithms can be deployed in a transparent manner. Analysis of the proposed algorithms showed them to be resilient to different traffic types and to work well in situations where various proportions of high, medium and low priority packets are used.

In this chapter only results of simulations were shown for Source Set I which modelled self-similar and long-range dependent traffic. This explains sudden fluctuations from the average trend in presented graphs. Calculating an average from many simulations or running longer simulations would help to minimize these effects. On the other hand such an approach would lead to unacceptably long experiments. It can be seen that trend lines measured for Source Sets II and III presented in Appendixes at the end of the thesis are smooth.

Chapter 8

Conclusions

The growing demand for broadband access to the Internet has led to new interest in access schemes based on optical fiber. Amongst them, the PON architecture is considered the most promising from technological and economic point of view. The scope of this thesis is focused on EPONs, as this technology has many advantages over its rival solutions. The adoption of the Ethernet protocol makes converging access, local and metropolitan networks into one system possible.

8.1 Contributions

An efficient bandwidth allocation method is crucial to maximizing the utilization of available resources. The main goal of this thesis was to design a framework that would enable more efficient support for the DiffServ paradigm and enable deployment of DiffServ-compatible bandwidth allocation schemes. Upon analysis of existing work in this area and published standards, a range of contributions has been made.

8.1.1 Improvements to the MPCP

It has been recognized that the MPCP does not provide enough functionality to implement a centralized allocation scheme where both inter and intra-ONU scheduling are performed by the OLT. This is a disadvantage as the OLT has up to date information about the overall network status and resource utilization. Moreover, some bandwidth allocation methods [58–61] rely on centralized allocation. Implementing these methods in networks based on the current standard, would result in proprietary solutions. Interoperability between the ONU and OLTs from different vendors would be problematic to achieve.

In this dissertation a new format of the GATE message is introduced which uncouples the bandwidth allocation process from the underlying infrastructure. Hence, equipment from diverse sources can coexist in the same domain. With the new format both centralized and distributed methods can be used in the same network while backwards compatibility is maintained.

8.1.2 Cyclic Polling with Adaptive Cycle Length

Based on the centralized approach a new scheme called Cyclic Polling with Adaptive Cycle Length was proposed. This new scheme is a compromise between IPACT [41, 50, 51] proposed by Kramer et al. and Polling Cycle introduced by Choi et al. in [53]. In the adaptive polling transmission windows are assigned consecutively and the total length of the polling cycle can be determined after all ONUs were serviced. This results in a variable length of the polling cycle and for small loads average delays experienced by packets in the system are shorter. At the same time the size of the control channel is variable and load dependent.

Conversely, in the cyclic polling approach all transmission windows are assigned at the beginning of the polling cycle. This means that the OLT can determine the size of the polling cycle.

Combining these two methods together gives the OLT an opportunity to adjust the size of the control channel to current conditions. If the volume of medium and low priority traffic in the downstream direction is substantially larger than in the upstream direction, there is no need for large control channel. In such situation, the OLT can increase the length of the polling cycle. Hence, less bandwidth is dedicated to carry control packets and more bandwidth can be dedicated to the customers' payload.

Performance of the proposed approach was compared with cyclic polling. Measured values demonstrate that the approach proposed in this thesis is more efficient.

8.1.3 Grant Multiplexing Scheme

A Grant Multiplexing scheme was introduced in this dissertation and results of simulation experiments were shown. The method implemented in networks supporting centralized bandwidth allocation and utilizing the new GATE message allowed to position time slots dedicated to QoS sensitive classes at the beginning of the cycle. Such approach helps reducing the impact of variable size of transmission windows allocated to medium and low priority traffic on positioning of transmission windows allocated to privileged traffic. In the result the smaller jitter was measured for packets belonging to high priority classes. The results show that the improvements depend on the mixture of high, medium and low priority traffic. Nonetheless the recorded jitter was at least ten times shorter when compared with a system without Grant Multiplexing scheme.

8.1.4 SLA-DBA and A-DBA Algorithms

One of the aims of this work was to design efficient bandwidth allocation algorithms that will enable EPONs to support the DiffServ architecture. Two new algorithms were introduced in this dissertation. Their performance was tested for different mixes and types of traffic. The results were recorded in normal conditions as well as in situations where network was heavily congested. The detailed experiments helped in analysis of performance of the proposed algorithms.

The SLA-DBA algorithm proved to be an efficient solution in normal conditions when the total load was not larger than the link capacity. In heavily congested networks, the performance of the SLA-DBA algorithm was highly dependent on the chosen SLA. For SLA 1 the recorded average delays were shorter than for any other algorithm and non-compliant sources received much worse QoS. In case of SLA 3, the difference between compliant and non-compliant sources was difficult to notice. Moreover, for some particular SLAs and traffic mixes the bandwidth allocation process was shown to be unstable.

The A-DBA method was based on a completely different approach. In normal traffic conditions the performance of the A-DBA algorithm was as good as the performance measured for the SLA-DBA algorithm. The experiments carried out in the presence of heavy traffic showed that the A-DBA algorithm was better in maintaining agreed levels of QoS.

8.1.5 EPON Simulator

Extensive simulation experiments were carried out on a purpose-built EPON simulator. Although not as sophisticated in its architecture as other popular simulation tools, its structure was tailored to a specific purpose and high performance was achieved. The results produced by this simulator were validated through comparison with the TDMA system. The results showed good correlation between the output of the simulator and the analytical model.

8.2 Future Work

In this dissertation some necessary modifications to the MPCP were outlined without which, the implementation of DiffServ aware EPONs would be hard to achieve. Although results of simulation experiments were shown, the ultimate test would require implementing the proposed algorithms in real network with realistic traffic. Growing interest in new broadband access technologies and particularly in EPONs gives hope that access will be possible to a working system.

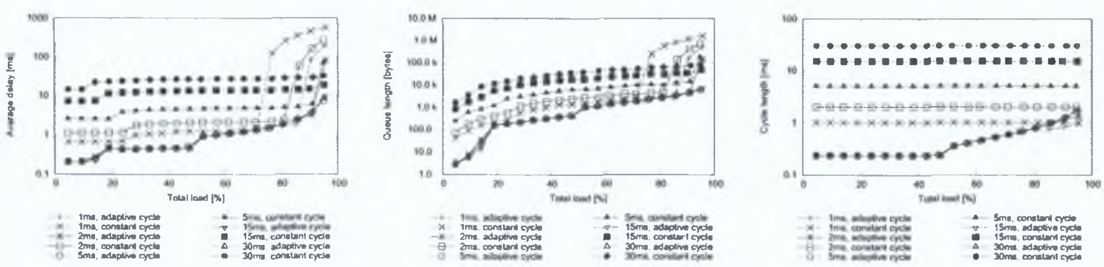
The algorithms presented in this thesis proved to be efficient and enabled EPONs to support classes of traffic with different QoS. It is possible that improved solutions can be found. An approach, where bandwidth allocation is based on the actual time that Ethernet frames or packets are spending in the queues, is particularly interesting. Knowing the delays experienced by packets and parameters agreed in the SLAs, the OLT would vary the size of allocated time slots to satisfy the needs of all sources.

The introduction of PONs with WDM access is another huge area for research. In such a network, the OLT will have to make decisions not only about allocating time slots to ONUs but also about choosing an appropriate wavelength to carry the signal. As the complexity of such a system is much higher, the problems faced by the bandwidth allocation algorithm are harder to solve.

It is envisaged that EPONs will eventually become part of a larger system and ultimately in a network based on the DiffServ paradigm it will be possible to guarantee bandwidth from one end station to another. To realize such a project substantial resources and the co-operation of many research, and industry bodies would be necessary.

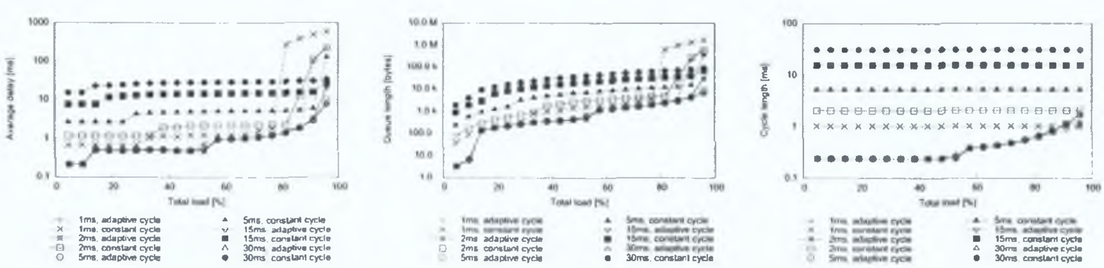
Appendix A

Cyclic Polling with Adaptive Cycle Length



(a) Average delay (b) Average queue length (c) Average cycle length

Figure A.1: The Comparison of Cyclic Polling with Adaptive and Constant Polling Time. Traffic Type 2.

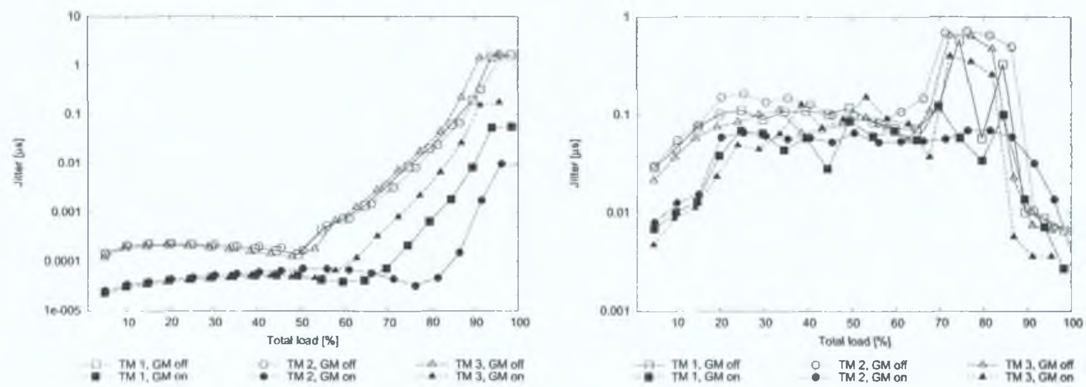


(a) Average delay (b) Average queue length (c) Average cycle length

Figure A.2: The Comparison of Cyclic Polling with Adaptive and Constant Polling Time. Traffic Type 3.

Appendix B

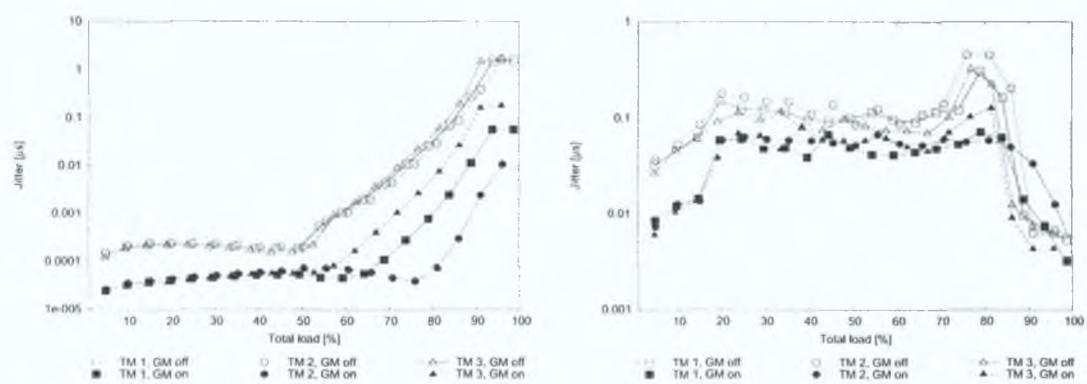
The Results for Grant Multiplexing Scheme



(a) adaptive cycle length

(b) constant cycle length

Figure B.1: Jitter Comparison. Source Set II.



(a) adaptive cycle length

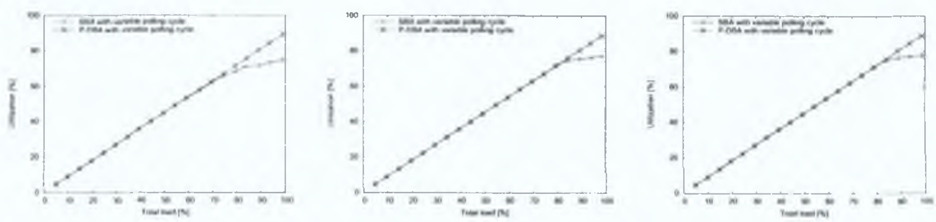
(b) constant cycle length

Figure B.2: Jitter Comparison. Source Set III.

Appendix C

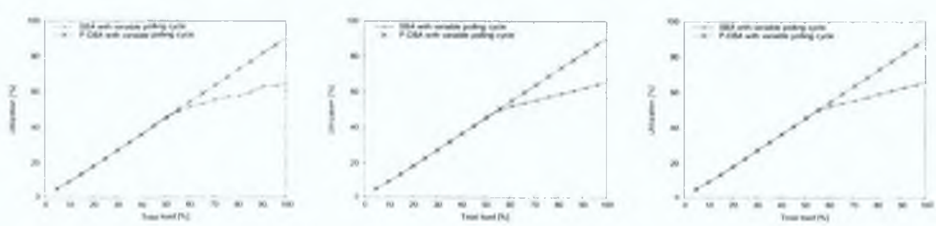
The SBA and P-DBA Algorithm Performance Comparison

C.1 Throughput and Cycle Length Comparison



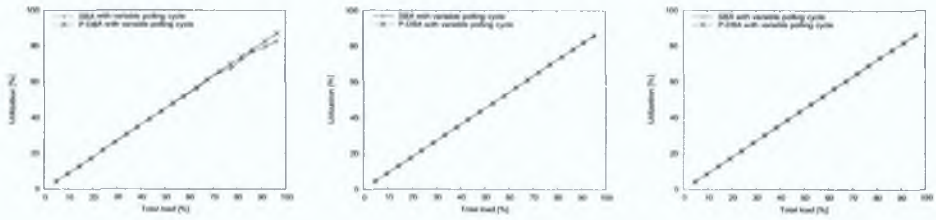
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure C.1: Throughput Comparison. Traffic Mix 1. Maximum Cycle Length5 ms.



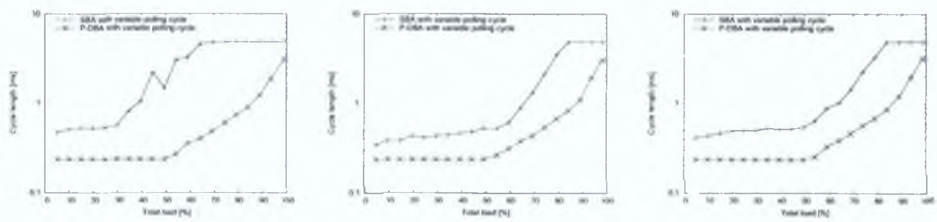
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure C.2: Throughput Comparison. Traffic Mix 2. Maximum Cycle Length 5 ms.



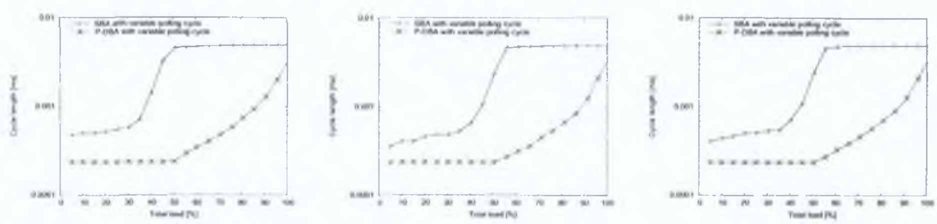
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure C.3: Throughput Comparison. Traffic Mix 3. Maximum Cycle Length 5 ms.



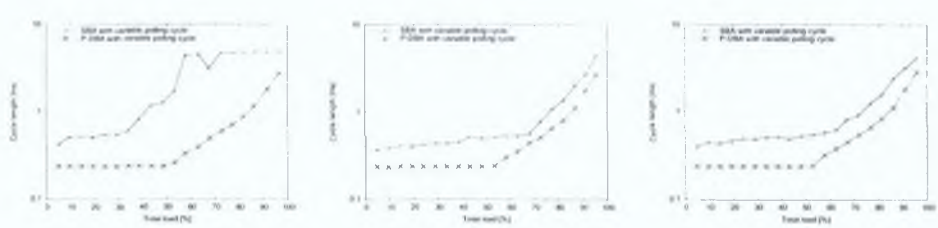
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure C.4: Cycle Length Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

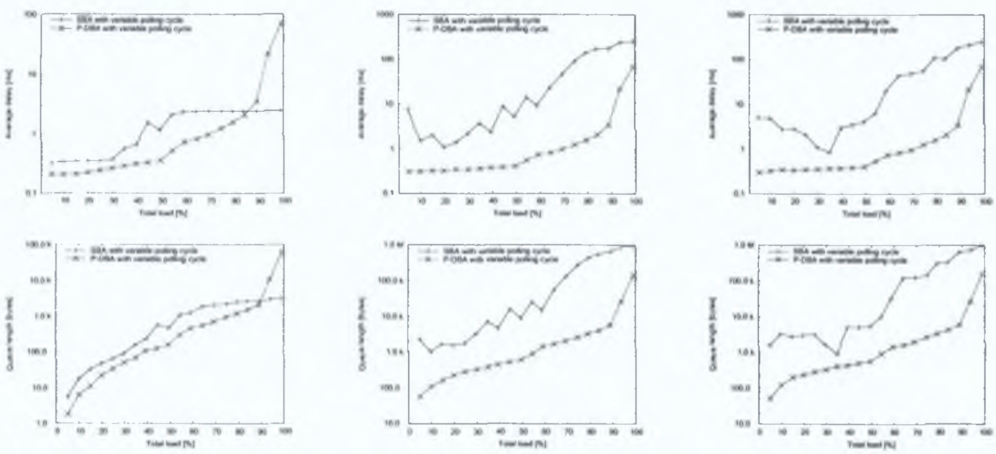
Figure C.5: Cycle Length Comparison. Traffic Mix 2. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

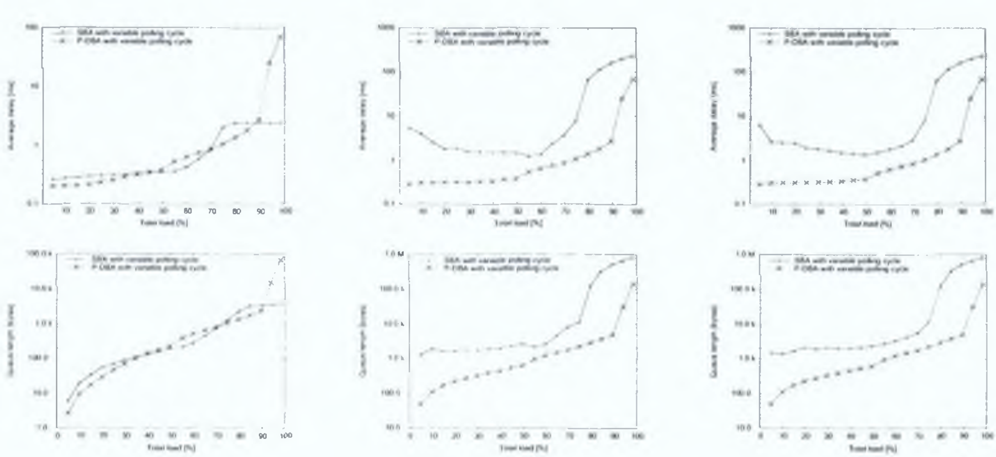
Figure C.6: Cycle Length Comparison. Traffic Mix 3. Maximum Cycle Length 5 ms.

C.2 Average Delay and Queue Length Comparison



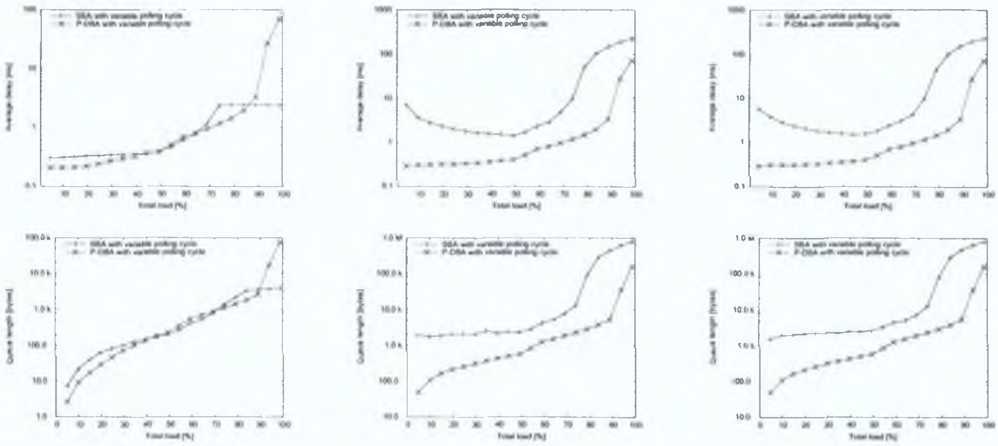
(a) EF (b) AF (c) BE

Figure C.7: Algorithm Performance Comparison. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



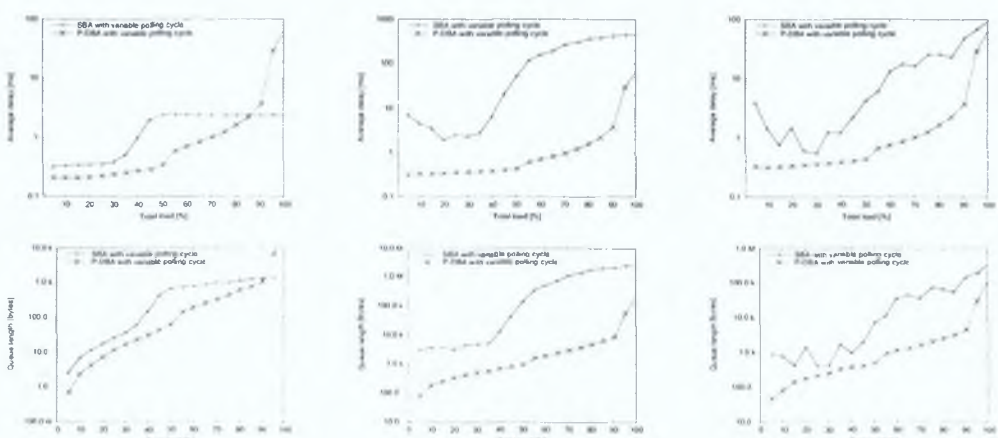
(a) EF (b) AF (c) BE

Figure C.8: Algorithm Performance Comparison. Traffic Mix 1. Source Set II. Maximum Cycle Length 5 ms.



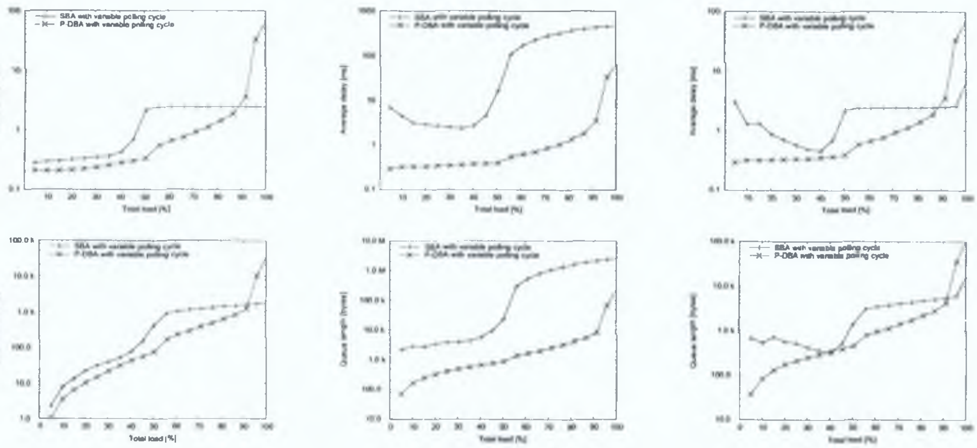
(a) EF (b) AF (c) BE

Figure C.9: Algorithm Performance Comparison. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



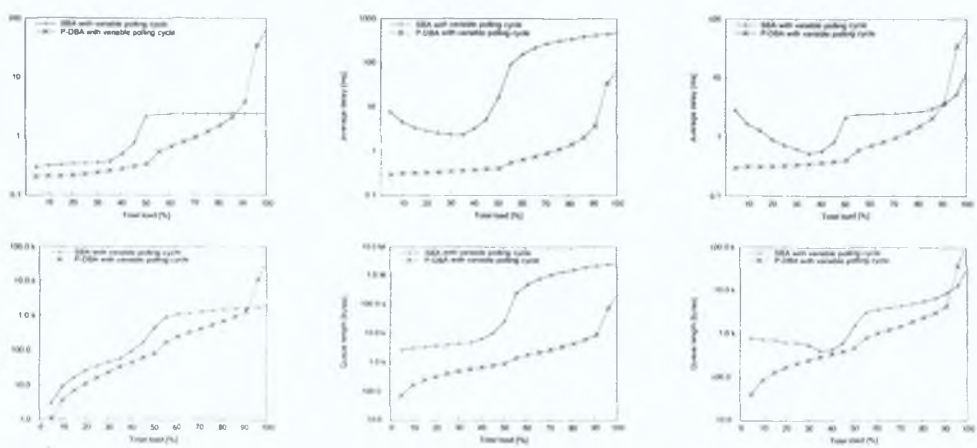
(a) EF (b) AF (c) BE

Figure C.10: Algorithm Performance Comparison. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



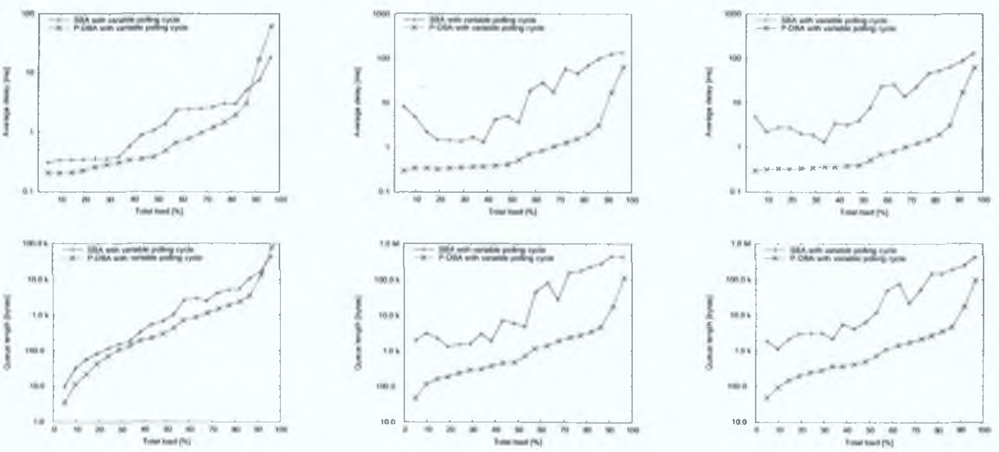
(a) EF (b) AF (c) BE

Figure C.11: Algorithm Performance Comparison. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



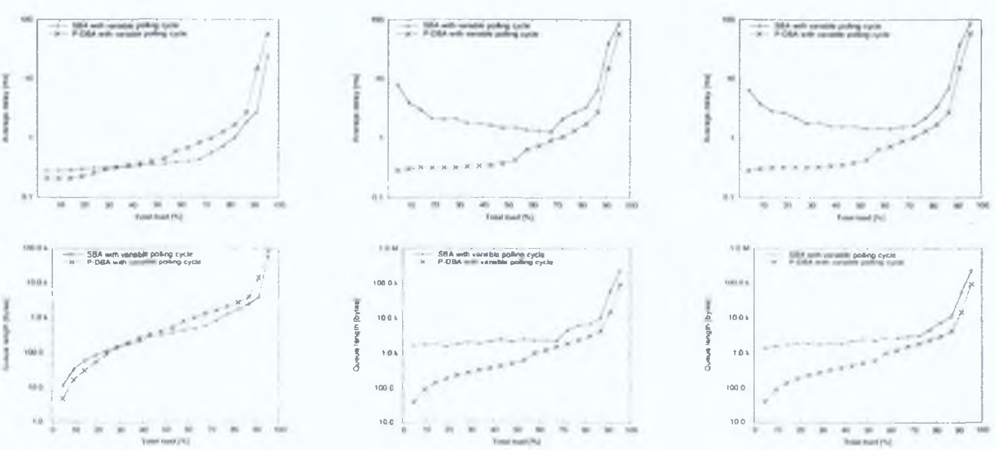
(a) EF (b) AF (c) BE

Figure C.12: Algorithm Performance Comparison. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



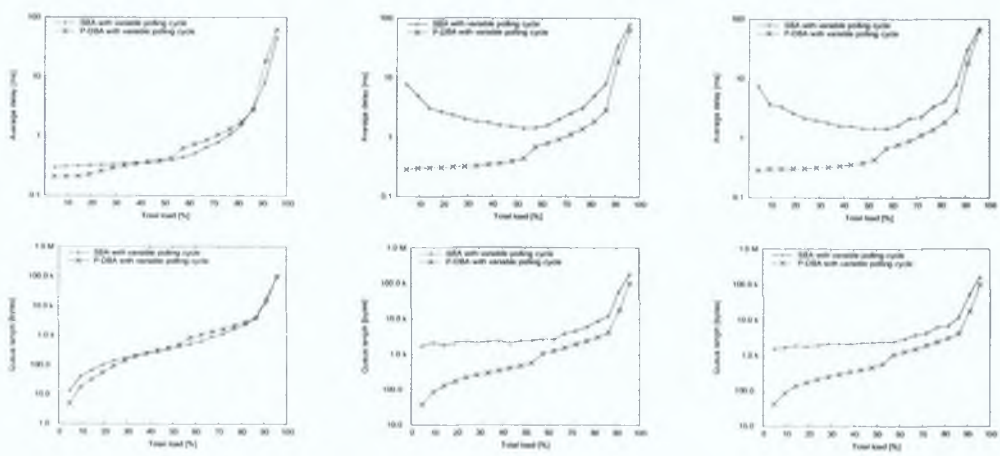
(a) EF (b) AF (c) BE

Figure C.13: Algorithm Performance Comparison. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

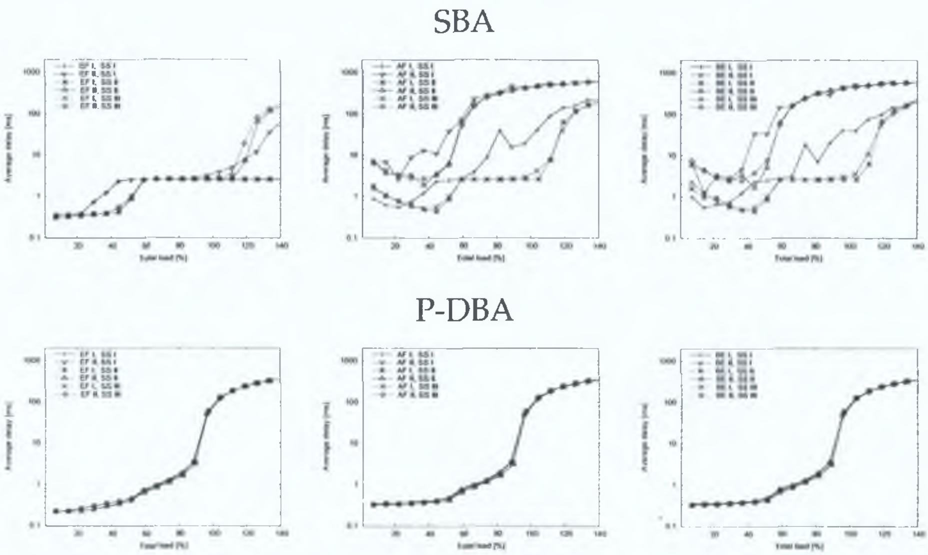
Figure C.14: Algorithm Performance Comparison. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

Figure C.15: Algorithm Performance Comparison. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

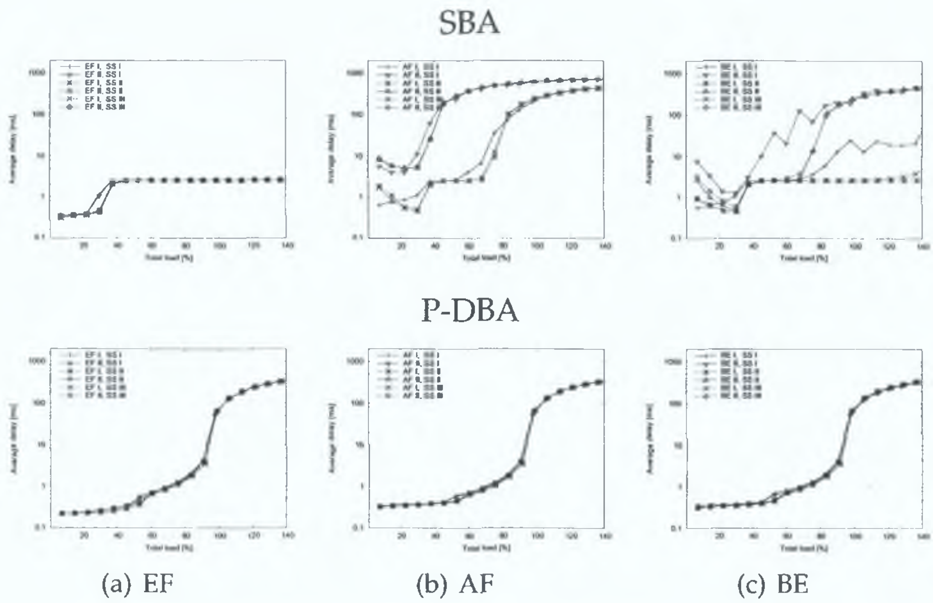
C.3 Performance Comparison in Congested Networks



(a) EF (b) AF (c) BE

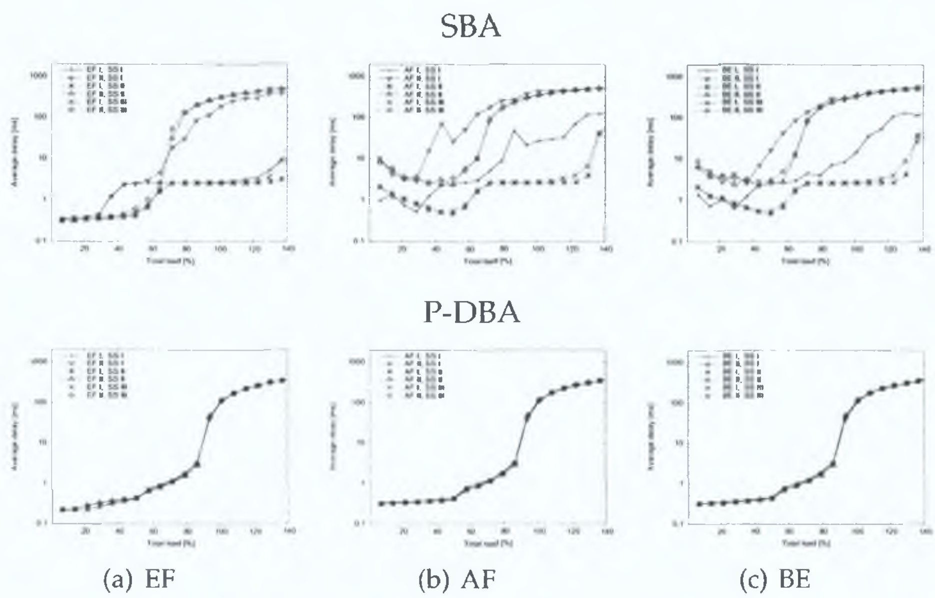
I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure C.16: Average Delay Comparison. Congested Network. Traffic Mix 1.



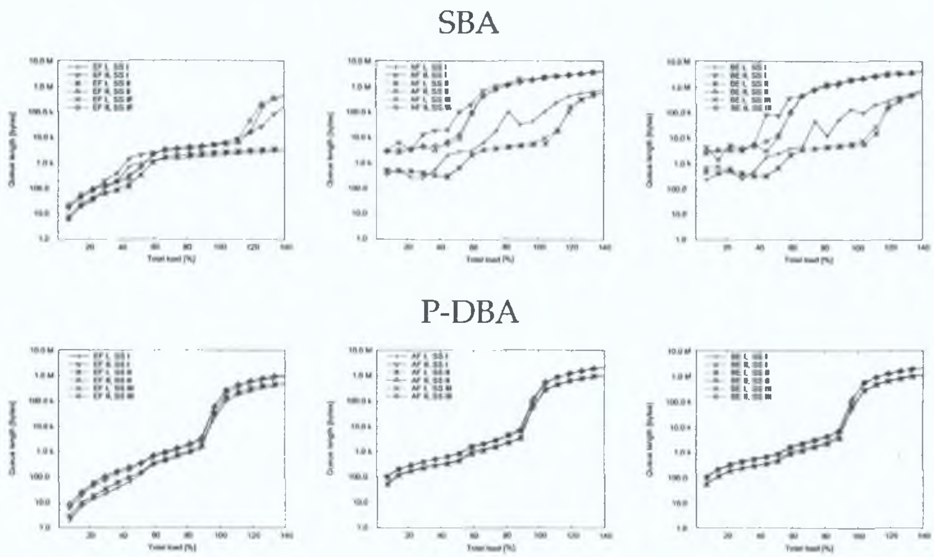
I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure C.17: Average delay comparison. Congested Network. Traffic Mix 2.



I ONUs sending at the normal rate II ONUs sending at the doubled rate

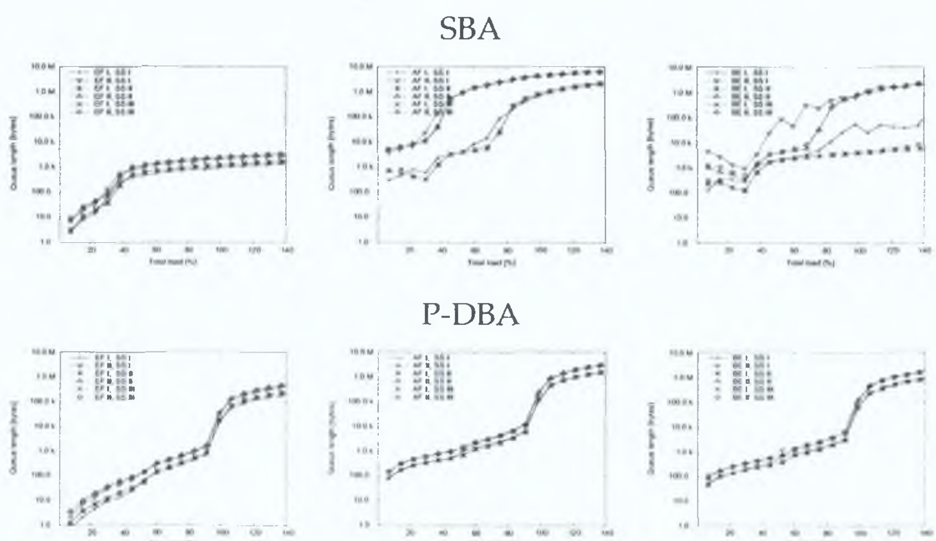
Figure C.18: Average Delay Comparison. Congested Network. Traffic Mix 3.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure C.19: Queue Length Comparison. Congested Network. Traffic Mix 1.

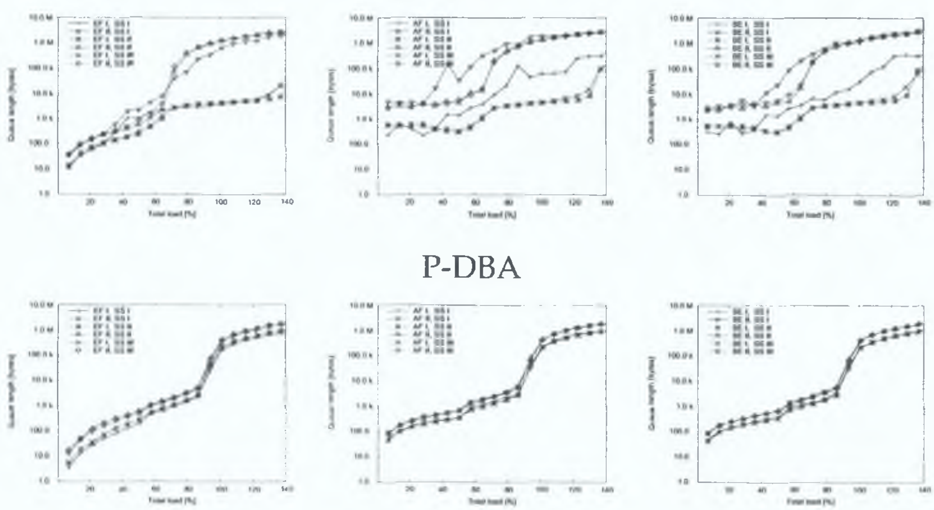


(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure C.20: Queue Length Comparison. Congested Network. Traffic Mix 2.

SBA



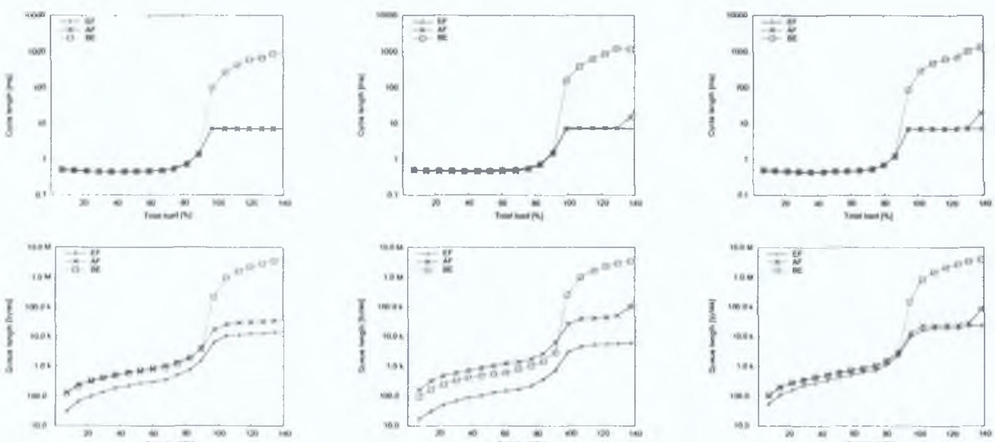
(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure C.21: Queue Length Comparison. Congested Network. Traffic Mix 3.

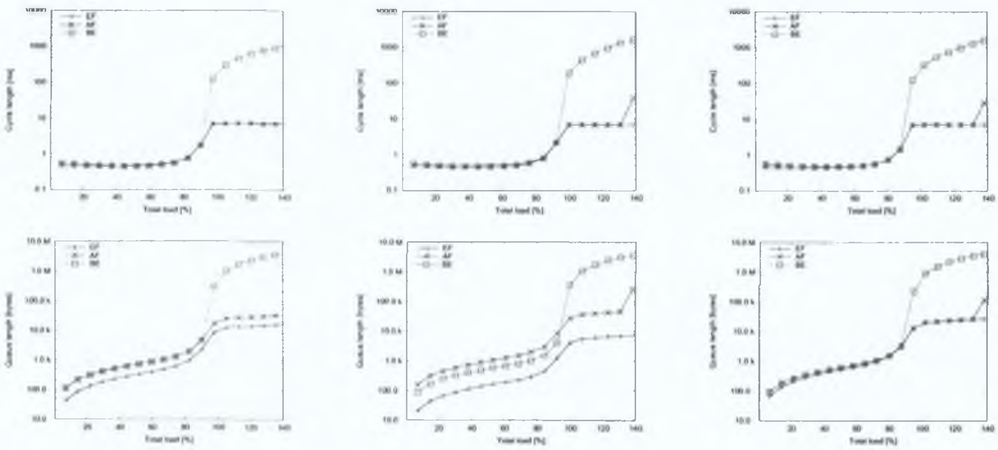
Appendix D

The Performance of the SP-DBA Algorithm



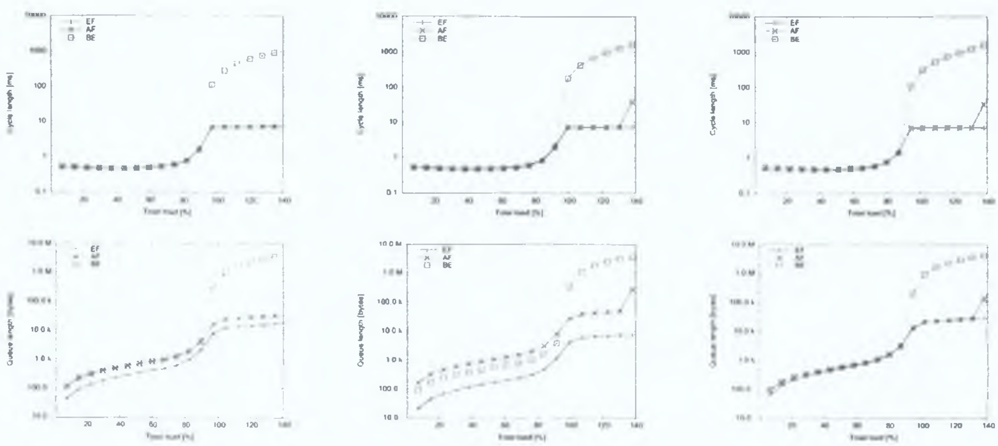
(a) TM 1 (b) TM 2 (c) TM 3

Figure D.1: SP-DBAPerformance Comparison. Source Set I. Maximum Cycle Length 5 ms.



(a) TM 1 (b) TM 2 (c) TM 3

Figure D.2: SP-DBAPerformance Comparison. Source Set II. Maximum Cycle Length 5 ms.



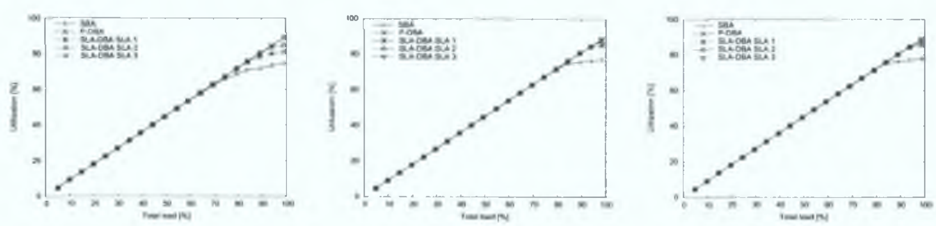
(a) TM 1 (b) TM 2 (c) TM 3

Figure D.3: SP-DBAPerformance Comparison. Source Set III. Maximum Cycle Length 5 ms.

Appendix E

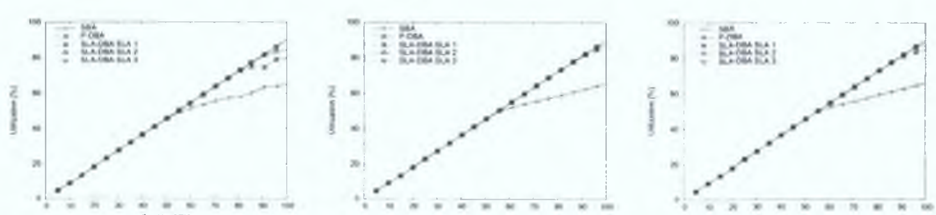
The Performance of the SLA-DBA Algorithm

E.1 Throughput and Cycle Length Comparison



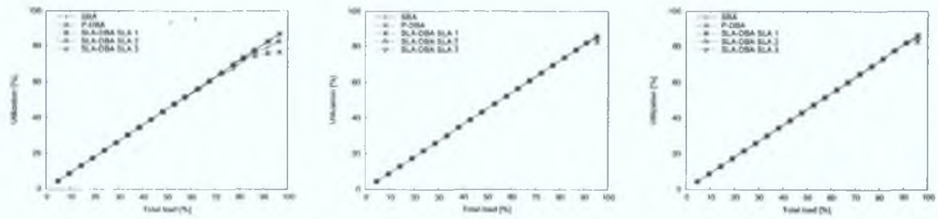
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure E.1: Throughput Comparison. Traffic Mix 1. Maximum Cycle Length5 ms.



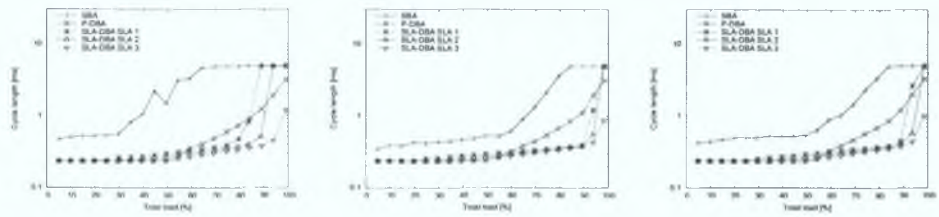
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure E.2: Throughput Comparison. Traffic Mix 2. Maximum Cycle Length 5 ms.



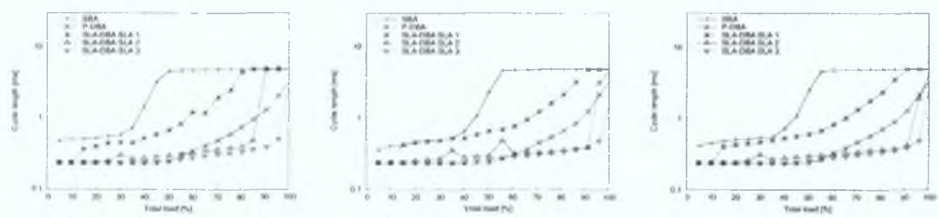
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure E.3: Throughput Comparison. Traffic Mix 3. Maximum Cycle Length 5 ms.



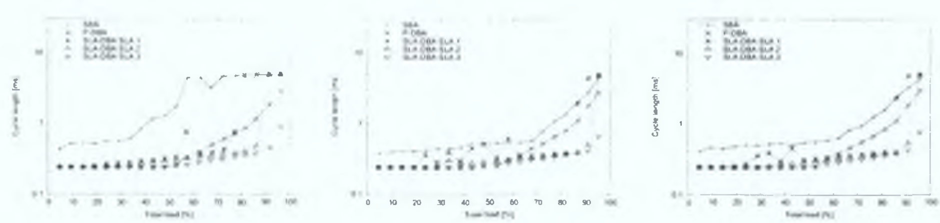
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure E.4: Cycle Length Comparison. Traffic Mix 1. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

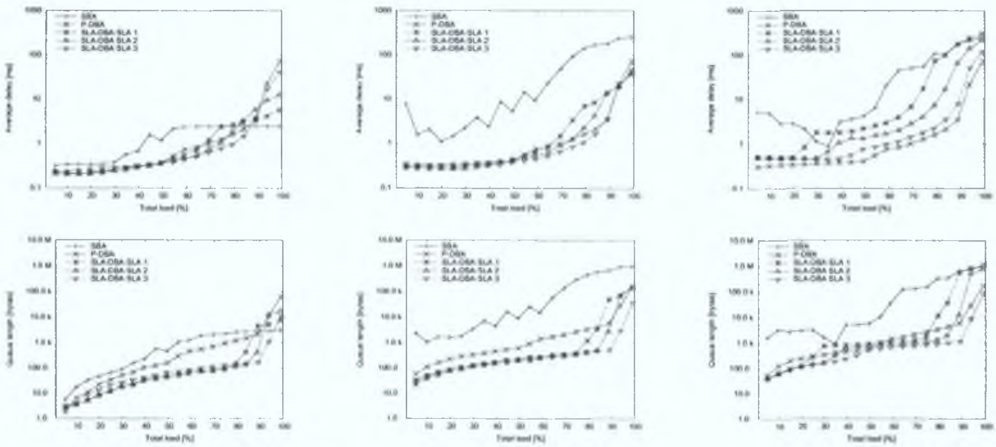
Figure E.5: Cycle Length Comparison. Traffic Mix 2. Maximum Cycle Length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

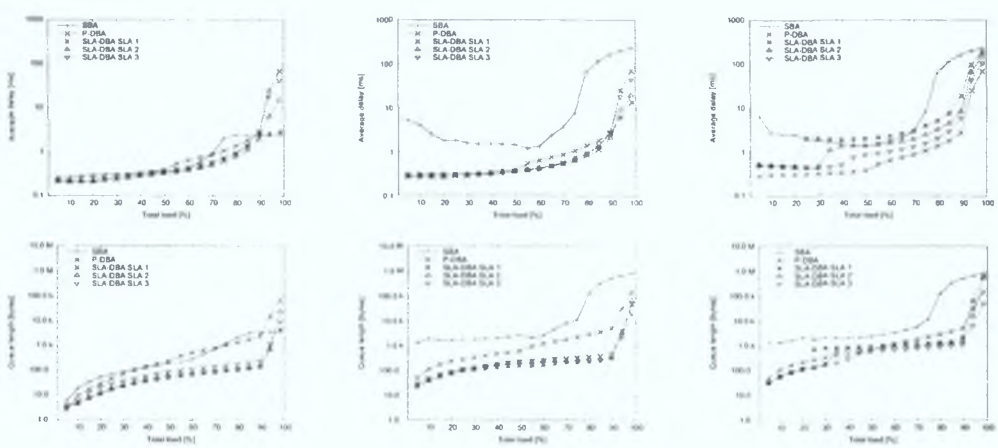
Figure E.6: Cycle Length Comparison. Traffic Mix 3. Maximum Cycle Length 5 ms.

E.2 Performance in Normal Conditions



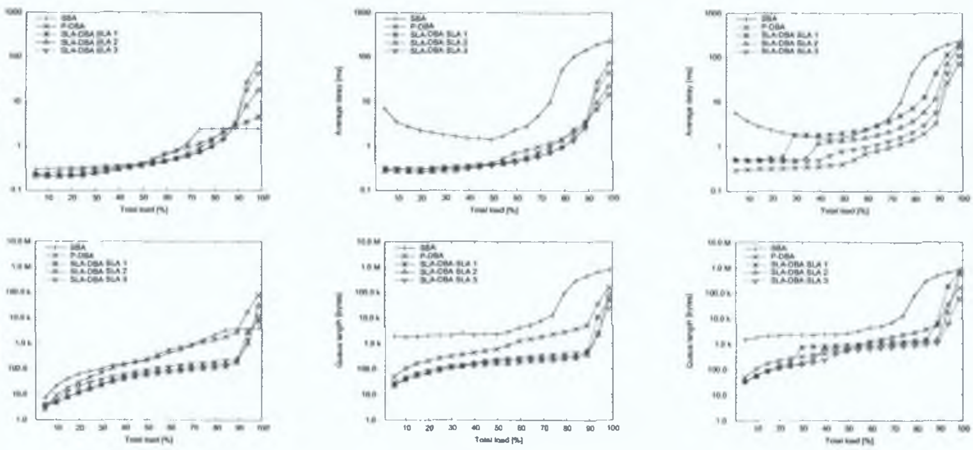
(a) EF (b) AF (c) BE

Figure E.7: Algorithm Performance Comparison. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



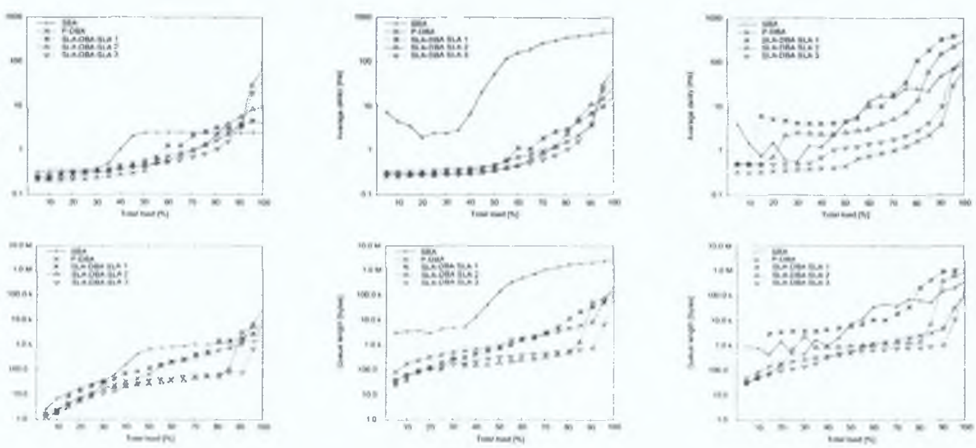
(a) EF (b) AF (c) BE

Figure E.8: Algorithm Performance Comparison. Traffic Mix 1. Source Set II. Maximum Cycle Length 5 ms.



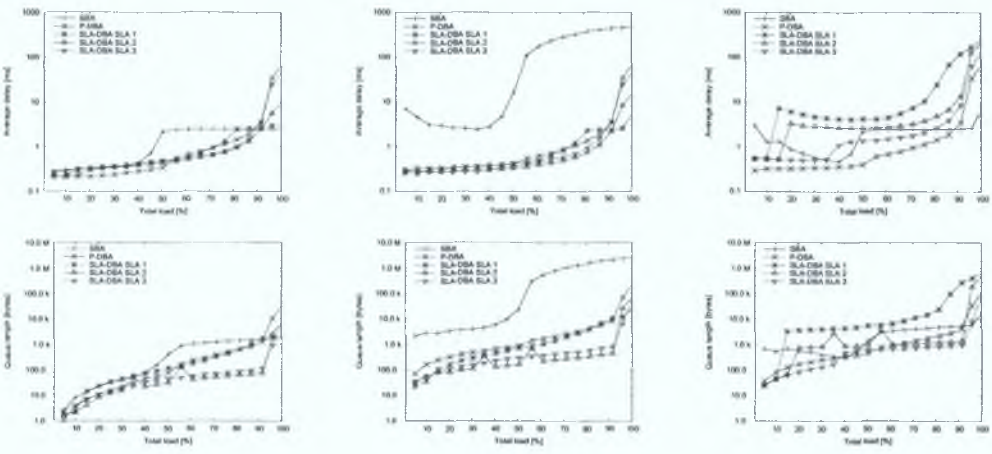
(a) EF (b) AF (c) BE

Figure E.9: Algorithm Performance Comparison. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



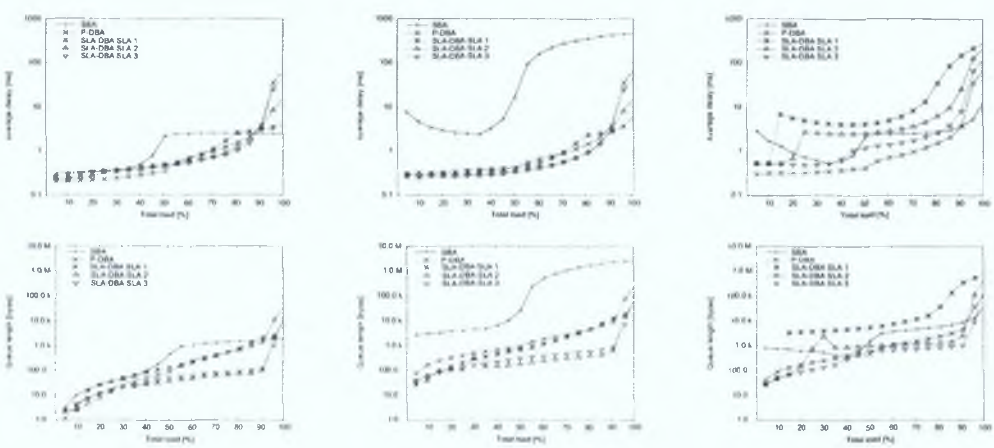
(a) EF (b) AF (c) BE

Figure E.10: Algorithm Performance Comparison. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



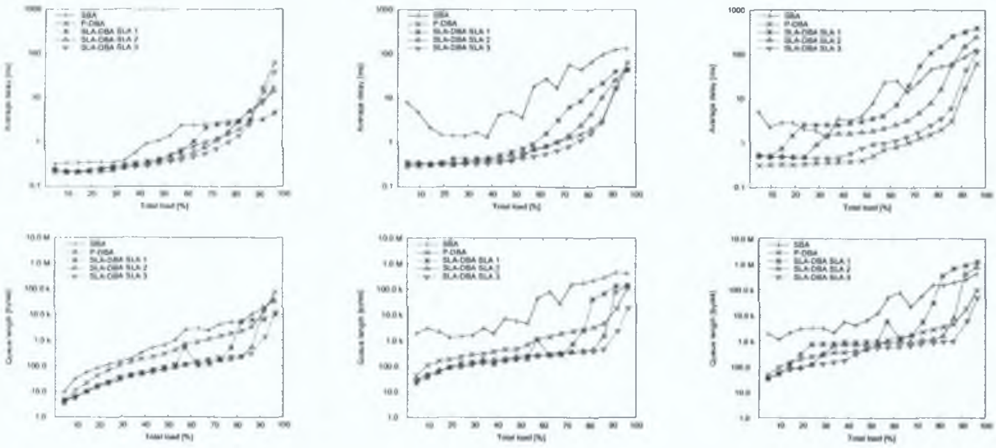
(a) EF (b) AF (c) BE

Figure E.11: Algorithm Performance Comparison. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



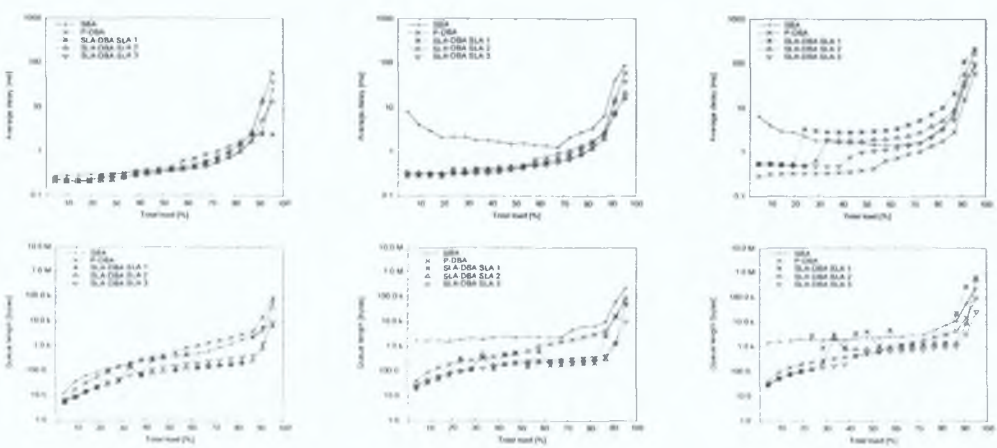
(a) EF (b) AF (c) BE

Figure E.12: Algorithm Performance Comparison. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



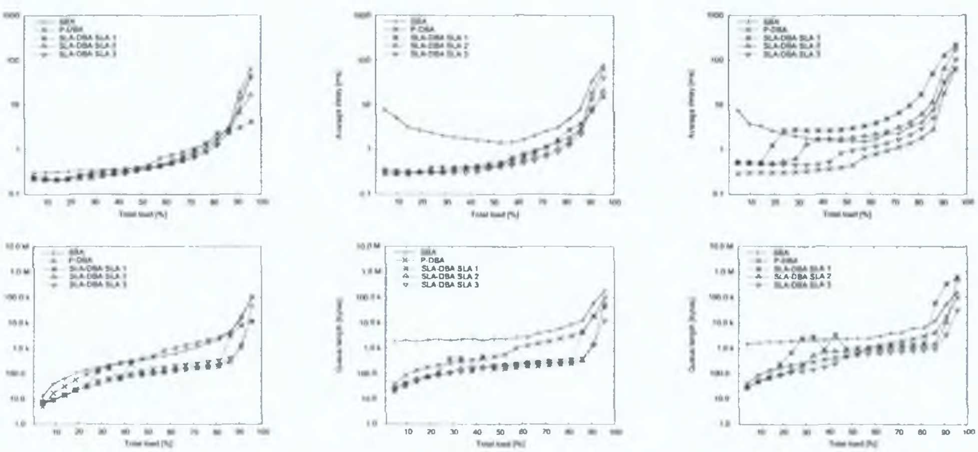
(a) EF (b) AF (c) BE

Figure E.13: Algorithm Performance Comparison. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

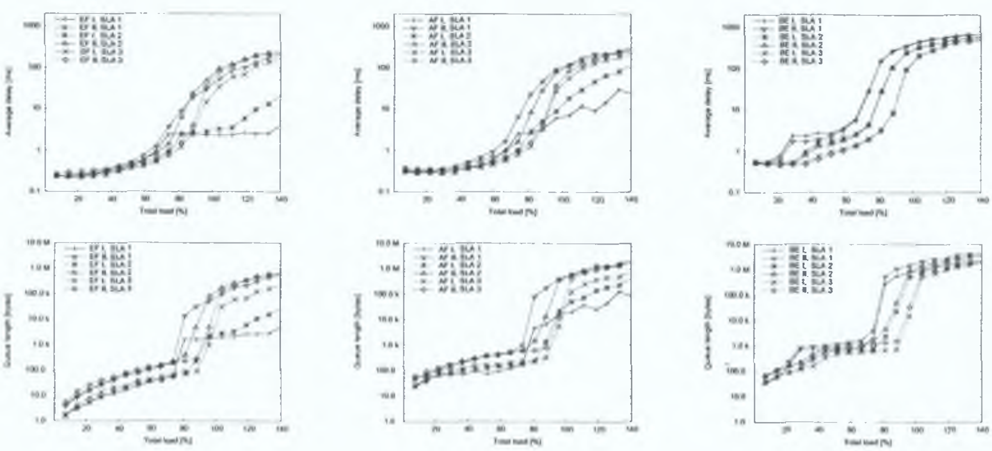
Figure E.14: Algorithm Performance Comparison. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

Figure E.15: Algorithm Performance Comparison. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

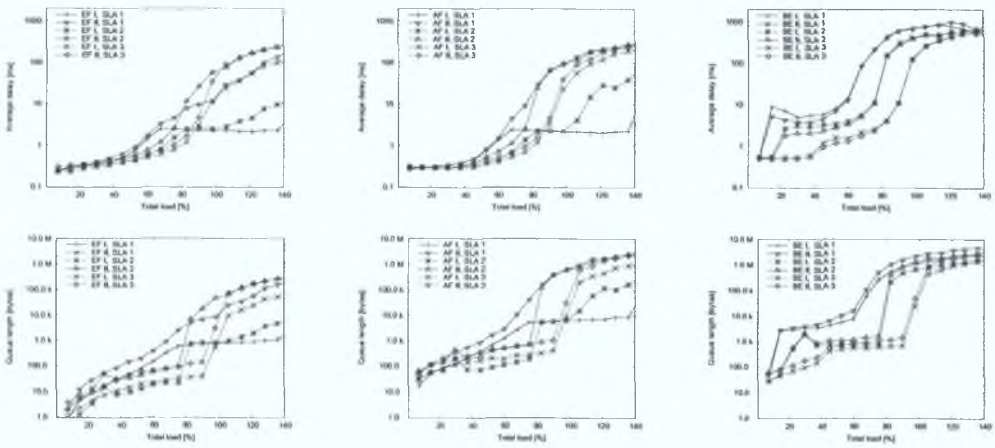
E.3 Performance in Congested Networks



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

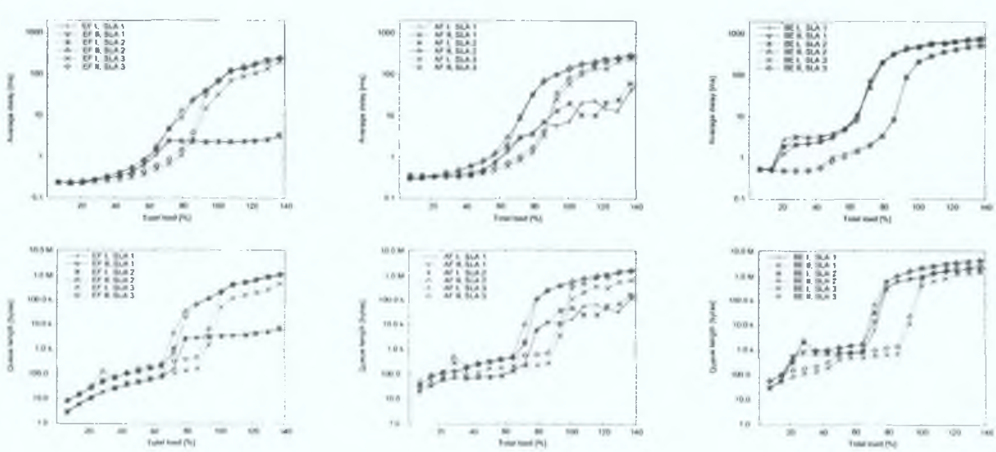
Figure E.16: Algorithm Performance Comparison . Congested Network. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

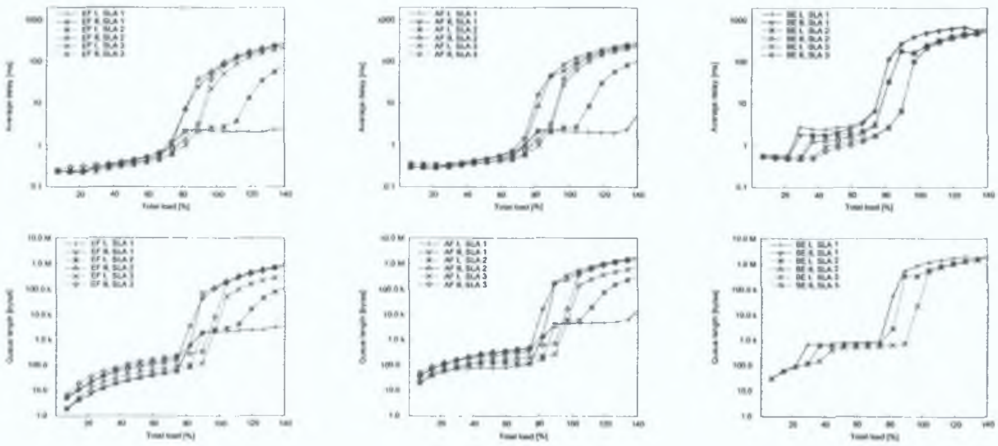
Figure E.17: Algorithm Performance Comparison . Congested Network. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

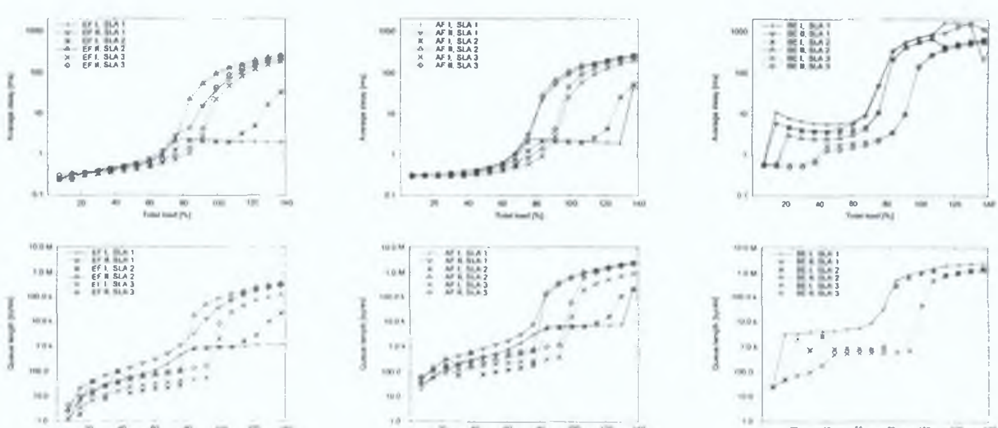
Figure E.18: Algorithm Performance Comparison . Congested Network. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

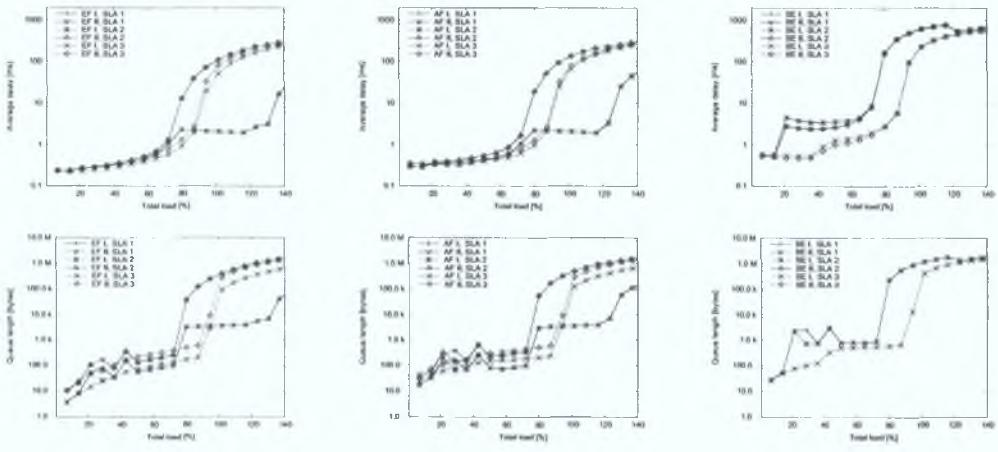
Figure E.19: Algorithm Performance Comparison . Congested Network. Traffic Mix 1. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

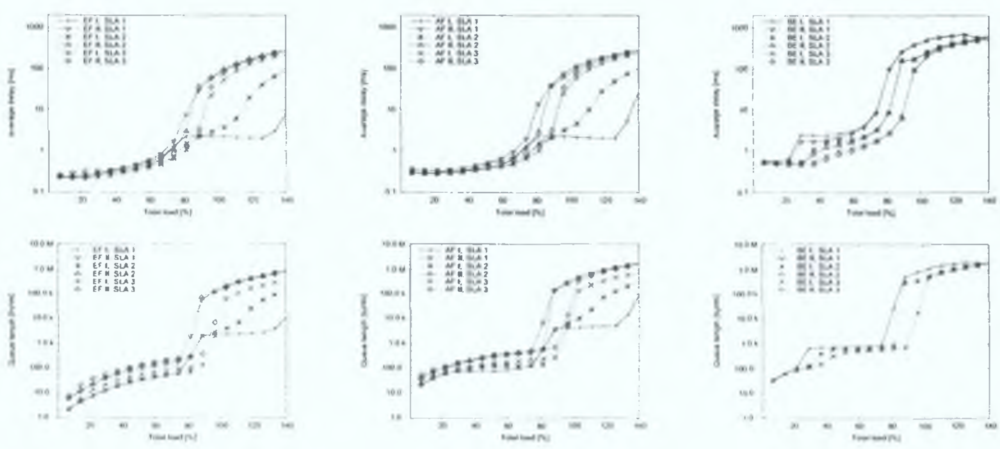
Figure E.20: Algorithm Performance Comparison . Congested Network. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

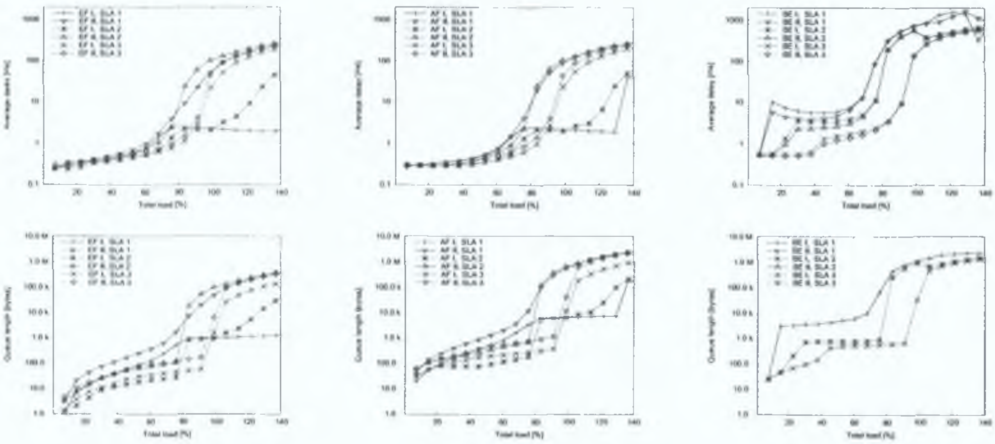
Figure E.21: Algorithm Performance Comparison . Congested Network. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

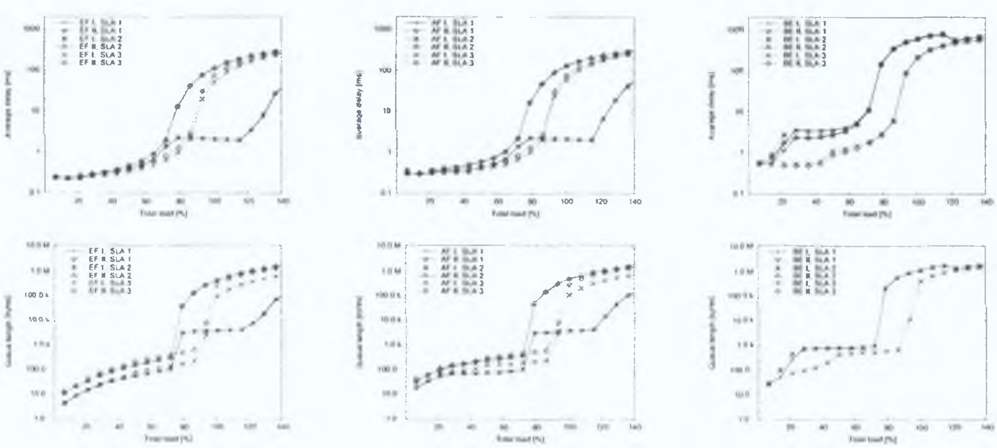
Figure E.22: Algorithm Performance Comparison . Congested Network. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure E.23: Algorithm Performance Comparison . Congested Network. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



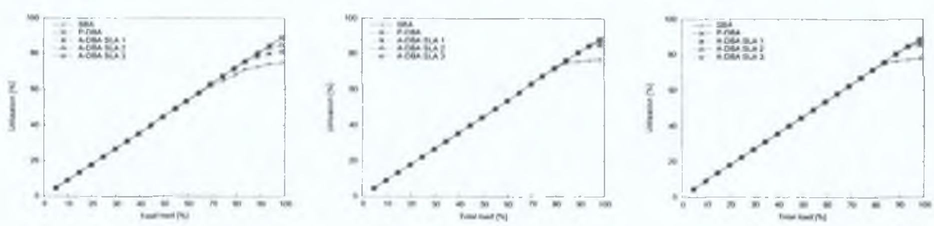
(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure E.24: Algorithm Performance Comparison . Congested Network. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

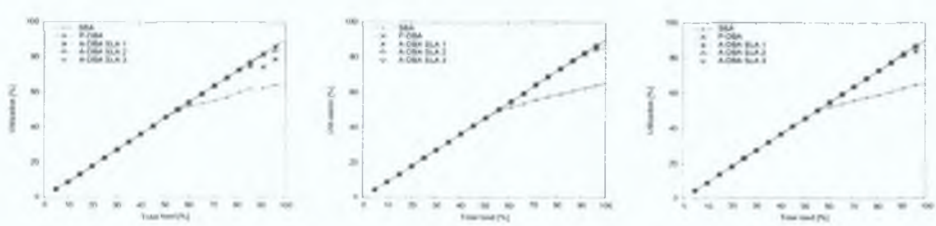
Appendix F
The Performance of the A-DBA Algorithm

F.1 Throughput and Cycle Length Comparison



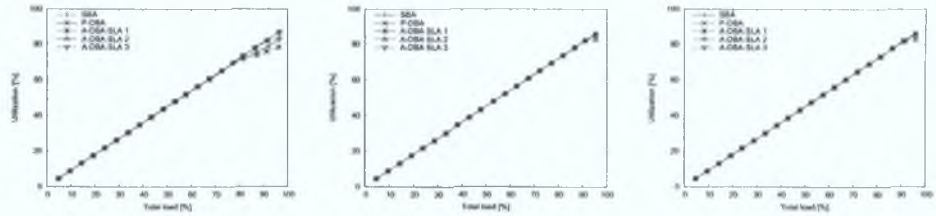
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure F.1: Throughput comparison. Traffic Mix I. Maximum cycle length5 ms.



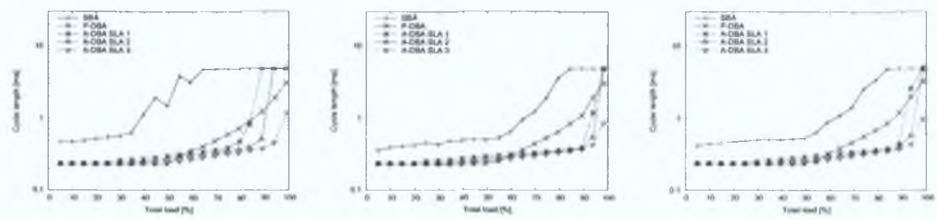
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure F.2: Throughput comparison. Traffic Mix II. Maximum cycle length 5 ms.



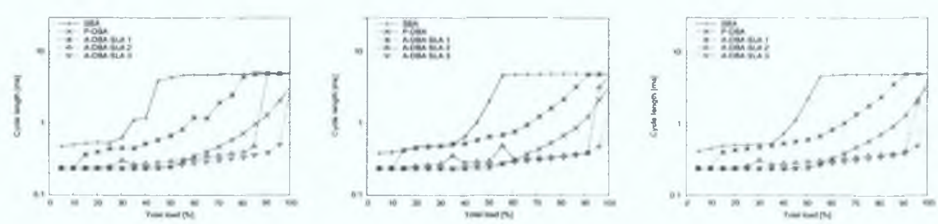
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure F.3: Throughput comparison. Traffic Mix III. Maximum cycle length 5 ms.



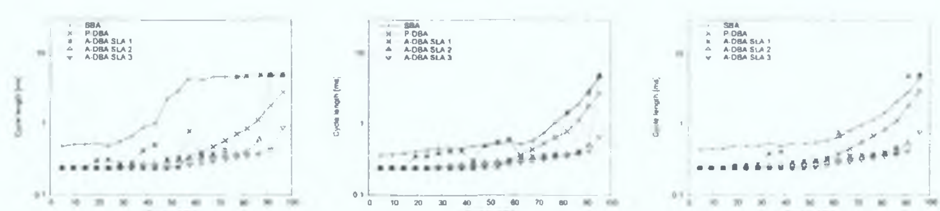
(a) Sources Set I (b) Sources Set II (c) Sources Set III

Figure F.4: Cycle length comparison. Traffic Mix I. Maximum cycle length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

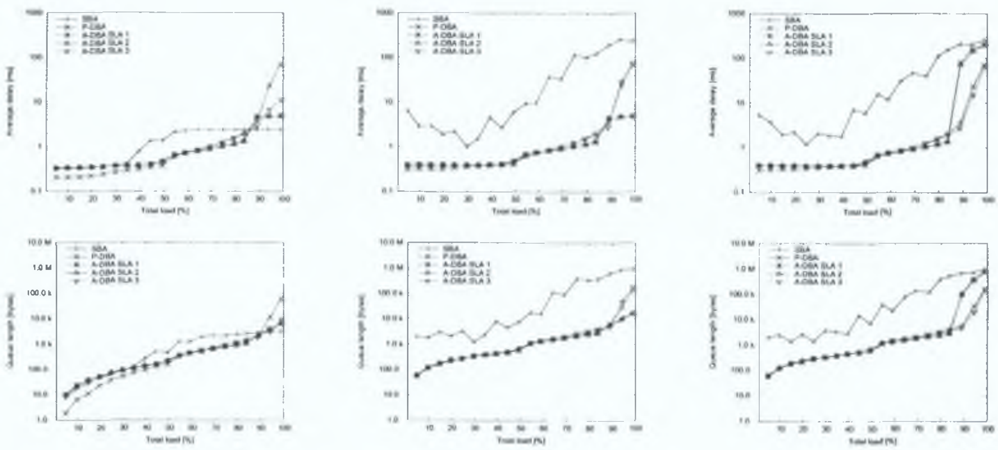
Figure F.5: Cycle length comparison. Traffic Mix II. Maximum cycle length 5 ms.



(a) Sources Set I (b) Sources Set II (c) Sources Set III

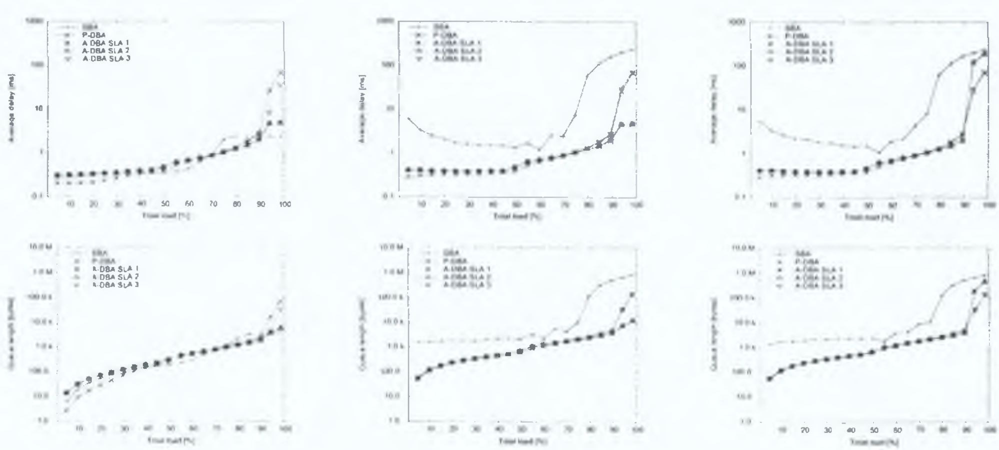
Figure F.6: Cycle length comparison. Traffic Mix III. Maximum cycle length 5 ms.

F.2 Performance under Normal Conditions



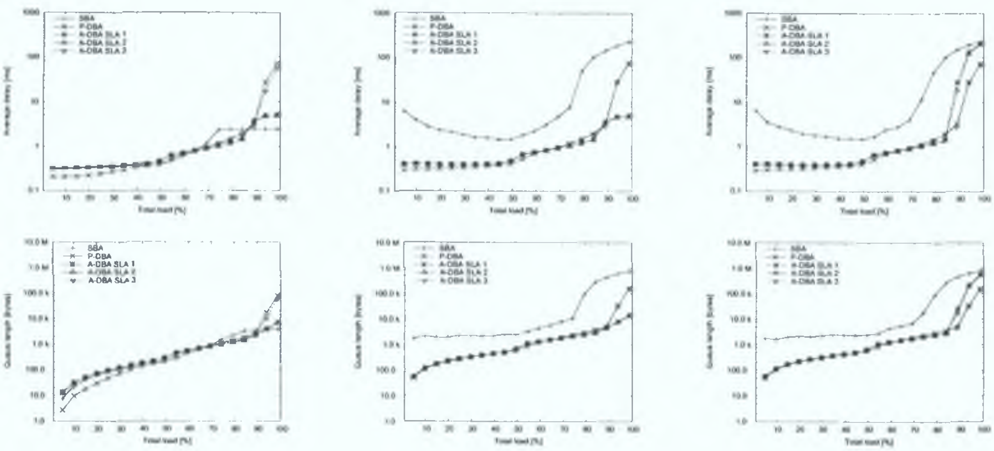
(a) EF (b) AF (c) BE

Figure F.7: Algorithms performance comparison. Traffic Mix I. Sources Set I. Maximum cycle length 5 ms.



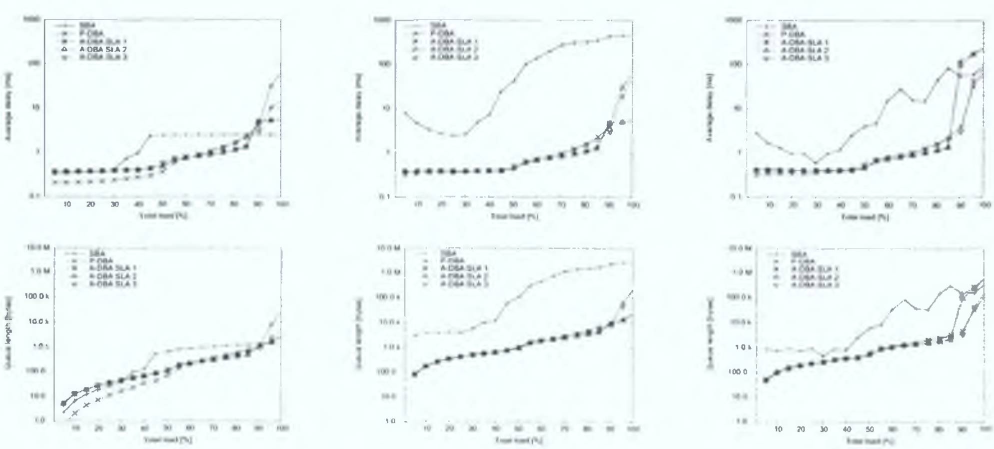
(a) EF (b) AF (c) BE

Figure F.8: Algorithms performance comparison. Traffic Mix I. Sources Set II. Maximum cycle length 5 ms.



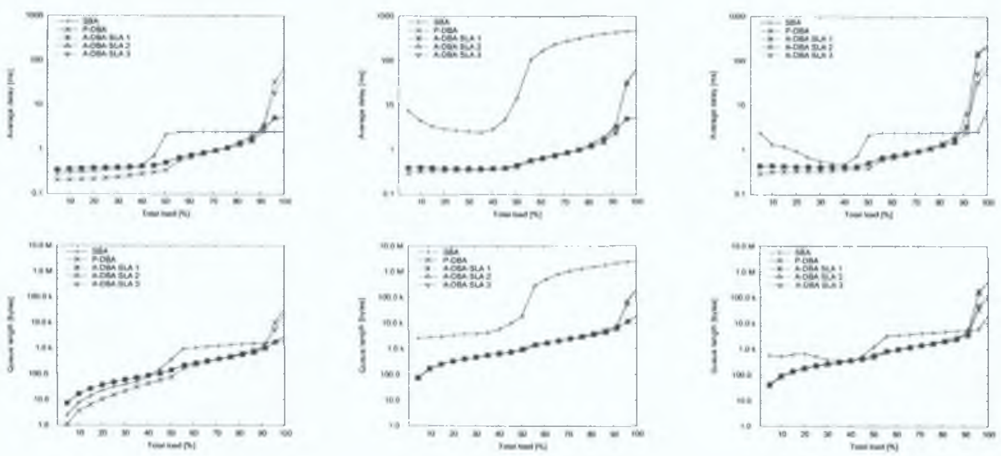
(a) EF (b) AF (c) BE

Figure F.9: Algorithms performance comparison. Traffic Mix I. Sources Set III. Maximum cycle length 5 ms.



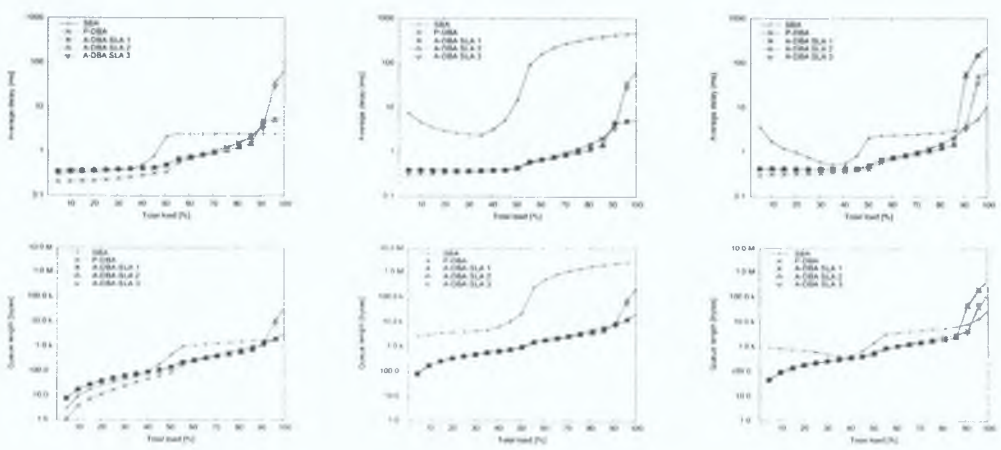
(a) EF (b) AF (c) BE

Figure F.10: Algorithms performance comparison. Traffic Mix II. Sources Set I. Maximum cycle length 5 ms.



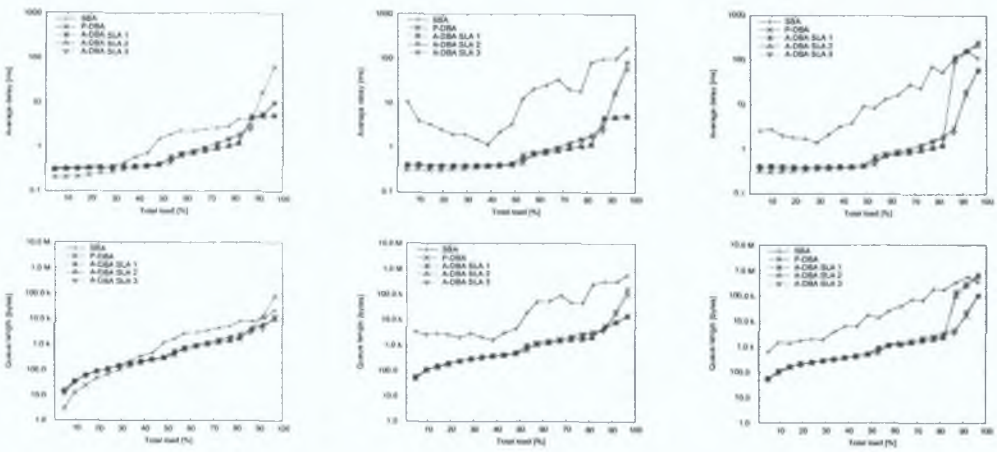
(a) EF (b) AF (c) BE

Figure F.11: Algorithms performance comparison. Traffic Mix II. Sources Set II. Maximum cycle length 5 ms.



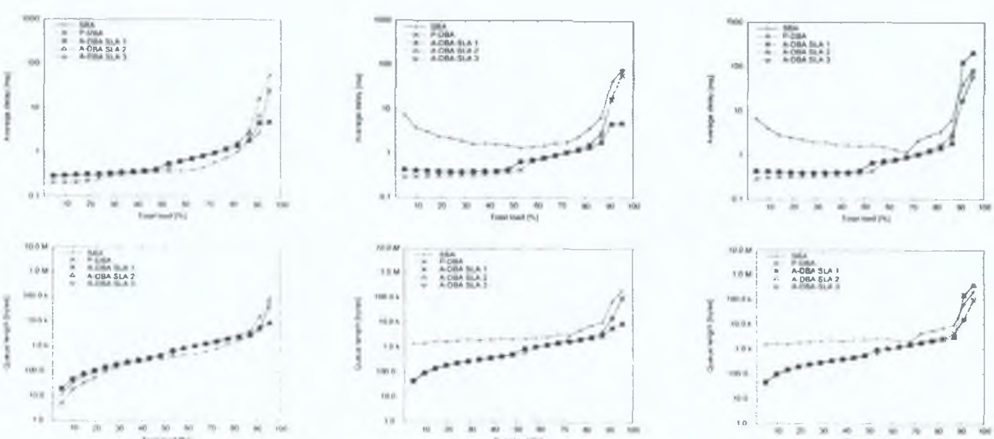
(a) EF (b) AF (c) BE

Figure F.12: Algorithms performance comparison. Traffic Mix II. Sources Set III. Maximum cycle length 5 ms.



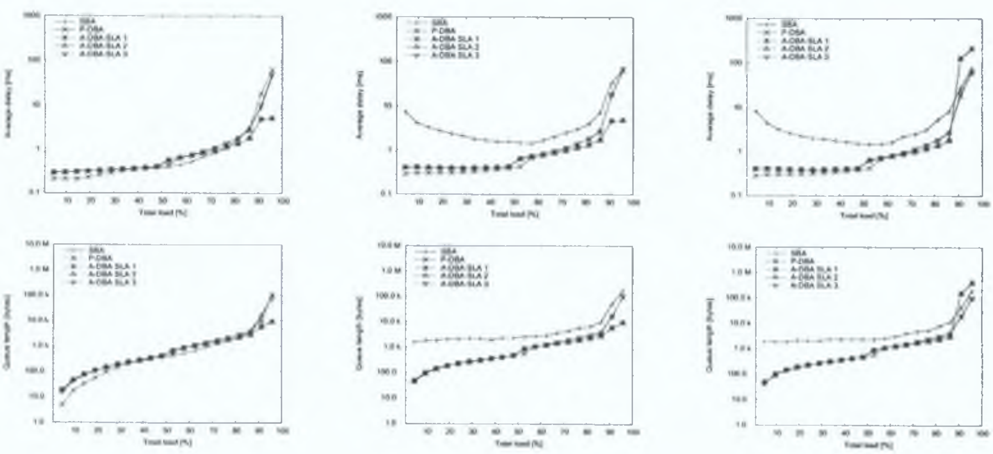
(a) EF (b) AF (c) BE

Figure F.13: Algorithms performance comparison. Traffic Mix III. Sources Set I. Maximum cycle length 5 ms.



(a) EF (b) AF (c) BE

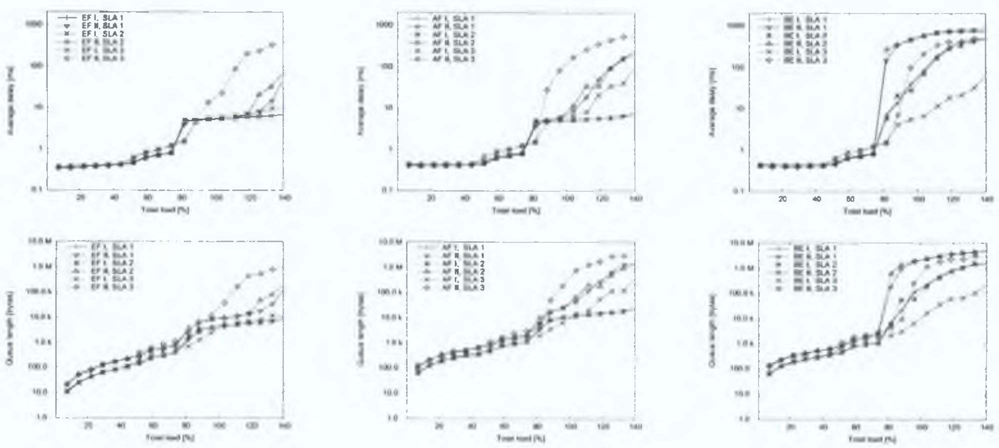
Figure F.14: Algorithms performance comparison. Traffic Mix III. Sources Set II. Maximum cycle length 5 ms.



(a) EF (b) AF (c) BE

Figure F.15: Algorithms performance comparison. Traffic Mix III. Sources Set III. Maximum cycle length 5 ms.

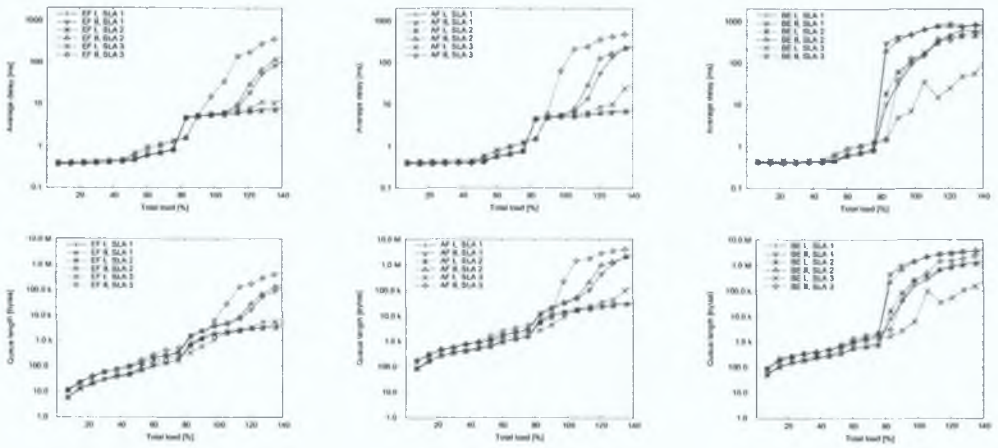
F.3 Performance in Congested Networks



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

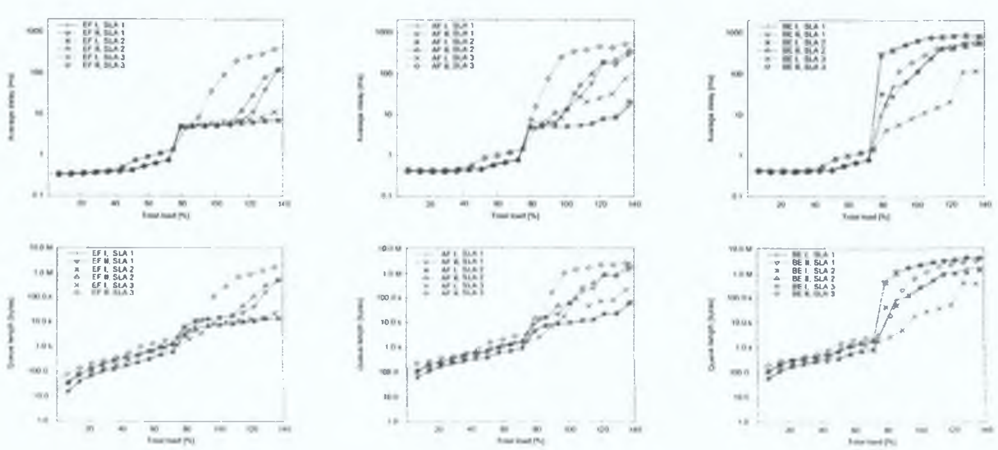
Figure F.16: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

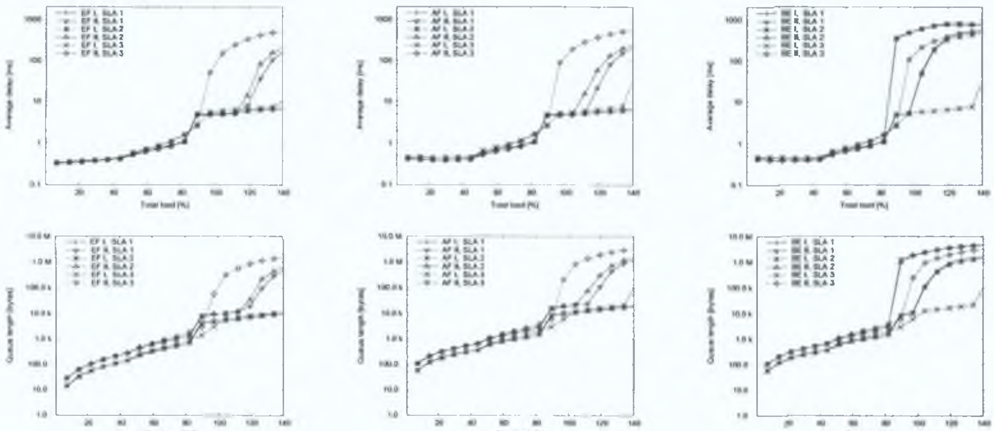
Figure F.17: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

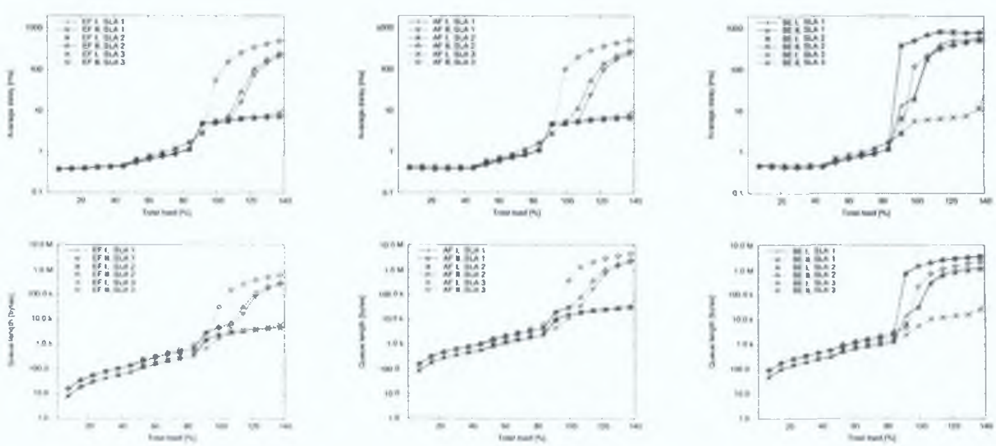
Figure F.18: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

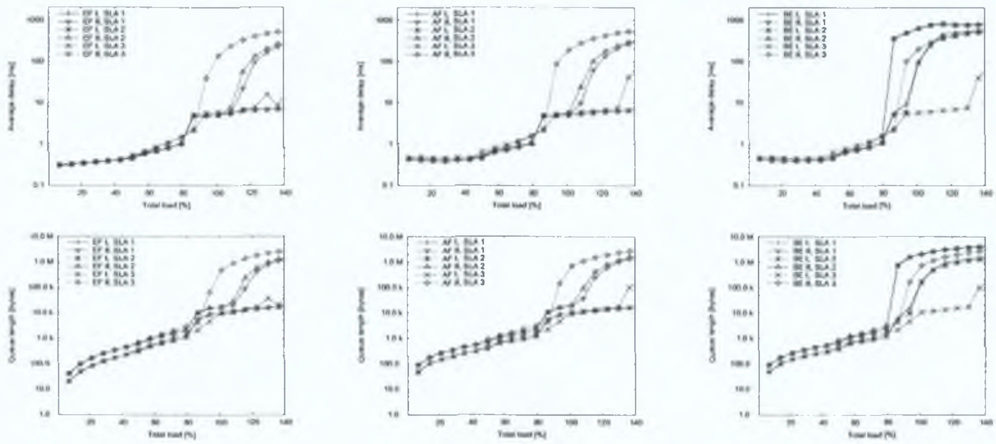
Figure F.19: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

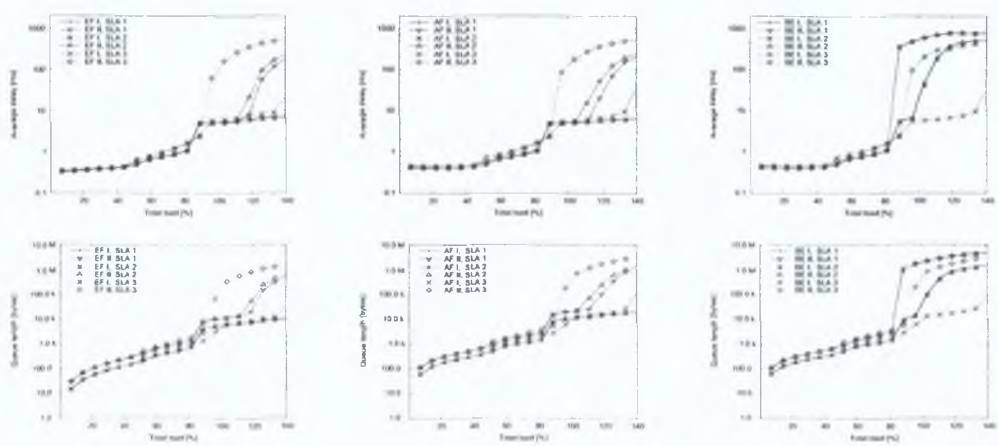
Figure F.20: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

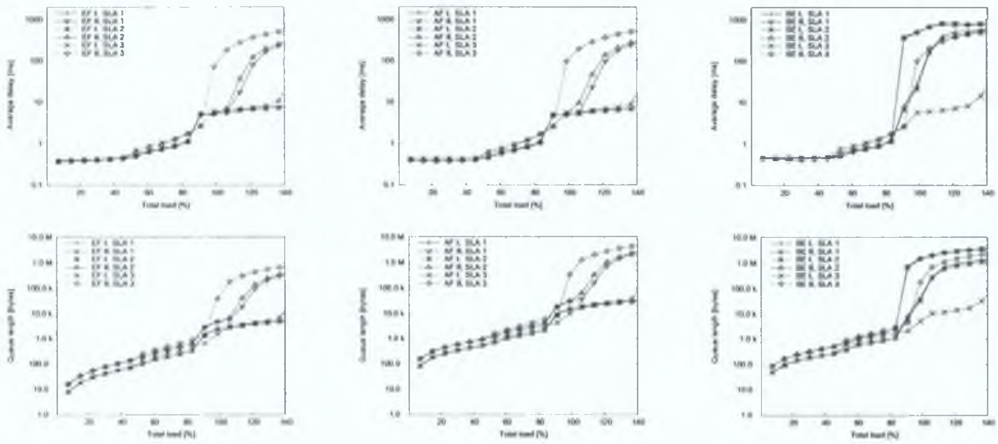
Figure F.21: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

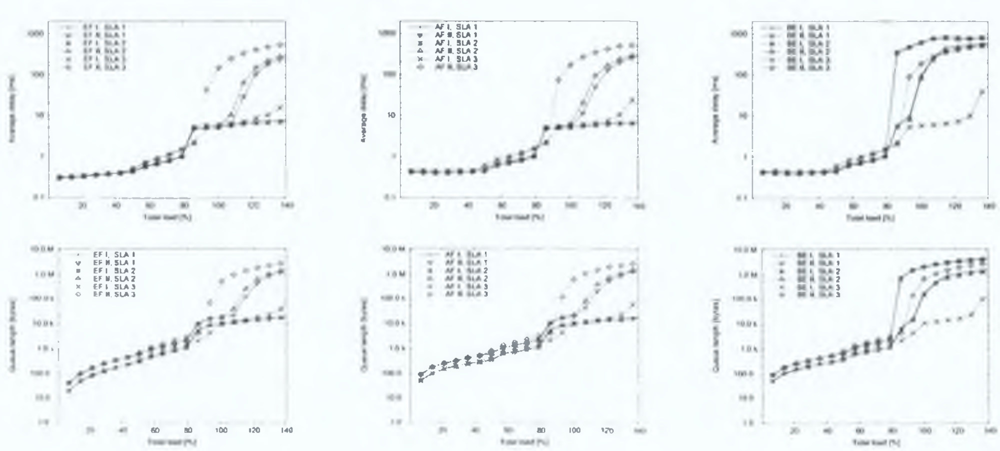
Figure F.22: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure F.23: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

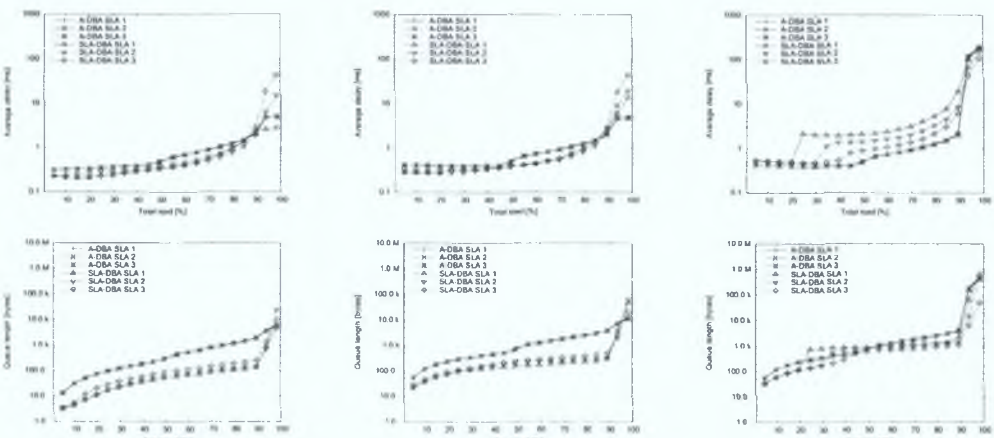
I ONUs sending at the normal rate II ONUs sending at the doubled rate

Figure F.24: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

Appendix G

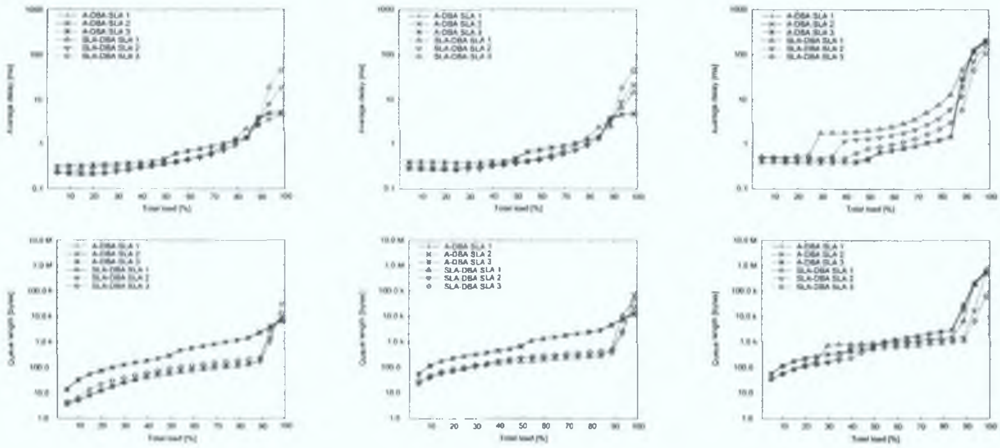
SLA-DBA and A-DBA Comparison

G.1 Normal Conditions



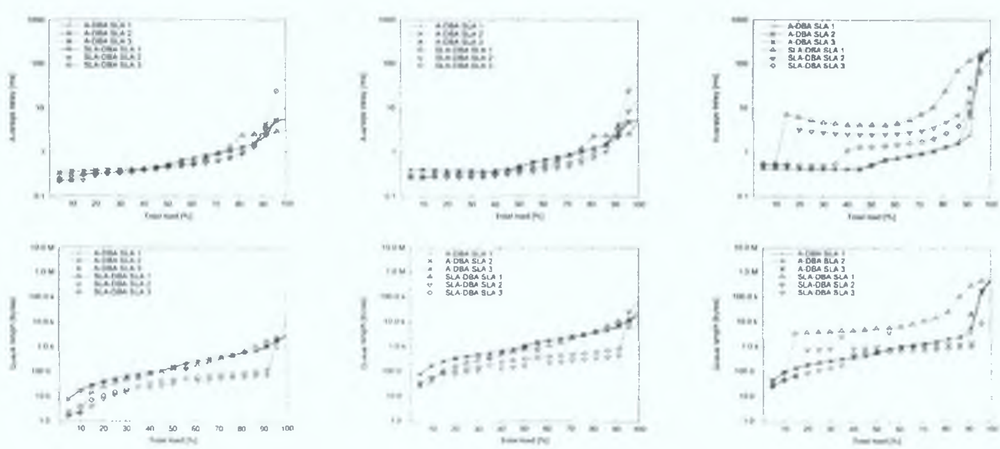
(a) EF (b) AF (c) BE

Figure G.1: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 1. Source Set II. Maximum cycle length 5 ms.



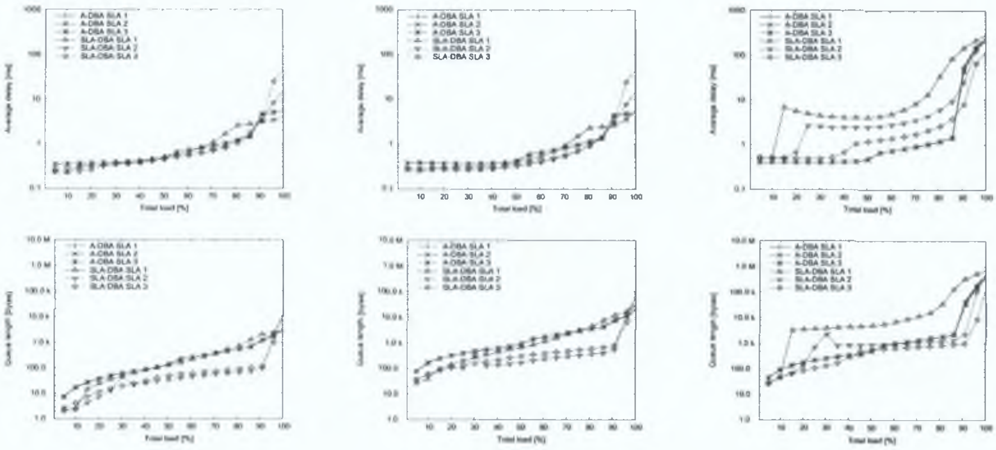
(a) EF (b) AF (c) BE

Figure G.2: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



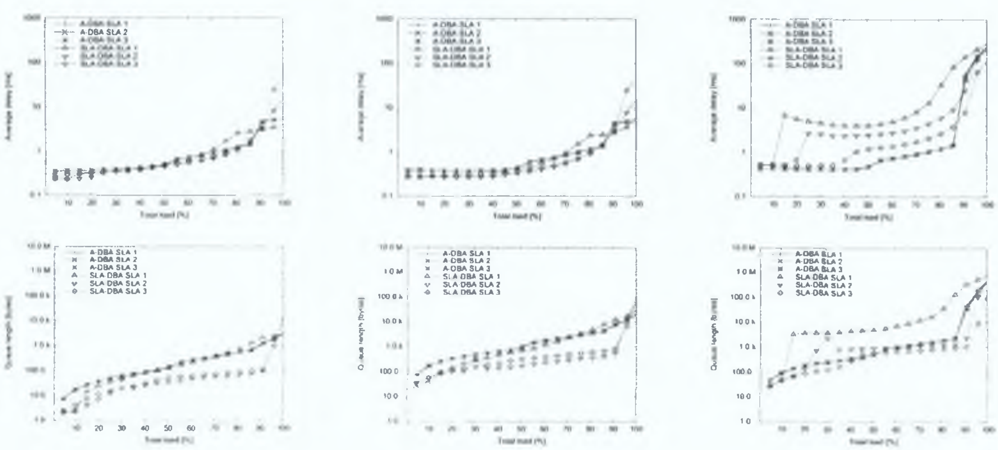
(a) EF (b) AF (c) BE

Figure G.3: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



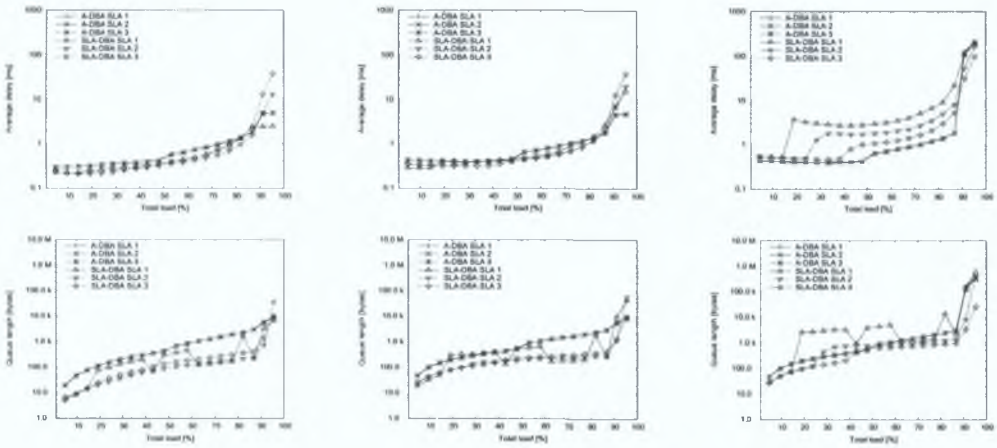
(a) EF (b) AF (c) BE

Figure G.4: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



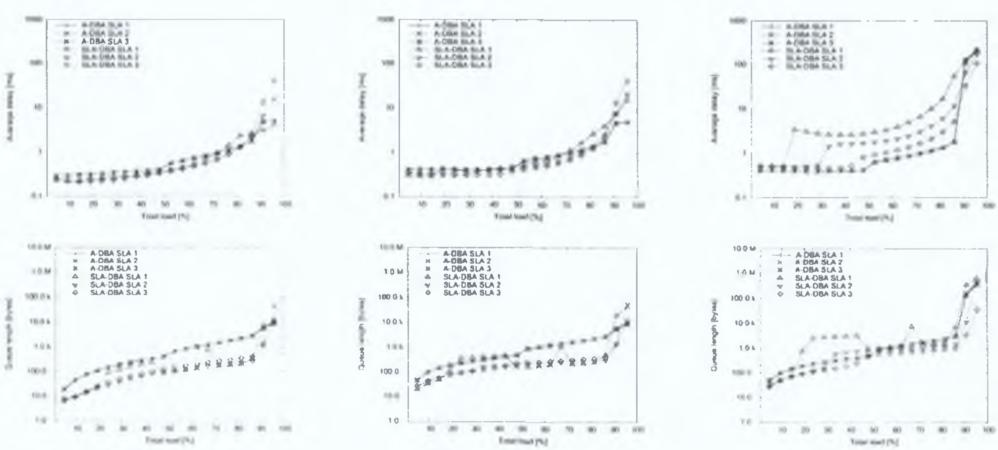
(a) EF (b) AF (c) BE

Figure G.5: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



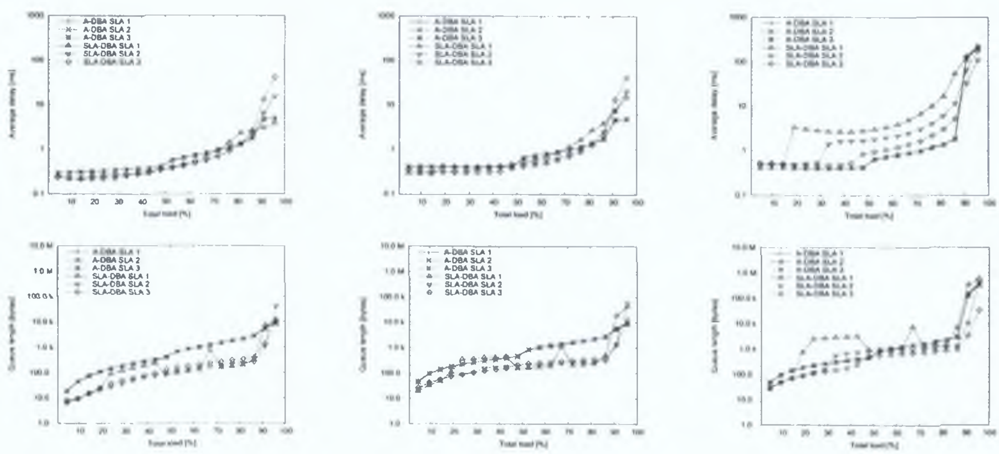
(a) EF (b) AF (c) BE

Figure G.6: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

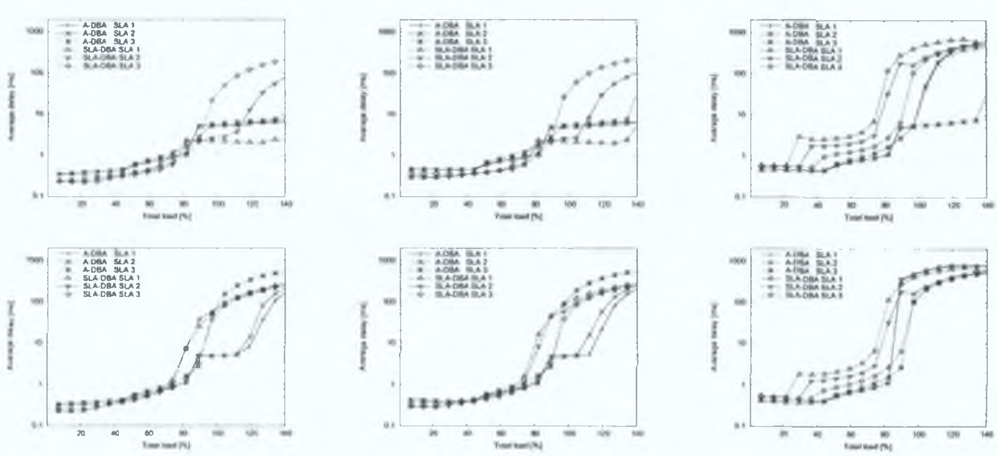
Figure G.7: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

Figure G.8: Algorithm Performance Comparison. Normal Conditions. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

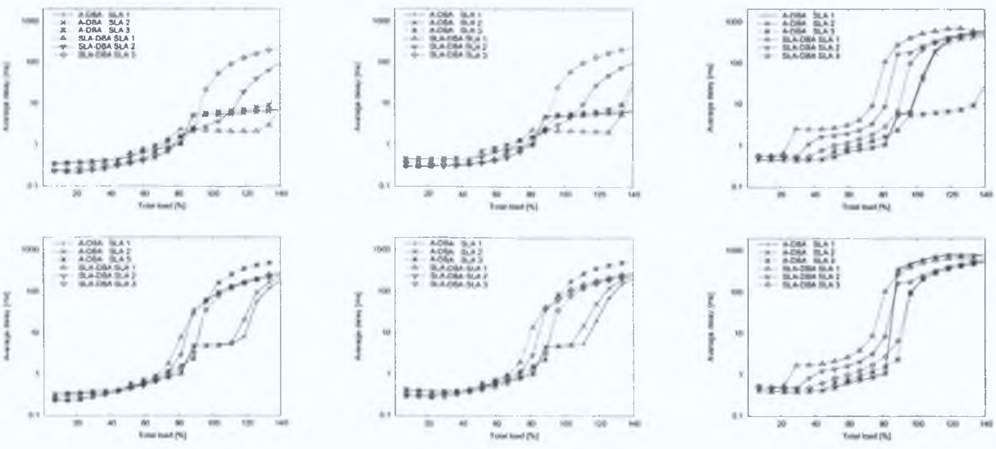
G.2 Congested Networks



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

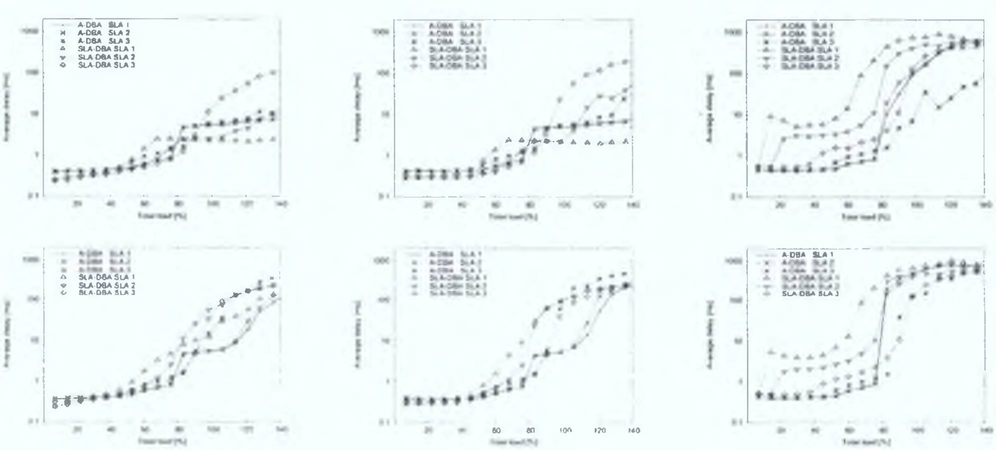
Figure G.9: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

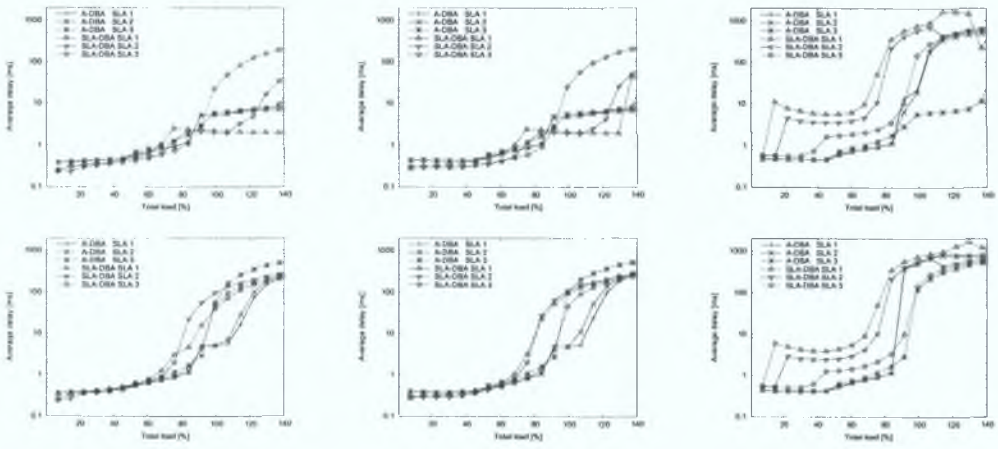
Figure G.10: Algorithm Performance Comparison. Congested Network. Traffic Mix 1. Source Set III. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

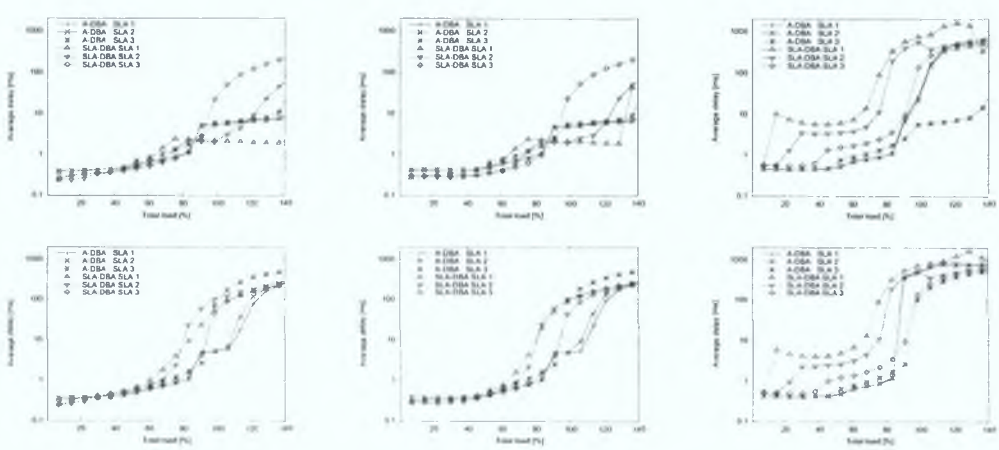
Figure G.11: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

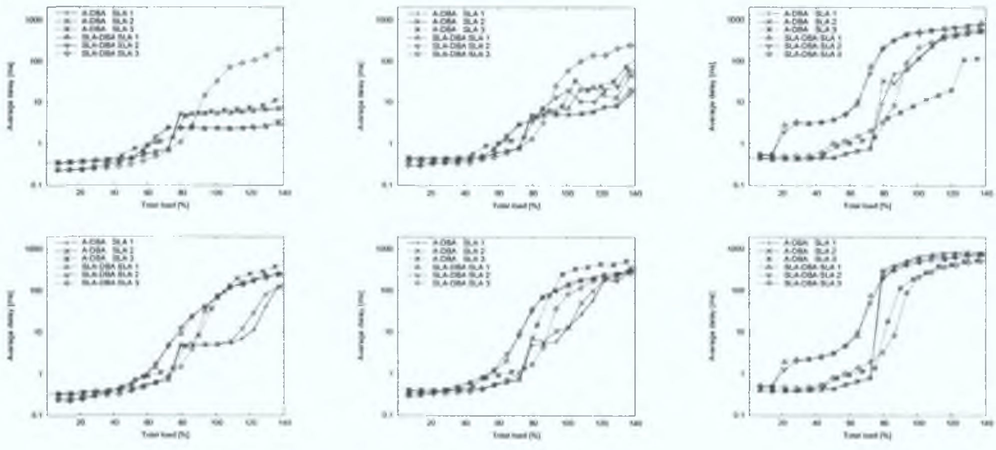
Figure G.12: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

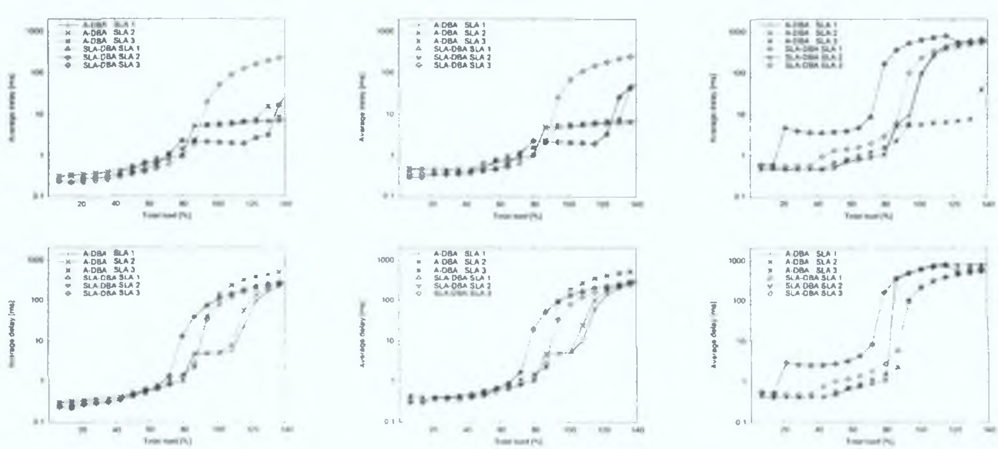
Figure G.13: Algorithm Performance Comparison. Congested Network. Traffic Mix 2. Source Set III. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

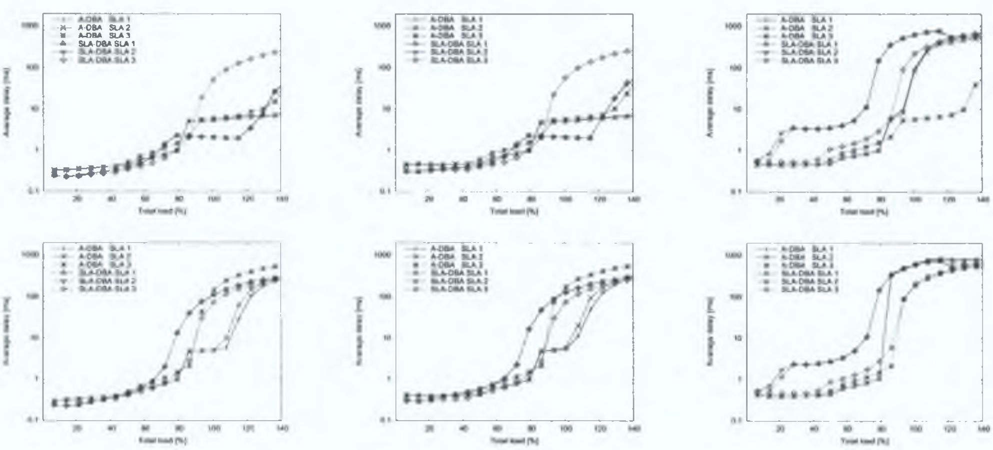
Figure G.14: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set I. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

Figure G.15: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set II. Maximum Cycle Length 5 ms.



(a) EF (b) AF (c) BE

top ONUs sending at the normal rate bottom ONUs sending at the doubled rate

Figure G.16: Algorithm Performance Comparison. Congested Network. Traffic Mix 3. Source Set III. Maximum Cycle Length 5 ms.

List of Figures

List of Tables

Abbreviations

AES Advanced Encryption Standard

AF Assured Forwarding

A-DBA Adaptive Dynamic Bandwidth Allocation

ATM Asynchronous Transfer Mode

BPON Broadband Passive Optical Network

BE Best Effort

BG Bandwidth Guaranteed

CRC Cyclic Redundancy Check

DA Destination Address

DBA Dynamic Bandwidth Allocation

DDBA Delta DBA

DiffServ Differentiated Services

DSL Digital Subscriber Loop

EF Expedited Forwarding

EFMA Ethernet in the First Mile Alliance

EPON Ethernet Passive Optical Network

FCA Fiber Cable Attenuation

FCS Frame Check Sequence

FSAN Full Service Network Access

FTTH Fiber-to-the-Home

GFP Generic Framing Procedure

GIANT GigaPON Access Network

GPON Gigabit Passive Optical Network

HSSR Hybrid Slot-Size/Rate DBA

IFG Interframe Gap

IEEE Institute of Electric and Electronic Engineers

IETF Internet Engineering Task Force

IPACT Interleaved Polling with Adaptive Cycle Time

ISDN Integrated Services Digital Network

ITU International Telecommunications Union

LCG Linear Congruential Method

LAN Local Area Network

LLC Logical Link Control

MAC Media Access Control

MPCP Multi-Point Control Protocol

NOBEL Next Generation Optical Networks for Broadband European Leadership

MUSE Multi Service Access Everywhere

not-BG Bandwidth Not Guaranteed

OAM Operations and Maintenance

OLT Optical Line Terminal

ONU Optical Network Unit

OSI Open Systems Interconnection

Opcode Operation Code

PON Passive Optical Network

QoS Quality of Service

RED Random Early Detection

SA Source Address

SBA Static Bandwidth Allocation

SL Splitter Loss

SLA Service Level Agreement

SLA-DBA Service Level Agreement aware Dynamic Bandwidth Allocation

SP-DBA Strict Priority Dynamic Bandwidth Allocation

SS Source Set

P-DBA Proportional Dynamic Bandwidth Allocation

RTT Round Trip Time

TLBA Two-Layer Bandwidth Allocation

T-CON Transmission Container

TDMA Time Division Multiplexing Access

TM Traffic Mixture

WDM Wavelength Division Multiplexing

Index

Bibliography

[1] J.R. Stern, J.W. Ballance, D.W. Faulkner, S. Hornung, and D.B. Payne, "Passive Optical Local Networks for Telephony Applications and Beyond," *Electronics Letters*, vol. 23, no. 24, pp. 1255–1257, Nov. 1987.

[2] Donald. E.A. Clarke and Tetsuya Kanada, "Broadband: The Last Mile," *IEEE Communications Magazine*, vol. 31, no. 3, pp. 94–100, Mar. 1993.

[3] Yih-Kang Maurice Lin, Dan R. Spears, and Mih Yin, "Fiber-Based Local Access Network Architectures," *IEEE Communications Magazine*, vol. 27, no. 10, pp. 64–73, Oct. 1989.

[4] Borgar T. Olsen, Alcibiade Zaganiaris, Kjell Stordahl, L. Aa. Ims, D. Myhre, T. Øverli, M. Tahkokorpi, I. Welling, M. Drieskens, J. Kraushaar, J. Mononen, M. Lähteenoja, S.Markatos, M. De Bortoli, U. Ferrero, M. Ravera, S. Balzaretti, F. Fleuren, N. Gieschen, M. De Oliveira Duarte, and E. de Castro, "Techno-Economic Evaluation of Narrowband and Broadband Access Network Alternatives and Evolution Scenario Assessment," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 6, Aug. 1996.

[5] OECD – Organisation For Economic Co-Operation and Development, "OECD Communications Outlook," June 2003.

[6] Yoichi Maeda, "Broadband Access Deployment in NTT," http://www.fsanweb.org/Atlanta_Workshop.asp, Mar. 2003, FSAN PON Workshop, Atlanta, USA.

[7] Yoichi Maeda, "Overwiew of Optical Broadband in Japan," in *Proceedings of ECOC*, Sept. 2004, vol. 1.

[8] Bob Whitman, "Fibre Access Deployment Worldwide: Market Drivers, Politics and Technology Choices," in *Proceedings of ECOC*, Sept. 2004, pp. 6–9.

[9] Jeon Hongbeom, "Broadband access, service developments and trends in Korea," in *Proceedings of ECOC*, Sept. 2004, vol. 5.

[10] Chu Hwan Yim, "Convergence Acess Network in Korea," http://www.itu.int/ITU-T/worksem/asna/index.html, June 2004, All Star Network Access Workshop, Geneva, Switzerland.

[11] Jae-Woo Yang, "Broadband convergence Network (BcN) - a Korean National Broadband Project," in *Proceedings of OECC/COIN*, July 2004, pp. 50–51.

[12] Yong-Kyung Lee and Dongmyun Lee, "Broadband Access in Korea: Experience and Future Perspective," *IEEE Communications Magazine*, vol. 41, no. 12, pp. 30–36, Dec. 2003.

[13] Michael Kunigonis, "International Broadband Deployments. Lessons learned around the globe.," http://www.apectelwg.org/apecdata/telwg/28tel/bw/telwg28-BB-12.pdf, Oct. 2003, APEC Telecommunications and Information Working Group, Taipei, Taiwan.

[14] Commision Of The European Communities, "eEurope 2005: An information society for all," http://www.europa.eu.int/information_society/eeurope/2005/index_en.htm, June 2002.

[15] Andrew Houghton, "Supporting the Rollout of Broadband in Europe: Optical Network Research in the IST Program," *IEEE Communications Magazine*, vol. 41, no. 9, pp. 58–64, Sept. 2003.

[16] IDATE, "Ftth situation in europe," www.idate.org, Jan. 2005.

[17] Jonathan Doran, "Mass-Market Fiber Remains Distant on the European Horizon," www.yankeegroup.com, Dec. 2004.

[18] John A. Jay, "An Overview of International Fiber to the Home Deployment," in *Proceedings of FTTH Conference*, Oct. 2002.

[19] Lawrence K. Vanston, "The Local Exchange Network in 2015," www.tfi.com, 2001.

[20] Sam Paltridge, "Industry in Crisis (?)," in *Proceedings of CEPT Conference*, 2003.

[21] Dana Bisaro, "FTTH-"HFC vs. PON"," in *Proceedings of FTTH Conference*, Oct. 2002.

[22] Patrick T. Garvey, "Economics of FTTH," in *Proceedings of FTTH Conference*, Oct. 2004.

[23] J. Rosenberg, H. Salama, and M. Squire, "RFC 3219 – Telephony Routing over IP (TRIP)," Jan. 2002, http://www.ietf.org/rfc/rfc3219.txt.

[24] D. Ferrari, "RFC 1193 – Client requirements for real-time communication services," Nov. 1990, http://www.ietf.org/rfc/rfc1193.txt.

[25] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W.Weiss, "RFC 2475 – An Architecture for Differentiated Services," Dec. 1998, http://www.ietf.org/rfc/rfc2475.txt.

[26] K. Nichols and B. Carpenter, "RFC 3086 – Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification," Apr. 2001, http://www.ietf.org/rfc/rfc3086.txt.

[27] Paul W. Shumate and Richard K. Snelling, "Evolution of Fiber in the Residential Loop Plant," *IEEE Communications Magazine*, vol. 29, no. 3, pp. 68–74, Mar. 1991.

[28] Thomas R. Rowbotham, "Local Loop Developments in the U.K.," *IEEE Communications Magazine*, vol. 29, no. 3, pp. 50–59, Mar. 1991.

[29] ITU-T, "G.983.1 - Broadband Passive Optical Networks (BPON): General characteristics," June 1999.

[30] David Faulkner, Rajendrakumar Mistry, Tom Rowbotham, Kenji Okada, Wsewolod Warzanskyj, Albert Zylbersztejn, and Yves Picault, "The Full Services Access Networks Initiative," *IEEE Communications Magazine*, vol. 35, no. 4, pp. 58–68, Apr. 1997.

[31] Yoichi Maeda, Kenji Okada, and David Faulkner, "FSAN OAN-WG and future issues for broadband optical access networks," *IEEE Communications Magazine*, vol. 39, no. 12, pp. 126–132, Dec. 2001.

[32] Ingrid Van de Voorde and Gert Van der Plas, "Full Service Optical Access Networks: ATM Transport on Passive Optical Networks," *IEEE Communications Magazine*, vol. 35, no. 4, pp. 70–75, 1997.

[33] Hiromi Ueda, Kenji Okada, Brian Ford, Glenn Mahony, Stephen Hornung, David Faulkner, Jacques Abiven, Sophie Durel, Ralph Ballart, and John Erickson, "Deployment status and common technical specifications for a B-PON system," *IEEE Communications Magazine*, vol. 39, no. 12, pp. 134–141, Dec. 2001.

[34] David Kettler, Hank Kafka, and Dan Spears, "Driving Fiber to the Home," *IEEE Communications Magazine*, vol. 38, no. 11, pp. 106–110, Nov. 2000.

[35] ITU-T, "G.984.1 - Gigabit-capablePassive Optical Networks (GPON): General characteristics," Mar. 2003.

[36] Lothar Zier, Wolfgang Fischer, and Frank Brockners, "Ethernet-based public communication services: Challenge and opportunity," *IEEE Communications Magazine*, vol. 42, no. 3, pp. 88–95, Mar. 2004.

[37] The Institute of Electrical and Electronics Engineers, "IEEE 802.3ah - IEEE Standard for Local and Metropolitan Area Networks. Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications. Amendment: Media Access Control Parameters, Physical Layers, and Management Parameters for Subscriber Access Networks," Sept. 2004.

[38] Ivan Andonovic and Deepak Uttamchandani, Eds., *Principles Of Modern Optical Systems*, chapter 11, Artech House, 1989.

[39] ITU-T, "G.983.4 - A Broadband optical access system with increased service capability using dynamic bandwidth assignment," Nov. 2001.

[40] Frank J. Effenberger, Hiroshi Ichibangase, and Haruo Yamashita, "Advances in Broadband Passive Optical Networking Technologies," *IEEE Communications Magazine*, vol. 39, no. 12, pp. 118–124, Dec. 2001.

[41] G.Kramer, B.Mukherjee, and G.Pesavento, "Interleaved Polling with Adaptive Cycle Time IPACT: A Dynamic Bandwidth Distribution Scheme in an Optical Access Networks," *Photonic Network Communication*, vol. 3, no. 3, July 2001.

[42] Jyh-Horng Wen and Jin-Fu Chang, "T Effect of Multipath Interference on the Performance of Packet Radios," *IEEE Transactions on Communications*, vol. 38, no. 6, pp. 740–743, June 1990.

[43] ITU-T, "G.984.2 - Gigabit-capablePassive Optical Networks (GPON): Physical Media Dependend (PMD) layer specification," Mar. 2003.

[44] Yu-Li Hsueh, Matthew S. Rogge, Wei-Tao Shaw, Leonid G. Kazovsky, and Shu Yamamoto, "SUCCESS-DWA: A Highly Scalable and Cost-effective Optical Access Network," *IEEE Communications Magazine*, vol. 42, no. 8, pp. 24–30, Aug. 2004.

[45] Dong Jae Shin, Dae Kwang Jung, Hong Seok Shin andJin Wook Kwon, Seong-taek Hwang, Yunje Oh, and Changsup Shim, "Hybrid WDM/TDM-PON With Wavelength-Selection-Free Transmitters," *IEEE/OSA Journal of Lightwave Technology*, vol. 23, no. 1, pp. 187–195, Jan. 2005.

[46] The Institute of Electrical and Electronics Engineers, "IEEE 802 - IEEE Standard for Local and Metropolitan Area Networks: Overview and Architecture," Mar. 2002.

[47] The Institute of Electrical and Electronics Engineers, "ANSI/IEEE 802.1D - IEEE Standard for Local and Metropolitan Area Networks: Media Access Control (MAC) Bridges," 1998.

[48] The Institute of Electrical and Electronics Engineers, "IEEE 802.1Q - IEEE Standard for Local and Metropolitan Area Networks: Virtual Bridged Local Area Networks," May 2003.

[49] G.Kramer, B.Mukherjee, and A.Maislos, "Ethernet Passsive Optical Networks (EPON)," in *IP Over WDM: Building the Next Generation Internet*, S. Dixit, Ed., chapter 8. John Wiley & Sons, Inc, Mar. 2003.

[50] G.Kramer, B.Mukherjee, S.Dixit, Y.Ye, and R.Hirth, "On supporting differentiated classes of service in Ethernet passive optical network," *OSA Journal of Optical Networking*, vol. 1, no. 8/9, pp. 280–298, 2002.

[51] G.Kramer, B.Mukherjee, and G.Pesavento, "IPACT: A Dynamic Bandwidth Distribution Scheme in an Ethernet PON(EPON)," *IEEE Communications Magazine*, vol. 40, no. 2, pp. 66–73, 2002.

[52] Kae Hsiang Kwong, David Harle, and Ivan Andonovic, "Dynamic Bandwidth Allocation Algorithm for Differentiated Services over WDM EPONs," in *The Ninth International Conference on Communications Systems*, Sept. 2004, pp. 116–120.

[53] Su-il Choi, "Cyclic Polling-Based Dynamic Bandwidth Allocation for Differentiated Classes of Service in Ethernet Passive Optical Networks," *Photonic Network Communications*, vol. 7, no. 1, pp. 87–96, 2004.

[54] Ch.M.Assi, Y.Ye, S.Dixit, and M.A.Ali, "Dynamic Bandwidth Allocation for Quality-of-Service Over Ethernet PONs," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 9, pp. 1467–1477, Nov. 2003.

[55] Ch.M.Assi, Y.Ye, and S.Dixit, "Support of QoS in IP-Based Ethernet-PON," in *Proceedings of IEEE GLOBECOM*, Dec. 2003, vol. 22, pp. 3737–3741.

[56] D. Nikolova, B. Van Houdt, and C. Blondia, "Dynamic bandwidth allocation algorithms in EPON:a simulation study," in *OptiComm*, 2003, pp. 369–380.

[57] Yeon-Mo Yang, Ji-Myong Nho, and Byung-Ha Ahn, "An Enhanced Burst-Polling based Delta Dynamic Bandwidth Allocation Scheme for QoS over E-PONs," in *ACM Next-Generation Residential Broadband Challenges*, Oct. 2004.

[58] Ho-Sook Lee, Tae-Whan Yoo, Ji-Hyun Moon, and Hyeong-Ho Lee, "A Two-Step Scheduling Algorithm to Support Dual Bandwidth Allocation Ppolicies in an Ethernet Passive Optical Network," *ETRI Journal*, vol. 26, no. 2, pp. 185–188, Apr. 2004.

[59] Yongquing Zhu, Maode Ma, and Tee Hiang Cheng, "A novel multiple access scheme for Ethernet Passive Optical Networks," in *GLOBECOM 2003 - IEEE Global Telecommunications Conference*, Dec. 2003, vol. 22, pp. 2649–2653.

[60] Fu-Tai An, Yu-Li Hsueh, Kyeong Soo Kim, Ian M. White, and Leonid G. Kazovsky, "A New Dynamic Bandwidth Allocation Protocol with Quality of Service in Ethernet-based Passive Optical Networks," in *OFC*, Mar. 2003, vol. 1, pp. 134–135.

[61] Jing Xie, Shengming Jiang, and Yuming Jiang, "A Dynamic Bandwidth Allocation Scheme for Differentiated Services in EPONs," *IEEE Communications Magazine*, vol. 42, no. 8, pp. 32–33, Aug. 2004.

[62] Kyuho Son, Hyungkeun Ryu, Song Chong, and Taewhan Yoo, "Dynamic Bandwidth Allocation Schemes to Improve Utilization under Non-Uniform Traffic in Ethernet Passive Optical Networks," in *Proceedings of IEEE ICC*, 2004.

[63] G.Kramer and B.Mukherjee, "Ethernet PON: design and analysis of an optical access network," *Photonic Network Communication*, vol. 3, no. 3, pp. 307–319, July 2001.

[64] Sally Floyd and Van Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 3, pp. 397–413, 1993.

[65] Mark Wurtzler, "Analysis and Simulation of Weighted Random Early Detection (WRED) Queues," M.S. thesis, University of Kansas, Lawrence, Kansas, USA, 2003.

[66] N. Ghani, A. Shami, C.Assi, and Y. Raja, "Intra-ONU Bandwidth Scheduling in Ethernet Passive Optical Networks," *IEEE Communications Letters*, vol. 8, no. 11, pp. 683–685, Nov. 2004.

[67] H. Zhang, "Service Disciplines for Guaranteed Packet Performance Service in Packet-Switching Networks," in *Proceedings of the IEEE*, Oct. 1995.

[68] Pawan Goyal, Harrick M. Vin, and Haichen Cheng, "Start-time fair queueing: A scheduling algorithm for integrated services packet switching networks," in *IEEE/ACM Transactions on Networking*, Oct. 1997, vol. 5, pp. 690–704.

[69] N. Ghani and J. W. Mark, "Hierarchical Scheduling for Integrated ABR/VBR Services in ATM Networks," in *Proceedings of IEEE GLOBECOM*, Oct. 1997.

[70] Sherif R. Sherif, Antonis Hadjiantonis, Georgios Ellinas, Chadi Assi, and Mohamed A. Ali, "A Novel Decentralized Ethernet-Based PON Access Architecture for Provisioning Differentiated QoS," *IEEE/OSA Journal of Lightwave Technology*, vol. 22, no. 11, pp. 2483–2497, Nov. 2004.

[71] Joseph L. Hammond and Peter J.P. O'Reilly, *Performance Analysis of Local Computer Networks*, Addison-Wesley Publishing Company, Mar. 1988.

[72] Yishay Mansour and Boaz Patt-Shamir, "Jitter Control in QoS Networks," *IEEE/ACM Transactions on Networking*, vol. 9, no. 4, pp. 492–502, Aug. 2001.

[73] Walter Willinger, Murad S. Taqqu, Robert Sherman, and Daniel V. Vilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," *IEEE/ACM Transactions on Networking*, vol. 5, no. 1, pp. 71–86, Feb. 1997.

[74] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the Self-Similar Nature of Ethernet Traffic (extended version)," *IEEE/ACM Transactions on Networking*, vol. 2, no. 1, pp. 1–15, Feb. 1994.

[75] Vern Paxson and Sally Floyd, "Wide area traffic: the failure of Poisson modeling," *IEEE/ACM Transactions on Networking*, vol. 3, no. 3, pp. 226–244, June 1995.

[76] Jerry Banks, Ed., *Handbook of Simulation. Principles, Methodology, Advances, Applications, and Practice*, John Wiley & Sons, Inc., 1998.

[77] Jerry Banks, John S. Carson, II, and Barry L. Nelson, *Discrete-Event System Simulation*, Prentice Hall, 1999.

[78] Reuven Y. Rubinstein and Benjamin Melamed, *Modern Simulation and Modelling*, John Wiley & Sons, Inc., 1998.

[79] James F. Kurose and Hussein T. Mouftah, "Computer-Aided Modelling Analysis, and Design of Communication Networks," *IEEE Journal on Selected Areas in Communications*, vol. 6, no. 1, pp. 130–145, Jan. 1998.

[80] Hussein T. Mouftah and Rene P. Sturgeon, "Distributed Discrete Event Simulation for Communication Networks," *IEEE Journal on Selected Areas in Communications*, vol. 8, no. 9, pp. 1723–1734, Dec. 1990.

[81] Russel C.H. Cheng, *Handbook of Simulation. Principles, Methodology, Advances, Applications, and Practice*, chapter 5, John Wiley & Sons, Inc., 1998.

[82] Donald E. Knuth, *The Art of Computer Programming*, vol. 2, chapter 1, Addison-Wesley, 1997.

[83] Luc Devroye, *Non-Uniform Random Variate Generation*, chapter 2, Springer-Verlag, 1986.

[84] Makoto Matsumoto and Takkuji Nishimura, "Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator," *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, pp. 3–30, Jan. 1998.

[85] Pei-Chi Wu, "Random Number Generation with Primitive Pentanomials," *ACM Transactions on Modeling and Computer Simulation*, vol. 11, no. 4, pp. 346–351, Oct. 2001.

[86] Leonard Kleinrock, *Queueing Systems*, vol. 1, John WIley & Sons, Inc., 1975.

[87] Jan Beran, Robert Sherman, Murad S. Taqqu, and Walter Willinger, "Long-Range Dependence in Variable-Bit-Rate Video Traffic," *IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1566–1579, February/March/April 1995.

[88] Mark E. Crovella and Azer Bestavros, "Self-Similarity in World Wide Web Traffic: Evidence and Possible Causes," *IEEE/ACM Transactions on Networking*, vol. 5, no. 6, pp. 835–846, Dec. 1997.

[89] Kihong Park and Walter Willinger, Eds., *Self-Similar Network Ttraffic and Performance Evaluation*, John Wiley & Sons, Inc., 2000.

[90] Jean M. McManus and Keith W. Ross, "Video-on-Demand Over ATM: Constant-Rate Transmission and Transport," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 6, pp. 1087 – 1098, Aug. 1996.

[91] Soung C. Liew and Derek Chi yin Tse, "A control-theoretic approach to adapting VBR compressed video for transport over a CBR communications channel," *IEEE/ACM Transactions on Networking*, vol. 6, no. 1, pp. 42–55, Feb. 1998.

[92] The Institute of Electrical and Electronics Engineers, "IEEE 802.3 - IEEE Standard for Local and Metropolitan Area Networks. Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications.," Mar. 2003.

[93] The Institute of Electrical and Electronics Engineers, "IEEE 802.3ae - IEEE Standard for Local and Metropolitan Area Networks. Part 3: Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications. Amendment: Media Access Control (MAC) Parameters, Physical Layers, and Management Parameters for 10 Gb/s Operation," Aug. 2002.