Ethernet Passive Optical Network Dynamic Bandwidth Allocation Study

by

Du Zhao

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved November 2011 by the
Graduate Supervisory Committee:

Martin Reisslein, Chair
Michael McGarry
John Fowler

ARIZONA STATE UNIVERSITY

December 2011

ABSTRACT

Fiber-Wireless (FiWi) network is the future network configuration that uses optical fiber as backbone transmission media and enables wireless network for the end user. Our study focuses on the Dynamic Bandwidth Allocation (DBA) algorithm for EPON upstream transmission. DBA, if designed properly, can dramatically improve the packet transmission delay and overall bandwidth utilization. With new DBA components coming out in research, a comprehensive study of DBA is conducted in this thesis, adding in Double Phase Polling coupled with novel Limited with Share credits Excess distribution method. By conducting a series simulation of DBAs using different components, we found out that grant sizing has the strongest impact on average packet delay and grant scheduling also has a significant impact on the average packet delay; grant scheduling has the strongest impact on the stability limit or maximum achievable channel utilization. Whereas the grant sizing only has a modest impact on the stability limit; the SPD grant scheduling policy in the Double Phase Polling scheduling framework coupled with Limited with Share credits Excess distribution grant sizing produced both the lowest average packet delay and the highest stability limit.

DEDICATION

To my parents and all my family.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude and respect to all my committee members, especially my advisor, Dr. Martin Reisslein and Dr. Michael McGarry, for their invaluable guidance and support through my research, and Dr. Fowler, whose class gives me fundamental basis of this research. I would also like to thank my lab mates, who give me immediate help and encouragement when I have problems. I would also like to acknowledge the contribution from my family and friends who give me unconditional support.

TABLE OF CONTENTS

Page

LIST OF FIGURES vi

CHAPTER

1 INTRODUCTION 1
  1.1 Fiber-Wireless (FiWi) Networks 1
    1.1.1 RoF Networks and R&F Networks 2
    1.1.2 Example Testbeds 7
    1.1.3 SuperMan 9
  1.2 Ethernet Passive Optical Network (EPON) 12
    1.2.1 EPON Architecture 12
      1.2.1.1 Standard PON Architecture 11
      1.2.1.2 Alternative PON Architectures 13
    1.2.2 Downstream and Upstream Transmission 15
    1.2.3 Multi-Point Control Protocol (MPCP) 17
  1.3 Thesis Outline 18
2 DYNAMIC BANDWIDTH ALLOCATION 19
  2.1 Grant Scheduling Framework .20
    2.1.1 Offline Scheduleing Framework 22
    2.1.2 Online Scheduling Framework 23
    2.1.3 Online Just-In-Time Two-step Scheduling Framework . . .24
    2.1.4 Multi-Group Scheduling Framework 26
  2.2 Grant Sizing 28
    2.2.1 Fixed Size 28
    2.2.2 Gated Size 29

CHAPTER Page

2.2.3 Limited Grant Size 29

2.2.4 Limited Grant Sizing with Excess Distribution 30

2.2.5 Exhaustive Grant Sizing Using Queue Size Prediction . . 33

2.3 Grant Scheduling Policy 34

2.3.1 Inter ONU Scheduling 34

2.3.2 Intra ONU Scheduling 35

3 SIMULATION TOOL INTRODUCTION AND EXPERIMENT DESCRIPTION 37

4 RESULT ANALYSIS 40

4.1 Packet Delay 41

4.2 Stability Limit (i.e., Maximum Achievable Channel Utilization) 47

5 CONCLUSION AND FUTURE WORK 52

REFERENCE 54

LIST OF FIGURES

Figure Page

1 Radio-over-SMF Network Downlink Using EAMs for Different Radio Client Signals [25] 4

2 Simultaneous Modulation and Transmission of FTTH Baseband Signal and RoF RF Signal Using An External Integrated Modulator Consisting of Three Mach-Zehnder[13] 5

3 Georgia Institute of Technology RoF Field Demonstration of SD/HD Video Delivery Using 2.4 GHz and 60GHz Millimeter-wave Transmissions[3] 7

4 UC Davis R&F Testbed Integration of EPON and WMN for Voice, Video, and Data Traffic[16] 8

5 SuperMAN Architecture Integrating Next-generation WiFi Technologies with WDM EPON and Next-generation WiMAX Technologies with RPR[17] 9

6 Network Architecture of a PON with One Optical Line Terminal (OLT) and N = 5 Optical Network Units (ONUs), Each with A Different Round-Trip Time (RTT)[18]. 13

7 Broadcast PON Architecture [18] 14

8 Two-Stage PON Architecture [18] 15

9 EPON Downstream Transmission. 15

10 EPON Upstream Transmission 16

11 MPCP Operation: Two-way Messaging Assignment of Time Slots for Upstream Transmission Between ONU and OLT[18] 17

12 DBA Components Structure. 20

13 Offline Scheduling Framework [18] 23

14 Online Scheduling Framework [18] 24

15 Double Phase Polling Framework [2] 27

16 Multi-Thread Polling Framework [24] 27

17 Eponsim Simulation Manager Interface 37

Figure Page

18 Average Packet Delay for DBAs in EPON with Different Propagation Delay Ranges. 41

19 Stability Limit for DBAs in EPON with Different Propagation Delay Ranges. 47

CHAPTER 1

INTRODUCTION

1.1 Fiber-Wireless (FiWi) Networks

During recent years research in access network, Fiber-Wireless (FiWi) networks has become rapidly mature and given rise to new powerful access network solutions and examples. Hybrid FiWi access network is implemented by integrating wireless access technologies (WiFi, Wimax, and cellular) with optical fiber communication infrastructure that has been installed closer and closer to individual home and business users over the last few years.

Optical fiber can provide a bandwidth far greater than any other known transmission medium. A single strand of fiber reaches a total bandwidth of 25 000 GHz more than 1000 times of the total bandwidth of radio on the planet Earth. Besides, optical fiber has some more desirable properties such as low attenuation, longevity, and low maintenance costs which will eventually render fiber the medium of choice in wired first/last mile access networks [17]. In addition, the Plastic Optical Fiber (POF) technology is very suitable for optical home networks due to its low-cost simple wiring.

With all these advantages, especially the huge bandwidth fiber optical network can offer, Fiber to the Home (FTTH) is technically possible, which is also expected to be the major event in optical communication field. With FTTH network, the first/last mile bandwidth bottleneck between end users and high-speed backbone networks is going to be resolved. As a result, high bandwidth require services people have been dreaming for long, such as online HDTV, remote medical care, fast file transfer will soon be in normal life.

Wireless communication is another hot direction of future. Cellulars are becoming more and more functional and portable as well, in the meantime, lower and lower power consumption. Optical and wireless technologies are complementary in our age and will coexist over several decades [17]. Integrating them both, i.e. FiWi network has a prosperous future for the high-bandwidth as well as high-mobility network access it can provide.

1.1.1 RoF Networks and R&F Networks

Radio over Fiber (RoF) networks as a method to combine optical fiber networks and wireless networks, have been considered for many years. In RoF networks, Radio Frequencies (RFs), supporting various wireless applications such as Wireless Local Area Networks (WLAN) and microcellular radio systems, are carried on optical fiber links between a Central Office (CO) and multiple low-cost Remote Antenna Units (RAUs). For instance, a distributed antenna system connected to the base station of a microcellular radio system via optical fibers was proposed in [14]. To efficiently support time-varying traffic between the central station and its attached base stations, a centralized dynamic channel assignment method is applied at the central station of the proposed fiber optic microcellular radio system. To avoid having to equip each radio port in a fiber optic microcellular radio network with a laser and its associated circuit to control the laser parameters such as temperature, output power, and linearity, a cost-effective radio port architecture deploying remote modulation can be used[26].

RoF networks are attractive since they provide transparency against modulation techniques and are able to support various digital formats and wireless standards in a cost-effective manner [25]. It was experimentally demonstrated in [25] that RoF networks are capable to concurrently transmit Wideband Code Division Multiple Access (WCDMA), IEEE 802.11a/g Wireless Local Area Network (WLAN), Personal Handyphone System (PHS), and Global System for Mobile communications (GSM) signals. Figure1.1 illustrates the method researched in [25] for two different radio client signals transmitted by the central station on a Single-Mode Fiber (SMF) downlink to a base station and onward to a mobile user or vehicle. At the CO, both radio client signals are first converted to a higher frequency by a frequency converter, and then the two RF signals are fed into two different Electro Absorption Modulators (EAMs) and modulate the optical carrier wavelength emitted by two separate laser diodes. The two optical signals are then combined through an optical combiner onto the SMF downlink. At the base station, a photodiode converts the incoming optical signal to the electrical domain before the signal is amplified and radiated through an antenna to a mobile user where two separate frequency converters are used to retrieve the two different radio client signals. While SMFs are typically found in outdoor optical networks, many constructions have already preinstalled Multi-Mode Fiber (MMF) cables. Multi-Mode Fiber (MMF)-based networks can be cost-effective by installing low-cost Vertical Cavity Surface Emitting Lasers (VCSELs). In[12], various MMF combined with Commercial Off-The-Shelf (COTS) components were experimentally tested to validate the feasibility of indoor radio-over-MMF networks for the in-building coverage of second-generation GSM and third-generation cellular radio networks as well as IEEE 802.11a/b/g WLAN and Digital Enhanced Cordless Telecommunication Packet Radio Service (DECT PRS).

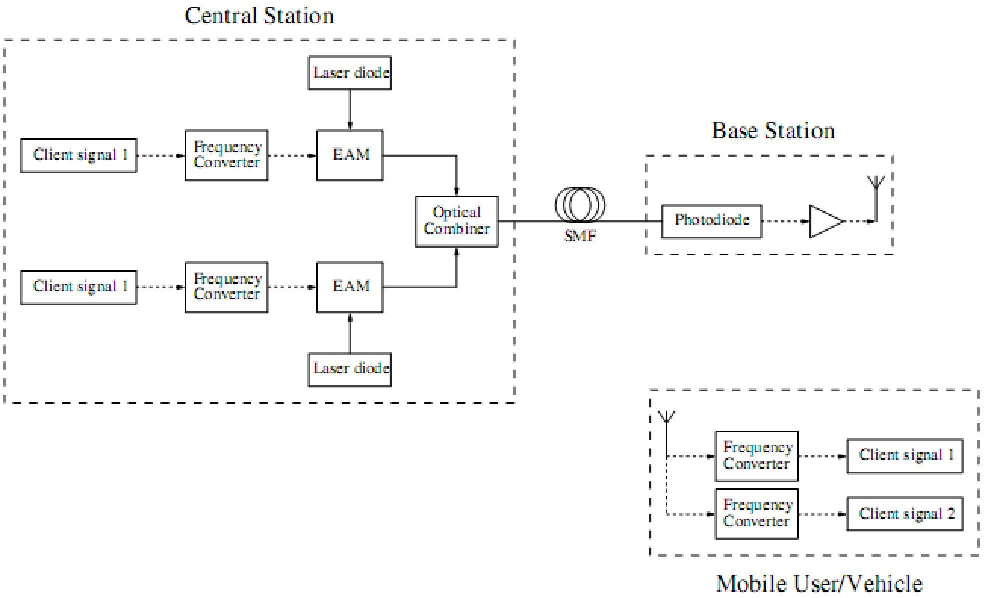


Figure 1.1: Radio-over-SMF Network Downlink Using EAMs for Different Radio Client Signals [25]

To realize future multiservice access networks, integrating RoF systems with existing optical access networks is significant. In [13], a novel approach for simultaneous modulation and transmission of both RoF RF and FTTH baseband signals using a single external integrated modulator was experimentally established, as shown in Figure1.2. The external integrated modulator contains three different Mach-Zehnder Modulators (MZMs) 1, 2, and 3. MZM 1 and MZM 2 are implanted in the two arms of MZM 3. The RoF RF and FTTH baseband signals independently use MZM 1 and MZM 2 to modulate the optical carrier generated from a common laser diode. Then the optical wireless RF and wired-line baseband signals are combined at MZM 3. After propagation over an SMF downlink, the two signals are separated an optical filter and forwarded to the wireless and FTTH application, respectively.

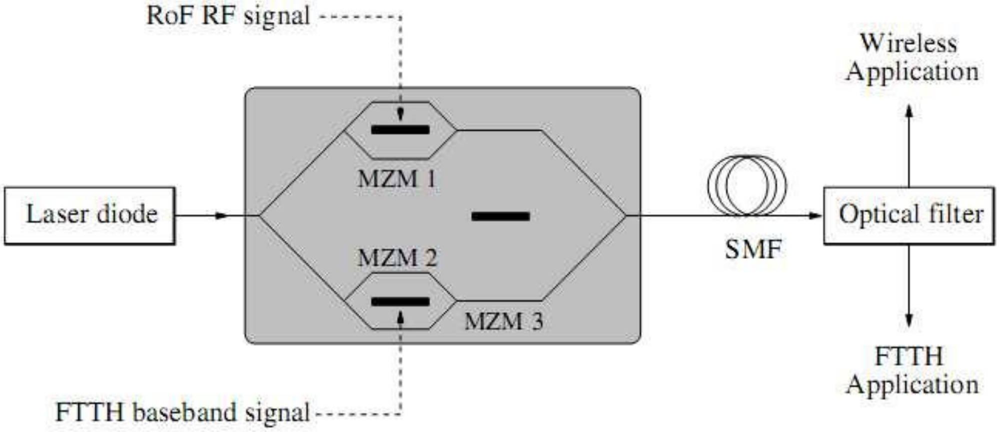


Figure 1.2: Simultaneous Modulation and Transmission of FTTH Baseband Signal and RoF RF Signal Using An External Integrated Modulator Consisting of Three Mach-Zehnder [13]

It was experimentally proved that a 1.25 Gb/s baseband signal and a 20-GHz 622 Mb/s RF signal can be simultaneously modulated and transmitted over 50 km standard SMF with acceptable performance penalties. The research projects in [13] successfully demonstrated the feasibility and maturity of low-cost multiservice RoF networks. It was shown that RoF networks can have an optical fiber range of up to 50 km. However, adding an optical distribution system in wireless networks can lead to severe impact on the performance of Medium Access Control (MAC) protocols [15]. The additional propagation delay may surpass certain timeouts of wireless MAC protocols, resulting in failed network performance. More precisely, MAC protocols based on centralized polling and scheduling (e.g.,IEEE 802.16 WiMAX) are less affected by enlarged propagation delays because of their capability to take longer latency into account between the CO and wireless Subscriber Stations (SSs). However, in distributed MAC protocols, such as the widely utilized Distributed Coordination Function (DCF) in IEEE 802.11a/b/g WLANs, the extra propagation delay between wireless STAtions (STAs) and Access Points (APs) brings severe challenges. Due to the acknowledgment (ACK) timeout, optical fiber can only be used in WLAN-based RoF networks with a maximum length to guarantee appropriate process of DCF[17].

The aforementioned limits of WLAN-based RoF networks can be prevented in Radio-and-Fiber (R&F) networks[8]. While RoF networks use optical fiber as an analog transmission medium between a CO and one or more RAUs with the CO being in charge of controlling access to both optical and wireless media, in R&F networks, access to the optical and wireless media is controlled separately from each other by using two different MAC protocols in the optical and wireless media, with protocol translation taking place at their interface. As a result, wireless MAC frames do not need to travel along the optical fiber to be processed at the central control station, but simply traverse their related access point and remain in the WLAN. In WLAN-based R&F networks, access control is done locally inside the WLAN without concerning any central control station, thus avoiding the undesirable impact of fiber propagation delay on the network throughput. R&F networks are well suitable to construct WLAN-based FiWi networks of extended coverage without bringing stringent limits on the size of the optical backhaul, contrasted with RoF networks that bound the length of fibers to a couple of kilometers. Note that this holds for distributed MAC protocols only such as DCF, but not for MAC protocols that use centralized polling and scheduling, e.g., WiMAX.

1.1.2 Example Testbeds

A. RoF Testbed

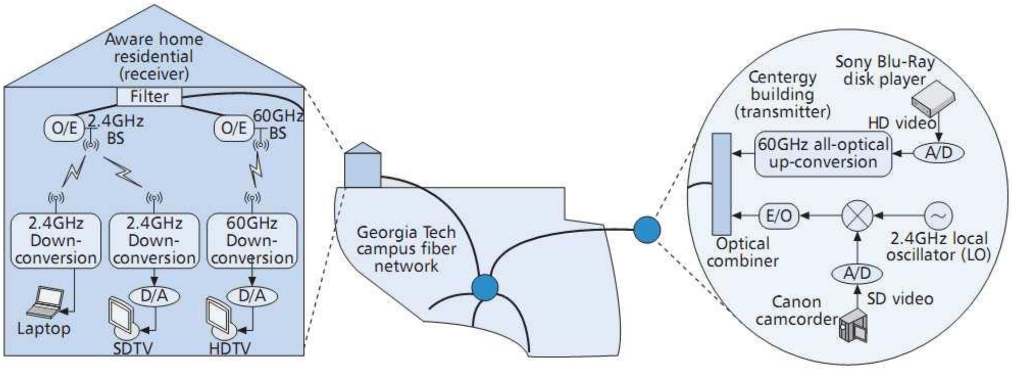


Figure 1.3: Georgia Institute of Technology RoF Field Demonstration of SD/HD Video Delivery Using 2.4 GHz and 60 GHz Millimeter-wave Transmissions[3]

Figure.3 shows the RoF testbed designed at the Georgia Institute of Technology for the field trial demonstration of 270 Mb/s Standard Definition (SD) and 1.485 Gb/s High Definition (HD) real-time video stream delivery using 2.4 and 60 GHz millimeter-wave transmissions over 2.5 km SMF between the Centergy building (transmitter) and the aware home residential building (receiver)[3]. All-optical upconversion is used at the transmitter to generate a 60 GHz millimeter-wave signal (modulated by phase modulation) and to send the HD video signal at 1554 nm. As shown in Figure1.3, electrical mixing and double-sideband optical modulation techniques are used to upconvert the SD video 2.4 GHz radio signal before optical transmission at 1550 nm. PIN photodiodes are used at the receiver to perform O/E conversion of the filtered optical signals. The experimental results demonstrate a very good Bit Error Rate (BER) performance of the received video signals.

B. R&F

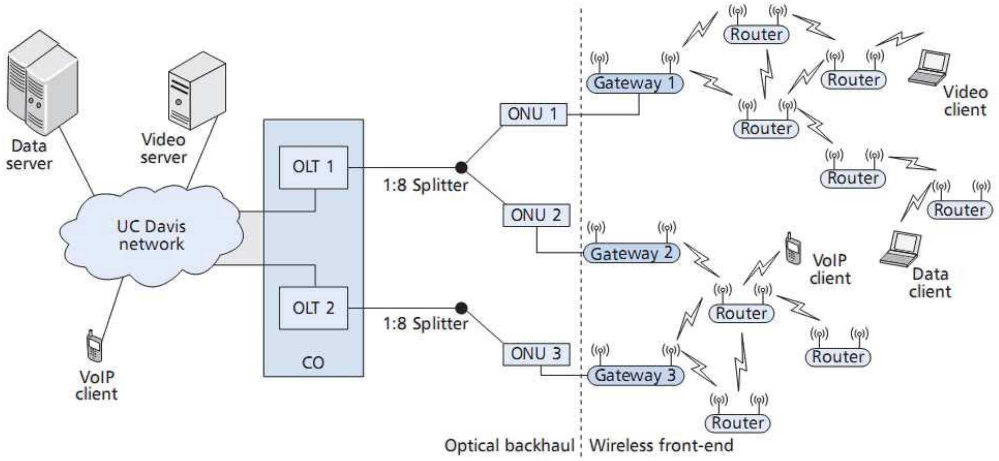


Figure 1.4: UC Davis R&F Testbed Integration of EPON and WMN for Voice, Video, and Data Traffic[16]

Testbed Figure1.4 shows the University of California (UC) Davis R&F testbed integration of two EPONs and an IEEE 802.11g WLAN-based WMN with a maximum transmission rate of 54 Mb/s for voice, video, and data traffic[16]. In this architecture optical protection is presented by using full PON duplication. Programmability was implemented by using a separate Linux PC connected to each ONU, and open source firmware in each wireless gateway and router. The results indicate that the quality of video transmissions abruptly deteriorates for an increasing number of wireless hops. In fact, the video client received a blank screen after four wireless hops. The experimental results evidently demonstrate that running EPON and WMN networks independently gives poor FiWi network performance. More involved testbeds of integrated FiWi network architectures is necessary considering hybrid access control protocols, integrated path selection algorithms, and advanced resilience techniques.

1.1.3 SuperMan

In[17], the authors proposed a SuperMan network structure which uses next-generation low-cost WiFi technologies combined with WDM-enhanced EPON access networks while integrating WiMAX with optical Metropolitan Area Network (MAN) technologies.

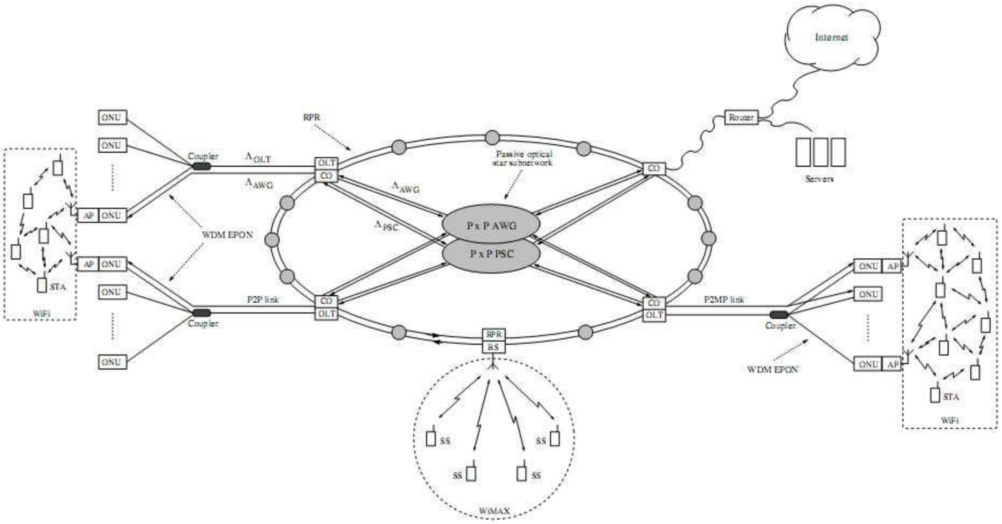


Figure 1.5: SuperMAN Architecture Integrating Next-generation WiFi Technologies with WDM EPON and Next-generation WiMAX Technologies with RPR[17]

Figure1.5 shows the network architecture of SuperMAN. It builds on our all-optically integrated Ethernet-based access-metro network, described in detail in [21], and stretched by optical-wireless interfaces with next-generation WiFi and WiMAX networks. More specifically, the optical part of SuperMAN consists of an IEEE 802.17 Resilient Packet Ring (RPR) metro network that are linked with multiple WDM EPON access networks attached to a subset of RPR nodes. RPR is an optical dual-fiber bidirectional ring network that targets at combining

Ethernet's statistical multiplexing gain, low equipment cost, and simplicity with SONET/SDH's carrier-class functionalities of high availability, reliability, and profitable TDM (voice) support. In RPR, destination stripping is adopted to increase spatial reuse of bandwidth and thus raise the capacity of the network. Each of the attached WDM EPONs has a tree topology with the OLT, the root of tree, collocated with one of the P COs. No particular WDM architecture is enforced on the ONUs, therefore the decision is made by economics, state-of-the-art transceiver manufacturing technology, traffic demands, and service provider preferences. The recommended WDM extensions to the IEEE 802.3ah Multi-Point Control Protocol (MPCP), described at length in [10], ensure backward compatibility with legacy TDM EPONs and enable the OLT to schedule transmissions to and receptions from ONUs on any supported wavelength channel. The optical access-metro network lets low-cost PON technologies follow low-cost Ethernet technologies from access networks into metro networks by interconnecting the P collocated OLTs/COs with a passive optical star subnetwork whose hub consists of an a thermal wavelength-routing $P \times P$ Arrayed Waveguide Grating (AWG) in parallel with a wavelength-broadcasting $P \times P$ Passive Star Coupler (PSC)[17]. In each WDM EPON, two different sets of wavelengths, λ_{OLT} and λ_{AWG}, are used, where λ_{OLT} is used for upstream and downstream transmissions between ONUs and the corresponding OLT locating in the same WDM EPON, and the second set, λ_{AWG}, contains wavelengths that optically bypass the collocated OLT/CO and allow ONUs locating in diffrent WDM EPONs to communicate all-optically with each other in a single hop across the

AWG of the star subnetwork, provided the ONUs are equipped with transceivers operating on these wavelengths. Similar to IEEE 802.3ah EPON, the optical part of SuperMAN is not limited to any specific Dynamic Bandwidth Allocation (DBA) algorithm. The DBA algorithms for WDM EOPNs need to be modified to adapt to SuperMAN.

1.2 Ethernet Passive Optical Network (EPON)

In the integrating of FiWi network, Ethernet Passive Optical Network (EPON) is an attractive network architecture to act as the fiber network part for its cost efficiency and its structural advantage. PON uses a passive optical splitter/combiner to create a shared fiber medium in the physical plant. Sharing the fiber medium reduces cost in the physical fiber deployment, and using passive components in the physical plant decreases power costs compared to active components. These reduced costs make PON a preferable choice for access networks. Compared to Asynchronous Transfer Mode (ATM) PON (APON), EPON uses Ethernet, in the frames of which, 90 percent of data traffic generates and terminates. Thus using EPON can reduce the unnecessary adaptation to transfer data between LAN and access network. Furthermore, ATM’s fixed data unit causes needless segmentation and reassembly at the terminating points of networks. Delay time as well as error rate is accordingly increased by the segmentation and reassembly. Therefore, EPON rather than APON is more suitable for data dominated networks.

1.2.1 EPON Architecture

1.2.1.1 Standard PON Architecture

Employing a Passive Optical Network (PON) between service providers and customers can provide a cost efficient, high bandwidth and flexible infrastructure. In a PON, a shared fiber medium is created using a passive optical splitter/combiner in the physical plant[18]. A PON generally utilizes a tree topology, where one Optical Line Terminal (OLT), locating at the central office of the service provider, links to several Optical Network Units (ONUs) in the field. The OLT connects to the ONUs using a feeder fiber that is successively split by a 1:N optical splitter/combiner for the ONUs to share the optical fiber, as shown in Figure1.6. The transmission direction of downstream is defined from OLT to ONU and operates as a broadcast medium. The transmission direction of upstream is defined from the ONUs to the OLT and is not a broadcast medium in upstream. The EPON is a multi-point-to-point [11] medium, where the ONUs cannot sense each other's transmission because the upstream optical signal is only received by the OLT. Yet, ONUs share the same fiber; thereby, their transmissions can collide, and collision detection and prevention is needed.

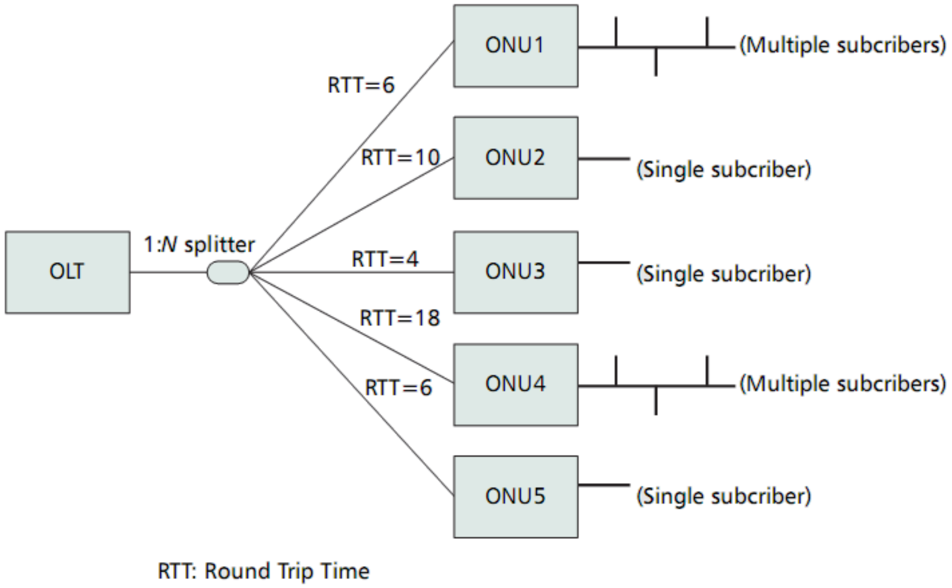


Figure 1.6: Network Architecture of a PON with One Optical Line Terminal (OLT) and N = 5 Optical Network Units (ONUs), Each with A Different Round-Trip Time (RTT) [18]

1.2.1.2 Alternative PON Architectures

Alternative PON architectures like Broadcast PON and Two-stage PON are introduced with more detail in [20, 22, 23] respectively. Broadcast PON requires reflection of the upstream signal back to the ONUs, as illustrated in Figure1.7, downstream OLT to ONUs transmissions are copied by splitter “d” to all ONUs, while each upstream ONU to OLT transmission is reflected by splitter “ru” back to all ONUs, thus creating a broadcast network for both upstream and downstream transmissions. The dashed lines represent the extra fibers used to carry the reflected upstream signal back to the ONUs[18]. Splitters between OLT and ONUs are deployed to create a broadcast network that can take advantage of a decentralized medium access control protocol (e.g., Carrier Sense Multiple Access with Collision Detection [CSMA/CD]). However there are economic drawbacks to this approach.

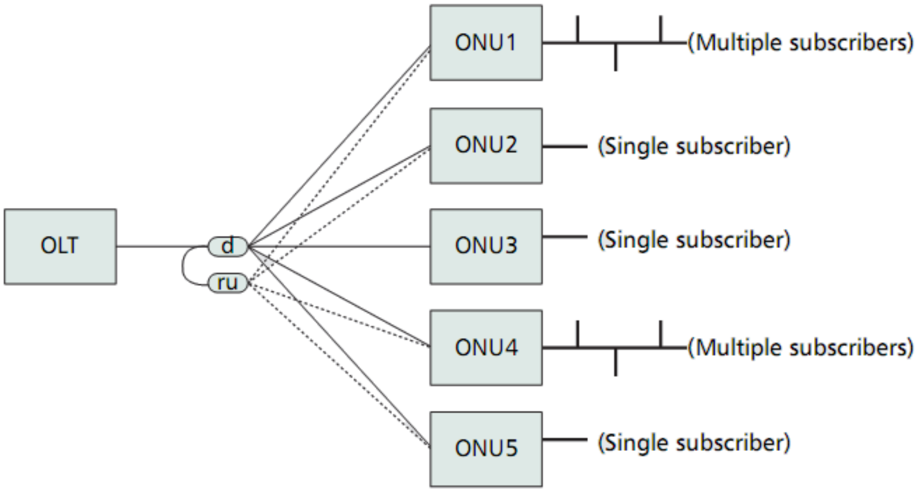


Figure 1.7: Broadcast PON Architecture[18]

Two-stage PON architecture can enable a PON to adapt larger number of ONUs than a single-stage PON. Two-stage PONs can increase the reach of the PON. In the first stage, some ONUs act as sub-OLTs for second-stage ONUs, as illustrated in Figure1.8, certain ONUs act as sub-OLTs that regenerate the optical signal for ONUs in a Second stage, thereby allowing for an increase in the total number of served ONUs[18]. These sub-OLTs restore the optical signal from both upstream and downstream, as well as collect the traffic of their child ONUs. This allows a single OLT in a central office to reach a larger number of ONUs since the sub-OLTs act as optical switches. When increasing the number of ONUs, diminishing optical power budget needs concerns.

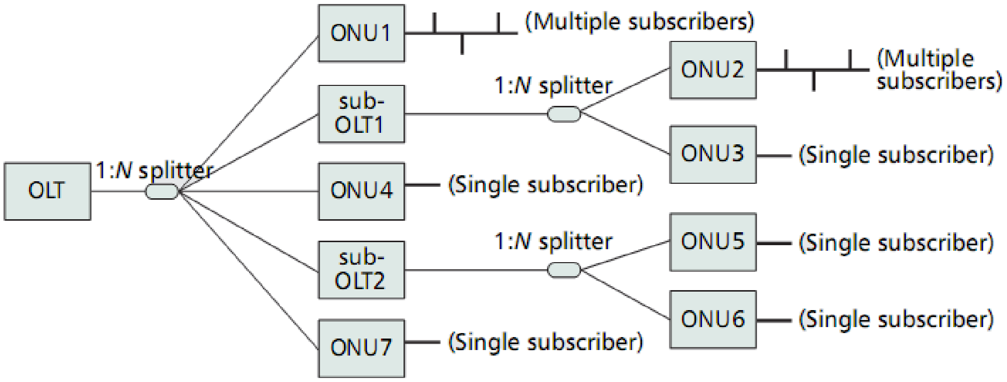


Figure 1.8: Two-Stage PON Architecture[18]

1.2.2 Downstream and Upstream Transmission

In the downstream direction, i.e. from OLT to ONUs, Ethernet packets transmitted pass through a 1×N passive splitter or a set of cascaded splitters before arriving at each ONU. The value of N is typically between 4 and 64 (limited by the available optical power budget). This situation is basically a shared medium network. Because of Ethernet's broadcasting nature, in the downstream direction, PON architecture fits very well with the Ethernet: packets are broadcasted by the OLT and selectively accepted at their destination ONUs.

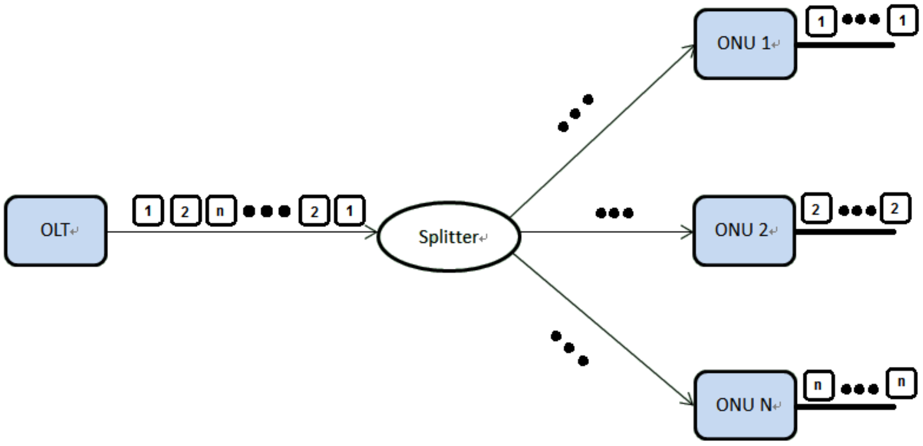


Figure 1.9: EPON Downstream Transmission

In the upstream direction i.e. from ONU to OLT, Ethernet packets from any ONU will reach the only OLT, but not the rest of the ONUs because of the optical combiner. In this case, the situation of EPON is similar to a point-to-point architecture. However, different from a point-to-point network, data packets in EPON from different ONUs transmitted simultaneously would collide with each other. Therefore, in the upstream direction, EPON needs to employ some arbitration mechanism to avoid data collisions and fairly share the channel capacity among ONUs.

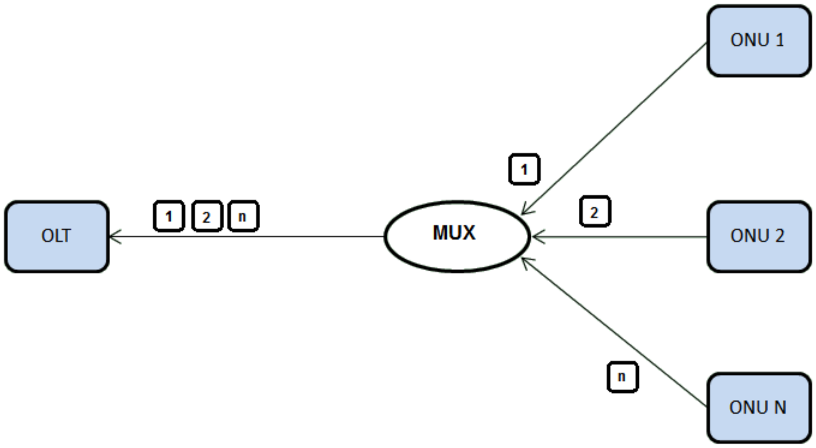


Figure 1.10: EPON Upstream Transmission

As can be seen in Figure1.10, all ONUs are synchronized to a common time reference, and each ONU is allocated a timeslot which is capable of carrying several Ethernet packets. The performance of an EPON highly depends on a bandwidth allocation scheme. The possible timeslot allocation schemes range from static allocation (fixed TDMA) to dynamic adjustment of the slot size based on instantaneous queue load in every ONU (statistical multiplexing scheme). Choosing the best allocation scheme, however, is not a trivial task.

1.2.3 Multi-Point Control Protocol (MPCP)

IEEE802.3ah task force defined Multi-Point Control Protocol (MPCP) as a new function of the MAC control layer to assist medium access control. The MPCP contains five messages: REGISTER REQ, REGISTER, REGISTER ACK, REPORT and GATE, among which REGISTER REQ, REGISTER, and REGISTER ACK are used for the discovery and registration of new ONUs; REPORT and GATE are used for facilitating centralized medium access control. The REPORT message reports the current queue length at an ONU to OLT. The OLT then uses a DBA to make decisions of each ONU’s grant sizing and grant scheduling and send this information in a Gate message to each ONU.

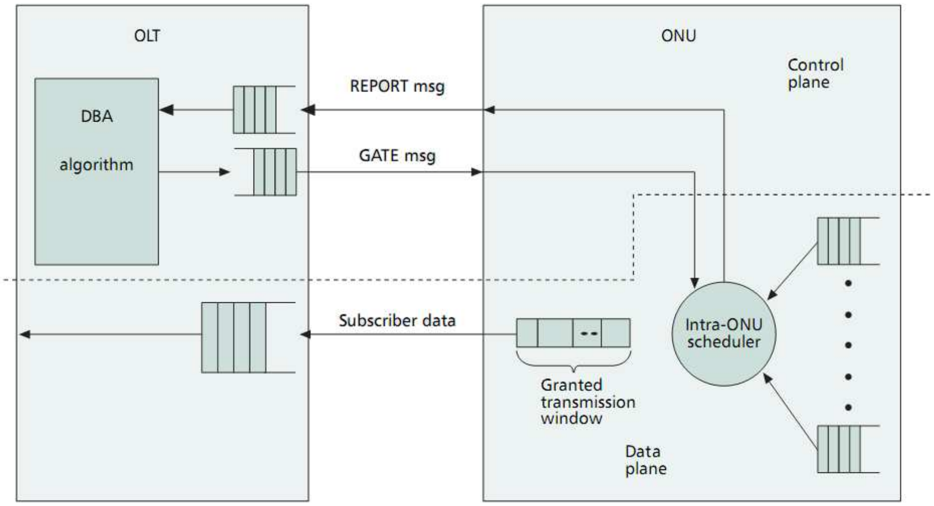


Figure 1.11: MPCP Operation: Two-way Messaging Assignment of Time Slots for Upstream Transmission Between ONU and OLT[18]

MPCP has two modes of operation:

(1) Auto-discovery mode: to discover newly activated ONUs, the MPCP should initiate the discovery procedure periodically. The auto-discovery mechanism is used to detect newly connected ONUs and learn their round-trip

delays and MAC addresses. Both the OLT and ONUs implement the discovery process, which is driven by the discovery agent. Auto-discovery employs four MPCP messages: GATE, REGISTER_REQ, REGISTER, and REGISTER_ACK. These messages are carried in MAC control frames.

(2) Bandwidth assignment mode: to maintain the communication between OLT and ONUs, the MPCP should provide periodic granting for each ONU. It employs GATE and REPORT messages.

1.3 Thesis Outline

The remainder of this thesis is organized as follows. In section II, EPON components study and categorization is presented. In section III, simulation tool and experiment conditions are introduced. In section IV, experiment results and analysis are presented and recommended DBA is given. Section V concludes this paper and pointes out future work. Limited bandwidth with accumulating excess credits grant sizing method is first time introduced. Some of the DBA combinations are at the first time simulated and studied in this paper such as Double Phase Polling scheduling framework coupled with Limited Gate with Excess distribution and Shortest Propagation Delay first grant scheduling policy, Double Phase Polling scheduling framework coupled with Limited Gate with Share Credit Excess distribution and Shortest Propagation Delay which proved to be the best performance DBA algorithm in simulation regarding average packet delay and bandwidth stability limit.

CHAPTER 2

DYNAMIC BANDWIDTH ALLOCATION

To EPON's performance regarding packet delay, network utilization, QoS etc., EPON Dynamic Bandwidth Allocation (DBA) is the key factor. DBA generally is defined as the process of providing statistical multiplexing among ONUs[18]. In access network, each link is established between a single subscriber or a small network of subscribers with very bursty data traffic sources, for instance, web browse data, packetized video data, etc. As a result of the traffic's bursty nature, the bandwidth requirements usually vary widely with time. Thus, statically allocating bandwidth to every subscriber in PON topology would lead to inefficient bandwidth utilization and large packet delay. Dynamic bandwidth allocation method that can instantaneously assign bandwidth requirements is needed to achieve high bandwidth utilization and lower packet delay. Typically, a DBA algorithm is polling process operated at Optical Line Terminal (OLT), which responsible for providing statistical multiplexing. The ONUs would report their right away queue sizes in a control frame and propagate this bandwidth requirement information to the OLT through the PON. The OLT calls for instantaneous bandwidth requirement information from each Optical Network Unit (ONU) to make access decisions before send them back to ONUs. However because of the propagation delays in network, it is impossible to have the precise bandwidth require information. Typical value of the propagation delay is up to 100μsec, much greater than 12.3μsec, the transmission time of the maximum Ethernet frame size.

Though DBA design for upstream transmission in EPON has received significant attention, a comprehensive framework for classifying DBA mechanisms and elements has not been formed.

In the research of DBA, we classify DBA mechanisms with three factors:

- Grant scheduling framework(α), which is characterized by the event triggering a bandwidth allocation;
- Grant sizing(β), which determines the size of the upstream transmission window allocated to an UNU;
- Grant scheduling policy(γ), which determines the temporal order of several simultaneously scheduled transmission windows.

With this classification, each DBA can be noted as a triple of (α, β, γ) [19].

Scheduling Framework	**Grant Sizing**	**Grant Scheduling**

Figure 2.1: DBA Components Structure

2.1 Grant Scheduling Framework

The grant scheduling framework determines when the OLT will make access decisions and send transmission grants to the ONUs. We can differentiate the scheduling frameworks according to the event that triggers the production of a grant schedule.

To better characterize the difference of various scheduling framework, let N denote the number of ONUs, k represent the k^{th} ONU in a granting cycle, t_g be the guard time between ONU transmissions, t_G be the transmission time of a GATE message, $t_{start}(k, j)$ be the start time of the granted transmission window of the k^{th} ONU in granting cycle j, $t_{poll}(k, j)$ be the polling time for the k^{th} ONU in granting cycle j, $G(k, j)$ be the length of the granted transmission window of the k^{th} ONU in granting cycle j, τ (i) be the symmetric propagation delay between the OLT and ONU i, and [k,j] be an operator that returns the number, i, of the k^{th} ONU in granting cycle j, $t_{end}(k, j)$ be the time that the grant to the k^{th} ONU for granting cycle j ends.

Then we have:

$$t_{end}(k, j) = t_{start}(k, j) + G(k, j) \quad (2.1)$$

Let $t_a(k, j)$ be the time the upstream channel is free when the OLT schedules the granted transmission window for the k^{th} ONU in granting cycle j. The upstream channel becomes free a guard time after the end of the granted transmission window of the previous ONU,

$$t_a(k, j) = \begin{cases} t_{end}(N, j-1) + t_g, k = 1 \\ t_{end}(k-1, j) + t_g, k \neq 1 \end{cases} \quad (2.2)$$

Let $t_{sched}(k, j)$ be the time when the OLT schedules the granted transmission window the k^{th} ONU for cycle j. The start time of the granted transmission window the k^{th} ONU would be

$$t_{start}(k,j) = \max\{(t_{sched}(k,j) + t_{poll}(k,j)), t_a(k,j)\} \qquad (2.3)$$

Different scheduling framework schedules, polls and orders ONUs differently. As a result, the values of $t_{start}(k,j)$, $t_{poll}(k,j)$, and [k,j] would differ.

2.1.1 Offline Scheduling Framework

In offline scheduling framework, every cycle j, OLT collects all REPORT messages from each ONU during previous cycle j-1 , before making bandwidth allocations. Therefore,

$$t_{sched}(k,j) = t_{end}(N, j-1), \forall k \qquad (2.4)$$

and all ONUs are polled together, so

$$t_{poll}(k,j) = k \cdot t_G + 2 \cdot \tau \cdot ([k,j]) \qquad (2.5)$$

where [k,j] is determined by grant scheduling policy (discussed in 2.3).

The benefit for offline scheduling is when OLT making decisions about each ONU’s grant size (transmission window duration), it can integrate the knowledge of every ONU’s packet queue sizes and allocate bandwidth to ONUs accordingly, using advanced grant sizing methods. For instance it may grant a large window slot to the ONUs whose queue size is large while grant small window slot to ONUs whose buffered data is not much. An advance grant sizing method, excess bandwidth distribution, can only be applied when OLT has all the Report messages.

The disadvantage for offline scheduling framework is there would be an idle period for OLT between the last grant size message is sent and the data of the first ONU has not received. There is no way to avoid this idle time though it can be reduced by using Short Propagation Delay First (SPD) scheduling policy which will be introduced in later section.

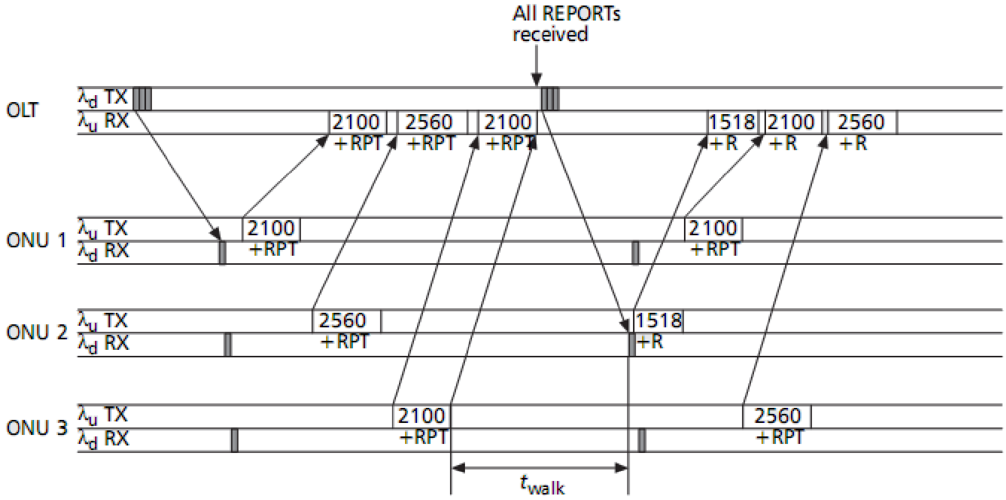


Figure 2.2: Offline Scheduling Framework[18]

2.1.2 Online Scheduling Framework

In online scheduling framework, OLT will immediately grant transmission window to ONU k in granting cycle j, as long as it receives the REPORT message during granting cycle j-1. Therefore,

$$t_{sched}(k, j) = t_{end}(k, j-1), \forall k \qquad (2.6)$$

and because only one ONU is polled at $t_{sched}(k, j)$, we have

$$t_{poll}(k, j) = t_G + 2 \cdot \tau \cdot ([k, j]) \qquad (2.7)$$

Online scheduling framework would not have the OLT idle time problem because pooling of different ONUs in a cycle is overlapped with other polling cycles, while the disadvantage of online scheduling however, is when OLT making bandwidth allocations, it only has the knowledge of the reporting ONU, thus advanced scheduling policies that needs more information from other ONUs cannot applied under this framework.

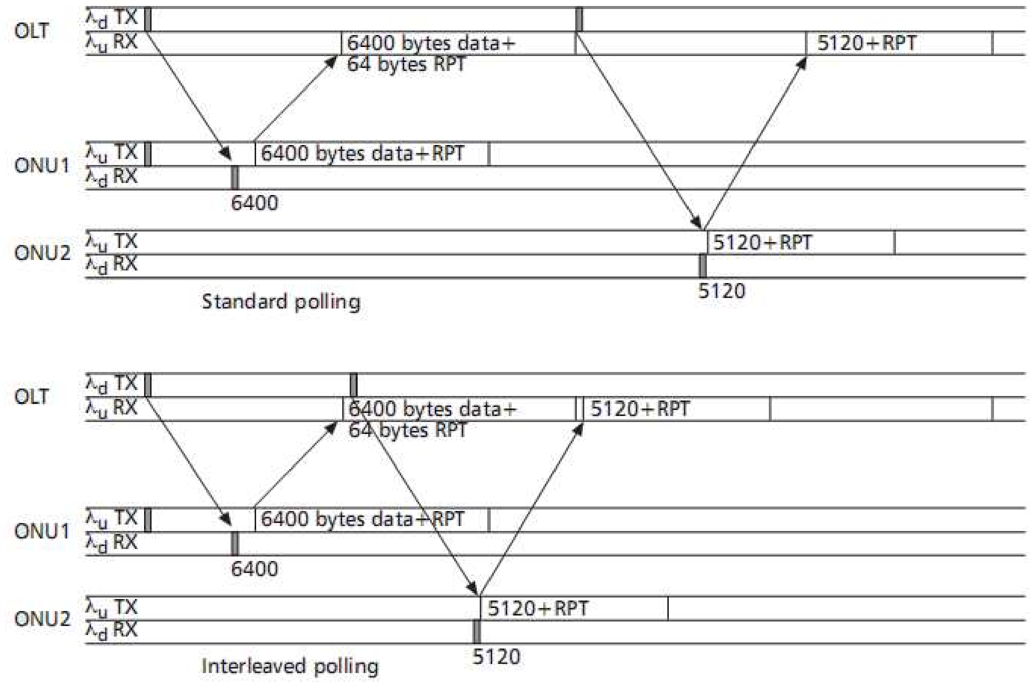


Figure 2.3: Online Scheduling Frameowrk[18]

2.1.3 Online Just-In-Time Two-step Scheduling Framework

Online just-in-time two-step scheduling framework is a compromise between online scheduling framework and offline scheduling framework. In online just-in-time scheduling framework, the ONUs are divided into two sets every granting cycle. The schedule for the first set of ONUs is produced a polling time before the REPORT message is received from the last ONU, i.e. N^{th}, during granting cycle j-1. Let $t_{sched}(j)$ be this time for franting cycle j, $S_1(j)$ be the first set of ONUs for granting cycle j, $S_2(j)$ be the second set of ONUs for granting cycle j . Therefore,

$$t_{sched}(j) = [(t_{sched}(N, j-1) + t_g) - \min\{t_{poll}(1, j), \cdots, t_{poll}(N, j)\}] \quad (2.8)$$

$$t_{sched}(i, j) = t_{sched}(j), \forall i \in S_1(j) \quad (2.9)$$

This choice for $t_{sched}(j)$ allows the first ONU, in cycle j to utilize the upstream transmission channel as soon as it is available. At $t_{sched}(j)$, the OLT will divide the ONUs into the two sets: $S_1(j)$, ONUs whose REPORTs have been received before the schedule trigger time $t_{sched}(j)$, and $S_2(j)$, the ONUs whose REPORTs have not yet been received at $t_{sched}(j)$. The partitioning of the ONUs can be expressed as,

$$S_1(j) = \{t_{end}(i, j-1) \le t_{sched}(j), \forall i\} \quad (2.10)$$

$$S_2(j) = \{i \notin S_1(j), \forall i\} \quad (2.11)$$

After the ONUs in $S_1(j)$ are scheduled, the OLT waits until all of the REPORT messages have been received for the ONUs in $S_2(j)$ and then schedules those ONUs. Therefore,

$$t_{sched}(i, j) = \max_{i \in S_2(j)}\{t_{end}(i, j-1)\}, \forall i \in S_2(j) \quad (2.12)$$

Online just-in-time scheduling framework does not leave an idle time for OLT, also gives OLT more information to make better scheduling decisions than online scheduling framework. To enable this framework, we need to make sure that the Gate message or messages are transmitted by OLT at least a round trip propagation delay time before OLT receiving messages from ONU. Because propagation delays of each ONU may be different with the others, using the largest RTT in the EPON as the time OLT send Grant messages before OLT is available can ensure zero OLT idle time.

Because OLT does not have full knowledge of all the ONUs Report messages, some prediction grant sizing methods may be combined together with online just-in-time scheduling framework to achieve a better result. Typical prediction grant sizing methods are constant prediction where the unknown ONU's queue size is predicted as a constant, linear prediction where unknown ONU's queue size is predicted based on the actual data received during the previous waiting time, and higher order prediction where the linear predictor has its weights updated by means of the Least Mean Square (LMS) algorithm.

2.1.4 Multi-Group Scheduling Framework

Multi-Group scheduling framework is another type of compromise between online scheduling framework and offline scheduling framework. In multi-group scheduling framework, ONUs are divided into multiple groups and are granted group by group using a grant sizing and grant scheduling policy within each group. Multi-group polling scheme can also eliminate the gap time of OLT that happened in offline scheduling framework by sending Grant messages group by group.

A typical multi-group scheme is called Double Phase Polling (DPP)[2], where the group number is 2, the first half ONUs in group 1 and the second half ONUs in group 2 and each group is scheduled separately. The schedule for the entire granting cycle j for a group is produced when the REPORT message is received from the last ONU for that group during granting cycle j-1. Therefore,

$$t_{sched}(k, j) = \begin{cases} t_{end}(N/2, j-1), k \le N/2 \\ t_{end}(N, j-1), k > N/2 \end{cases} \qquad (2.13)$$

When OLT is receiving the Report messages from group 1, it calculate the grant assignments to group 2 at the same time; when group 1 ONUs finish transmission of data and Reports, group 2 ONUs data and Reports would arrived. Group1 and group2 ONUs alternatively transmitting so that OLT idle time is eliminated and OLT has enough information to apply advanced grant sizing methods like excess bandwidth grant sizing.

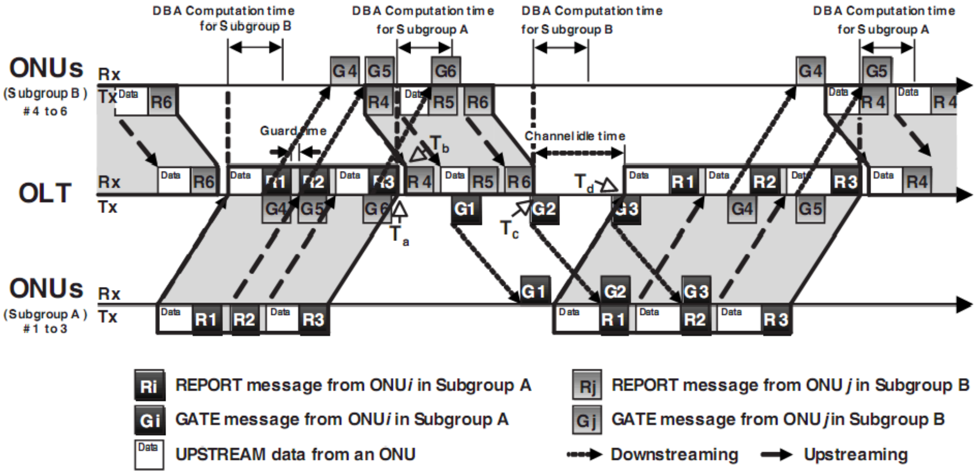


Figure 2.4: Double Phase Polling Framework[2]

Another way to do multi-group is called Multi-Thread Polling (MTP) scheme[24], where ONUs send Report messages after a period of time without necessarily after receiving the Grant message. In this way, the OLT idle time is also removed and OLT grant ONUs by the information it has. In this case, the groups are not fixed.

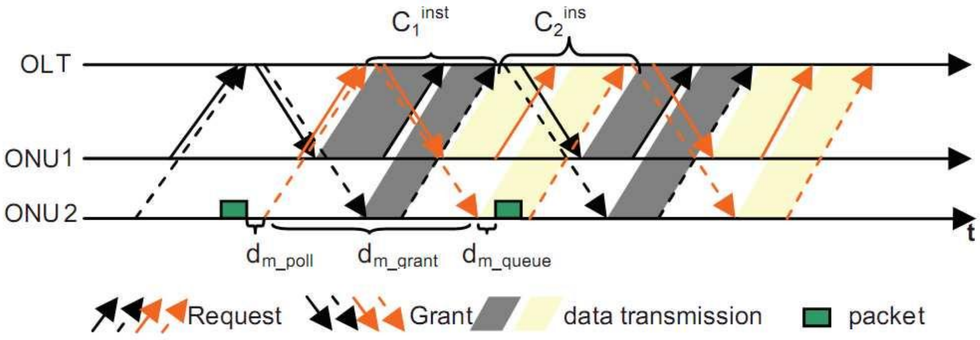


Figure 2.5: Multi-Thread Polling Framework [24]

2.2 Grant Sizing

Grant sizing is the problem of how much amount of bandwidth should be granted to each ONU when OLT doing bandwidth allocation. Grant sizing can be categorized into 5 major types: Fixed grant sizing, Gated grant sizing, Limited grant sizing, Limited grant sizing with excess distribution, Exhaustive grant sizing with queue size prediction.

To better express different grant sizing method, we are going to use the following notations: G(i,j) be the length of the granted transmission window of the ONU i in granting cycle, R(i,j) represents the current queue size reported from ONU i in cycle j; $G_{max}(i)$ represents the maximum grant size for ONU i, P(i,j) be the predicted traffic queued up for ONU i in the period between when the REPORT message sent from ONU to OLT and the end of the granted window sent from OLT to ONU i in granting cycle j. A common equation for grant sizing would then be:

$$G(i,j) = f[R(i,j), G_{max}(i)] + E(i,j) + P(i,j) \quad (2.14)$$

Different techniques use different function f() and different value of E(i,j) and P(i,j)

2.2.1 Fixed Size

Fixed grant sizing uses

$$G(i,j) = G_{max} \quad (2.15)$$

as the common function. In every cycle, each ONU gets a fixed length transmission window. Fixed grant sizing is the simplest way to implement and OLT has least amount of calculation compare to other grant sizing methods. But due to the bursty nature of network communication, this grant sizing method works very bad because it is very inflexible. Simulation results and analysis in research give the conclusion that fixed grant sizing scheme performs much worse than the other schemes mentioned below.

2.2.2 Gated Size

Gated grant sizing uses

$$G(i,j) = R(i,j) \quad (2.6)$$

as the common function. Grant size for ONU i is equal to its reported requirement. Gated grant sizing can offer lower average delay than fixed sizing but no guarantee of fairness between ONUs because there is no limit to the grant size and an ONU may easily monopolize the upstream channel if it happens to generate a large amount of traffic. In-depth analysis and simulations of this scheme from related work proved this point.

2.2.3 Limited Grant Size

Limited grant sizing uses

$$G(i,j) = \min(R(i,j), G_{max}(i)) \quad (2.17)$$

as the general function. Grant size to ONU k is set to the smaller value between maximum limit bandwidth and requested bandwidth. This scheme is an upgraded version of gated sizing. It puts an upper bound to the request transmission window duration and thus avoids any ONU seizing the shared bandwidth for too long. It can provide a rather fair bandwidth distribution among ONUs. Simulation results in[11] indicate that using limited grant sizing achieves same level of delay as gated grant sizing method. However, using limited grant sizing, guard times between grants may lead to a drop of bandwidth utilization especially when a small maximum value is set to $G_{max}(i)$ due to the reduce of the cycle length. Another drawbacks include queue underserved when $R(i,j) > G_{max}(i)$. Ethernet frame caused lost since usually the end of the grant doesn't fit into an Ethernet frame.

2.2.4 Limited Grant Sizing with Excess Distribution

Let $E_{total}(l,i)$ be the total excess bandwidth credits for polling group l in cycle j, U(l,i) be the set of underloaded ONUs within polling group l during grant cycle j, and O(l.i) be the set of overloaded ONUs within polling group l during granting cycle j. Limited grant sizing with excess distribution improves limited grant sizing by redistribute extra bandwidth that is left in overload ONUs where $R(i,j) < G_{max}(i)$ to the underloaded ONUs. The general function is

$$G(i,j) = \begin{cases} R(i,j), R(i,j) < G_{max}(i) \\ G_{max}(i) + E(i,j), R(i,j) > G_{max}(i) \end{cases} \quad (2.18)$$

where

$$E(i,j) = f(E_{total}(j)) \quad (2.19)$$

and

$$E_{total}(j) = \sum_{i \in O} (G_{max}(i) - R(i,j)) \quad (2.20)$$

Excess distribution needs excess division algorithm and excess allocation algorithm. In excess division, total excess bandwidth from the overloaded ONUs for which $R(i,j) > G_{max}(i)$, is divided among underloaded ONUs where $R(i,j) < G_{max}(i)$.

There are many ways to redistribute the excess, i.e. the function from $E_{total}(j)$ to get E(i,j). For instance, demand driven excess distribution [6],

$$E(i, j) = \frac{R(i, j)}{\sum_{i \in O} R(i, j)} \cdot E_{total}(j) \quad (2.21)$$

where each ONU's share of the excess is only depend on the its request size. It at most times is not fair; Equitable excess distribution [1], M denotes the number of overloaded ONUs in the granting cycle,

$$E(i, j) = \frac{1}{M} \cdot E_{total}(j) \quad (2.22)$$

divides $E_{total}(j)$ equally among overloaded ONUs; And weighted excess distribution [14], uses ONU priority weights w(i) to decide how much excess bandwidth an overloaded ONU gets.

$$E(i, j) = \frac{w(i)}{\sum_{i \in O} w(i)} \cdot E_{total}(j) \quad (2.23)$$

The afore mentioned excess division methods may lead to waste bandwidth since for some ONUs, granted bandwidth with excess may be larger than its original requested bandwidth, R(i,j) < G(i,j) + E(i,j). A better algorithm avoiding bandwidth waste is: when total ONU bandwidth demand is less than available bandwidth, assign each ONU what it requests no matter it is overloaded or underloaded, otherwise assign the excess according to the amount of unfulfilled bandwidth of each overloaded ONUs

$$G(i, j) = \begin{cases} R(i, j), E_{demand}(j) < E_{total}(j) \\ G_{max}(i) + \dfrac{E_{demand}(i, j)}{\sum_{i \in O} E_{demand}(i, j)} \cdot E_{total}(j), E_{demand}(j) > E_{total}(j) \end{cases} \quad (2.24)$$

where

$$E_{demand}(i, j) = R(i, j) - G_{max}(i) \quad (2.25)$$

and

$$E_{demand}(j) = \sum_{i \in O} E_{demand}(i, j) \quad (2.26)$$

When $R(i, j) < G_{max}(i) + E(i, j)$, the unused excess bandwidth would be ignored. To take advantage of the unused excess bandwidth, there is an improved excess distribution sizing algorithm which takes the excess bandwidth from previous cycle, i.e. cycle j-1, and adds it to the current cycle's excess bandwidth. This sharing should be done in a way that does not permit excess bandwidth credits to circulate for more than a cycle. With this in mind, we allow the unused excess bandwidth credits that were accumulated but unused within a group during a granting cycle to be forwarded to the other group. Credits that were forwarded by another group but unused by the group that received them cannot be forwarded back to that group. This is especially useful in multi-group scheduling framework. Let $E_{share}(1, j)$ be the excess credits to be shared by group 1 with group 2, $E_{share}(2, j)$ be the excess credits to be shared by group 2 with group 1. Then we have the total excess credits for group 1 and 2 during granting cycle j are,

$$E_{total}(1, j) = \sum_{i \in U(1,j)} (G_{max}(i) - R(i, j)) + E_{share}(2, j-1) \quad (2.27)$$

and

$$E_{total}(2, j) = \sum_{i \in U(2,j)} (G_{max}(i) - R(i, j)) + E_{share}(1, j) \quad (2.28)$$

respectively.

Let $E_{used}(l, j)$ be the total credits used by group i during granting cycle j, define

$$E_{used}(l, j) = \sum_{i \in O(i,j)} E(i, j) \qquad (2.29)$$

then if

$$E_{used}(l, j) \le \sum_{i \in U(i,j)} (G_{max}(i) - R(i, j)) \qquad (2.30)$$

we get

$$E_{share}(l, j) = \sum_{i \in U(i,j)} (G_{max}(i) - R(i, j)) - E_{used}(l, j) \qquad (2.31)$$

otherwise,

$$E_{share}(l, j) = 0 \qquad (2.32)$$

2.2.5 Exhaustive Grant Sizing Using Queue Size Prediction

Exhaustive grant sizing using queue size prediction, in which queue size prediction is used to estimate the traffic generated between REPORT message sent from ONU and gated transmission window starts. This estimation is useful since when REPORT message is sent, additional traffic may come afterwards and thereby increase the actual requirement. Generally the traffic is not coming in a constant rate, but in a bursty variable bit rate. Thus predicting it is challenging. Basic strategies for queue size prediction of bursty source are constant credit in which the OLT adds a constant to the grant size, and linear credit, assuming that coming traffic has a linear relation to the requirement. Simulation results from [15] where a one step back linear predictor is used shows that packet delay for expedited forwarding traffic is lowered compared to fixed grant sizing and limited grant sizing. Higher order linear predictor is discussed in [9]. The disadvantage is the possible throughput decrease when prediction is more than actual traffic due to the bursty nature of data traffic in Ethernet.

2.3 Grant Scheduling Policy

Grant scheduling problems in EPON Medium Access Control level concern both inter ONU scheduling and intra ONU scheduling. Inter ONU scheduling concerns scheduling grants among different ONUs, while Intra ONU scheduling concerns scheduling different queues at each ONU for transmission within the granted transmission window. This division of scheduling if referred to as hierarchical scheduling.

2.3.1 Inter ONU Scheduling

Inter ONU scheduling is performed at OLT in conjunction with scheduling framework and grant sizing. In order to arrange the scheduling order of ONUs, OLT generally needs to wait until all the REPORT queue lengths from ONUs are received before making decisions. This means an offline framework is required and there is usually an idle time for OLT between cycles.

There are many scheduling orders that has been proposed and studied. For example, Shortest Processing Time first (SPT) scheduling orders ONUs according to their processing times and schedules the ONU with short processing time first; Largest Processing Time first is similar to SPT but orders ONUs reversely; Largest Number of Frames first (LNF) or Shortest Number of Frames first (SNF) order ONUs by their number of Ethernet frames piled up with larger or shorter number of frames would be scheduled earlier; Earliest Arrival First (EAF) scheduling ONUs by their REPORT message arrival time and early arrival ONUs would be scheduled first; Shortest Propagation Delay first (SPD) or Largest Propagation Delay first (LPD) order ONUs by their propagation delay time to OLT, in other words, by their distances to OLT and the shortest propagation delay ONU or largest propagation delay ONU would be scheduled first.

Among these scheduling policies, SPD has been proved mathematically to offer the shorter idle OLT time between cycles than other scheduling policies under the same conditions in [27]. The idea of SPD is if we grant the ONU whose distance to OLT is the shortest, the OLT idle time between the last grant and the first coming of data would be minimized since the propagation delay can be no less than the shortest ONU's propagation delay. Thus SPD can provide a shorter average delay and a higher bandwidth utilization compare to other grant scheduling policies when combined with offline grant framework. The simulation results shows that using SPD grant scheduling policy can greatly improve the performance of a DBA.

2.3.2 Intra ONU Scheduling

Intra ONU scheduling is concerned with scheduling between multiple queues at one ONU given the Granted transmission window. If the number of queues in an ONU is relatively small, this intra-ONU scheduling can be performed at the OLT along with inter ONU scheduling. This can keep the cost of ONUs at minimum, but have potential scaling problem because the number of queues may increase. As the number of queues increases, the complexity of scheduling would be a big concern at OLT and usually intra ONU scheduling is typically made hierarchical with the inter-ONU scheduling at the root of the hierarchy in the OLT and one level of branches.

The ONU would perform intra-ONU scheduling. To keep the low cost of ONU, the complexity of intra ONU scheduling should not be high. Thus, key factor in designing intra ONU scheduling algorithm is to achieve quality of service guarantees with low complexity.

Generally there are ways to schedule multiple queues with different priorities: Strict Priority (SP) scheduling and Weighted Fair Queuing (WFQ) scheduling.

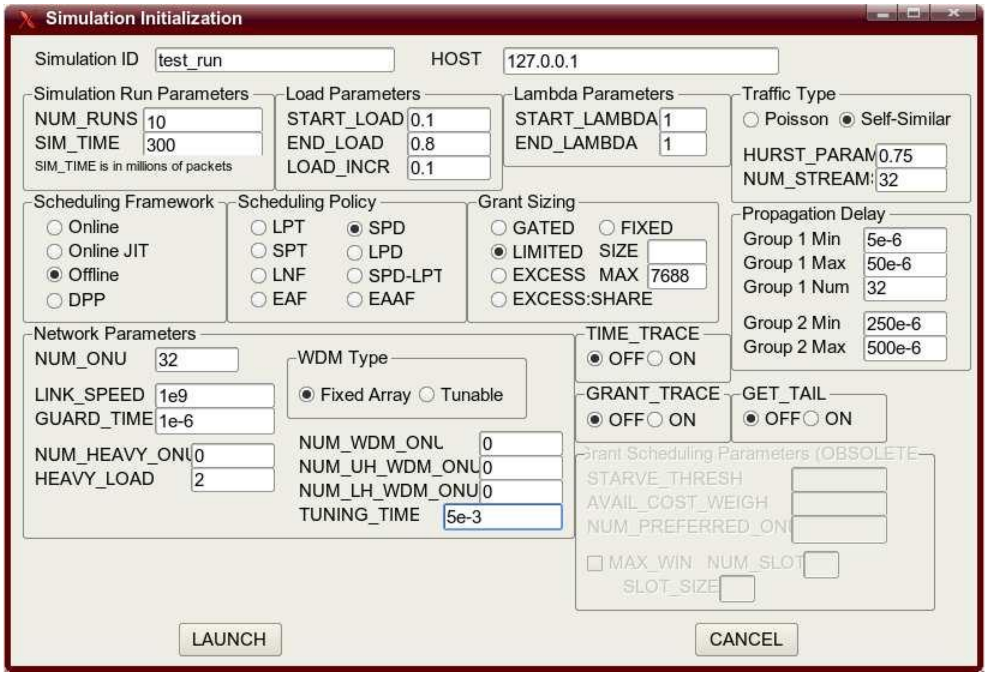
SP scheduling serves high priority order queues first regardless of how much queue lengths are for high priority and low priority queues. Therefore, it may create a situation that lower priority queues continuous wait without a grant. This is regard as unfair because ideally, the scheduler should guarantee a minimal portion of the bandwidth for every priority queue.

WFQ [15] is a packet approximation of ideal traffic model, Generalized Processor Sharing (GPS), whose deviation from the ideal case is bounded by the maximum packet size. WFQ calculates the start time of a packet under the ideal GPS system and based on this start time, computes the finish time under ideal GPS. Then, packets are transmitted in the order of the calculated finish time. The calculations of the ideal GPS times can be computationally intensive for ONUs. A few schemes have been proposed to simplify these calculations at the expense of approximation accuracy to the ideal GPS.

CHAPTER 3

SIMULATION TOOL INTRODUCTION AND EXPERIMENT DESCRIPTION

The simulation tool – Eponsim was programed in C using CSIM discrete event simulation libraries. It has integrated most of the existing DBAs and can combine different framework with different grant sizing methods and grant scheduling policies. Every parameter in the process is adjustable: number of ONUs, number of channels, distances from ONU to OLT etc.



Simulation ID test_run
HOST 127.0.0.1

Simulation Run Parameters
NUM_RUNS 10
SIM_TIME 300
SIM_TIME is in millions of packets

Load Parameters
START_LOAD 0.1
END_LOAD 0.8
LOAD_INCR 0.1

Lambda Parameters
START_LAMBDA 1
END_LAMBDA 1

Traffic Type
○ Poisson ● Self-Similar
HURST_PARAM 0.75
NUM_STREAM: 32

Scheduling Framework
○ Online
○ Online JIT
● Offline
○ DPP

Scheduling Policy
○ LPT ● SPD
○ SPT ○ LPD
○ LNF ○ SPD-LPT
○ EAF ○ EAAF

Grant Sizing
○ GATED ○ FIXED
● LIMITED SIZE
○ EXCESS MAX 7688
○ EXCESS:SHARE

Propagation Delay
Group 1 Min 5e-6
Group 1 Max 50e-6
Group 1 Num 32
Group 2 Min 250e-6
Group 2 Max 500e-6

Network Parameters
NUM_ONU 32
LINK_SPEED 1e9
GUARD_TIME 1e-6
NUM_HEAVY_ONU 0
HEAVY_LOAD 2

WDM Type
● Fixed Array ○ Tunable
NUM_WDM_ONL 0
NUM_UH_WDM_ONU 0
NUM_LH_WDM_ONU 0
TUNING_TIME 5e-3

TIME_TRACE
● OFF ○ ON

GRANT_TRACE
● OFF ○ ON

GET_TAIL
● OFF ○ ON

rant Scheduling Parameters (OBSOLETE
STARVE_THRESH
AVAIL_COST_WEIGH
NUM_PREFERRED_ONI
☐ MAX_WIN NUM_SLOT
SLOT_SIZE

Figure 3.1: Eponsim Simulation Manager Interface

We simulated an EPON with a channel capacity, C, of 1 Gbps and N = 32 ONUs. We varied the maximum propagation delay to represent three different EPON reaches: (1) 1 km to 10 km (5 µsec to 50 µsec), (2) 1 km to 50 km (5 µsec to 250 µsec), and (3) 1 km to 100 km (5 µsec to 500 µsec) (in [7] the authors illustrated the feasibility of these ranges in practical EPON architectures). A quad modal packet size distribution was used for all simulation experiments: 60% 64 bytes, 4% 300 bytes, 11% 580 bytes, and 25% 1518 bytes. We set the guard time, t_g , to 1 μsec , and $G_{max}(i) = 7688$ bytes or 61.5 μsec, ∀ i (i.e., $\sum_{i=1}^{32}(G_{max}(i) + t_g) = 2m\sec$).

The DBAs we conduct simulations with are series of combinations of different scheduling framework, scheduling policies and grant sizing methods, expressed as (α, β, γ):

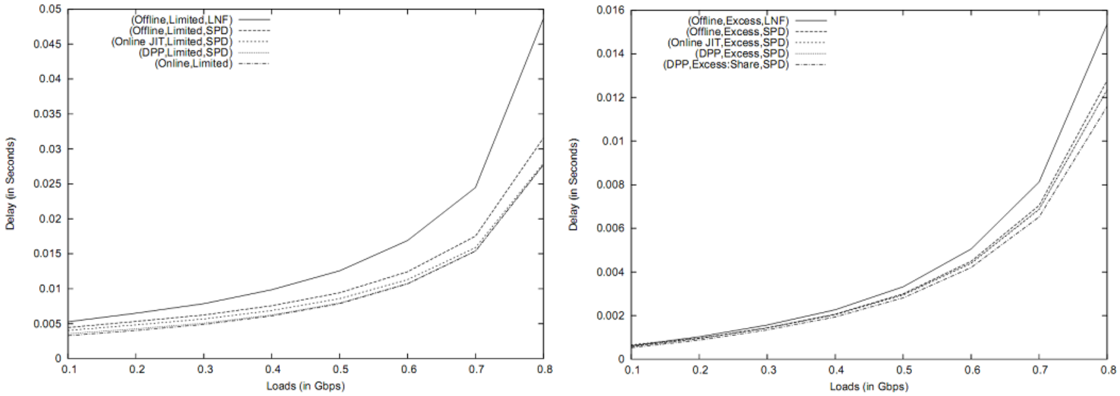
- (Online, Limited)
- (Offline, Limited, LNF)
- (Offline, Limited, SPD)
- (JIT, Limited, SPD)
- (DPP, Limited, SPD)
- (Offline, Excess, LNF)
- (Offline, Excess, SPD)
- (DPP, Excess, SPD)
- (DPP, Excess: Share, SPD)

These simulations contains four previous mentioned scheduling framework, limited and excess grant sizing methods which are more advanced than gated and fixed grant sizing, and SPD and LNF scheduling policies. Since SPD has been mathematically proved to be the optimal scheduling policy under offline framework, simulating all the scheduling frameworks would have little meaning. So we just compare one other scheduling framework to it. The goal of these simulation experiments is to determine: (1) which component of a DBA algorithm has the largest impact on the average packet delay and stability limit measures, and (2) which combination of components of a DBA algorithm provides the lowest average packet delay and highest stability limit.

Chapter 4

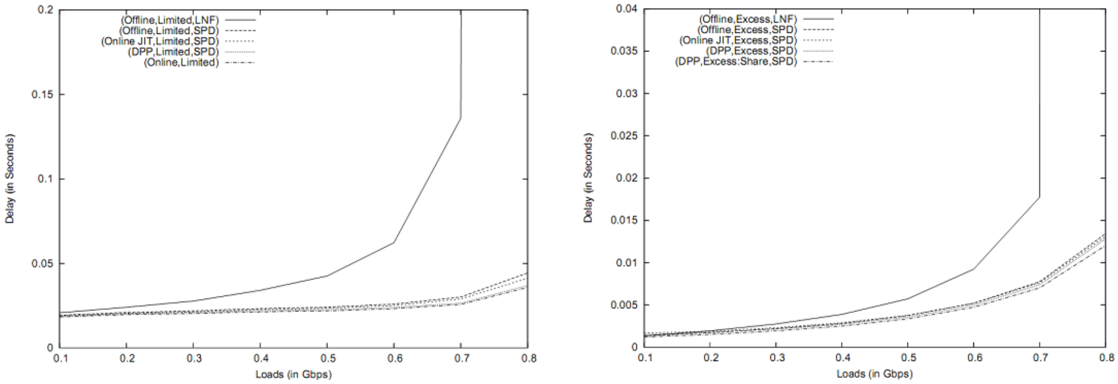
RESULT ANALYSIS

The simulation results are given here in a plot fashion.



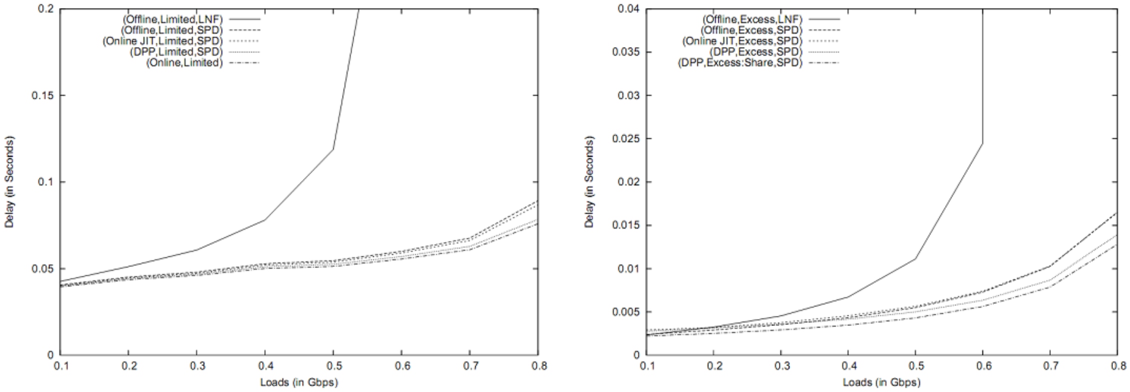
(a) β=Limited, 50μs max. prop. delay (i.e. up to 10km)

(b) β=Excess, 50μs max. prop. delay (i.e. up to 10km)



(c) β=Limited, 250μs max. prop. delay (i.e. up to 50km)

(d) β=Excess, 250μs max. prop. delay (i.e. up to 50km)



(e) β=Limited, 500μs max. prop. delay (i.e. up to 100km)

(f) β=Excess, 500μs max. prop. delay (i.e. up to 100km)

Figure 4.1: Average Packet Delay for DBAs in EPON with Different Propagation Delay Ranges

Figure 4.1 shows the average packet queuing delay values for the nine different DBA algorithms for the three different EPON reach configurations. Those DBA algorithms that used Limited grant sizing are in the left column of the figure and those that used Limited with Excess grant sizing are in the right column.

From the plotted simulation results comparison, we can observe that:

1. The SPD scheduling policy results in much lower average queuing delay than the LNF scheduling policy. The difference increases dramatically with increasing propagation delays.
2. The SPD scheduling policy narrows the average packet queuing delay performance gap between the Offline and Online scheduling frameworks.
3. The DPP scheduling framework results in an average queuing delay very close to the Online scheduling framework.
4. Limited with Excess Distribution grant sizing results in a very significant reduction in average queuing delay compared to Limited grant sizing. Our new method of sharing excess credits among polling groups within the DPP scheduling framework reduces average queuing delay even further compared to not sharing the excess credits.

Observation 1

Since the SPD scheduling policy minimizes the cycle length, the time between transmission grants for each ONU will be shortened and this will lead to reduced queuing delay at each ONU. Let $t_q(i,t)$ be the queuing delay for a packet that arrives at ONU k at time t, $\gamma(i,t)$ be the queued workload at ONU i ahead of the packet that arrived at time t at the start of the granted transmission window in which it is serviced, and R be the upstream transmission rate. For a GATED service discipline,

$$t_q(i,t) = (\min\{t_{start}(i,j) \mid t_{end}(i,j-1) > t\} - t) + \frac{\gamma(i,t)}{R} \quad (4.1)$$

The first term could be reduced by reducing the time between $t_{start}(i,j)$ and $t_{end}(i,j-1)$. Let $t_{idle}(i,j)$ be the upstream channel idle time right before the granted transmission window of the k^{th} ONU in granting cycle j. By definition,

$$t_{idle}(k,j) = \begin{cases} t_{start}(k,j) - t_{end}(N,j-1), k = 1 \\ t_{start}(k,j) - t_{end}(k-1,j), k > 1 \end{cases} \quad (4.2)$$

Using t_{idle}, the value of the time difference between $t_{start}(i,j)$ and $t_{end}(i,j-1)$ is:

$$t_{start}(i,j) - t_{end}(i,j-1) = \sum_{k=i+1}^{N}[t_{idle}(k,j-1) + G(k,j-1)] + \sum_{k=1}^{i-1}[t_{idle}(k,j) + G(k,j)] + t_{idle}(i,j) \quad (4.3)$$

It is clear from this equation that reducing the values of t_{idle} will reduce the differences between $t_{start}(i,j)$ and $t_{end}(i,j-1)$, therefore will reduce the values of $t_q(i,t)$. Grant times will be determined by the work load so reducing the values of t_{idle} is the only tool in which the DBA algorithm can reduce the queuing delay.

For a Limited Service discipline,

$$t_q(i,t) = (\min\{t_{start}(i,j) \mid t_{end}(i, j - \left\lfloor \frac{\beta(i,t)}{G_{max}(i)} \right\rfloor) > t\} - t) + \frac{\gamma(i,t)}{R} \quad (4.4)$$

We now turn our attention to an analysis of the values of t_{idle}. Looking at Eq. 2.2 we see that if $t_{start} = t_a$ for some arbitrary ONU in some arbitrary granting cycle then $t_{idle} = t_g$, a mandatory guard time. However, if $t_{start} = t_{sched} + t_{poll}$, then for an offline scheduling framework,

$$t_{idle}(k, j) = \begin{cases} t_G + 2 \cdot \tau([1, j), k = 1 \\ t_G + (2 \cdot \tau(1, j) - 2 \cdot \tau([k-1, j])) - G(k-1, j), k > 1 \end{cases} \quad (4.5)$$

For DPP scheduling framework,

$$t_{idle}(k, j) = \begin{cases} [t_{end}(N/2, j-1) - t_{end}(N, j-1)] + t_G + 2 \cdot \tau([1.j]), k = 1 \\ [t_{end}(N, j-1) - t_{end}(N/2, j-1)] + t_G + 2 \cdot \tau([N/2.j]), k = N/2+1 \\ t_G + (2 \cdot \tau([k, j]) - 2 \cdot \tau([k-1, j])) - G(k-1, j), k \neq 1, k \neq N/2-1 \end{cases} \quad (4.6)$$

The first square bracketed term in both Eqs. 4.6 when k=1 and k=N/2+1 are clearly negative and will diminish the value of the other two terms in both equations.

SPD by ordering ONUs in ascending order by their propagation delay minimizes the propagation delay of the first ONU, and minimizes the propagation delay difference between adjacent ONUs. As a result, SPD reduces t_{idle} further than other scheduling methods which will cause the propagation delay to the first ONU and the differences between the propagation delays of adjacent ONUs to often be larger.

Observation 2

By reducing t_{idle}, SPD allows t_{idle} to approach the minimum of t_g, a mandatory guard time, which is the same minimum for both the Offline and

Online scheduling frameworks. This narrows the performance gap between the Offline and Online Scheduling frameworks.

Observation 3

For the Online scheduling framework,

$$t_{idle}(k, j) = \begin{cases} [t_{end}(1, j-1) - t_{end}(N, j-1)] + t_G + 2 \cdot \tau \cdot ([1, j]), k = 1 \\ [t_{end}(k, j-1) - t_{end}(k-1, j)] + t_G + 2 \cdot \tau \cdot ([k, j]), k > 1 \end{cases} \quad (4.7)$$

Comparing Eq. 4.7 for online scheduling framework to Eq. 4.6 for DPP scheduling framework, it is clear that the first bracketed terms in each equation are negative. Therefore unless the polling time is greater than the first bracketed term, t_{start} will be determined by the channel available time, t_g and t_{idle} will be a mandatory guard time, t_g.

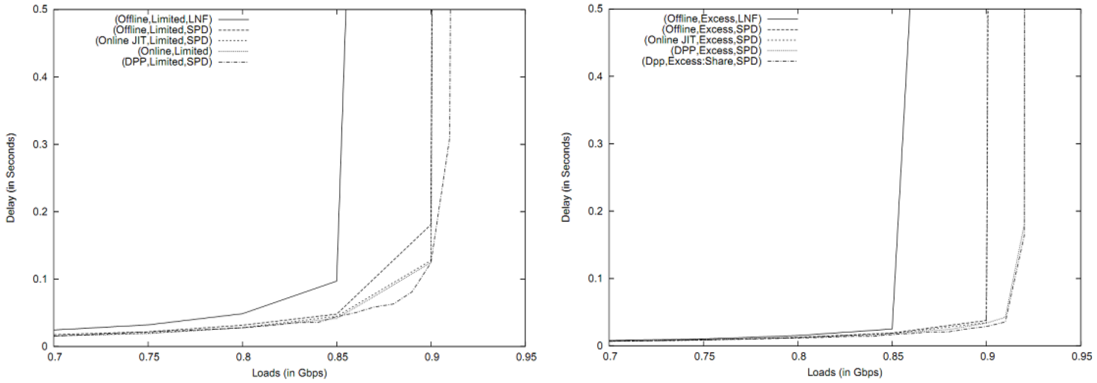
From the experimental data, it appears that the DPP scheduling framework is allowing for t_{idle} to be determined by t_g as frequently as the Online scheduling framework resulting in similar values for t_{start}; thereby closing the gap with the Online scheduling framework.

Observation 4

Limited with Excess Distribution grant sizing provides lower average queuing delay because, as can be seen by comparing Eq. 2.18 to Eq. 2.17, grant sizes are greater than or equal to Limited grant sizing when Limited with Excess Distribution grant sizing is used. Increased grant sizes means that more queued packets can be dequeued and transmitted during the next granting cycle after they are REPORTed. As a result, average queuing delay will be lower when using Limited with Excess Distribution grant sizing. Looking at the experimental data we can quantify the difference. For a presented load of 0.7 Gbps and an EPON reach up to 100 km (DPP, Limited, SPD) yields an average queuing delay of 62.8 msec whereas (DPP, Excess:Share, SPD) yields an average queuing delay of 7.8 msec, a dramatic 87 % decrease. For the same presented load, (Offline, Excess, SPD) yields an average queuing delay of 10.2 msec; (DPP, Excess:Share, SPD) provides a 24 % decrease. Also for the same presented load, (DPP, Excess, SPD) yields an average queuing delay of 8.7 msec; (DPP, Excess:Share, SPD) provides a 13 % decrease.

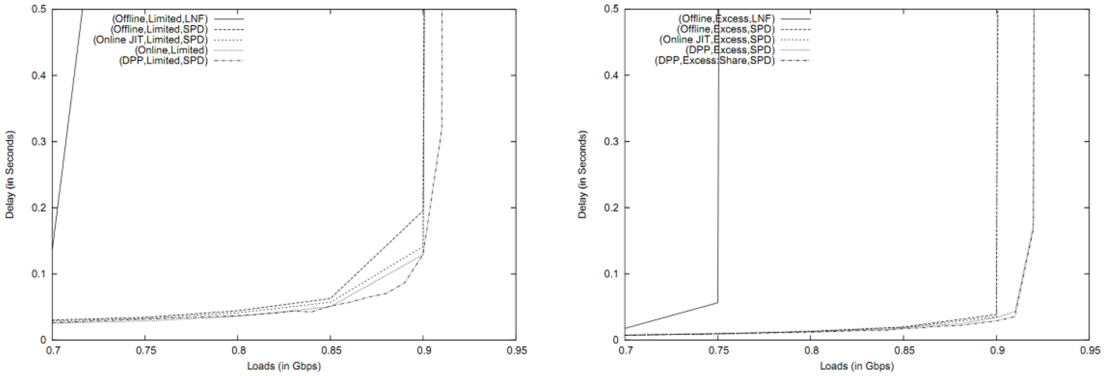
Although Limited with Excess Distribution grant sizing allows for large grant sizes to overloaded ONUs, when the load goes very high, there would be no underloaded ONUs and therefore Limited with Excess Distribution grant sizing turns to Limited grant sizing at very high loads.

In summary, the grant sizing has the largest impact on the average packet delay. Further, the grant scheduling also has a significant impact and the SPD grant scheduling policy can significantly reduce the average packet delay performance gap between the offline and online scheduling frameworks. When SPD grant scheduling is coupled with Limited with Excess distribution grant sizing, the average packet delay performance is better than with the online scheduling framework which cannot take advantage of conventional excess distribution techniques.



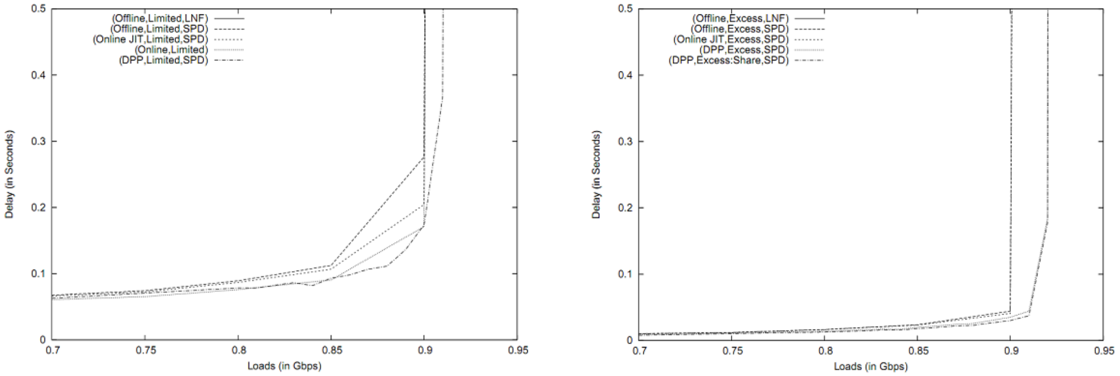
(a) β=Limited, $50\mu s$ max. prop. delay (i.e. up to 10km)

(b) β=Excess, $50\mu s$ max. prop. delay (i.e. up to 10km)



(c) β=Limited, $250\mu s$ max. prop. delay (i.e. up to 50km)

(d) β=Excess, $250\mu s$ max. prop. delay (i.e. up to 50km)



(e) β=Limited, $500\mu s$ max. prop. delay (i.e. up to 100km)

(f) β=Excess, $500\mu s$ max. prop. delay (i.e. up to 100km)

Figure 4.2: Stablity Limit for DBAs in EPON with Different Propagation Delay Ranges

Figure 4.2 shows the delay values for high loads to determine the point at which the delay becomes asymptotically unstable (i.e., the point at which the maximum achievable channel utilization is reached) for the ten different DBA algorithms for the three different EPON reach configurations. Those DBA algorithms that used Limited grant sizing are in the left column of the figure and those that used Limited with Excess grant sizing are in the right column. We can observe from Figure 4.2 that:

1. The SPD scheduling policy results in a much higher stability limit than LNF for the Offline scheduling framework. The difference increases dramatically with increasing propagation delays.
2. The DPP scheduling framework provides a stability limit that is identical to that provided by the Online scheduling framework when using Limited grant sizing. Additionally, the impact of the scheduling policy on the stability is unnoticeable.
3. The Limited with Excess Distribution grant sizing scheme provides the highest stability limit. (Offline, Excess, SPD) provides a stability limit slightly greater than (Online, Limited) and (DPP, Excess, SPD)/(DPP, Excess:Share, SPD) provide the highest stability limit.

Observation 1

As was illustrated in our discussion of Observation 1 in Section 4.1, SPD scheduling by ordering ONUs in ascending order of propagation delay minimizes the value of the propagation delay of the first ONU and the values of propagation delay differences between adjacent ONUs. As a result, SPD scheduling minimizes the values of t_{idle} for an Offline and DPP scheduling framework. As a further result, the channel utilization is minimized each granting cycle.

Let $\eta(j)$ be the channel utilization in granting cycle j.

$$\eta(j) = \frac{\sum_{k=1}^{N} G(k,j)}{\sum_{k=1}^{N} (t_{idle}(k,j) + G(k,j))} \qquad (4.8)$$

It is clear from Eq. 4.8 that a DBA with less t_{idle} will have larger channel utilization. SPD minimizes t_{idle} for the Offline Online JIT and DPP scheduling frameworks and therefore will maximize channel utilization for these scheduling frameworks.

Observation 2

As was illustrated in our discussion of Observation 3 in Section 4.1, when using DPP t_{idle} is determined by the mandatory guard time, t_g as often as the Online scheduling framework. As a result, both scheduling frameworks will produce similar channel utilization that will produce the same stability limit. Additionally, the impact of the adjacent propagation delay differences is diminished resulting in the scheduling policy having an unnoticeable impact on the stability limit.

Observation 3

It is clear from Eq. 4.8 that a grant sizing scheme that produces larger grant sizes will result in higher channel utilization. Limited with Excess distribution grant sizing, Eq. 2.18, clearly illustrate that it will produce grant sizes that are greater than or equal to those produced by Limited grant sizing. As a result, Limited with Excess distribution grant sizing will result in a channel utilization that is greater than or equal to that of Limited grant sizing.

Figure 4.2 shows the delay values comparison of the nine DBA algorithms in three different EPON reach configurations in high traffic loads to observe bandwidth stability limit which is when the maximum channel utilization is reached.

Let $t_{idle}(k, j)$ be the upstream channel idle time right before the grant transmission window the k^{th} ONU in grant cycle j.

$$t_{idle}(k, j) = \begin{cases} t_{start}(k, j) - t_{end}(N, j-1), k = 1 \\ t_{start}(k, j) - t_{end}(k-1, j), k > 1 \end{cases} \qquad (4.9)$$

Channel utilization then can be calculated by Eq 4.8.

From the equation of channel length calculation we can see that longer grant sizes result in higher channel utilization. From limited grant sizing equation and limited with excess distribution equation, it is clear that limited with excess distribution grant sizing will produce grant sizes that are greater than or equal to those produced by Limited grant sizing. As a result, Limited with Excess distribution grant sizing will result in a channel utilization that is greater than or equal to that of Limited grant sizing. Further, a grant sizing scheme that results in larger grant sizes will service more packets in one cycle resulting in a smaller average packet delay. Therefore, Limited with Excess distribution grant sizing will result in an average packet delay that is less than or equal to that of Limited grant sizing.

In summary from these observations, the grant scheduling has the largest impact on the stability limit while the grant sizing has only a modest impact on the stability limit. Further, when SPD grant scheduling is coupled with Limited with Excess distribution grant sizing, the stability limit is the same as with the online scheduling framework. SPD grant scheduling optimally minimizes the granting cycle length for offline and DPP scheduling framework.

CHAPTER 5

CONCLUSION AND FUTURE WORK

In conclusion, a comprehensive introduction and background knowledge of Fiber-Wireless (FiWi) network and Ethernet Passive Optical Network (EPON) are given and an exhaustive study and novel categorization and notation of EPON Dynamic Bandwidth Allocation (DBA) components are provided.

After a comprehensive simulations and analysis, we have obtained the conclusion that grant sizing has the strongest impact on average packet delay. Specifically, Limited with Excess distribution grant sizing produces significantly lower average packet delay compared to Limited grant sizing. We have also noted that the grant scheduling also has a significant impact on the average packet delay. When SPD grant scheduling is coupled with Limited with Excess distribution grant sizing, the average packet delay is less than the LNF coupled with Limited with Excess distribution grant sizing. We have found that grant scheduling has the strongest impact on the stability limit or maximum achievable channel utilization whereas the grant sizing only has a modest impact on the stability limit. Of the nine DBA algorithms examined the SPD grant scheduling policy in the Double Phase Polling scheduling framework coupled with Limited with Share credits Excess distribution grant sizing produced both the lowest average packet delay and the highest stability limit.

Future work includes adding multi-thread polling framework combined with different grant sizing and grant scheduling algorithm into the comparison and obtain a more thorough results to DBA study. Also this paper has been focused on EPONs with a single upstream wavelength channel. A promising avenue of future investigation is to develop a similar notational framework and conduct comprehensive performance evaluations for WDM EPONs with multiple upstream channels.

REFERENCES

[1] Xiaofeng Bai, Abdallah Shami, and Chadi Assi. On the fairness of dynamic bandwidth allocation schemes in ethernet passive optical networks. *Computer Communications*.

[2] Sam Yeoul Choi, Sangho Lee, Tae-Jin Lee, Min Young Chung, and Hyunseung Choo. Double-phase polling algorithm based on partitioned onu subgroups for high utilization in epons. *Optical Communications and Networking, IEEE/OSA Journal of*, 1(5):484 –497, oct. 2009.

[3] P. Chowdhury, B. Mukherjee, S. Sarkar, G. Kramer, and S. Dixit. Hybrid wireless-optical broadband access network(woban): prototype development and research challenges. *Network, IEEE*, 23(3):41–48, may-june 2009.

[4] T.-S. Chu and M.J. Gans. Fiber optic microcellular radio. *Vehicular Technology, IEEE Transactions on*, 40(3):599 –606, aug 1991.

[5] Bao Linh Dang and I. Niemegeers. Analysis of ieee 802.11 in radio over fiber home networks. In *Local Computer Networks, 2005.30th Anniversary. The IEEE Conference on*, pages 744 –747, nov. 2005.

[6] A.R. Dhaini, C.M. Assi, M. Maier, and A. Shami. Dynamic wavelength and bandwidth allocation in hybrid tdm/wdm epon networks. *Lightwave Technology, Journal of*, 25(1):277 –286, jan. 2007.

[7] T. Holmberg. Analysis of epons under the static priority scheduling scheme with fixed transmission times. In *Next Generation Internet Design and Engineering, 2006. NGI '06.2006 2nd Conference on*, pages 8 pp. –199, 0-0 2006.

[8] Zhensheng Jia, Jianjun Yu, G. Ellinas, and Gee-Kung Chang. Key enabling technologies for optical-wireless networks: Optical millimeter-wave generation, wavelength reuse, and architecture. *Lightwave Technology, Journal of*, 25(11):3452–3471, nov. 2007.

[9] G. Kramer, A. Banerjee, N.K. Singhal, B. Mukherjee, Sudhir Dixit, and Yinghua Ye. Fair queueing with service envelopes (fqse): a cousin-fair hierarchical scheduler for subscriber access networks. *Selected Areas in Communications, IEEE Journal on*, 22(8):1497–1513, oct. 2004.

[10] G. Kramer, B. Mukherjee, and G. Pesavento. Ipact a dynamic protocol for an ethernet pon (epon). *Communications Magazine, IEEE*, 40(2):74–80, feb 2002.

[11] Glen Kramer, Biswanath Mukherjee, and Gerry Pesavento. Ethernet pon (epon): Design and analysis of an optical access network, 2001.

[12] C. Lethien, C. Loyez, and J.-P. Vilcot. Potentials of radio over multimode fiber systems for the in-buildings coverage of mobile and wireless lan applications. *Photonics Technology Letters, IEEE*, 17(12):2793 – 2795, dec. 2005.

[13] Chun-Ting Lin, J. Chen, Peng-Chun Peng, Cheng-Feng Peng, Wei-Ren Peng, BiShiou Chiou, and S. Chi. Hybrid optical access network integrating fiber-to-the-home and radio-over-fiber systems. *Photonics Technology Letters, IEEE*, 19(8):610 –612, april15, 2007.

[14] Yuanqiu Luo and N. Ansari. Bandwidth allocation for multiservice access on epons. *Communications Magazine, IEEE*, 43(2):S16 –S21, feb. 2005.

[15] Yuanqiu Luo and Nirwan Ansari. Limited sharing with traffic prediction for dynamic bandwidth allocation and qos provisioning over ethernet passive optical networks. *J. Opt. Netw.*, 4(9):561–572, Sep 2005.

[16] M. Maier, M. Herzog, and M. Reisslein. Stargate: the next evolutionary step toward unleashing the potential of wdm epons [topics in optical communications]. *Communications Magazine, IEEE*, 45(5):50 –56, may 2007.

[17] Martin Maier, Navid Ghazisaidi, and Martin Reisslein. The audacity of fiberwireless (fiwi) networks. In *AccessNets'08*, pages 16–35, 2008.

[18] M. Mcgarry, M. Reisslein, and M. Maier. Ethernet passive optical network architectures and dynamic bandwidth allocation algorithms. *Communications Surveys Tutorials, IEEE*, 10(3):46 –60, quarter 2008.

[19] M.P. McGarry, M. Reisslein, F. Aurzada, and M. Scheutzow. Impact of epon dba components on performance. In *Computer Communications and Networks (ICCCN), 2011 Proceedings of 20th International Conference on*, pages 1 –5, 31 2011-aug. 4 2011.

[20] M.P. McGarry, M. Reisslein, C.J. Colbourn, M. Maier, F. Aurzada, and M. Scheutzow. Just-in-time scheduling for multichannel epons. *Lightwave Technology, Journal of*, 26(10):1204 –1216, may15, 2008.

[21] M.P. McGarry, M. Reisslein, and M. Maier. Wdm ethernet passive optical networks. *Communications Magazine, IEEE*, 44(2):15 – 22, feb. 2006.

[22] A. Shami, X. Bai, N. Ghani, C.M. Assi, and H.T. Mouftah. Qos control schemes for two-stage ethernet passive optical access networks. *Selected Areas in Communications, IEEE Journal on*, 23(8):1467 – 1478, aug. 2005.

[23] S.R. Sherif, A. Hadjiantonis, G. Ellinas, C. Assi, and M.A. Ali. A novel decentralized ethernet-based pon access architecture for provisioning dif-

ferentiated qos. *Lightwave Technology, Journal on*, 22(11):2483–2497, nov. 2004.

[24] H. Song, B.-W. Kim, and B. Mukherjee. Multi-thread polling: a dynamic bandwidth distribution scheme in long-reach pon. *Selected Areas in Communications, IEEE Journal on*, 27(2):134 –142, february 2009.

[25] Pak Kay Tang, Ling Chuen Ong, A. Alphones, B. Luo, and M. Fujise. Per and evm measurements of a radio-over-fiber network for cellular and wlan system applications. *Lightwave Technology, Journal of*, 22(11):2370–2376, nov. 2004.

[26] Jingshown Wu, Jiunn-Shyen Wu, and H.-W. Tsao. A fiber distribution system for microcellular radio. *Photonics Technology Letters, IEEE*, 6(9):1150–1152, sep 1994.

[27] J. Zheng and H.T. Mouftah. Adaptive scheduling algorithms for ethernet passive optical networks. *Communications, IEE Proceedings-*, 152(5):643–647, oct. 2005.