

國立交通大學

電信工程學系碩士班

碩士論文

乙太被動光纖擷取網路之上鏈路排程機制

Uplink Scheduling Schemes in EPON Access Network



研究生：彭崇禎

指導教授：張仲儒 教授

鄭瑞光 教授

中華民國九十三年六月

乙太被動光纖擷取網路之上鏈路排程機制

Uplink Scheduling Schemes in EPON Access Network

研究生：彭崇禎
指導教授：張仲儒 教授
鄭瑞光 教授

Student: Chun-Chen Peng
Advisor: Dr. Chung-Ju Chang
Dr. Ray-Guang Cheng

國立交通大學

電信工程學系碩士班



A Thesis
Submitted to Institute of Communication Engineering
College of Electrical Engineering and Computer Science
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in
Electrical Engineering
June 2004
Hsinchu, Taiwan, Republic of China

中華民國九十三年六月

乙太被動光纖擷取網路之上鏈路 排程機制

研究生：彭崇禎

指導教授：張仲儒教授
鄭瑞光教授

國立交通大學電信工程學系碩士班

中文摘要

寬頻服務在家庭及小型公司的需求激增已經成為擷取網路技術日益進步的主要因素。而如今，乙太被動光纖擷取網路(EPON)更被認為是下一代擷取網路的明日之星，原因是它整合了既有的光纖基礎建設以及目前已發展成熟並且成本低廉的乙太設備。除此之外，另一個讓乙太被動光纖網路備受注目的原因是它解決了寬頻服務所帶來對頻寬需求激增的問題，像是 IP 電話、隨選視訊(VoD)等服務無不催促著網路操作者致力於發展能提供全服務性質的擷取網路。在本篇論文中，我們提出了一個能適用於乙太被動光纖擷取網路的上鏈路排程方法，我們針對及時性的服務(例如語音服務、視訊服務等)設計了以延遲為考量的排程機制；另外，也針對非及時性的服務(例如一般資料服務)而設計了兩個以公平性作考量的排程機制，分別為 Hybrid LQF-QLP 機制以及 Hybrid EQL-QLP 機制。模擬的結果顯示我們所提出的排程演算法確實能夠讓語音封包的平均延遲限制在一個可接受的範圍之內，並同時對於資料服務考量了封包延遲與封包阻隔機率的公平性。

再者，我們提出了一個利用預測的方式來排程的機制，其中預測器是採用了移動平均(Moving Average)的技巧。我們發現當加入一個預測器之後，最大週期時間將可延長並且能提高系統的封包流量。模擬的結果顯示這種預測的方式的確能夠有效的提升系統的效能。

Uplink Scheduling Schemes in EPON Access Network

Student: Chun-Chen Peng

Advisor: Dr. Chung-Ju Chang
Dr. Ray-Guang Cheng

Institute of Communication Engineering
National Chiao Tung University

Abstract

Rapid deployment of broadband services in the residential and small business area has played an important role in the evolution of access networks. Currently, Ethernet passive optical networks (EPON) are being considered as a promising solution for the next generation access network, due to the convergence of low-cost Ethernet equipment and low-cost of fiber infrastructure. In addition, the growing demand of broadband services such as IP telephony, video on demand has urged the network operator to accelerate the deployment of full-service access networks. In this thesis, we proposed a delay-considered scheduling scheme for real-time services, i.e. voice and video service, and two fairness-considered scheduling schemes, i.e. Hybrid LQF-QLP scheme and Hybrid EQL-QLP scheme, to support non-real-time data service. The goal of the scheduling algorithm is to meet the delay bound of voice service, and to simultaneously maintain the fairness of both packet delay and packet blocking probability for non-real-time data service. Simulation results show that the proposed scheduling method can meet our goal.

In addition, we proposed a prediction-based scheduling method, in which we adopt a Moving Average technique. We find that by implementing a predictor, the maximum cycle time can be extended and the system throughput can be improved. Simulation results show that the proposed scenario can improve performance well.

致 謝

回首過去這兩年的碩士生涯，內心滿是無限的感動及回憶。研究的過程中常有許多的靈機一動，但也遭遇過無數的希望破滅；嘗過收穫時的喜悅卻不知經歷過多少低潮及失意；兩年同聚的歡笑也終將面臨別離時的難捨，這所有的酸甜苦辣都是寶貴的人生歷練，豐富了我的生命、增長了我的智慧，我有幸能參與其中，並期許自己能以此為成長的跳板，百尺竿頭，更進一步。

碩士論文能順利的完成，我必須要感謝許許多多幫助過我的貴人。首先必須感謝恩師張仲儒教授以及學長兼指導教授鄭瑞光老師的殷勤指導、批評指正以及無數的金玉良言。除了讓我學到嚴謹的治學態度與研究方法之外，兩位老師更開啟了一扇“超越自我”的大門，砥礪我們時時刻刻不忘充實自己，使自己能禁得起時代的考驗。

在實驗室的前輩們中，我先要感謝惠我良多也是令我十分尊敬的立峯學長，沒有學長在百忙之中犧牲自己的時間並熱心參與討論、傾囊相授，學弟的研究過程中勢必將困難許多；再來要感謝聰明絕頂、妙語如珠的家慶、義昇、正文及詠翰學長，總是在我最需要幫助的時候給予我許多寶貴的經驗及建議；另外還要感謝寧佑、凱盟、至永與小白學長在我剛進實驗室時親切的給予提攜與照顧，讓我能很快的適應研究生的生活。除了學長們的提攜、照顧之恩，志明、朕逢、凱元、宗軒及立忠學弟們對我的幫助也是功不可沒，感謝你們替我分擔許多壓力並與我分享生活點滴。當然，在實驗室中我還要感謝與我朝夕相處、一同奮戰的俊憲、皓棠夥伴，與你們同甘共苦的這段日子會是我碩士生涯中最重要的回憶。

最後，我要感謝女友、大哥、父母及其他親友對我的關懷與照顧，你們的支持是讓我勇往直前的最大的動力，沒有這份根植於心中的親情力量，我將永遠無法開拓自己的疆野。我願將這碩士文獻給所有幫助過我、愛護我的人，希望你們與我一起分享這得來不易的成果及心中的喜悅。

彭崇禎 謹誌

民國九十三年

Contents

| | |
|--|-------------|
| 中文摘要..... | i |
| Abstract..... | ii |
| Acknowledgement..... | iii |
| Contents..... | iv |
| List of Figures..... | vi |
| List of Tables..... | viii |
| Chapter 1 Introduction..... | 1 |
| Chapter 2 Uplink Scheduling in EPON Access Network..... | 7 |
| 2.1 Introduction..... | 7 |
| 2.2 Background of EPON..... | 9 |
| 2.3 System Model..... | 12 |
| 2.3.1 System Architecture..... | 12 |
| 2.4 Scheduling Algorithm..... | 17 |
| 2.4.1 Hybrid EQL-QLP scheme..... | 21 |
| 2.4.2 Hybrid LQF-QLP scheme..... | 22 |
| 2.5 Simulation Result..... | 24 |
| 2.5.1 Traffic Source Models..... | 24 |
| 2.5.2 Simulation Environment..... | 28 |
| 2.5.3 Simulation Result and Discussion..... | 30 |
| 2.6 Concluding Remarks..... | 39 |
| Chapter 3 Prediction-Based Scheduling Algorithms..... | 40 |
| 3.1 Introduction..... | 40 |
| 3.2 System Model..... | 41 |
| 3.3 Predictor..... | 43 |
| 3.4 Prediction-based Scheduling Algorithm..... | 45 |
| 3.4.1 Hybrid PEQL-PQLP Scheme..... | 46 |
| 3.4.2 Hybrid PLQF-PQLP scheme..... | 47 |
| 3.5 Simulation Result..... | 49 |
| 3.5.1 Simulation Environment..... | 49 |
| 3.5.2 Simulation Result and Conclusion..... | 50 |

| | |
|-----------------------------------|-----------|
| 3.6 Concluding Remarks..... | 59 |
| Chapter 4 Conclusion | 60 |
| Appendix A | 62 |
| Appendix B | 63 |
| Appendix C | 64 |
| Bibliography | 68 |
| Vita | 71 |



List of Figures

| | |
|--|----|
| Figure 1.1: An EPON network..... | 1 |
| Figure 1.2: Downstream transmissions in EPON | 2 |
| Figure 1.3: Upstream transmissions in EPON | 3 |
| Figure 2.1: GATE message and REPORT message frame format..... | 10 |
| Figure 2.2: Round-trip time measurement | 11 |
| Figure 2.3: The EPON system architecture | 12 |
| Figure 2.4: Functional modules of a scheduler..... | 13 |
| Figure 2.5: IPACT-based polling procedural | 15 |
| Figure 2.6: The case that channel is fully filled with real-time voice packets..... | 19 |
| Figure 2.7: The average packet delay of voice service | 20 |
| Figure 2.8: Synthetic self-similar traffic generator..... | 26 |
| Figure 2.9: ON/OFF Parato-distributed source i..... | 26 |
| Figure 2.10: Average Packet Delay of Voice Service by adopting Hybrid LQF-QLP..... | 31 |
| Figure 2.11: Average packet blocking probability (high-loading ONUs) of data service by adopting Hybrid LQF-QLP..... | 32 |
| Figure 2.12: Packet blocking probability fairness index of data service by adopting Hybrid LQF-QLP | 32 |
| Figure 2.13: Average Packet Delay (high-loading ONUs) of Data Service by adopting Hybrid LQF-QLP | 33 |
| Figure 2.14: Packet delay fairness index of data service by adopting Hybrid LQF-QLP | 34 |
| Figure 2.15: Overall fairness index of data service by adopting Hybrid LQF-QLP..... | 34 |
| Figure 2.16: Average packet delay of voice service by adopting Hybrid EQL-QLP..... | 35 |
| Figure 2.17: Average packet blocking probability (high-loading ONUs) of data service by adopting Hybrid EQL-QLP..... | 36 |
| Figure 2.18: Packet blocking probability fairness index of data service by adopting Hybrid EQL-QLP | 36 |
| Figure 2.19: Average Packet Delay (high-loading ONUs) of Data Service by adopting Hybrid EQL-QLP | 37 |
| Figure 2.20: Packet delay fairness index of data service by adopting Hybrid EQL-QLP | 38 |
| Figure 2.21: Overall fairness index of data service by adopting Hybrid EQL-QLP | 38 |
| Figure 3.1: A prediction-based EPON model | 42 |
| Figure 3.2: The prediction-based scheduler architecture..... | 43 |
| Figure 3.3: The concept of prediction method..... | 44 |
| Figure 3.4: The System Throughput ($T_{\max} = 0.66\text{ms}$)..... | 51 |

| | |
|---|----|
| Figure 3.5: Average packet delay of Voice Service ($T_{\max} = 0.66\text{ms}$)..... | 52 |
| Figure 3.6: Packet blocking probability fairness index of data service ($T_{\max} = 0.66\text{ms}$) | 53 |
| Figure 3.7: Packet delay fairness index of data service ($T_{\max} = 0.66\text{ms}$) | 53 |
| Figure 3.8: Average Packet Delay of Voice Service | 54 |
| Figure 3.9: The System Throughput | 55 |
| Figure 3.10: Fairness index of packet blocking probability | 57 |
| Figure 3.11: Fairness index of packet delay..... | 58 |
| Figure 3.12: Overall fairness index of data service | 58 |
| Figure A.1: First example of EQL scheme | 65 |
| Figure A.2: Second example of EQL scheme | 66 |



List of Tables

| | |
|--|----|
| Table 2.1: Parameter descriptions used in scheduling algorithm..... | 18 |
| Table 2.2: System parameters used in the simulation environment..... | 29 |
| Table 3.1: System parameters used in prediction-based environment..... | 49 |



Chapter 1

Introduction

While the backbone network bandwidth grows tremendously, the access network still remains the bottleneck. Ethernet passive optical networks (EPONs), which represent the convergence of low-cost Ethernet equipment and low-cost fiber infrastructure, appear to be one of the best candidates for the next-generation access network [1]. The low-cost of fiber infrastructure and the high-speed Gigabit Ethernet equipment make EPON very attractive. In addition, the growing demand of broadband services such as IP telephony, video on demand (VoD) has urged the network operator to accelerate the deployment of full-service access networks (FSAN) [2].

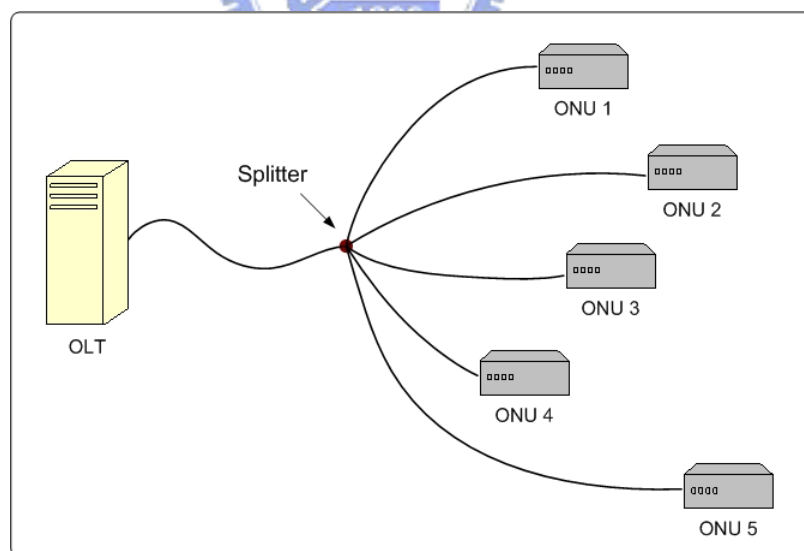


Figure 1.1: An EPON network

Figure 1.1 shows a tree-topology EPON network. The main components in EPON are optical line terminal (OLT) and optical network unit (ONU). The OLT resides in the central office (CO) and connects the optical access network to the metropolitan area network (MAN) or wide-area network

(WAN). The ONU is usually located at either the curb or the end-user location, and provides broadband video, data, and voice services.

Passive optical network is a point-to-multipoint optical network with no active elements in the signals' path from source to destination. The only interior elements used in a PON are passive optical components, such as optical fiber and splitters. An EPON is a PON that carries all data encapsulated in Ethernet frames and is backward compatible with existing IEEE 802.3 Ethernet standards. Moreover, Ethernet is an inexpensive technology that is ubiquitous with a variety of legacy equipment. Due to the growing demand of broadband services such as IP telephony, video on demand (VoD), an EPON with high bandwidth is a promising infrastructure in access network.

In the downstream direction, as shown in Figure 1.2, the OLT broadcasts Ethernet packets to all ONUs via a passive splitter. The ONU checks the MAC address of every incoming packet and discards the packet if the packet is not destined to the ONU.

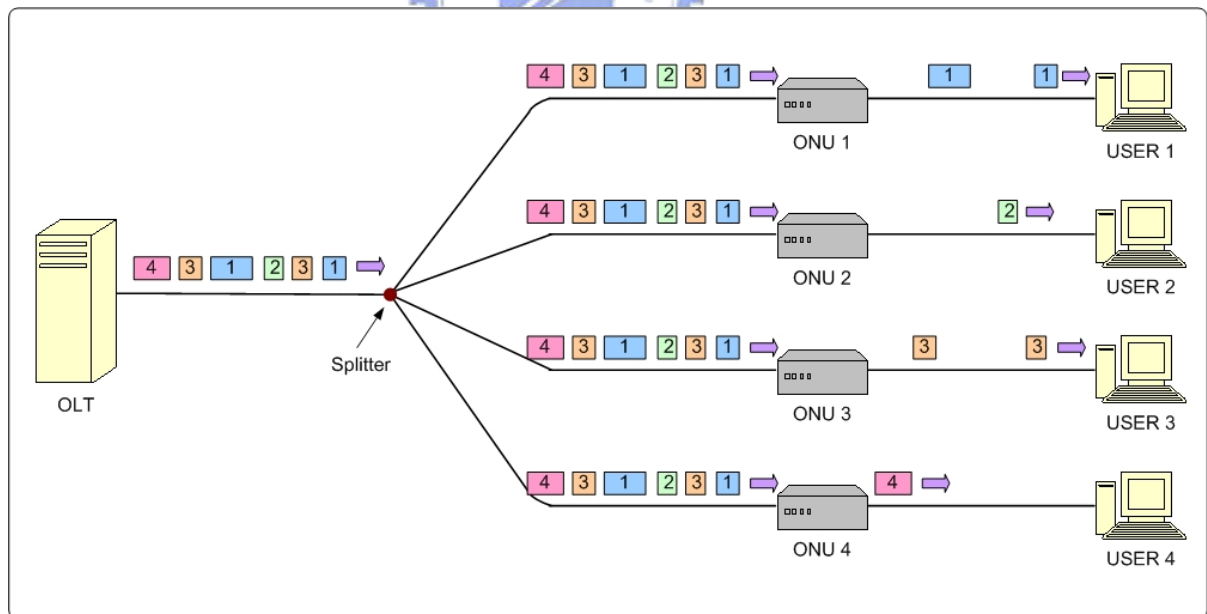


Figure 1.2: Downstream transmissions in EPON

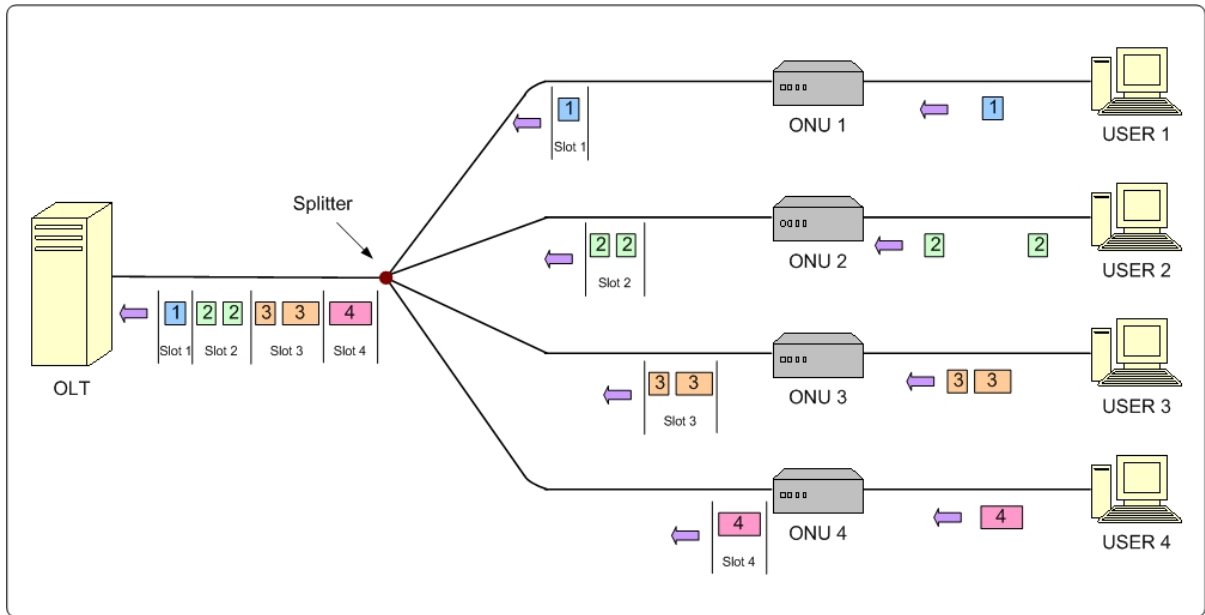


Figure 1.3: Upstream transmissions in EPON

In the upstream direction (Figure 1.3), a multiple access scheme should be adopted by the EPON to prevent the collision of packets originating from different ONUs. In wavelength division multiple access (WDMA) scheme [1], each ONU transmits data at different wavelengths. So it would require a tunable receiver at the OLT to receive multiple channels. The drawback of the WDMA scheme is that the WDM components are very expensive. So it is not a good cost-effect scheme in EPON environment. Contention-based multiple access scheme (similar to CSMA/CD) is difficult to implement, because of the directional property of the splitter and the problem to detect collisions in ONUs. The optical code division multiple access (OCDMA) scheme is an alternative solution of supporting multiple access in EPON [3]. It allows active users share the same wavelength band and transmit data simultaneously. OCDMA has several advantages, such as flexibility in network design and security communication capability. But due to no negative power for optical signals, it is more complicated to implement than the other schemes. Compare with the other schemes, TDMA schemes is the most cost-effective solution for EPON network, because it requires only one transceiver in the OLT and it is easy to implement.

By adopting TDMA scheme, all ONUs must be synchronized to a common clock. The OLT

allocates timeslots to ONUs. Each timeslot is able to carry several Ethernet packets. An ONU should buffer packets received from a subscriber and forward the packets to the OLT until its timeslot arrives. If the buffered frames are not able to fill the entire timeslot, idles frames are transmitted. In the OLT, a bandwidth allocation scheme should be provided to allocate timeslots to ONUs in a static or dynamic way.

There are many bandwidth allocation schemes proposed for EPON. In [4], a scheduling scheme of fixed timeslot assignment algorithm for EPON was presented. This method adopted a TDMA approach to deliver Ethernet packets. A fixed duration of the timeslot is allocated to each ONU. The advantages of this scheme are easy to implement and able to provide the service for multiple users in a single wavelength. However, the result also showed that a considerable amount of bandwidth was wasted because the unused bandwidth can not be shared to high-loading ONUs (no statistical multiplexing).

To increase the channel utilization, an OLT-based polling scheme was proposed in [5], called Interleaved Polling with Adaptive Cycle Time (IPACT). It adopted a polling-based scheme to deliver Ethernet packets from ONUs to OLT over the EPON access network. In IPACT, the duration of the timeslot is variable and the channel utilization is improved. Different scheduling algorithms were studied in [5], i.e. fixed, limited, gated, constant credit, linear credit, and elastic. The limited scheduling algorithm, in which the OLT grants an ONU the requested number of bytes with an upper bound, exhibits the best performance among these algorithms.

EPON is expected to support emerging IP-based multimedia traffic with diverse quality-of-service (QoS) requirements [6]. [7] proposed a limited service scheme to support QoS in EPON network. The proposed scheme adopts a gated transmission mechanism (MPCP, Multi-Point Control Protocol) with priority scheduling. A strict-priority (exhaustive) buffer management is adopted and thus, the packet delay of the real-time services is guaranteed. In the proposed scheme, there exists the effect of light-load penalty, where the queuing delay for some traffic classes

increases when the network load decreases. The authors proposed to adopt a two-stage buffering method or a CBR credit method to eliminate the effect. The two-stage buffering method eliminates the effect but it also increases the delays for higher priority classes. The CBR credit method can only partially eliminate the effect but it requires external knowledge of the arrival process.

A promising approach to support differentiated QoS is to employ a central controller that can dynamically allocate bandwidth to end users according to the required bandwidth. Thus, bandwidth management for fair bandwidth allocation among different ONUs will be an important issue for the MAC protocols in the emerging EPON based networks. In [8], [9], and [10], different dynamic bandwidth allocation (DBA) algorithms are proposed. In [8], the authors proposed a DBA algorithm for multimedia services over EPON. They use strict priority queueing and presented control message formats to implement. In [9], the authors proposed a new bandwidth guaranteed polling (BGP) scheme that allows the upstream bandwidth to be shared based on the service level agreement between each subscriber and the operator. In [10], the authors proposed a DBA algorithm with QoS support over EPON-based access network. The DBA algorithms described above can support the guaranteed requirement, i.e. assigned bandwidth, packet delay, and etc., of highest priority service. However, there is no mechanism to take care of non-real-time service, which usually has no delay criterion during packet transmission. So, if the loading is unbalanced but still under the requested bandwidth, the non-real-time data service may experience much of difference in packet delay and packet blocking probability between each ONU. In other words, one of the important issues we want to consider about the non-real-time service is the fairness in average packet delay and average packet blocking probability among all ONUs.

In [11] and [12], we find that Queue Length Proportional (QLP) scheme and Longest Queue First (LQF) scheme can achieve the fairness in packet delay and blocking probability, respectively. Based on the idea of QLP and LQF, we proposed a Hybrid LQF-QLP scheduling algorithm and a Hybrid EQL-QLP scheduling algorithm to simultaneously ensure the fairness of packet delay and

blocking probability. Details of these two proposed schemes are described in Chapter 2.

In Chapter 2, we obtain the relationship between the maximum cycle time and the delay criterion of the voice service. That is, the maximum cycle time should be lower than an upper bound in order to meet the QoS requirement of the voice service. This phenomenon is the same as the effect of two-stage buffering which mentioned in [7]. However, shorter maximum cycle time gives rise to the lower channel utilization, because of the irremovable overhead introduced by control messages and guard time. Thus, to improve the channel utilization, we proposed a prediction-based scheduler architecture, which can effectively extend the maximum cycle time. We denote the new scheduling algorithms used in this environment as prediction-based scheduling schemes, i.e. PEQL-PQLP, and PLQF-PQLP. We will discuss the detail in Chapter 3.

And finally, a concluding remark is given in Chapter 4.



Chapter 2

Uplink Scheduling in EPON Access Network

2.1 Introduction

Rapid deployment of broadband services in the residential and small business area has played an important role in the evolution of access networks. Currently, Ethernet passive optical networks (EPON) [1] are being considered as a promising solution for the next generation access network, due to the convergence of low-cost Ethernet equipment and low-cost of fiber infrastructure. In the upstream direction in EPON, a multiple access scheme should be adopted by the EPON to prevent the collision of packets originating from different ONUs.

In the upstream direction of EPON, a multiple access scheme is needed to prevent the collision of packets originating from different ONUs. In [5], a polling-based scheme, called Interleaved Polling with Adaptive Cycle Time (IPACT) was proposed, in which the next ONU is polled before receiving the data from previous ONU. As a result, the channel utilization is higher, compare with traditional TDMA scheme [4], which assign fixed timeslot to all ONUs. Different scheduling algorithms were studied in [5], i.e. fixed, limited, gated, constant credit, linear credit, and elastic. The limited scheduling algorithm, in which the OLT grants an ONU the requested number of bytes with an upper bound, exhibits the best performance among these algorithms.

In [7], the author adopted a limited service scheme to support QoS in EPON network. A strict-priority (exhaustive) buffer management is adopted and, thus, the packet delay of the real-time

services is guaranteed. In the proposed scheme, there exists the effect of light-load penalty, where the queuing delay for some traffic classes increases when the network load decreases. The authors proposed to adopt a two-stage buffering method or a CBR credit method to eliminate the effect. The two-stage buffering method eliminates the effect but it also increases the delays for higher priority classes. The CBR credit method can only partially eliminate the effect but it requires external knowledge of the arrival process.

The limited service scheme combined with strict-priority buffer management is a simple way to support different class of service. The QoS of the highest-priority service can be guaranteed. However, the channel utilization is limited if only some ONUs are under high-loading condition, but the others are not. The reason is that the high-loading ONUs can't be granted to transmit packets which data volume exceeds an upper bound. Moreover, there is no mechanism to take care of non-real-time service, which usually has no delay criterion during packet transmission. So, if the loading is unbalanced but still under the requested bandwidth, the non-real-time service may experience much of difference in packet delay and packet blocking probability between each ONU.

One of the important issues we want to consider about the non-real-time service is the fairness in average packet delay and average packet blocking probability among all ONUs. In [11] and [12], we find that Queue Length Proportional (QLP) scheme and Longest Queue First (LQF) scheme can achieve the fairness in packet delay and blocking probability, respectively. Based on the idea of QLP and LQF, we proposed a Hybrid LQF-QLP scheduling algorithm and a Hybrid EQL-QLP scheduling algorithm to simultaneously ensure the fairness of packet delay and blocking probability.

The chapter is outlined as follows. In Session 2.2, we introduce the key properties of EPON that would be used in our model. In Session 2.3, we will describe the detailed system model including the EPON system architecture and the scheduler functional block. In Session 2.4, two proposed scheduling algorithms are described. And finally, in Session 2.5, we show the simulation result and make a conclusion.

2.2 Background of EPON

The most important packets in EPON network are control packets (Control messages), they are specified in IEEE standard 802.3ah [16]. The work on EPON architecture in the IEEE 802.3ah task force is still in progress. The Multi-Point Control Protocol (MPCP) is chosen by this task force to facilitate the implementation of various bandwidth-allocation algorithms in EPONs. The bandwidth-allocation algorithm is performed based on two types of Ethernet packets, GATE and REPORT, defined by MPCP. A GATE message is sent from the OLT to an ONU, and it used to assign a transmission timeslot. A REPORT message is used by an ONU to convey its local conditions to the OLT to help the OLT make intelligent allocation decisions. These control messages are basic IEEE 802.3 frames, and the packet size of the control messages are all 64-bytes, which is the smallest packet size of IEEE 802.3 frame. The frame structures are shown in Figure 2.1.

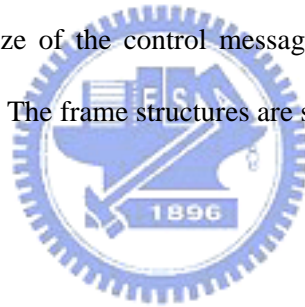




Figure 2.1: GATE message and REPORT message frame format

The definitions of each field list as follows:

- Header: The header field includes the information about Destination Address, Source Address, and Type. These are basic Ethernet packet field.
- Opcode: This field identifies the specific control message being encapsulated. The value of GATE message is 2, and the value of REPORT message is 3.
- Timestamp: The timestamp field conveys the content of the local time register at the time of transmission of the control message.
- Pad/Reserved: This field is used for the payload of control message.

- (e) Number of grants/Grant #n start time/Grant #n length: In GATE message, we can have at most 4 grants. Upon the number of grants, we must decide the start time and length of each grant.
- (f) Number of queue sets/REPORT bitmap/Queue #n report: Similar to GATE message, in REPORT message, at most 8 queues can request the bandwidth based on design purpose.

The Timestamp information in an ONU's REPORT message will be taken into account to get new round-trip-time (RTT) between this ONU and OLT. The measurement of RTT to the source ONU is shown in Figure 2.2.

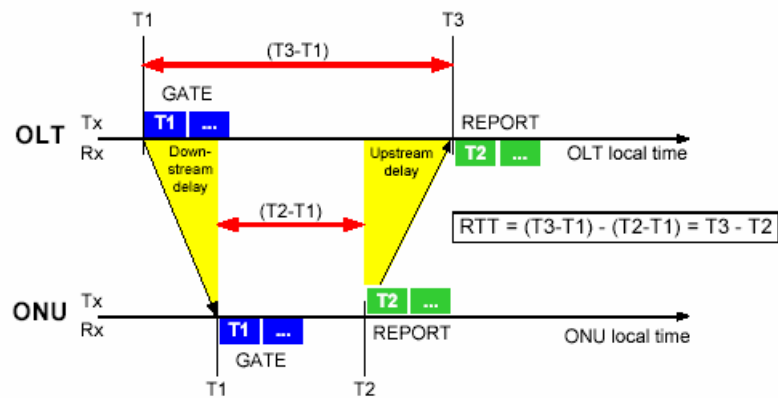


Figure 2.2: Round-trip time measurement

2.3 System Model

2.3.1 System Architecture

We consider an EPON access network consisting of an OLT and N ONUs as shown in Figure 2.3. In the OLT, there is a scheduler, which is responsible for determining the time and data volume that each ONU can transmit during a cycle, where the cycle is the duration between successive scheduling times.

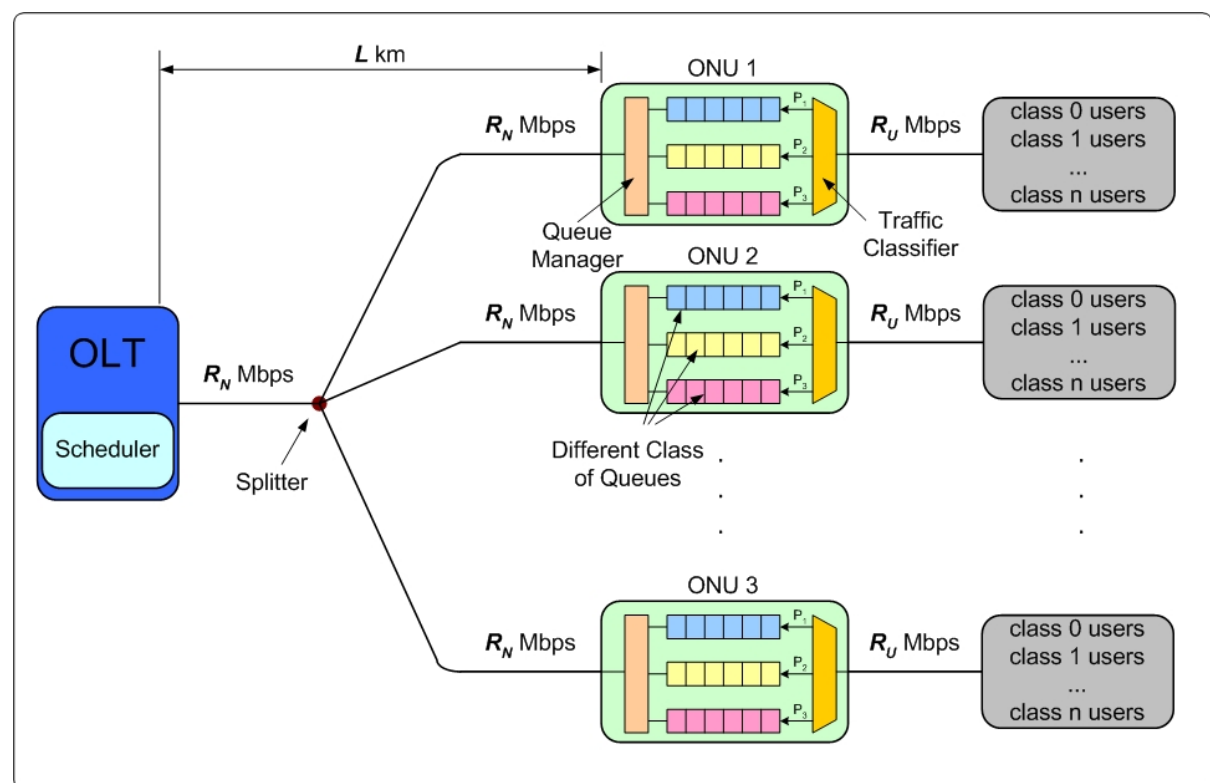


Figure 2.3: The EPON system architecture

In each ONU, three independent priority queues are adopted to support real-time voice, real-time video and non-real-time data service. Each class of service is assigned a specific priority, denoted as P_1, P_2 , and P_3 , respectively, where P_1 is the highest priority and P_3 is the lowest priority. Incoming packets are classified by traffic classifier and stored in the corresponding priority queue. The queue manager is responsible for receiving GATE message, transmitting appropriate

amount of packets from each queue to OLT, and generating REPORT message when the transmission of user information is finished.

The link rate from user equipment to an ONU is R_U Mbit/s. The link rate between an ONU and the OLT in upstream and downstream directions are the same and is assumed to be R_N Mbit/s.

The distance between the OLT and ONUs are different and it results in different round trip time (RTT) for each ONU. We denote L as the maximum distance in kilometer (km) between OLT and an ONU. Even in the same ONU, there still exists some small deviation of RTT compared with previously measured RTT. The RTT updating method can be referred to IEEE 802.3ah standard [16].

The interaction between OLT and ONUs is performed by two major control messages, i.e. GATE message and REPORT message. The GATE message is generated by OLT and is broadcasted to ONUs. It consists of the granted time and data volume. The REPORT message is generated by an ONU and is sent to the OLT. The message consists of the queue size of each service class.

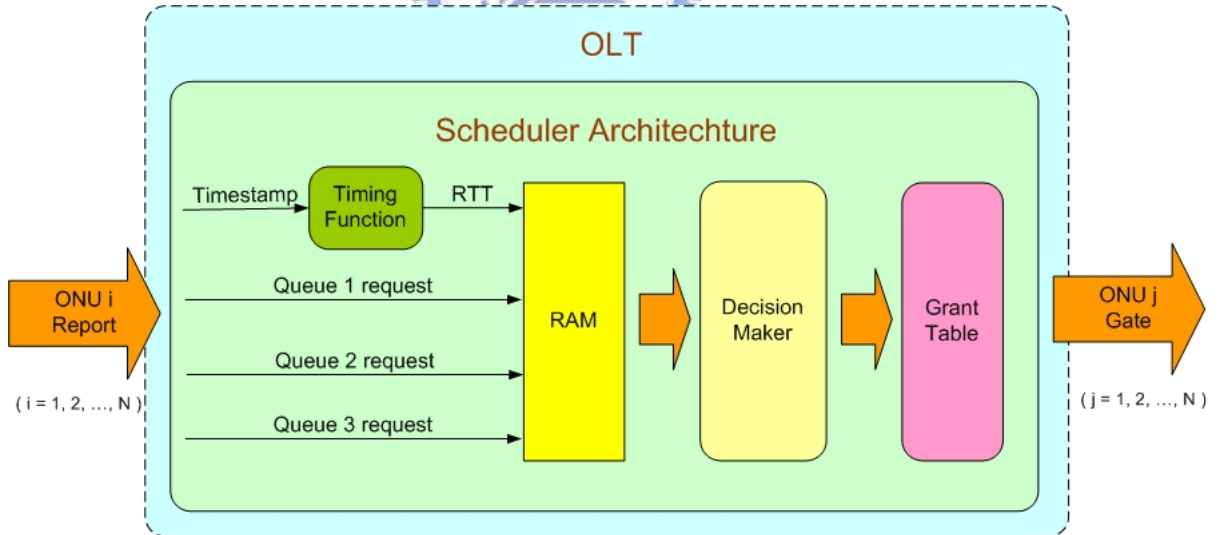


Figure 2.4: Functional modules of a scheduler

Now, we focus on the detailed operations of scheduler in OLT. Figure 2.4 depicts the functional modules of a scheduler. The functional modules include a Timing Function, a RAM, a Decision Maker, and a Grant Table. The Timing Function is used to calculate the round-trip-time of each ONU.

The RAM stores the information of the latest RTT and the amount of requested resource for each ONU. And the Decision Maker is responsible for determining the grant time and data volume of each class of service in every ONU. The proposed scheduling algorithm is implemented in it and we will discuss the detail scheduling operation in later session. And the last functional module in our scheduler is the Grant Table. It is a list that stores the sequence of transmission of all ONUs. It also depicts the granted data volume that each queue in every ONU can transmit during next cycle.

The input of scheduler is the REOPORT message, which is defined by MPCP, originated by each ONU. The REPORT message conveys the timestamp information and request data volume of each class of service of an ONU. When the OLT receives a REPORT message originated from ONU i , the timestamp information will pass to the Timing Function to calculate the latest RTT. Then, the latest RTT and the requested data volume of each class of service obtained from REPORT message will be stored in RAM. During a cycle, the scheduler continues receiving REPORT message from each ONU and storing the information in RAM. Before the end of a cycle, the Decision Maker will start up a scheduling operation. It will take all the information stored in RAM into account. By adopting different scheduling schemes, the amount of grand data volume of each queue in every ONU will be different. Finally, after scheduling, the Grant Table can be generated and OLT can send GATE message to each ONU based on the list of Grant Table.

The interaction between OLT and ONUs is based on IPACT, where the duration of the cycle is dynamically changed based on the requested data volumes reported by ONUs. By adopting IPACT-based polling scheme, there is no need to synchronize the ONUs to a common reference clock. Every ONU executes the same procedure driven by the GATE message received from the OLT. Figure 2.5 shows the IPACT-based polling procedural.

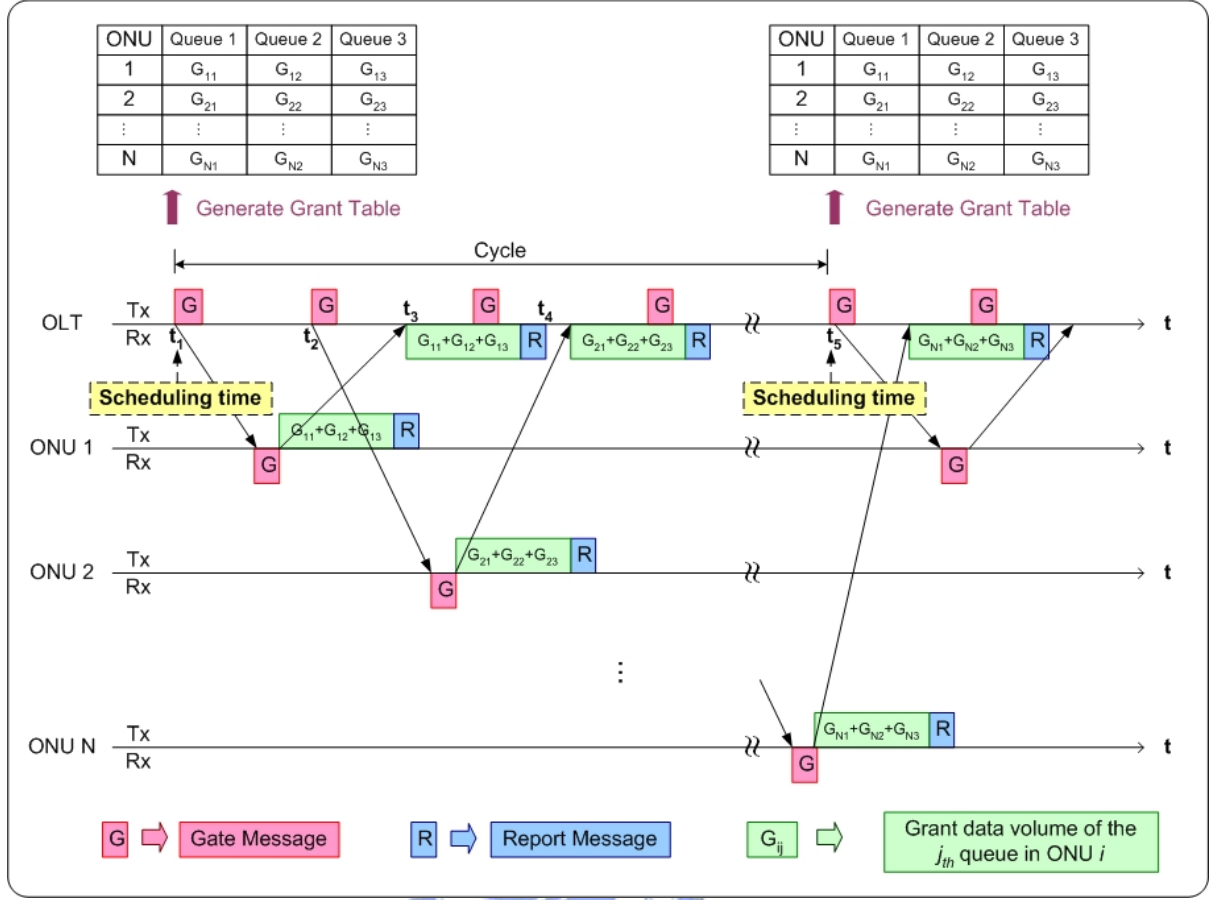


Figure 2.5: IPACT-based polling procedural

In Figure 2.5, The Grant Table is generated at scheduling time t_1 , in which the granted data volume of each queue, denoted as $G_{ij}, i = 1, \dots, N, j = 1, 2, 3,$ is determined based on the proposed scheduling algorithm. At time t_1 , the OLT sends a GATE message to ONU 1, allowing each queue in ONU 1 to send data volume G_{11}, G_{12} , and G_{13} . When ONU 1 finishes the user data transmission, it would generate REPORT message to inform OLT how much data volume remains in each queue. Before receiving the user data and REOPORT message form ONU 1, OLT sends GATE message to ONU 2 at time t_2 , which can be derived by the following equation.

$$\begin{aligned}
 t_2 &= t_4 + g - RTT_2 - G/R_N \\
 &= \left[t_3 + (G_{11} + G_{12} + G_{13} + R)/R_N \right] + g - RTT_2 - G/R_N \\
 &= \left[(t_1 + RTT_1 + G/R_N) + (G_{11} + G_{12} + G_{13} + R)/R_N \right] + g - RTT_2 - G/R_N,
 \end{aligned} \tag{2.1}$$

where g is the guard interval between the transmission of different ONU, RTT_i is the

round trip time of ONU i , G , and R are the GATE message and REPORT message, and R_N is the line rate of EPON. Based on equation (2.1), the times that OLT needs to send GATE message for the other ONUs can be derived. When OLT receives the REPORT message, for example at time t_4 , the requested data volume conveyed by the REPORT message will be stored in RAM until next scheduling time t_5 .

Intuitively, the uplink channel of OLT is almost fully utilized, besides the overhead introduced by REPORT messages and guard times. In other words, the ONUs, which have no data to transmit will not be allocated resource. That leads to a shortened cycle time and results in more frequent polling of active ONUs. Additionally, it is intuitively that the entire scheduling algorithm is located in the OLT. The ONUs do not need to negotiate or acknowledge new parameters, nor do they need to switch to new settings synchronously.

Note that the polling procedural is a little difference compared with traditional one. First, OLT can determine the granted data volume of each class of service in an ONU, and ONU also can request the resource of each service class. Second, the requested data volumes conveyed by REPORT messages would be stored in RAM until the scheduling time. After scheduling, the granted data volume that each queue in every ONU can transmit is determined. Third, the data volume that each queue can transmit in an ONU only depended on the corresponding granted data volume conveyed by GATE message.

2.4 Scheduling Algorithm

The proposed scheduling algorithm is a gated-based scheme, where the scheduler only use the information obtained before the scheduling time. The goal of the scheduling algorithms is to meet the delay bound of voice service, \bar{d} , and to simultaneously maintain the fairness of both packet delay and packet blocking probability for non-real-time data service. The delay bound of voice service and the overall fairness index, F , for non-real-time data service are defined as follows:

$$\bar{d} < d_{\max}, \quad (2.2)$$

$$F = x \cdot I_D + (1-x) \cdot I_B, \quad (2.3)$$

where \bar{d} is the average packet delay of voice service, d_{\max} is the delay criterion of voice service, x is a weighting factor ranged form 0 to 1, and I_D and I_B are the fairness index for average packet delay and average packet blocking probability, respectively. The I_D and I_B are defined below [20]

$$I_D = \frac{\left[\sum_{i=1}^N D_i \right]^2}{N \times \sum_{i=1}^N (D_i)^2}, \quad (2.4)$$

$$I_B = \frac{\left[\sum_{i=1}^N P_{B,i} \right]^2}{N \times \sum_{i=1}^N (P_{B,i})^2}, \quad (2.5)$$

where D_i and $P_{B,i}$ are the average packet delay and the average packet blocking probability of ONU i , respectively. The overall fairness index F described in equation (2.3) must be as close to 1 as possible.

Parameters used in the proposed scheduling algorithm are listed below:

| |
|--|
| <p>Q_{ij}: requested data volume (in bytes) of the j_{th} service class obtained from REPORT message sent by ONU i.</p> <p>G_{ij}: granted data volume (in bytes) of the j_{th} service class of ONU i encapsulated in GATE message sent by OLT.</p> <p>Q_{th}: the queue length threshold for data (class 3) service.</p> <p>L_{max}: the maximum data volume (in bytes) of real-time services each ONU can transmit during one cycle.</p> <p>T_{cycle}: the cycle time (in sec).</p> <p>T_{max}: the maximum cycle time (in sec).</p> <p>b: the guard time (in sec).</p> <p>R: the residual available capacity (in bytes) after allocating resource to real-time services.</p> |
|--|

Table 2.1: Parameter descriptions used in scheduling algorithm

Assume that at scheduling time, the RAM has stored the requested data volume of each service in every ONU, denoted as Q_{ij} , where $i = 1, \dots, N$, and $j = 1, 2, 3$. Then the granted data volume of each class of service, denoted as G_{ij} , where $i = 1, \dots, N$, and $j = 1, 2, 3$, is obtained by adopting proposed scheduling algorithm. In the proposed scheduling algorithm, the real-time services are allocated resource first based on the consideration of delay criterion. After that, the resource is fairly allocated to the non-real-time service in each ONU.

For real-time services, the granted data volumes, i.e. G_{i1} and G_{i2} , are obtained by

$$\begin{cases} G_{i1} = \min(L_{max}, Q_{i1}), \\ G_{i2} = \min(L_{max} - G_{i1}, Q_{i2}), \end{cases} \quad (2.6)$$

where $i = 1, 2, \dots, N$. For voice services, the granted data volumes equal to the requested data volume, but does not exceed L_{max} . Similarly, the granted data volumes for video service equal to the requested data volume, but does not exceed $L_{max} - G_{i1}$. In other words, the granted data volumes of real-time services for each ONU are bounded by L_{max} . And, the voice service has the highest priority to share the resource of L_{max} .

Before describing the scheduling method of data service, we must consider the relationship between the maximum cycle time, denoted as T_{\max} , and the delay bound of voice service, denoted as d_{\max} . The decision of T_{\max} must base on the criterion of packet delay requirement of voice service. Let us consider the special case, that is, the load of voice service is extremely high in every ONU and the uplink channel is filled with voice packets. Because all the packets are belong to voice service, the maximum data volume each ONU can transmit must be the same during a cycle. In other words, the duration that each ONU can transmits packets, denoted as $t_{\max,i}, i=1,2,\dots,N$, is also the same, as shown in Figure 2.6.

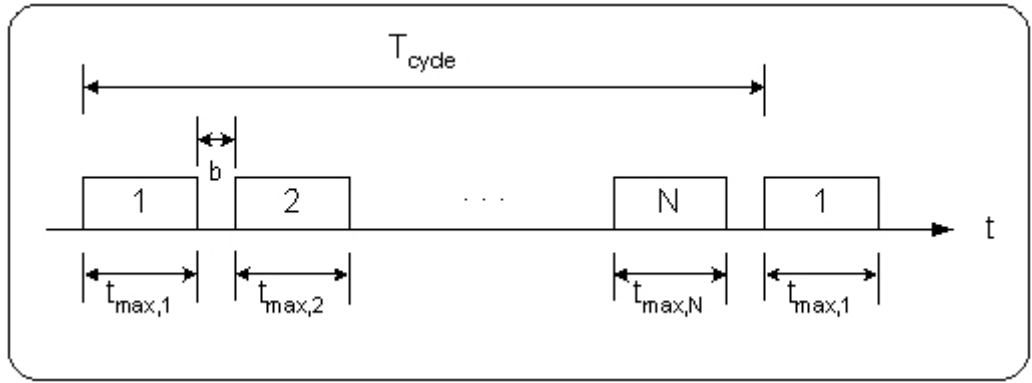


Figure 2.6: The case that channel is fully filled with real-time voice packets

Then, we can write $t_{\max,1} = t_{\max,2} = \dots = t_{\max,N} = \frac{L_{\max}}{R_N}$, where R_N is the uplink transmission rate. In this case, T_{cycle} is equal to T_{\max} , which can be derived by

$$T_{\max} = \sum_{i=1}^N (b + t_{\max,i}) = \sum_{i=1}^N \left(b + \frac{L_{\max}}{R_N} \right) = N \left(b + \frac{L_{\max}}{R_N} \right), \quad (2.7)$$

where b is the guard interval. Additionally, the average packet delay of voice packets can be derived. Figure 2.7 shows the arrival and departure of new-coming packets in ONU i .

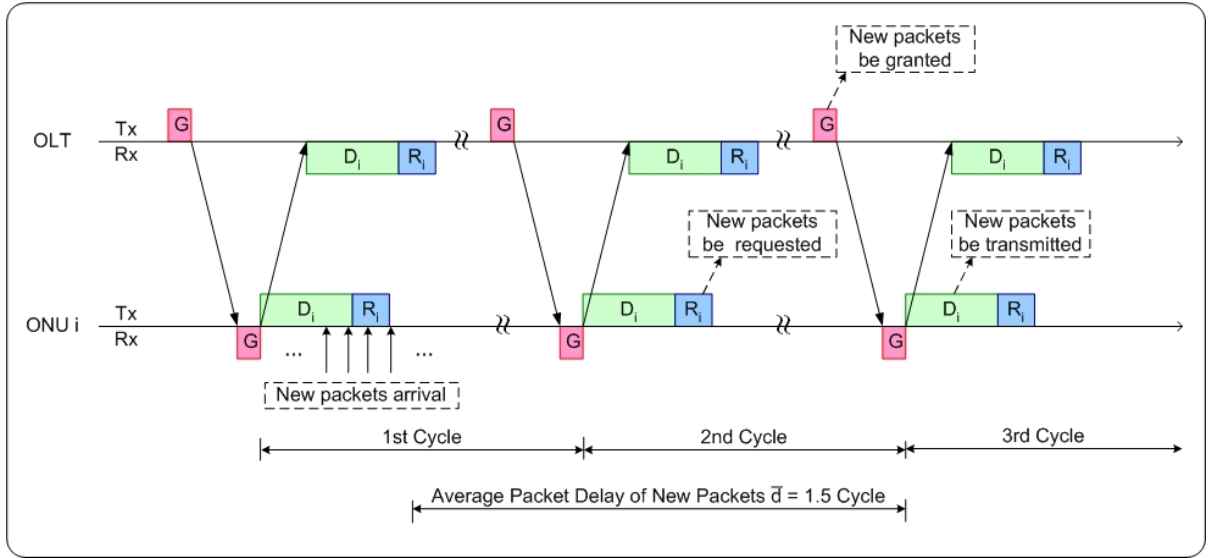


Figure 2.7: The average packet delay of voice service

The ONU i must report the amount of data volume that accumulates during the first cycle to the OLT at the end of transmission in second cycle. Then, the OLT sends a GATE message, which indicates that the requested data volume could be transmitted in third cycle, to the ONU i . Generally, the requested data volume of voice service reported in second cycle can be completely allowed to transmit in third cycle. Thus, we can derive that the average packet delay of voice service by

$$\begin{aligned} \bar{d} &= E[\text{packet delay in 1st Cycle}] + E[\text{packet delay in 2nd Cycle}] \\ &= 0.5 \text{ cycle time} + 1 \text{ cycle time} = 1.5 \text{ cycle time}, \end{aligned} \quad (2.8)$$

where the voice packets arrive at ONU i in first cycle follows uniform distribution. Thus, the packet delay would be 0.5 cycle in first cycle. Moreover, the maximum average packet delay of voice service should be bounded in a delay criterion. Thus, the equation (2.2) can be rewritten by

$$\bar{d} \leq 1.5T_{\max} \leq d_{\max}. \quad (2.9)$$

When the granted data volumes of real-time services are determined, the residual available capacity, R , for non-real-time data service can be derived by

$$R = (T_{\max} - N \cdot b) \times R_N - \sum_{i=1}^N (G_{i1} + G_{i2}). \quad (2.10)$$

If the summation of requested data volume of data service is smaller than the residual available

capacity R , all the requested packets are allowed to transmit in the next cycle, because the resource is adequate to assign to each data service. This situation would happen frequently under light-load condition. But under high-load condition, the resource is usually insufficient. Thus, we need an appropriate scheduling scheme to serve packets well.

As mentioned before, for non-real-time data service, the fairness of average packet blocking probability and the fairness of average packet delay among all ONUs are considered. We proposed a Hybrid EQL-QLP scheme and a Hybrid LQF-QLP scheme to achieve this goal. In Session 2.4.1, we introduce the Hybrid LQF-QLP scheme; and in Session 2.4.2, we introduce the Hybrid EQL-QLP scheme.

2.4.1 Hybrid EQL-QLP scheme

The first proposed scheduling scheme for non-real-time data service is Hybrid Equal Queue Length - Queue Length Proportional (Hybrid EQL-QLP) scheme. The inputs are the requested data volumes of data service, Q_{i3} , where $i=1,\dots,N$, and the residual available capacity R . The outputs are the granted data volumes G_{i3} , where $i=1,\dots,N$. The goal of Hybrid EQL-QLP scheme is to let the fairness index F , defined in equation (2.3), as close to 1 as possible. The granted data volumes of data service are determined by

$$G_{i3} = \begin{cases} Q_{i3}, & \sum_{j=1}^N Q_{j3} \leq R, \\ R \times \frac{Q_{i3}}{\sum_{j=1}^N Q_{j3}}, & \sum_{j=1}^N Q_{j3} > R, \max\{Q_{13}, Q_{23}, \dots, Q_{N3}\} \leq Q_{th}, \\ f_a(R, Q_{13}, Q_{23}, \dots, Q_{N3}), & \sum_{j=1}^N Q_{j3} > R, \max\{Q_{13}, Q_{23}, \dots, Q_{N3}\} > Q_{th}. \end{cases} \quad (2.11)$$

where $i=1,2,\dots,N$. If the summation of the requested data volumes of data service is smaller than the residual available capacity R , all the requested packets are allowed to transmit in the next cycle,

because the resource is adequate to assign to each data service. If the summation of the requested data volume of data service is larger than R , but all of them are smaller than Q_{th} , then the resource assigned to each data service queue is proportional to its requested data volume. However, if the summation of the requested data volumes of data service is larger than R , and one or more than one queue lengths are larger than Q_{th} , then the EQL method is adopted. The method is described as follows.

Define an index set $K = \{k_1, k_2, \dots, k_n\}$, where $K \in \{1, 2, \dots, N\}$ and index k_j satisfies $Q_{k_j,3} > avg$, $j = 1, 2, \dots, n$. Here, avg meets the equation

$$\sum_{j=1}^n (Q_{k_j,3} - avg) = R. \quad (2.12)$$

The avg can be rewritten as follows

$$avg = \frac{\left[\sum_{j=1}^n Q_{k_j,3} \right] - R}{n}. \quad (2.13)$$

Then the resource assign to each data service queue is

$$G_{i3} = f_a(R, Q_{13}, Q_{23}, \dots, Q_{N3}) = \begin{cases} Q_{i3} - avg, & i \in K, \\ 0, & i \notin K. \end{cases} \quad (2.14)$$

The resource allocated to each queue, which queue length is larger than avg , is the difference between its queue length and avg . However, there is no resource assigned to the queues, which queue lengths are smaller than avg . The examples of EQL scheme are given in Appendix C.

2.4.2 Hybrid LQF-QLP scheme

The EQL scheme tries to balance the queue sizes of non-real-time data service. It results in a situation that resource has more opportunity to be shared to every queue. If the queue sizes of some data service queue are very large, the LQF scheme is better to alleviate the packet blocking. Thus, we proposed the second scenario for non-real-time data service, named Hybrid Longest Queue First -

Queue Length Proportional (Hybrid LQF-QLP) scheme. The idea of this scheme is similar to previous one. Also, the inputs are the requested data volume of data service $Q_{i3}, i=1, \dots, N$, and the residual available capacity R , and the outputs are the granted data volume $G_{i3}, i=1, \dots, N$. The resource assignment is written as

$$G_{i3} = \begin{cases} Q_{i3}, & \sum_{j=1}^N Q_{j3} \leq R, \\ R \times \frac{Q_{i3}}{\sum_{j=1}^N Q_{j3}}, & \sum_{j=1}^N Q_{j3} > R, \max\{Q_{13}, Q_{23}, \dots, Q_{N3}\} \leq Q_{th}, \\ f_b(R, Q_{13}, Q_{23}, \dots, Q_{N3}), & \sum_{j=1}^N Q_{j3} > R, \max\{Q_{13}, Q_{23}, \dots, Q_{N3}\} > Q_{th}. \end{cases} \quad (2.15)$$

In equation (2.15), if the summation of queue lengths of data service is smaller than the residual available capacity R , all the requested packets are allowed to transmit in the next cycle, because the resource is adequate to assign to each data service. If the summation of queue lengths of data service is larger than R , but all of them are smaller than Q_{th} , then the resource assign to each data service queue is proportional to its requested data volume. However, if the summation of queue lengths of data service is larger than R , and one or more queue lengths are larger than Q_{th} , then the resource assigned to each data service is described as follows.

Define the permutation function $\pi: [1, N] \rightarrow [1, N]$, so that $Q_{\pi(1),3} \geq Q_{\pi(2),3} \geq \dots \geq Q_{\pi(N),3}$.

Note that the permutation function π is the index set with descent order of queue size mapping from the original index.

$$G_{\pi(i),3} = f_b(R, Q_{13}, \dots, Q_{N3}) = \begin{cases} \min\left[\max(Q_{\pi(i),3} - Q_{th}, 0), R\right] + R' \cdot P_{\pi(i)}, & i = 1, \\ \min\left[\max(Q_{\pi(i),3} - Q_{th}, 0), R - \sum_{j=1}^{i-1} G_{\pi(j),3}\right] + R' \cdot P_{\pi(i)}, & i = 2, \dots, N, \end{cases} \quad (2.16)$$

$$R' = \max\left(R - \sum_{j=1}^N \max(Q_{j3} - Q_{th}, 0), 0\right), \quad (2.17)$$

$$P_{\pi(i)} = \frac{Q_{\pi(i),3} - \max(Q_{\pi(i),3} - Q_{th}, 0)}{\sum_{j=1}^N [Q_{\pi(j),3} - \max(Q_{\pi(j),3} - Q_{th}, 0)]}. \quad (2.18)$$

In equation (2.16), the resource allocated to each data service queue consists of two parts of resource assignment. In the first part, the scheduler allocates the emergent queues, which queue occupancies are larger than queue length threshold, an amount of the difference between the queue occupancy and the queue length threshold. In addition, the longest queue has the highest priority to share the resource.

In second part, the scheduler allocates the remaining resource, denoted as R' , which is the available resource after allocation in the first part, as derived by equation (2.17), to each ONU again. However, different from the first part, the allocation is based on the proportion of remaining queue occupancies.

The goal of first part is to avoid packets blocking, because the queue with larger queue occupancy has higher priority to be served. Thus, the first part is benefit to the fairness of packet blocking probability. In addition, the goal of second part is to avoid that all the resource is allocated to the greedy queues. So, the second part is benefit to the fairness of packet delay. As a result, by adopting the Hybrid LQF-QLP scheme, the fairness of packet delay and packet blocking probability can be considered simultaneously. It fits the original goal of scheduling the non-real-time service.

2.5 Simulation Result

2.5.1 Traffic Source Models

We consider three kinds of services, i.e. real-time voice, real-time video and best-effort data. The priorities of these services are specified by P_1, P_2 , and P_3 , where class P_1 service has the highest priority and class P_3 service has the lowest priority.

Class P_1 service is used to emulate a T1 connection. The packet generation rate of P_1 service is assumed to be the constant bit rate (CBR). The T1 data arriving from the user is packetized in the ONU by placing 24 bytes of data in a packet. By adding the overhead such as Ethernet, UDP (User Data Protocol) and IP (Internet Protocol) headers in a packet, the packet results in a 70-byte frame and would be generated every 125μ s. Hence the data rate of class P_1 service is 4.48 Mbps.

Class P_2 service is used to emulate VBR video streams that exhibit properties of self-similarity and long-range-dependence (LRD). The packet size of this class of service is uniformly distribution and ranges from 64 to 1518 bytes. Class P_3 service has the lowest priority. It is used for non-real-time data transfer. The network does not guarantee the delivery or the delay of packets for this service. This class of service is also self-similar and LRD traffic with uniformly-distributed packet size ranged from 64 to 1518 bytes.

There is an extensive study, such as [17],[18],[19], and etc, showing that most network traffic flows can be characterized by self-similarity and long-range dependence (LRD). The characteristics of self-similar and LRD are described in Appendix A.

To generate self-similar traffic, we used the method described in [18], where the resulting traffic is an aggregation of multiple streams. The structure of the synthetic self-similar traffic generator is shown in Figure 2.8. Each source is performed by ON/OFF Parato-distributed model. The design of the number of sources (K) in a traffic generator is based on the experiment result discussed in [5]. It shows that the burstiness of the traffic (Hurst parameter) does not change with K if the total load is fixed.

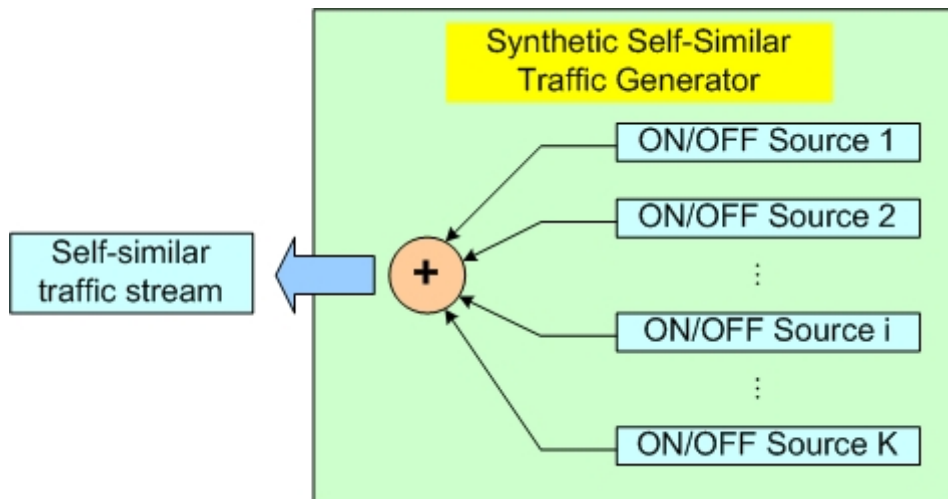


Figure 2.8: Synthetic self-similar traffic generator

The traditional ON-OFF source models assume finite variance distributions for the sojourn time in ON and OFF periods. As a result, the aggregation of large number of such sources will not have significant correlation, except possibly in the short range. An extension to such traditional ON-OFF models allows the ON and OFF periods to have infinite variance (high variability or Noah Effect). The superposition of many such sources produces aggregate traffic that exhibits long-range dependence (also called the Joseph Effect) [18].

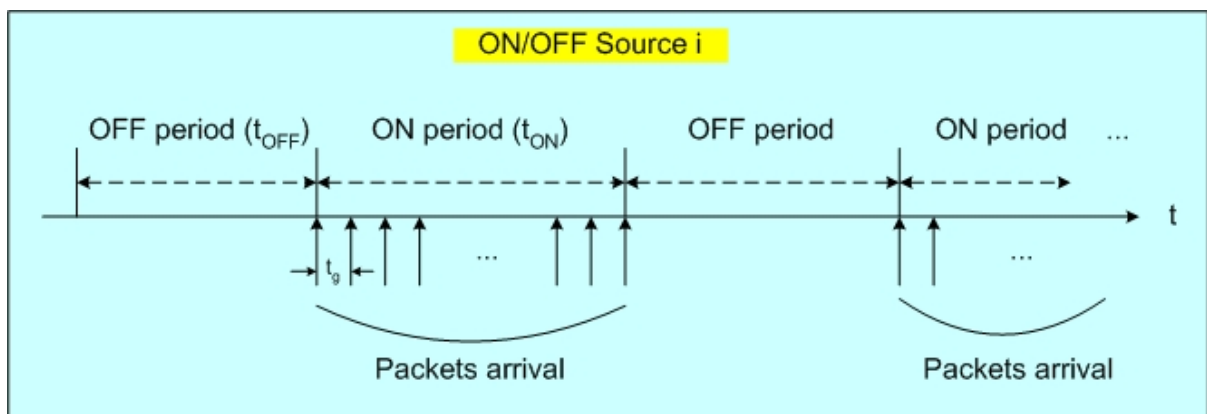


Figure 2.9: ON/OFF Parato-distributed source i

Now we discuss the detail of each source. Figure 2.9 shows the model of source i . The parameters of this model are described as follows:

The number of packets generated by source i , denoted as Np_i , during ON period follows Pareto distribution with a minimum of 1 and maximum of $2^{16}-1$. Pareto distribution can be defined as follows:

$$\begin{cases} f_X(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}, x \geq k, \\ \mu = \frac{k\alpha}{\alpha-1}, \alpha > 1, \\ \sigma^2 = \frac{k^2\alpha}{(\alpha-2) \cdot (\alpha-1)^2}, \alpha > 2, \end{cases} \quad (2.19)$$

where α is a shape parameter, and k is a location parameter. We set the shape parameter $\alpha = 1.4$. The choice of α was prompted by measurements on actual Ethernet traffic [19]. They reported the measured Hurst parameter of 0.8 for moderate network load. The relationship between the Hurst parameter and the shape parameter α is $H = (3-\alpha)/2$ [18]. Thus, $\alpha = 1.4$ should result in $H = 0.8$.

During ON period, the packet assumed to immediately follow the previous packet with minimum inter-packet gap t_g . We choose t_g equals to the standard preamble (8 bytes) of Ethernet packet.

Every source has a constant packet size from uniform distribution between 64 and 1518 (in bytes). We denote the packet size generated by source i is Ps_i . Then, the duration of ON period (t_{ON}) of source i can be described as

$$t_{ON} = (Ps_i + t_g) \times Np_i \times \text{bytetime}, \quad (2.20)$$

where bytetime depends on the line rate of EPON, i.e., $\text{bytetime} = 8 / \text{line rate}$.

OFF periods (intervals between the packet trains) also follow the Parato distribution with the shape parameter $\alpha = 1.2$. We used heavier tail for the distribution of the OFF periods represent a stable state in a network, i.e., a network can be in OFF state (no packet transmission) for an unlimitedly long time, while the durations of the ON Periods are ultimately limited by network resources and (necessarily finite) file sizes. To find the minimum value of OFF period (MIN_OFF),

we must consider the average load we want to generate for this self-similar traffic generator. Suppose every source in traffic generator has the same parameters, then the average load, denoted as $LOAD$, can be represent as

$$LOAD = K \times \frac{E[ON]}{E[ON] + E[OFF]}. \quad (2.21)$$

According to equation (2.19), we know that

$$E[ON] = MIN_ON \times \frac{\alpha}{\alpha - 1} = MIN_ON \times on_coef, \quad (2.22)$$

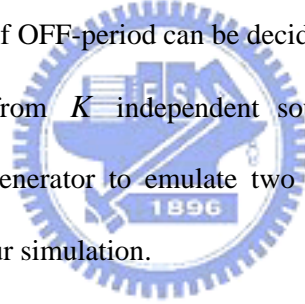
$$E[OFF] = MIN_OFF \times \frac{\alpha'}{\alpha' - 1} = MIN_OFF \times off_coef. \quad (2.23)$$

By adopting equation (2.21), (2.22) and (2.23), we can derive MIN_OFF as:

$$MIN_OFF = \frac{on_coef}{off_coef} \times MIN_ON \times \left(\frac{K}{LOAD} - 1 \right). \quad (2.24)$$

Then, the minimum value of OFF-period can be decided.

By aggregating streams from K independent sources, the realistic self-similar traffic is generated. We use the traffic generator to emulate two kinds of traffic sources, i.e., VBR video streams and best-effort data in our simulation.



2.5.2 Simulation Environment

The system parameters are described as follows:

| Parameter | Description | Value |
|-----------|-------------------------------|-----------|
| N | Number of ONUs | 16 |
| n | Number of queues in each ONU | 3 |
| R_u | Line rate of user-to-ONU link | 100 Mbps |
| R_N | Line rate between OLT and ONU | 1000 Mbps |
| Q | The size of each queue | 1 Mbytes |

| | | |
|-----------|---|------------|
| L_{max} | Maximum data volume of real-time service an ONU can transmit during one cycle | 5000 bytes |
| t_{max} | Maximum slot size of real time traffic | 0.04 ms |
| b | Guard time between adjacent slots | 5 μ s |
| T_{max} | Maximum cycle time | 0.72 ms |
| d_{max} | The delay bound of voice service | 1.5 ms |
| x | Waiting factor of overall fairness index | 0.5 |

Table 2.2: System parameters used in the simulation environment

The choice of d_{max} is based on the specification. International Telecommunication Standardization Sector (ITU-T) Recommendation G.114 “One-way transmission time” specifies 1.5ms one way propagation delay in access network (digital local exchange). The design of T_{max} also based on the delay bound of voice service. To keep the average delay within this bound, we set the parameter $t_{max} = 0.04$ ms, results in the maximum cycle time $T_{max} = 0.72$ ms.

And now the question is what is the appropriate setting of loads for all ONUs? In general, the loads supported to ONUs are all different. To simplify the simulation, we set the load of a group of ONUs fixed. And the loads of the others are adjusted between different experiments, but remain the same during one experiment. The loads are set as follows:

| |
|--|
| P_0 service: 4.48 Mbps \times 16(iid) |
| P_1 service: 15 Mbps \times 16(iid) |
| P_2 service: 15 Mbps \times 10(iid) + M \times 6(iid), where M = 15 Mbps ~ 80 Mbps |
| Average System Load \approx 600 Mbps ~ 1000 Mbps |

The average loads of all P_1 services, all P_2 services, and ten of P_3 services remain a fixed value in every experiment. But the average loads of six of P_3 services will change from 15 Mbps to 80 Mbps in different experiment. As a result, the average system load will change approximately from 600 Mbps to 1000 Mbps.

2.5.3 Simulation Result and Discussion

In this session, we show the performance of proposed schemes, i.e. Hybrid EQL-QLP scheme and Hybrid LQF-QLP scheme. Both schemes have queue length thresholds, denoted as Q_{th} , which are the conditions that whether change the mechanism to the other one or not. We define the Q_{th} by

$$Q_{th} = \alpha \times Q, \quad (2.25)$$

where α is ranged from 0 to 1 and the queue size Q is set to 1 Mbytes. If α equals to 0, then only EQL mechanism is used in Hybrid EQL-QLP scheme; and only LQF mechanism is used in Hybrid LQF-QLP scheme. On the contrary, if α equals to 1, then both schemes only use QLP mechanism.

In the following, we first introduce the simulation results of Hybrid LQF-QLP scheme. And then, the performance of Hybrid EQL-QLP scheme is observed. In both schemes, we will check whether the average packet delay of voice service is bounded in delay criterion or not. And we will see the performance of data service. We divide all ONUs into high-loading ONUs and low-loading ONUs. The mean rate of each class of service in low-loading ONUs remains fixed. In high-loading ONUs, the mean rate of real-time services also remains fixed, but the mean rate of best-effort service will change from low to high. We will see the performance of high-loading ONUs and low-loading ONUs individually. After considering the difference between high-loading ONUs and low-loading ONUs, the fairness of best-effort data among all ONUs can be obtained.

In Figure 2.10, we show the average packet delay of voice service by adopting Hybrid LQF-QLP scheme. We can see that all the cases, with different value of alpha, have average packet delays of voice service bounded in a value, which is lower than the specified delay criterion (1.5ms). The figure can verify that our system satisfies the basic delay requirement.

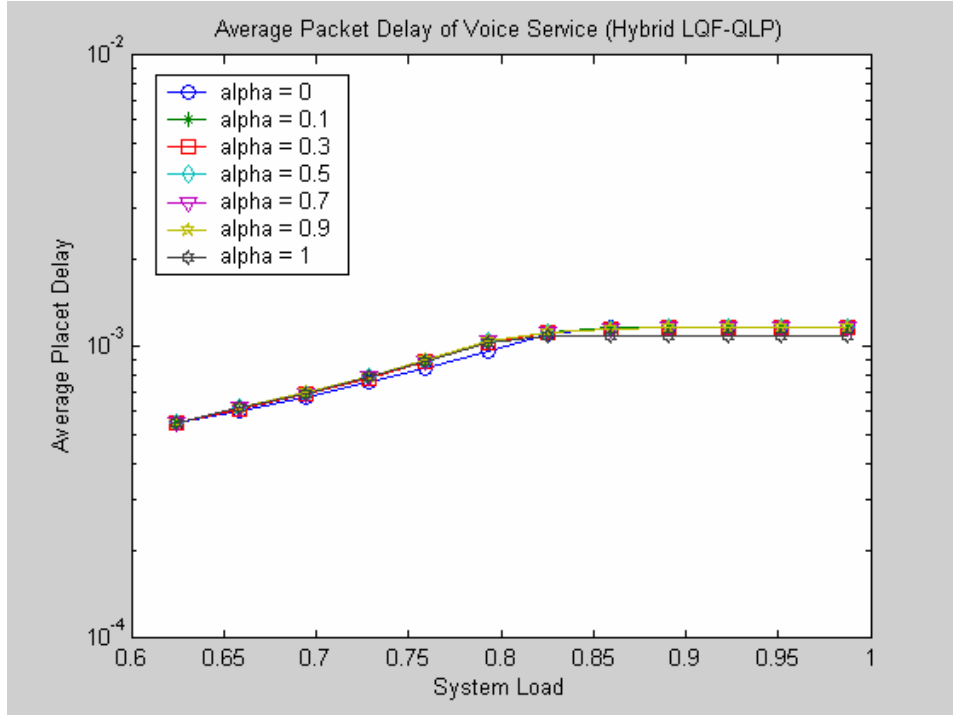


Figure 2.10: Average Packet Delay of Voice Service by adopting Hybrid LQF-QLP

Figure 2.11 and Figure 2.12 show the effect of Q_{th} in packet blocking probability by adopting Hybrid LQF-QLP scheme. In Figure 2.11, we can see that LQF mechanism has better performance in average packet blocking probability than QLP mechanism. The reason is that if LQF mechanism is adopted, high-loading ONUs will be assigned the higher priority than low-loading ONUs. So, most of the resource will be assigned to high-loading ONUs, and the packet blocking probability would be lower in high-loading ONUs. Intuitively, the packet blocking probability of low-loading ONUs would be higher by adopting LQF scheme than by adopting QLP scheme. Consequently, the difference of packet blocking probability between high-loading ONUs and low-loading ONUs would be smaller by adopting LQF scheme than by QLP scheme. Finally, we can see the fairness index of packet blocking probability is better in LQF scheme, as shown in Figure 2.12.

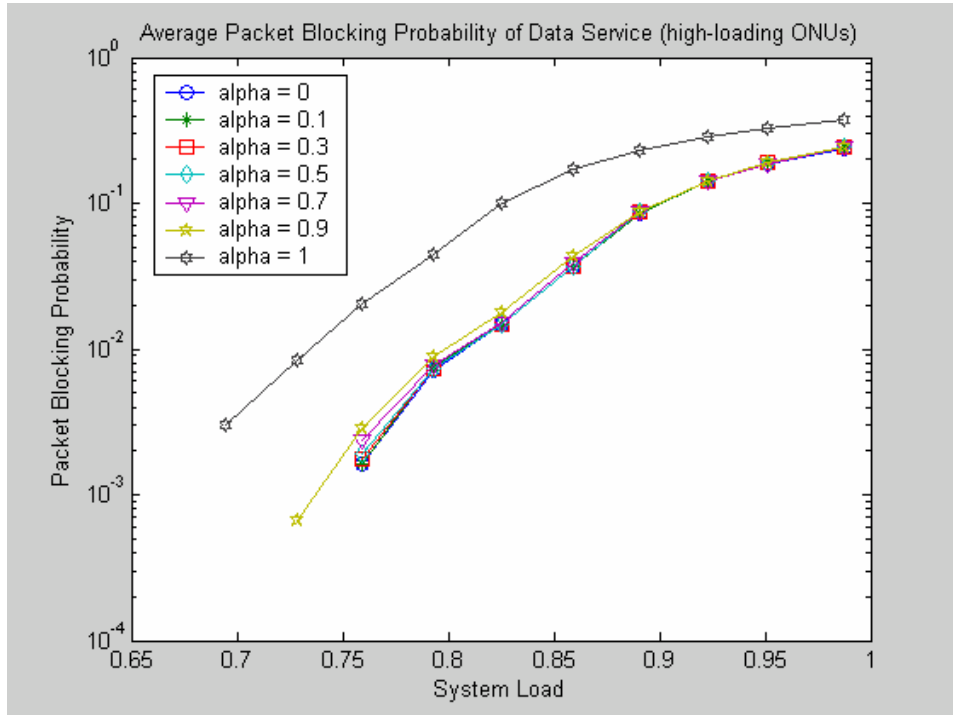


Figure 2.11: Average packet blocking probability (high-loading ONUs) of data service by adopting Hybrid LQF-QLP

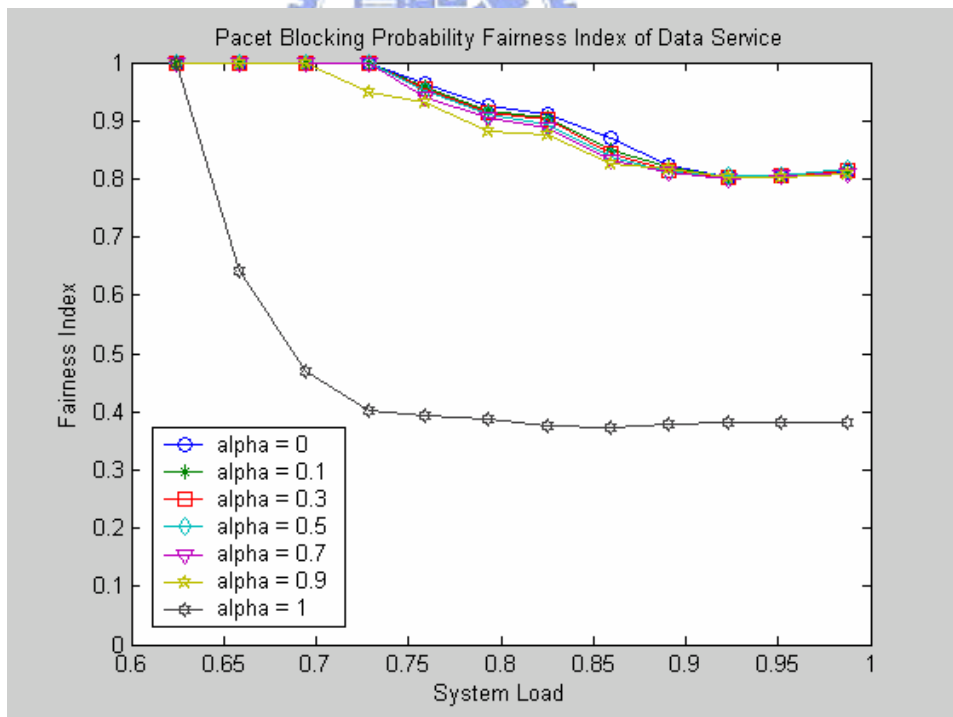


Figure 2.12: Packet blocking probability fairness index of data service by adopting Hybrid LQF-QLP

When we adjust Q_{th} from 0 to 1, the packet blocking probability would be higher in the

beginning of the curves, because the characteristic of Hybrid LQF-QLP scheme will approach to the characteristic of QLP mechanism as Q_{th} is close to 1. But if the system load approaches to saturate, the LQF scheme will dominate the performance of packet blocking probability. Because under saturation-load condition, the queue occupancy would increase greatly and almost exceed the condition of Q_{th} more. Thus, there is usually no resource after the first assignment of Hybrid LQF-QLP scheme. As a result, the performance of all cases will approaches to LQF mechanism.

And then, in Figure 2.13 and Figure 2.14, we show the effect of Q_{th} in average packet delay and it's fairness index. By adopting QLP, the resource assign to each ONU is proportional to the requested data volume. So, the variance of average packet delay is lower and the fairness of average packet delay would be higher. If we adopt LQF scheme, the high-loading ONUs get most of the resource. The average packet delay of high-loading ONUs would be decreased. But the average packet delay of low-loading ONUs will increase dramatically, even more than the average packet delay of high-loading ONUs. So the packet delay fairness will be lower than the others.

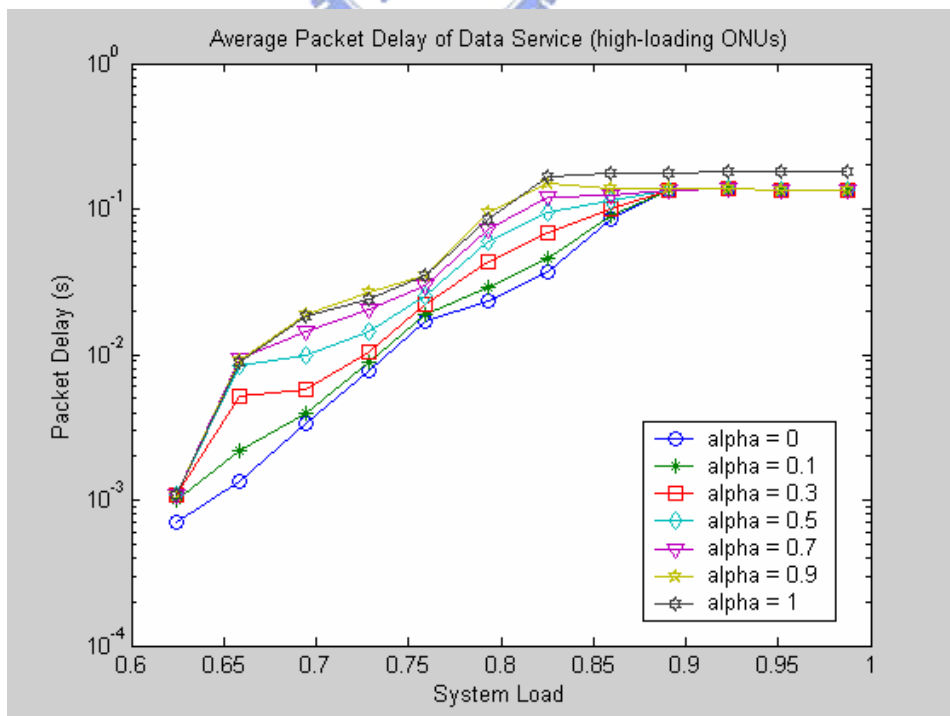


Figure 2.13: Average Packet Delay (high-loading ONUs) of Data Service by adopting Hybrid LQF-QLP

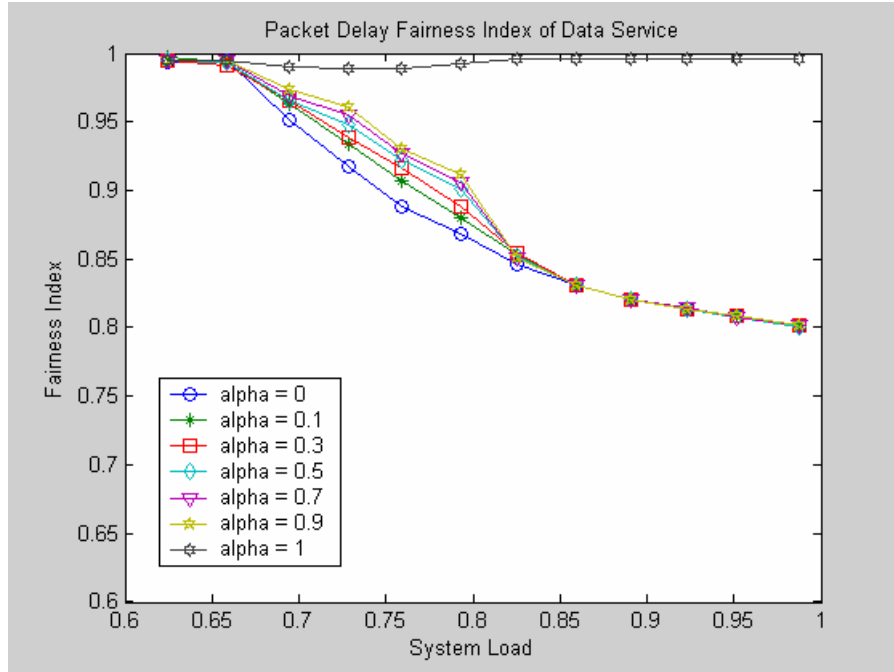


Figure 2.14: Packet delay fairness index of data service by adopting Hybrid LQF-QLP

The result of the overall fairness index F , defined in equation (2.3), is shown in Figure 2.15. We can see that if we consider fairness of packet delay and fairness of packet blocking probability together, the QLP scheme ($Q_{th} = 1$) and LQF scheme ($Q_{th} = 0$) would not be the best choices. In this figure, the hybrid scheme with $Q_{th} = 0.7$ exhibits the best performance.

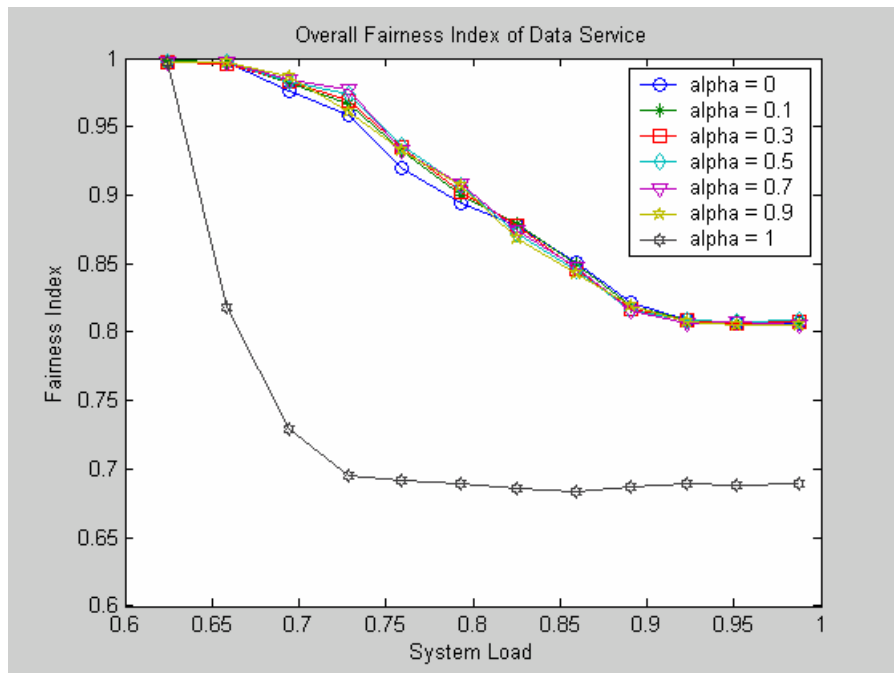


Figure 2.15: Overall fairness index of data service by adopting Hybrid LQF-QLP

And then, we show the effect of Q_{th} by adopting Hybrid EQL-QLP scheme. Similarly, we must guarantee the average packet delay of voice service. The result of average packet delay for voice service is shown in Figure 2.16. We can see that the average packet delay of voice service still bounds in 1.5ms.

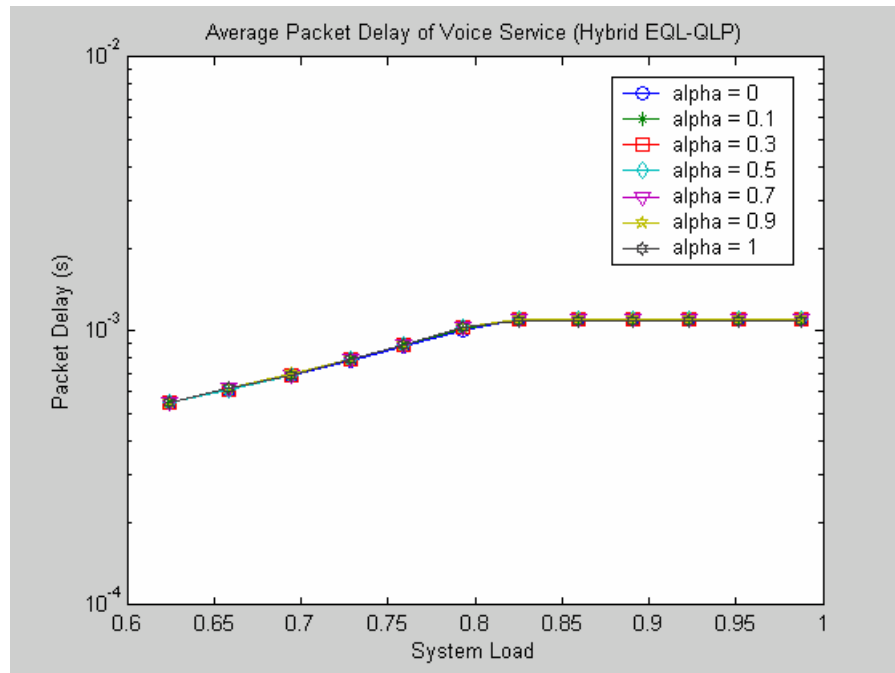


Figure 2.16: Average packet delay of voice service by adopting Hybrid EQL-QLP

In Figure 2.17 and Figure 2.18, we show the effect of Q_{th} in average packet blocking probability and fairness of packet blocking probability. The packet blocking probability of high-loading ONUs is higher in QLP than in EQL. It is also because that the most of the resource is assign to high-loading ONUs by adopting EQL scheme. The effect of Q_{th} in packet blocking probability is not obvious when the system load is increasing. Because when a system begins to block packets, the queue occupancies increase dramatically with system load. Under this condition, the EQL scheme will always be adopted. In addition, by adopting Hybrid EQL-QLP scheme, the assignment is independent of Q_{th} when any requested data volume exceeds Q_{th} . Thus, there is no obvious effect in packet blocking probability and fairness of packet blocking probability if we adjust the Q_{th} .

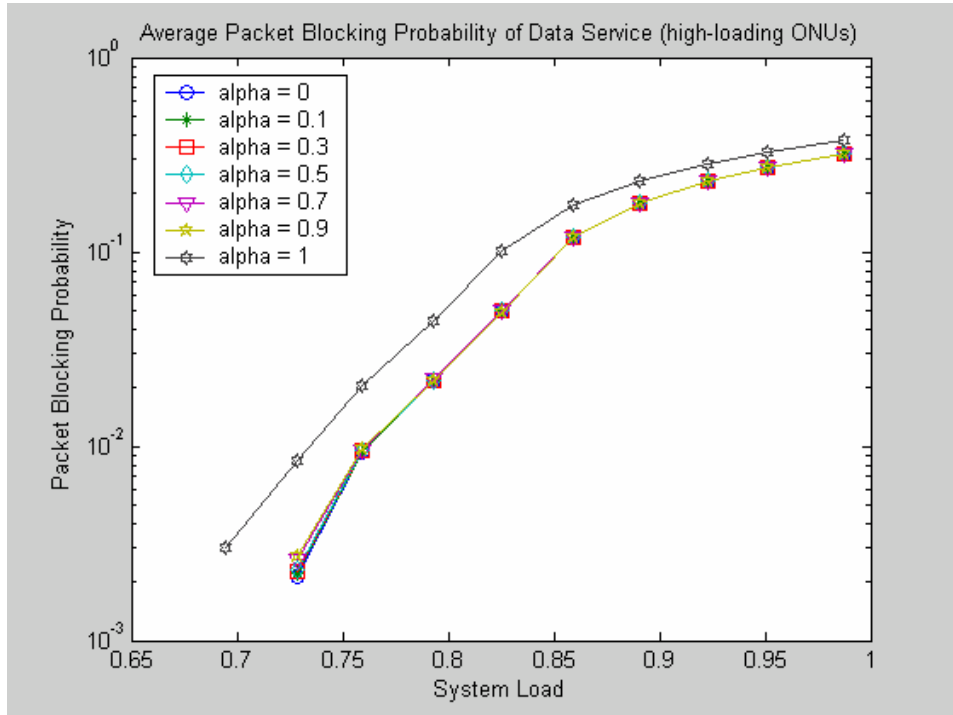


Figure 2.17: Average packet blocking probability (high-loading ONUs) of data service by adopting Hybrid EQL-QLP

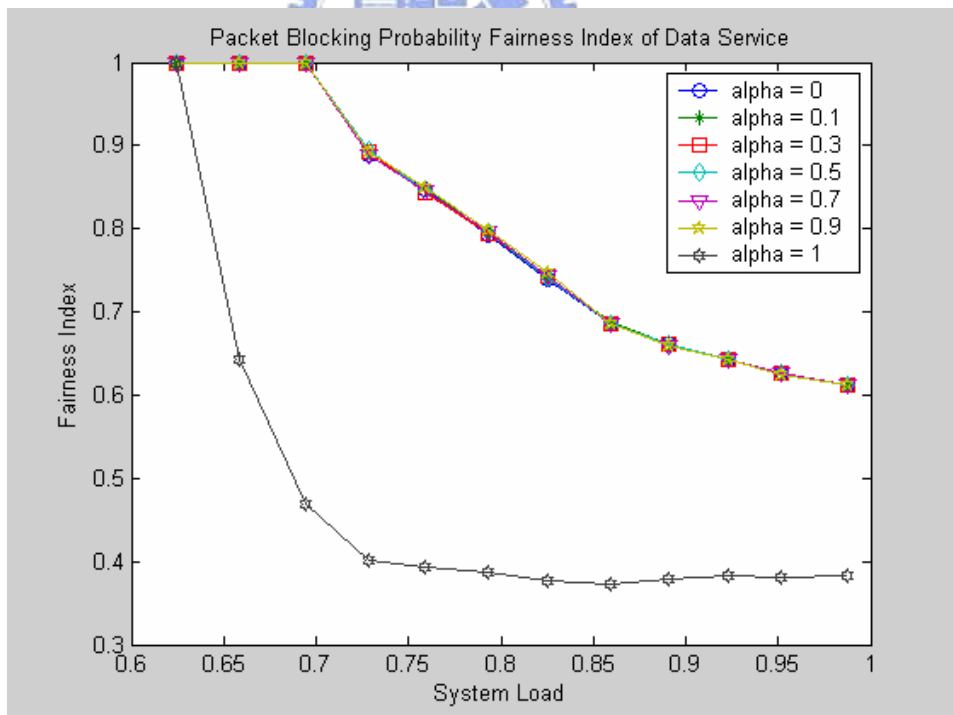


Figure 2.18: Packet blocking probability fairness index of data service by adopting Hybrid EQL-QLP

We show the result of average packet delay of high-loading ONUs and fairness of average packet delay in Figure 2.19 and Figure 2.20, respectively. Similarly, QLP has better performance in packet delay fairness, because the resource assigned to each user is proportional to the requested data volume. By adopting EQL scheme, the average packet delay of high-loading ONUs is lower, due to fact that most of the resource are assigned to them. But the average packet delay of low-loading ONUs would be increased very much, even exceeds the average packet delay of high-loading ONUs. So the fairness of packet delay would be worse than QLP. Then we focus the effect of Q_{th} in these two figures. For Hybrid EQL-QLP scheme, the design of Q_{th} decides when the algorithm must switch to EQL scheme. If the value of Q_{th} be set lower, then the probability that the algorithm switches from QLP to EQL is higher. It means that we have more opportunity to use EQL schemes. Then the performance will achieve the result of EQL scheme.

Finally, we can see the overall fairness index shown in Figure 2.21. The Hybrid EQL-QLP scheme has the best performance when $Q_{th} = 0.9$.

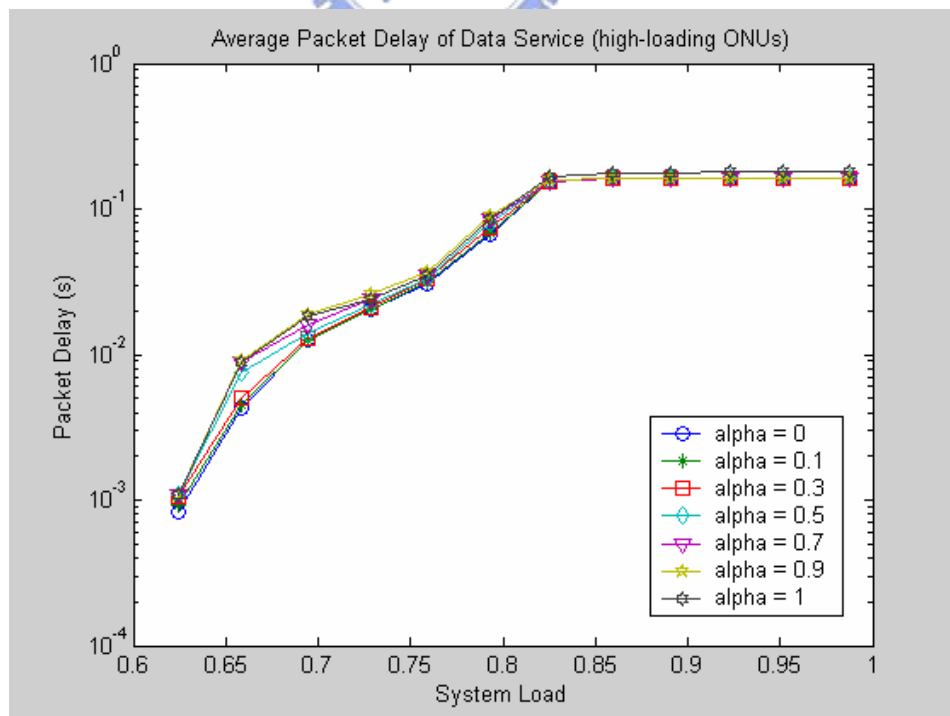


Figure 2.19: Average Packet Delay (high-loading ONUs) of Data Service by adopting Hybrid EQL-QLP

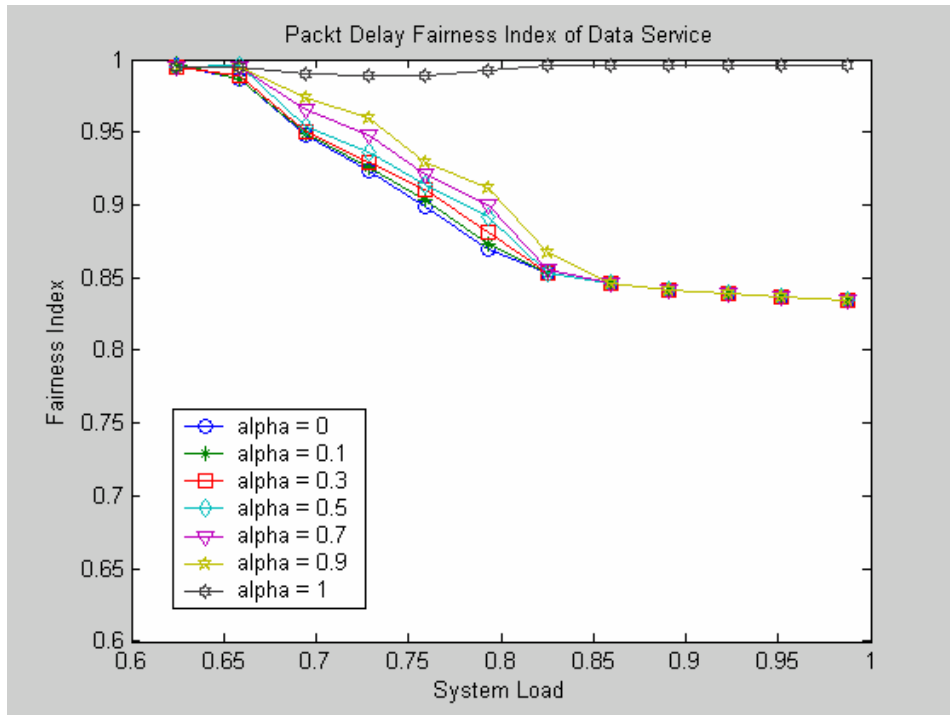


Figure 2.20: Packet delay fairness index of data service by adopting Hybrid EQL-QLP

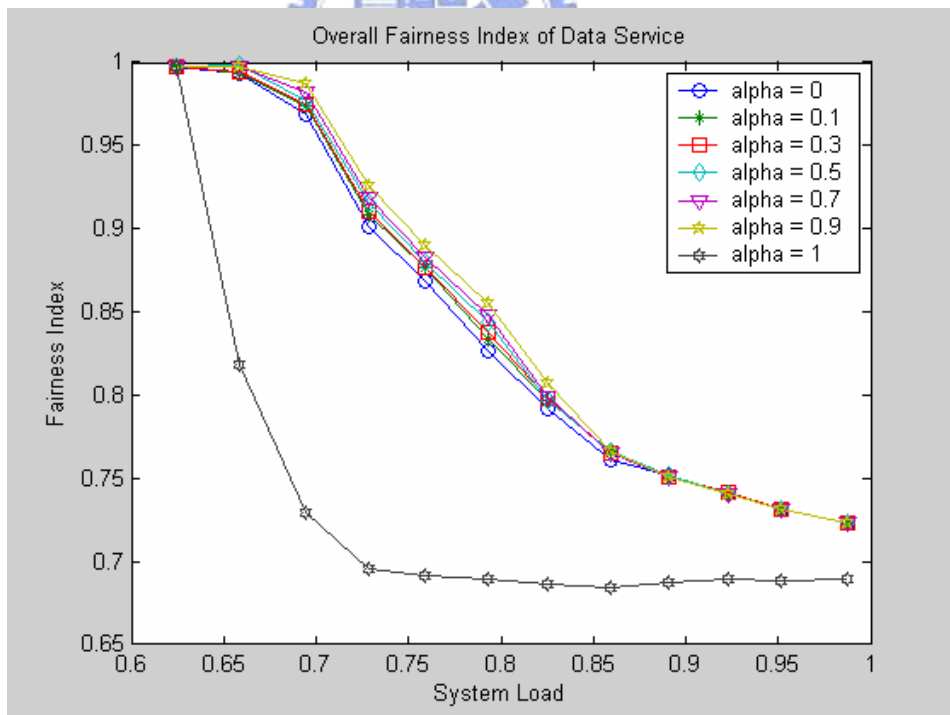


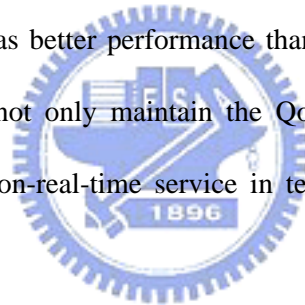
Figure 2.21: Overall fairness index of data service by adopting Hybrid EQL-QLP

2.6 Concluding Remarks

In this chapter, we proposed two scheduling scheme to obtain the overall fairness which combines the fairness of packet delay and fairness of packet blocking probability.

The proposed schemes are Hybrid LQF-QLP and Hybrid EQL-QLP. Each scheme is combined by two basic sub-schemes. We use a queue length threshold to be an adjusting parameter. In the simulation, we define three class of service, i.e. voice, video and data service. The basic requirement of the scheduling algorithm is to meet the delay bound of voice service. Under this condition, we try to improve the fairness of packet delay and fairness packet blocking probability, simultaneously.

Simulation results show that by adopting proposed schemes, the overall fairness of data service can be improved compare with traditional scheduling schemes, such as QLP, LQF. It also shows that the Hybrid LQF-QLP scheme has better performance than Hybrid EQL-QLP scheme. We conclude that the proposed scheme can not only maintain the QoS criterion of real-time service, but also support the good fairness for non-real-time service in terms of packet delay and packet blocking probability.



Chapter 3

Prediction-Based Scheduling Algorithms

3.1 Introduction

The scheduling architecture proposed in Chapter 2 can guarantee the delay criterion of voice service and support a good fairness to best-effort data service. In order to maintain the delay criterion of voice service, the maximum cycle time should be bounded [7]. We have shown the relationship between average packet delay of voice service and maximum cycle time (T_{\max}) in Chapter 2, where the average packet delay will achieve one and half of the cycle time. In [7], the authors proposed a CBR-credit method to eliminate a phenomenon called light-load penalty. The principle of CBR-credit method is to reserve more resource for CBR traffic. So that the granted transmission data volume of low priority service can be transmit without replacing by CBR traffic. In [10], the authors also use the method similar to CBR-credit to guarantee the packet delay of the highest priority service. The resource reserved for highest priority service is the same as the sum of data volume granted and reported in the last cycle. Thus, the average packet delay of highest priority service can be guaranteed. In [13], the authors proposed a prediction-based LQF scheduling algorithm. They proved that the packet blocking probability is lower than the case which adopts LQF scheduling algorithm only.

In IPACT, we know that if the overhead, such as control messages and guard times, remain fixed during every cycle, the throughput will increase with the maximum cycle time (T_{\max}). Thus, we want to increase the cycle time to improve the bandwidth efficiency. To resist the increasing packet delay, we add a predictor in our scheduler architecture. Based on the predicted value of queue occupancy, we can make a better optimization of our system.

In the following of this chapter, we first introduce the system model and the prediction-based scheduler architecture in Session 3.2. And we introduce the principle of the prediction in Session 3.3. Then, the prediction-based scheduling algorithm is expressed in Session 3.4. Finally, the simulation result is obtained in Session 3.5.

3.2 System Model

In the prediction-based system model, the EPON network still remain the same as non-prediction-based EPON network model, besides the scheduler is replaced by prediction-based scheduler, as shown in Figure 3.1. There are also an OLT and N ONUs supported in the prediction-based EPON architecture. And each ONU still can support three classes of services, denoted as P_1, P_2 , and P_3 , by equipping 3 priority queues. The line rate is R_N Mbps between OLT and ONUs and R_U Mbps between ONUs and user-site both in the downlink direction and uplink direction. In an ONU, the incoming packets will first be classified into different class of service and then stored in the corresponding priority queue. When ONU receive a GATE message, the queue manager can transmit packets based on the granted transmission data volume. And after transmitting user information, the queue manager will generate the REPORT message to request additional resource.

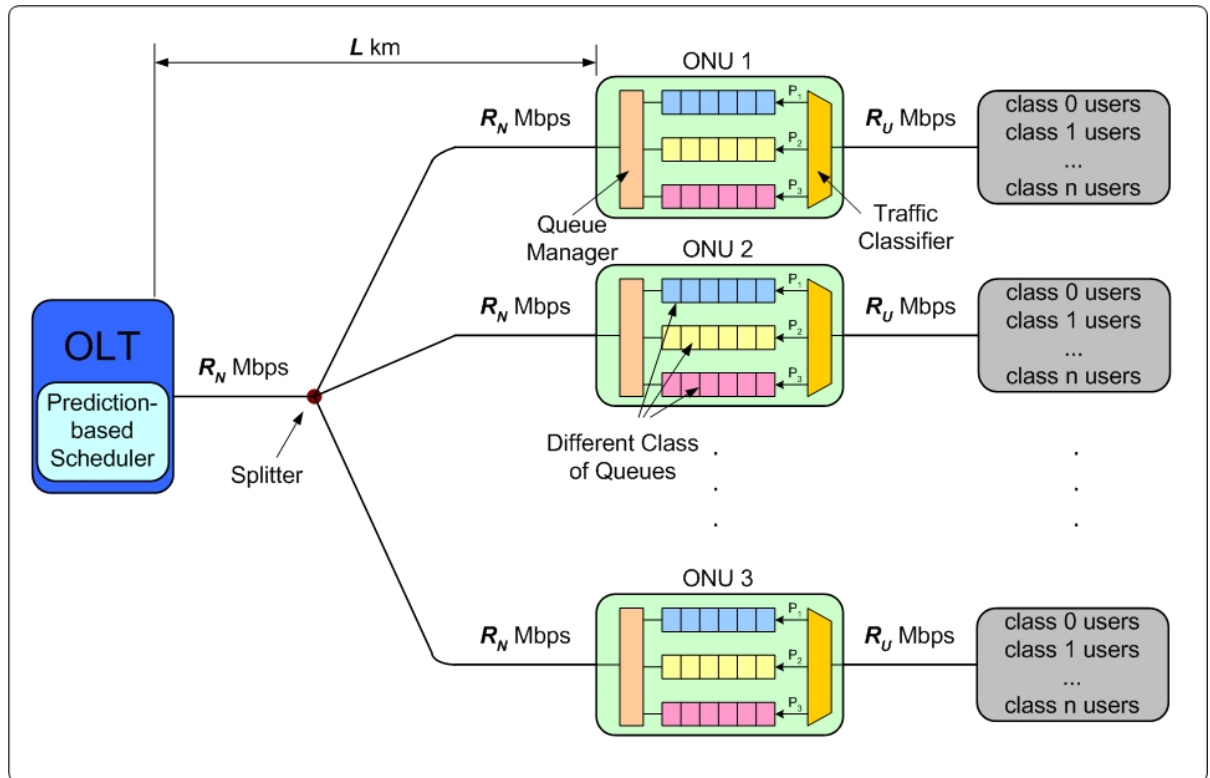


Figure 3.1: A prediction-based EPON model

The proposed prediction-based scheduler architecture is shown in Figure 3.2. We can see that there exists a predictor in our scheduler architecture. When a REPORT message arrive at OLT, it would be divided into two part of information, i.e. requested data volume of each class of service and the timestamp information of corresponding ONU. The request information will pass into predictor to predict the data volume of new coming packets which received by ONU after generating REPORT message. In other words, the predictor tries to guess the queue occupancies of each queue in corresponding ONU when this ONU receive the GATE message in the next cycle. We will describe the principle of our predictor in the later session. In the meantime, the timestamp information also passes into timing function to calculate the latest round-trip time (RTT). The method of updating the RTT is described in Figure 2.2. After the operation of timing function and predictor, the latest RTT and predicted queue occupancies will be stored in RAM. During a cycle, OLT will receive N REPORT messages from N ONUs, and they all will be stored until next scheduling time.

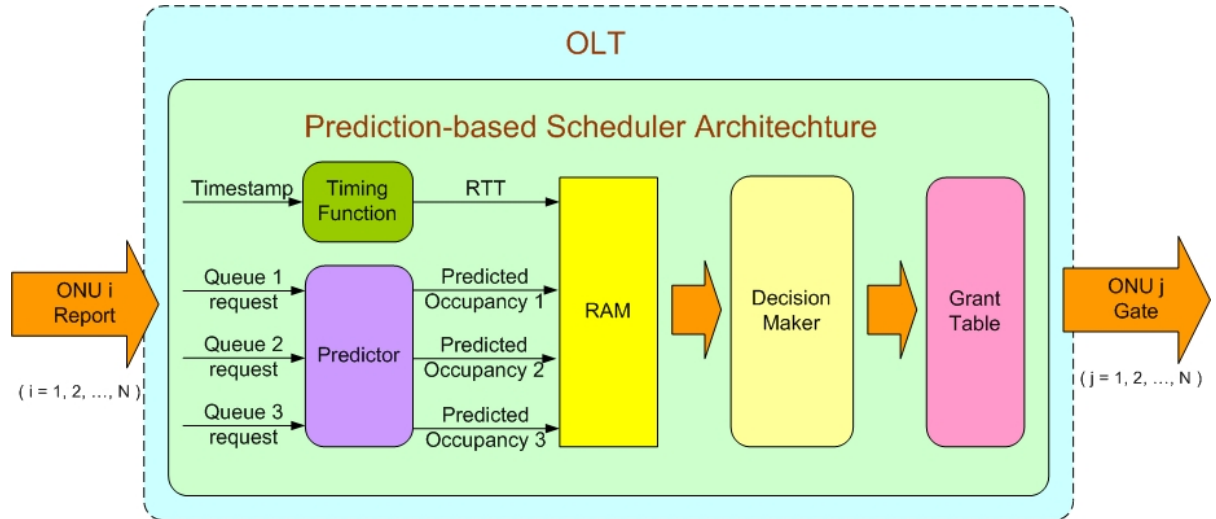


Figure 3.2: The prediction-based scheduler architecture

Before the beginning of the next cycle, the Decision Maker will calculate the timing and grant how much data volume each queue can transmit in the next cycle. The input is all the information stored in RAM, such as the latest RTT of each ONU and predicted queue occupancy of each queue in each ONU. After scheduling, the Grant Table would be updated. Finally, based on the Grant Table, the OLT can send GATE messages to ONUs so that ONUs know how much data volume each service can transmit.

In addition, we also adopt IPACT to be the interaction method between OLT and ONUs. In principle, IPACT uses an interleaved polling approach, where the next ONU is polled before the transmission from the previous one has arrived. The detail of IPACT-based polling procedure is introduced in Figure 2.5.

3.3 Predictor

The concept of prediction is shown in Figure 3.3. We see that before the beginning of Cycle $(n+1)$, we must do a scheduling operation so that the packets transmitted by ONU i (we assume that ONU i is the first candidate to transmit packets in a cycle) can be received in OLT just at the time

that last candidate finished its transition in Cycle (n).

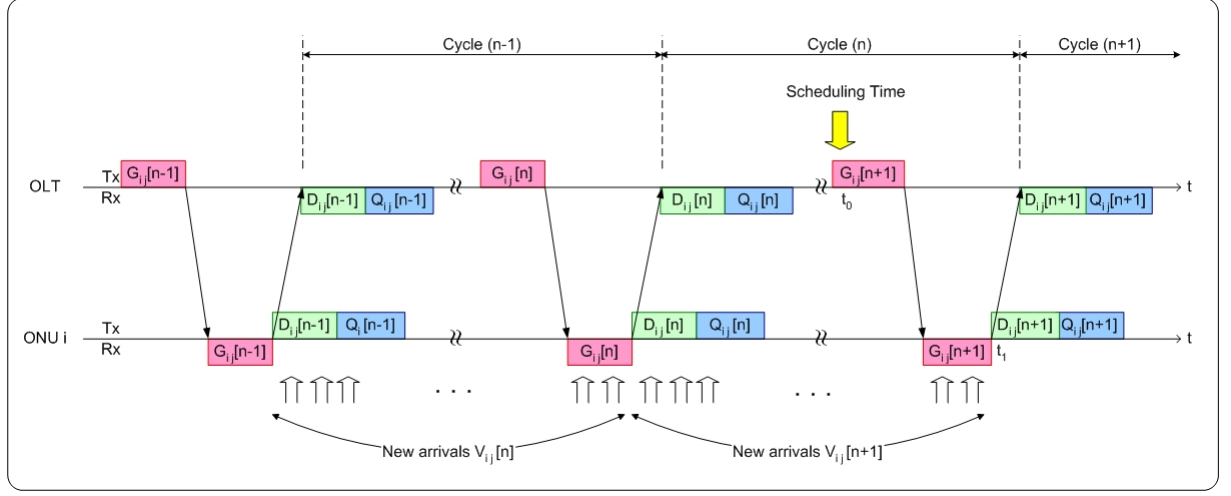


Figure 3.3: The concept of prediction method

Assume that OLT need do a scheduling operation at t_0 . During Cycle (n), the OLT has received user data, which data volume is denoted as $D_{ij}[n]$, and REPORT messages, in which the request data volume is denoted as $Q_{ij}[n]$, where $i=1, \dots, N$, and $j=1, 2, 3$. In addition, we denote $V_{ij}[n]$ as the data volume of arrival packets, where $i=1, \dots, N$, and $j=1, 2, 3$.

When OLT receives a REPORT message, in which the request data volume $Q_{ij}[n]$ is included, then the data volume of arrival packets $V_{ij}[n]$ can be obtained by

$$V_{ij}[n] = D_{ij}[n] + Q_{ij}[n] - Q_{ij}[n-1], \quad (3.1)$$

where i is ranged from 1 to N , and j is ranged from 1 to 3. Now we define the prediction order K as the number of samples we want to reference in the past history. Then prediction value of $V_{ij}[n+1]$, denoted as $V'_{ij}[n+1]$, can be obtained by

$$V'_{ij}[n+1] = \frac{1}{k} \sum_{m=0}^{K-1} V_{ij}[n-m]. \quad (3.2)$$

Finally, we define the predicted queue occupancy of class j service in ONU i as O_{ij} , and it

can be derived by

$$O_{ij}[n+1] = Q_{ij}[n] + V_{ij}'[n+1]. \quad (3.3)$$

When OLT receives a REPORT message, it will generate predicted queue occupancy for each class of service and store them in RAM. Then, at scheduling time t_0 , the scheduler can use the predicted queue occupancies to fairly assign the resource.

3.4 Prediction-based Scheduling Algorithm

Suppose that we have a set of predicted queue occupancies, i.e. O_{ij} , where $i = 1, \dots, N$, and $j = 1, 2, 3$, at scheduling time, we want to design the granted data volumes, i.e. G_{ij} , where $i = 1, \dots, N$, and $j = 1, 2, 3$, that each queue can transmit during next cycle. Also, we consider three kinds of service. The real-time services, i.e. voice and video service, have higher priority and the non-real-time service, i.e. data service, has lower priority. In our scheduling algorithm, we allocate the resource to real-time services as more as possible, that is

$$G_{i1} = \min(L_{\max}, O_{i1}), \quad (3.4)$$

$$G_{i2} = \min(L_{\max} - G_{i1}, O_{i2}). \quad (3.5)$$

Then we can derive the residual available capacity R as follows:

$$R = (T_{\max} - N \cdot b) \times R_N - \sum_{i=1}^N (G_{i1} + G_{i2}). \quad (3.6)$$

In the following session, we introduce the scheduling mechanisms for non-real-time data service. It is similar to the ones introduced in Chapter 2, besides the input of the scheduler is replaced by predicted queue occupancies. In Session 3.4.1, we explain the operation of Hybrid PEQL-PQLP scheme; and in Session 3.4.2, a detailed description of Hybrid PLQF-PQLP scheme is given.

3.4.1 Hybrid PEQL-PQLP Scheme

We rename the prediction-based Hybrid EQL-QLP scheme as Hybrid prediction-based Equal Queue Length – prediction-based Queue Length Proportion (Hybrid PLQF-PQLP) scheme. The goal of Hybrid PEQL-PLQF scheme is also to make the fairness index F , defined in equation (2.3), as close to 1 as possible. The idea of this scheme is similar to previous one. The difference is that if the predicted queue occupancy of any data service queues is over the queue length threshold, Q_{th} , the mechanism switches to the PEQL scheme to balance the queue occupancies. Otherwise, the PQLP mechanism is used.

Also, in the beginning of scheduling, the scheduler has the information of a set of predicted queue occupancies, i.e. O_{i3} , where $i = 1, \dots, N$. Then, the granted data volume of each data service queue is

$$G_{i3} = \begin{cases} O_{i3}, & \sum_{j=1}^N O_{j3} \leq R, \\ R \times \frac{O_{i3}}{\sum_{j=1}^N O_{j3}}, & \sum_{j=1}^N O_{j3} > R, \max\{O_{13}, O_{23}, \dots, O_{N3}\} \leq Q_{th}, \\ f_a(R, O_{13}, O_{23}, \dots, O_{N3}), & \sum_{j=1}^N O_{j3} > R, \max\{O_{13}, O_{23}, \dots, O_{N3}\} > Q_{th}. \end{cases} \quad (3.7)$$

If the summation of predicted queue lengths of data service is smaller than the residual available capacity R , all the data service queues are allowed to transmit packets which data volumes are at most to the predicted queue occupancy. If the summation of the predicted queue occupancies of data service is larger than R , but all of the them are smaller than Q_{th} , then the resource assign to each data service queue is proportional to its predicted queue occupancy. However, if the summation of the predicted queue occupancies of data service is larger than R , and one or more predicted queue occupancies are larger than Q_{th} , then the EQL method is adopted. The method is described as follows:

Define an index set $K = \{k_1, k_2, \dots, k_n\}$, where $K \in \{1, 2, \dots, N\}$ and index k_j satisfies

$O_{k_j,3} > avg$, $j = 1, 2, \dots, n$. Here, avg meets the equation

$$\sum_{j=1}^n (O_{k_j,3} - avg) = R. \quad (3.8)$$

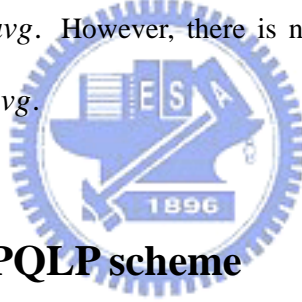
The avg can be rewritten as follows

$$avg = \frac{\left[\sum_{j=1}^n O_{k_j,3} \right] - R}{n}. \quad (3.9)$$

Then the resource assign to each data service queue is

$$G_{i3} = f_a(R, O_{13}, O_{23}, \dots, O_{N3}) = \begin{cases} O_{i3} - avg, & i \in K, \\ 0, & i \notin K. \end{cases} \quad (3.10)$$

The resource allocated to each queue, which queue length is larger than avg , is the difference between its queue length and avg . However, there is no resource assigned to the queues, which queue lengths are smaller than avg .



3.4.2 Hybrid PLQF-PQLP scheme

Also, we named the prediction-based Hybrid LQF-QLP as Hybrid prediction-based Longest Queue First – prediction-based Queue Length Proportion (Hybrid PLQF-PQLP). The goal of Hybrid PLQF-PQLP scheme is also to make the fairness index F , defined in equation (2.3), as close to 1 as possible. The inputs are the predicted queue occupancies, i.e. O_{i3} , where $i = 1, \dots, N$, and the residual available capacity R , and the outputs are the granted data volume, i.e. G_{i3} , where $i = 1, \dots, N$. The resource assignment is written as

$$G_{i3} = \begin{cases} O_{i3}, & \sum_{j=1}^N O_{j3} \leq R, \\ R \times \frac{O_{i3}}{\sum_{j=1}^N O_{j3}}, & \sum_{j=1}^N O_{j3} > R, \max\{O_{13}, O_{23}, \dots, O_{N3}\} \leq Q_{th}, \\ f_b(R, O_{13}, O_{23}, \dots, O_{N3}), & \sum_{j=1}^N O_{j3} > R, \max\{O_{13}, O_{23}, \dots, O_{N3}\} > Q_{th}. \end{cases} \quad (3.11)$$

Similarly, if the summation of predicted queue occupancy of data service is smaller than the residual available capacity R , all the requested packets are allowed to transmit in the next cycle, because the resource is adequate to assign to each data service. If the summation of predicted queue occupancy of data service is larger than R , but all of them are smaller than Q_{th} , then the resource assign to each data service queue is proportional to its requested data volume. However, if the summation of predicted queue occupancy of data service is larger than R , and one or more of them are larger than Q_{th} , then the resource assigned to each data service is described as follows.

Define the permutation function $\pi: [1, N] \rightarrow [1, N]$, so that $O_{\pi(1),3} \geq O_{\pi(2),3} \geq \dots \geq O_{\pi(N),3}$.

Note that the permutation function π is the index set with descent order of queue size mapping from the original index.

$$G_{\pi(i),3} = f_b(R, O_{13}, \dots, O_{N3}) = \begin{cases} \min\left[\max(O_{\pi(i),3} - Q_{th}, 0), R\right] + R' \cdot P_{\pi(i)}, & i = 1, \\ \min\left[\max(O_{\pi(i),3} - Q_{th}, 0), R - \sum_{j=1}^{i-1} G_{\pi(j),3}\right] + R' \cdot P_{\pi(i)}, & i = 2, \dots, N, \end{cases} \quad (3.12)$$

$$R' = \max\left(R - \sum_{j=1}^N \max(O_{j3} - Q_{th}, 0), 0\right), \quad (3.13)$$

$$P_{\pi(i)} = \frac{O_{\pi(i),3} - \max(O_{\pi(i),3} - Q_{th}, 0)}{\sum_{j=1}^N \left[O_{\pi(j),3} - \max(O_{\pi(j),3} - Q_{th}, 0)\right]}. \quad (3.14)$$

In equation (3.12), the resource allocated to each data service queue consists of two parts of resource assignment. In the first part, the scheduler allocates the emergent queues, which predicted

queue occupancies are larger than queue length threshold, an amount of the difference between the predicted queue occupancy and the queue length threshold. In addition, the longest queue has the highest priority to share the resource.

In second part, the scheduler allocates the remaining resource, denoted as R' , which is the available resource after allocation in the first part, as derived by equation (3.13), to each ONU again. However, different from the first part, the allocation is based on the proportion of remaining queue occupancies.

3.5 Simulation Result

3.5.1 Simulation Environment

The system parameters are described as follows:

| Parameter | Description | Value |
|-----------|---|-----------|
| N | Number of ONUs | 16 |
| n | Number of queues in each ONU | 3 |
| R_u | Line rate of user-to-ONU link | 100 Mbps |
| R_N | Line rate between OLT and ONU | 1000 Mbps |
| Q | The size of each queue | 1 Mbytes |
| b | Guard time between adjacent slots | 5 μ s |
| K | Number of samples we want to reference in the past history (prediction order) | 3 |

Table 3.1: System parameters used in prediction-based environment

The setting of loads for all ONUs is the same as described in session 2.5, that is

| |
|---|
| P_0 service: 4.48 Mbps \times 16(iid) P_1 service: 15 Mbps \times 16(iid) P_2 service: 15 Mbps \times 10(iid) + $M \times 6$ (iid), where $M = 15 \text{ Mbps} \sim 80 \text{ Mbps}$ Average System Load $\approx 600 \text{ Mbps} \sim 1000 \text{ Mbps}$ |
|---|

The average loads of all P_1 services, all P_2 services, and ten of P_3 services remain a fixed value in every experiment. But the average loads of six of P_3 services will change from 15 Mbps to 80 Mbps in different experiment. As a result, the average system load will change approximately from 600 Mbps to 1000 Mbps.

3.5.2 Simulation Result and Conclusion

During the simulation, we choose the $Q_{th} = 0.7$ in Hybrid PLQF-PQLP scheme and $Q_{th} = 0.9$ in Hybrid PEQL-PQLP scheme based on the result discussed in Session 2.5. First of all, we want to see the effect of adding a predictor in our scheduler, compare with the case that no predictor implemented in it. Figure 3.4 shows the system throughput of two proposed scheduling schemes. For each scheme, the prediction-based scheme is compared with non-prediction-based scheme. We can see that the system throughput in prediction-based schemes is a little bit worse than non-prediction-based scheme. It is because that the prediction results in a prediction error. In other words, the moving average prediction method can not perfectly match the behavior (variance) of self-similar traffic.

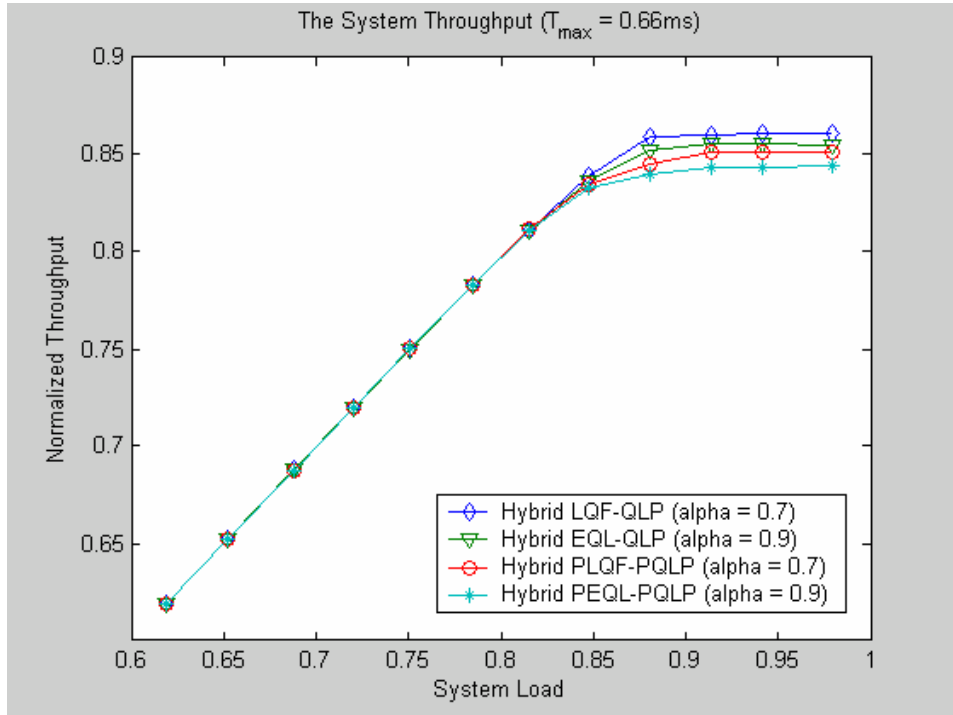


Figure 3.4: The System Throughput ($T_{\max} = 0.66\text{ms}$)

And then, the difference of the average packet delay of voice service between the prediction-based scheme and non-prediction-based scheme is shown in Figure 3.5. Also, we set the value of T_{\max} equal to 0.66ms in each scheme. The result shows that the average packet delay in prediction-based scheme can be reduced more, because the average packet delay of prediction-based scheme is half of the cycle, whereas the average packet delay is one and half of the cycle in non-prediction-based scheme.

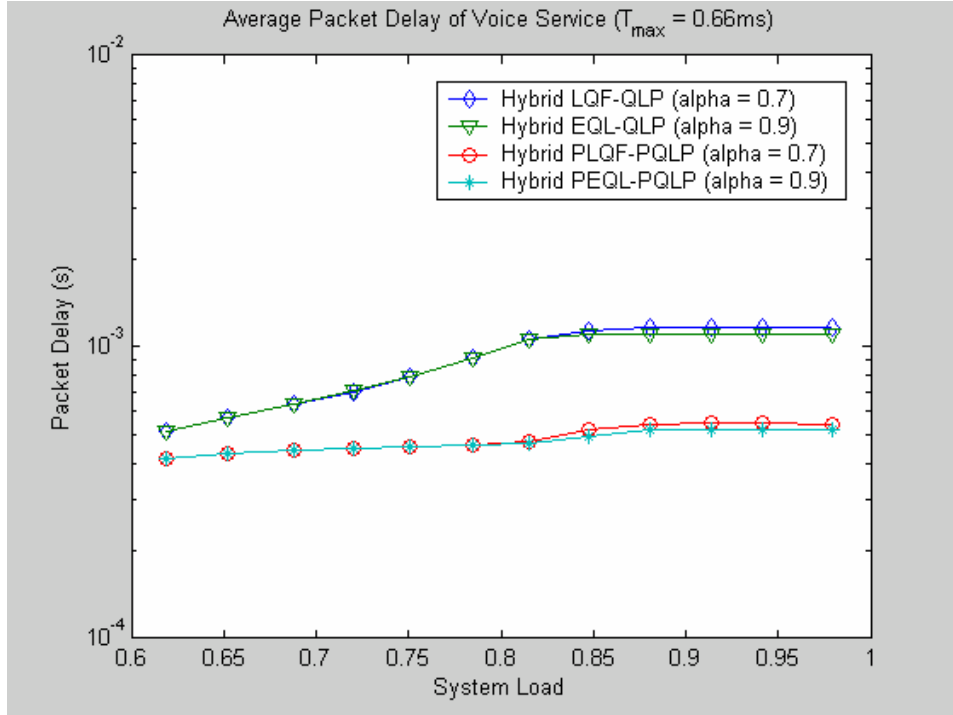


Figure 3.5: Average packet delay of Voice Service ($T_{\max} = 0.66\text{ms}$)

Then, the fairness of packet blocking probability and packet delay are illustrated in Figure 3.6 and Figure 3.7, respectively. In Figure 3.6, the result shows that the fairness in packet blocking probability can be improved in prediction-based schemes. By considering additional arrival packets in our schemes, the predicted queue occupancy may exceed the buffer size. Thus, high-loading ONUs can request more resource than the schemes with no predictor. As a result, the difference of packet blocking probability between high-loading ONUs and low-loading ONUs will be decreased. In Figure 3.7, the results show that the fairness is better in prediction-based schemes than in non-prediction-based scheme in the beginning of the curve, because the more real queue occupancies are taken into account. When the system load is high, the performance of prediction-based scheme is decreased, even lower than the schemes with no predictor. The reason is described as follows. The predicted queue occupancy will exceed the buffer size under high-loading condition. However, the arrival packets will be blocked when the buffer is overflowed. And the system ignores the packet delay of blocked packet. Thus, the effect of ignoring the packet delay of blocked packets results in a more unfair environment. If we combine two fairness indexes by the definition of overall fairness

index, we can find that the performance is better in the prediction-based environment than in the non-prediction-based environment.

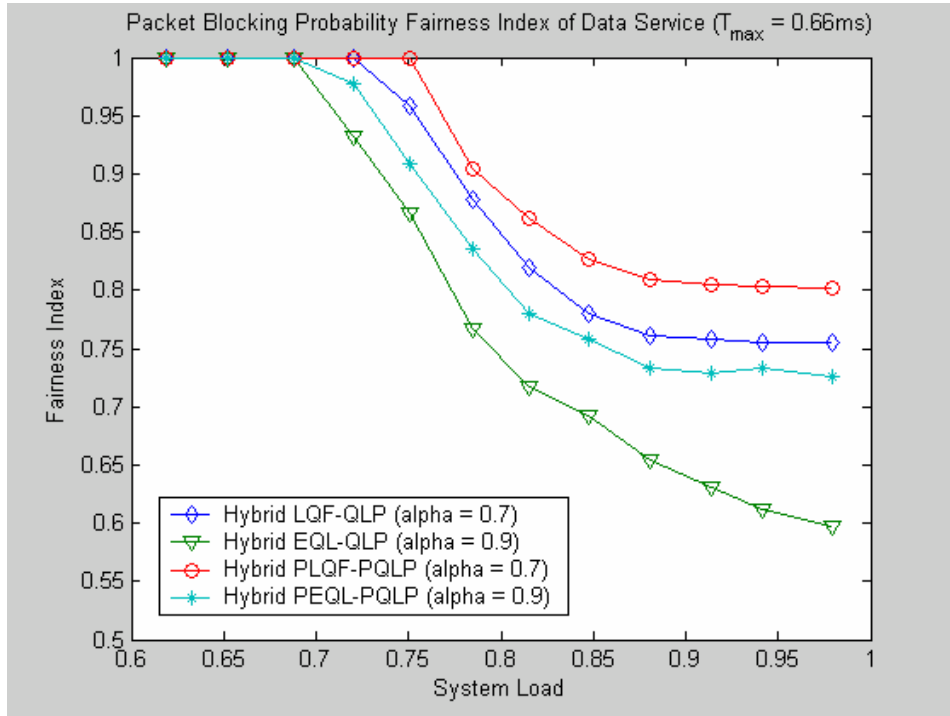


Figure 3.6: Packet blocking probability fairness index of data service ($T_{max} = 0.66ms$)

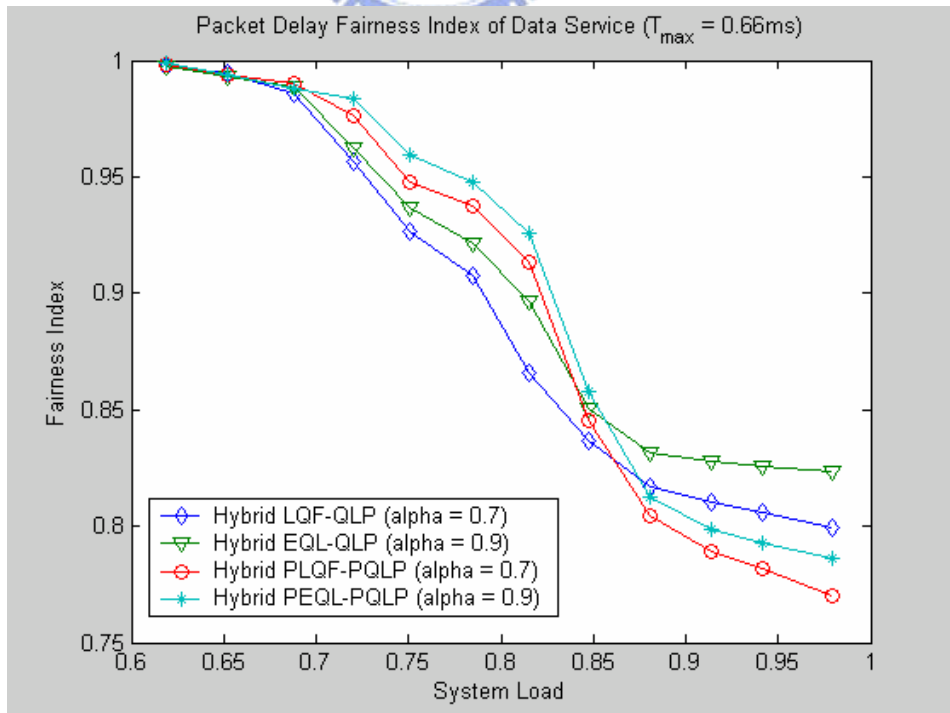


Figure 3.7: Packet delay fairness index of data service ($T_{max} = 0.66ms$)

As the result shown in Figure 3.5, the average packet delay of voice service in prediction-based scheme is much lower than the delay criterion. Thus, the maximum cycle time in prediction-based environment can be extended. Now we set the maximum cycle time to 2ms, and see the difference compared with $T_{\max} = 0.66$ ms.

In the beginning, we also want to check the average packet delay of voice service when we extend the maximum cycle time. The result is shown in Figure 3.8, we can see that both algorithms have the average packet delay lower than the specified delay bound (1.5ms). Under this basic requirement, then, we can observe the other performances.

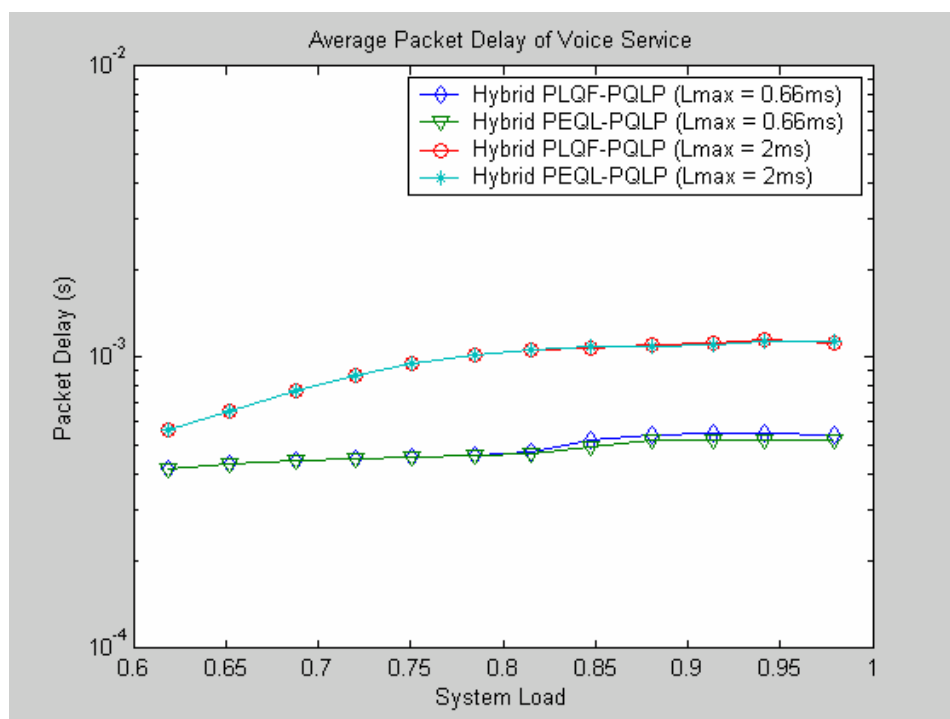


Figure 3.8: Average Packet Delay of Voice Service

In Figure 3.9, the performance of system throughput is observed, where we defined the system throughput as the ratio between the traffic load injected into ONUs and traffic load transmitted from ONUs to OLT. In this figure, we normalize the system throughput to maximum system bandwidth. In this figure, besides four proposed schemes, we also compare the performance with limited service, which is proposed in [7]. We can find that by adopting prediction-based schemes, the throughput can

be improved 7% more than non-prediction-based schemes. The reason is that by using prediction-based schemes, the T_{\max} can be extended to longer than non-prediction-based schemes. So the percentage of the overhead, such as control messages and guard, will be smaller. Additionally, we can see that the Hybrid (P)LQF-(P)QLP scheme has better performance in system throughput than Hybrid (P)EQL-(P)QLP scheme. It is because the packet truncation error is smaller in (P)LQF-(P)QLP scheme than in (P)EQL-(P)QLP scheme. In IPACT scheme, the total data volume, including real-time service and non-real-time service, an ONU can transmit during one cycle is constrained by L_{\max} . Even high-loading-ONUs tend to request resource exceeding L_{\max} , the OLT still grant these high-loading-ONUs to transmit data volume to L_{\max} . Because OLT grants a part of ONUs to transmit data volume L_{\max} , but grants the others smaller than L_{\max} , the average cycle time will be smaller than T_{\max} . As a result, the throughput would be lower than the proposed prediction-based hybrid schemes.

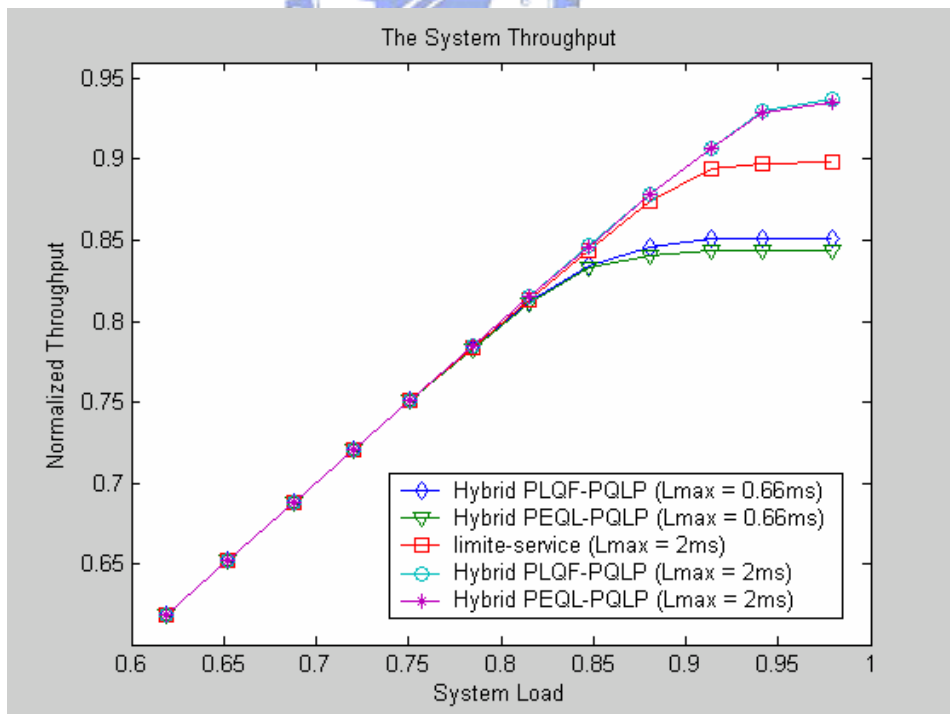


Figure 3.9: The System Throughput

The performance of packet blocking probability fairness index is shown in Figure 3.10. We compare two proposed prediction-based schemes and limited service together. We can see that the Hybrid PLQF-PQLP scheme has better performance in fairness of packet blocking probability than the other two schemes. By adopting Hybrid PLQF-PQLP scheme, the high-loading ONUs can always get the resource. So the packet blocking probability will be lower in high-loading ONUs and higher in low-loading ONUs. As a result, the difference of packet blocking probability between high-loading ONUs and low-loading ONUs is the smallest in these three schemes. By adopting Hybrid PEQL-PQLP scheme, if the system load is high, the scheduler tries to balance the queue occupancies of all queues. So, the low-loading ONUs have more chance to get the resource, and the difference of packet blocking probability between high-loading ONUs and low-loading ONUs would be higher than the case which adopts PLQF-PQLP scheme. By adopting limited-service scheme, the packet blocking probability is independent between high-loading ONUs and low-loading ONUs. The packet blocking probability of high-load ONUs will increase with system load, however, the packet blocking probability would remain a small value even the system load is increased. Thus, the difference of packet blocking probability between high-loading ONUs and low-loading ONUs is large and the performance of fairness is low.

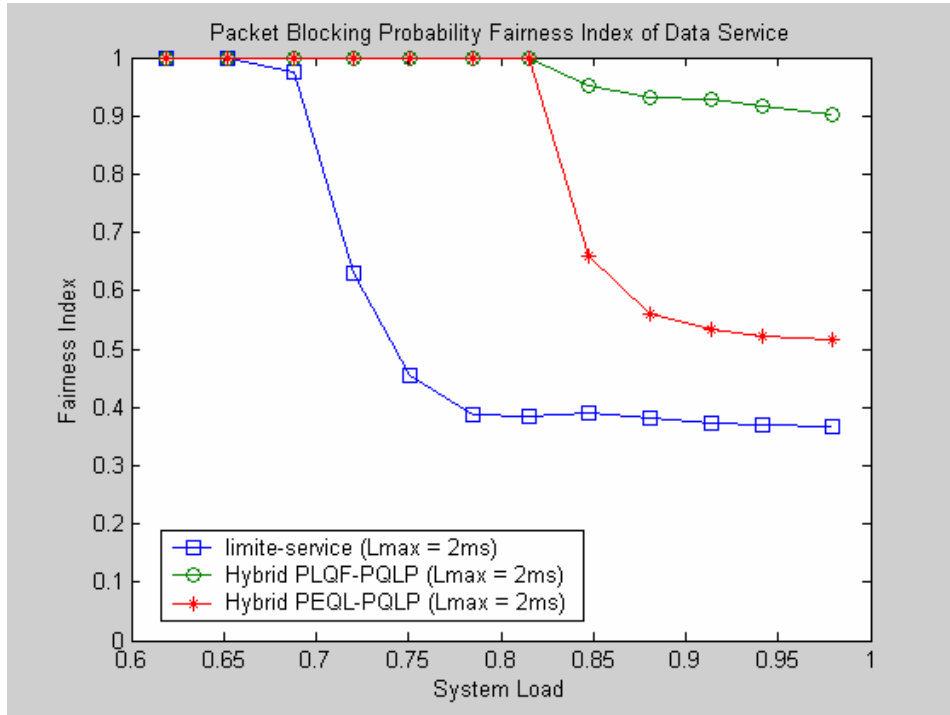


Figure 3.10: Fairness index of packet blocking probability

We show the performance of delay fairness index in Figure 3.11. We can see that PEQL-PQLP scheme is fairer than PLQF-PQLP scheme. This is because the scheduler can assign more resource to low-loading ONUs if we adopt PEQL-PQLP scheme. For limited-service scheme, the average packet delay is also independent between high-loading ONUs and low-loading ONUs. Thus, the performance is not good in packet delay fairness.

Finally, combined the performance of packet delay fairness and packet blocking probability fairness, we can get the overall fairness index, as shown in Figure 3.12. We can find that the PLQF-PQLP scheme is better than the other two schemes.

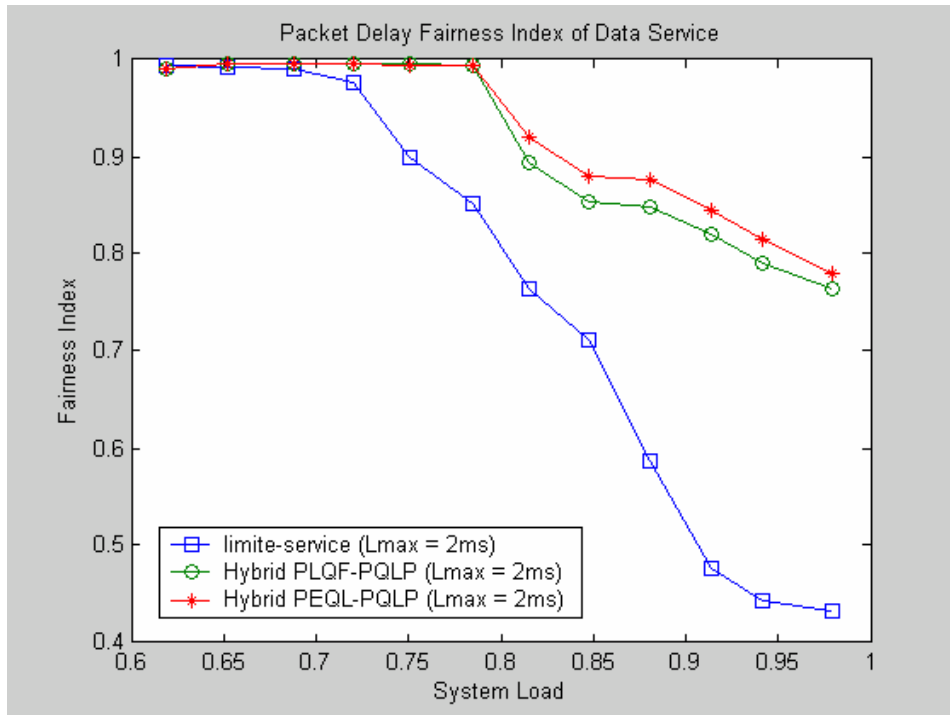


Figure 3.11: Fairness index of packet delay

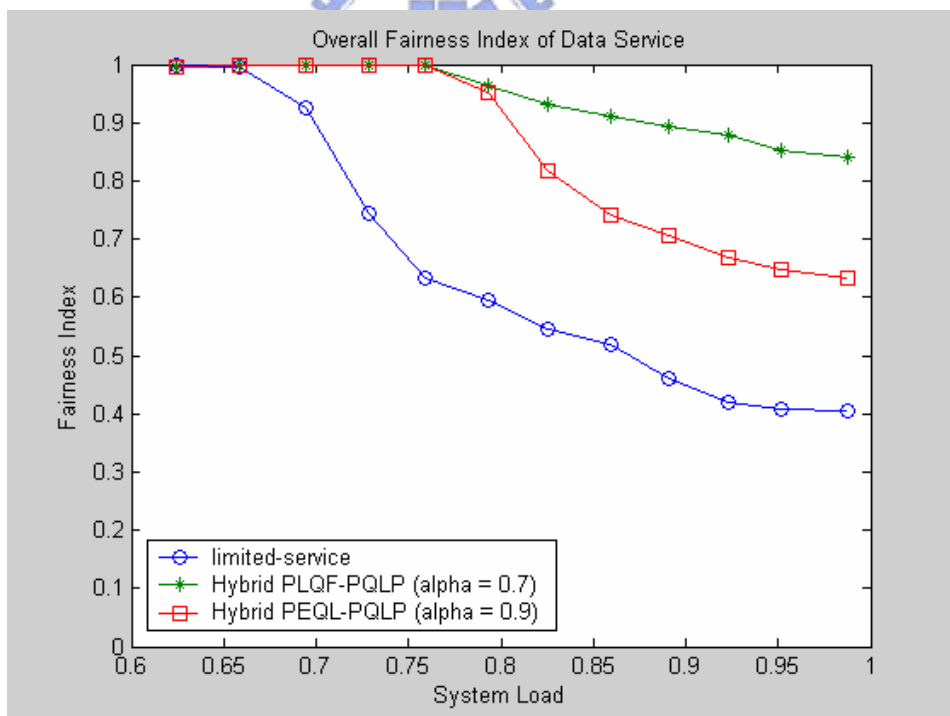


Figure 3.12: Overall fairness index of data service

3.6 Concluding Remarks

In this Chapter, we proposed a prediction-based scheduler architecture. A moving average method is chosen to estimate the number of arrived packets during a cycle. The results show that the prediction error results in a slightly decreased throughput, but will decrease the average packet delay especially in voice service. With the decreased delay, the maximum cycle time can be extended to improve the performance of system throughput.

In the simulation, we can see that the moving average method can not perfectly estimate the behavior of self-similar traffic. Because the variation of self-similar traffic is large, the predictor will over estimate frequently. We believe that there exist better predictors that can reduce the prediction error in this model.



Chapter 4

Conclusion

In this Thesis, we first proposed a scheduling method that is suitable for EPON access network in Chapter 2. Three classes of service are considered, i.e. real-time voice, real-time video, and non-real-time data service. For real-time service, the delay-considered scheduling is introduced, the average delay of voice packets and is considered. And then, for non-real-time service, the fairness-considered scheduling method is discussed. We proposed two scheduling algorithms to obtain the overall fairness which combines the fairness of packet delay with fairness of packet blocking probability. The proposed algorithms are Hybrid EQL-QLP and Hybrid LQF-QLP scheme. Each scheme is combined by two basic sub-schemes. We use a queue length threshold to be an adjusting parameter. The basic requirement of our scheme is to maintain the delay bound of voice service. Under this condition, we try to improve the fairness of packet delay and fairness packet blocking probability for non-real-time data service.

Simulation results show that, by adopting proposed scheme, the average packet delay of voice service can guaranteed and the overall fairness of data service can be improved compare with traditional scheduling schemes, such as QLP, LQF. It also shows that the Hybrid LQF-QLP scheme has better performance than Hybrid EQL-QLP scheme. We conclude that the proposed scheme can not only maintain the QoS criterion of voice service, but also support the good fairness for non-real-time service in terms of packet delay and packet blocking probability.

In Chapter 3, we proposed a prediction-based scheduler architecture. A moving average method is chosen to estimate the number of arrived packets during a cycle. The results show that the prediction error results in a slightly decreased throughput, but will decrease the average packet delay especially in voice service. With the decreased delay, the maximum cycle time can be extended to

improve the performance of system throughput.

In the simulation, we can see that the moving average method can not perfectly estimate the behavior of self-similar traffic. Because the variation of self-similar traffic is large, the predictor will over estimate frequently. However, the prediction of real-time traffic results in an extended cycle time. When the cycle time is extended, the system throughput can be improved more, compare with non-prediction-based scheme. Thus, we believe the moving average is a cost-effective solution to improve the system throughput in EPON environment.



Appendix A

LRD and Self-Similar

The characteristics of LRD and self-similarity can be specified as follows: Let $X(t)$ be a wide-sense stationary stochastic process, i.e., a process with a stationary mean $\mu = E[X(t)]$, a stationary and finite variance $\sigma^2 = E[(X(t) - \mu)^2]$, and a stationary auto-covariance function $\gamma(k) = E[(X(t) - \mu)(X(t+k) - \mu)]$, that depends only on k and not on t . For each m , let $X^{(m)}(t)$ denote a new time series obtained by averaging the original series $X(t)$, i.e.,

$$X^{(m)}(t) = \frac{1}{m} (X(mt - m + 1) + X(mt - m + 2) + \dots + X(mt)), \quad (\text{A.1})$$

and let $\gamma^{(m)}(k)$ be the auto-covariance function of $X^{(m)}(t)$. Then, the process $X(t)$ is said to be exactly (second-order) self-similar if the auto-covariance function is preserved across different time scales, i.e., $\gamma^{(m)}(k)$ equals to $\gamma(k)$ for all m and k . The process $X(t)$ is said to be asymptotically self-similar if $\gamma^{(m)}(k) \rightarrow \gamma(k)$, as $m \rightarrow \infty$. The measure of a process' self-similarity is a Hurst parameter H , where $\frac{1}{2} < H < 1$. The self-similarity can be viewed as an ability of an aggregated process to “preserve” the burstiness of the original process, i.e., the property of slowly decaying variance

$$\text{var}[X^{(m)}(t)] \sim m^{2H-2}. \quad (\text{A.2})$$

The property of long-range dependence refers to a non-summable auto-correlation function $\rho(k)$, where $\rho(k) = \gamma(k)/\sigma^2$:

$$\sum_{k=-\infty}^{\infty} \rho(k) = \infty. \quad (\text{A.3})$$

The long-range dependence results from a heavy-tailed distribution of the corresponding stochastic process. In a heavy-tail distribution, the decay obeys power law:

$$P[X > x] \sim cx^{-\alpha}, \quad \text{as } x \rightarrow \infty \quad \text{and} \quad 1 < \alpha < 2. \quad (\text{A.4})$$

As a result, the probability of an extremely large observation in LRD process is non-negligible. On the other hand, this means that extremely large bursts of data (packet trains) and extremely long periods of silence (inter-arrival times) will occur from time to time.

Appendix B

Reference Scheduling Algorithms

The reference schemes are observed and compared with the proposed algorithms. These reference schemes are limited-service [7], Queue Length Proportional (QLP) [11], and Longest Queue First (LQF) [12]. Before presenting the proposed algorithms, we want to introduce these schemes first.

In the beginning of session 2.4, we define the queue occupancies of class j service in ONU i as Q_{ij} and the granted data volume of class j service in ONU i as G_{ij} . By using equation (2.10), the residual available capacity R for best-effort data service can be derived. Then, by adopting limited-service, the grant transmission data volume of best-effort data service will be:

$$G_{i3} = \min(L_{\max} - G_{i1} - G_{i2}, Q_{i3}), \quad i = 1, 2, \dots, N. \quad (\text{A.5})$$

The limited-service discipline grants the requested number of bytes, but no more than the constraint of $(L_{\max} - G_{i1} - G_{i2})$. That means the overall data volume, including real-time packets and no-real-time packets, that an ONU can transmit in one cycle will not exceed the boundary of L_{\max} . If QLP scheme is adopted, then the grant transmission data volume assigned to each best-effort data service is proportional to the requested data volume, that is

$$G_{i3} = R \times \frac{Q_{i3}}{\sum_{j=1}^N Q_{j3}}, \quad i = 1, 2, \dots, N. \quad (\text{A.6})$$

The goal of QLP scheme is to guarantee a fair service. By adopting QLP scheme, the average packet delay would be the same. And last, we want to show the assignment method by adopting LQF scheme. If we sort the requested data volume $\{Q_{13}, Q_{23}, \dots, Q_{N3}\}$ from maximum value to minimum value and get the new sequence denoted as $\{q_1, q_2, \dots, q_N\}$, where $q_1 = \max\{Q_{13}, Q_{23}, \dots, Q_{N3}\}$, and $q_N = \min\{Q_{13}, Q_{23}, \dots, Q_{N3}\}$. Assume $Q_{i3} = q_j$, $i, j \in (1, N)$, then

$$G_{i3} = \begin{cases} Q_{i3}, & q_j \leq \left(R - \sum_{k=1}^{j-1} q_k \right), \\ \left(R - \sum_{k=1}^{j-1} q_k \right), & 0 \leq \left(R - \sum_{k=1}^{j-1} q_k \right) < q_j, \\ 0, & \text{otherwise.} \end{cases} \quad (\text{A.7})$$

By adopting LQF scheme, the queue with the largest queue occupancy is assigned the highest priority, and the scheduler always serves longest queue first. So, we can imagine that the goal of LQF is to reduce the queue length in order to cater to imminent bursty traffic.

Appendix C

Examples of EQL scheme

Assume that all the queue lengths are larger than the value Avg , as show in Figure A.1, Then, the capacity allows transmitting for each queue is:

$$G_{i3} = Q_{i3} - \frac{\left(\sum_{j=1}^N Q_{j3} \right) - R}{N}, \quad (\text{A.8})$$

where R is the available residual resource for data service. Obviously, after allocating bandwidth, the queue occupancies are the same, i.e. $Q_{i3} = Avg$, where $i = 1, 2, \dots, N$, and packet blocking probability is decreased. The average queue occupancy Avg is

$$\text{Avg} = \frac{\sum_{j=1}^N Q_{j3} - R}{N}. \quad (\text{A.9})$$

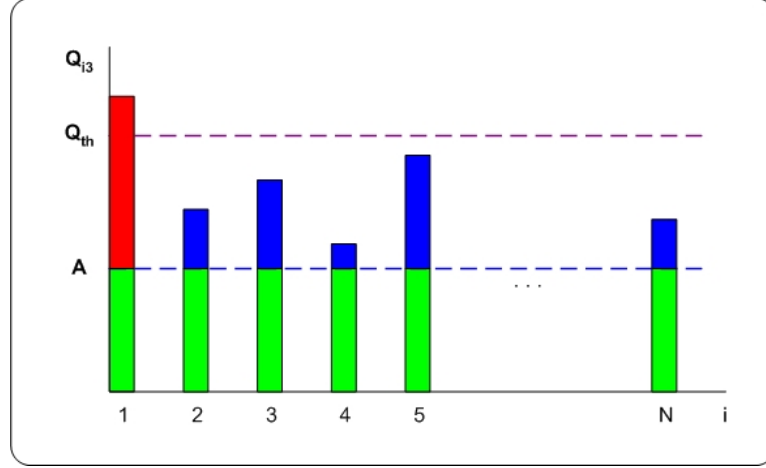


Figure A.1: First example of EQL scheme

Moreover, under a condition that the occupancies between each queue are quite different, it is possible that we can not achieve the goal of equal occupancies, as shown in Figure A.2. Suppose the queue with shortest occupancies is Q_{ik} , the optimal allocating method is to pick out the shortest one and allocate total available resource to the other ones. That is

$$G_{i3} = \begin{cases} Q_{i3} - \frac{\left(\sum_{j \neq k, j=1}^N Q_{j3} \right) - R}{N-1}, & i = 1, 2, \dots, N, i \neq k, \\ 0, & i = k. \end{cases} \quad (\text{A.9})$$

And after allocating bandwidth, the new average queue occupancy Avg' is:

$$\text{Avg}' = \frac{\sum_{j \neq k, j=1}^N Q_{j3} - R}{N-1}. \quad (\text{A.10})$$

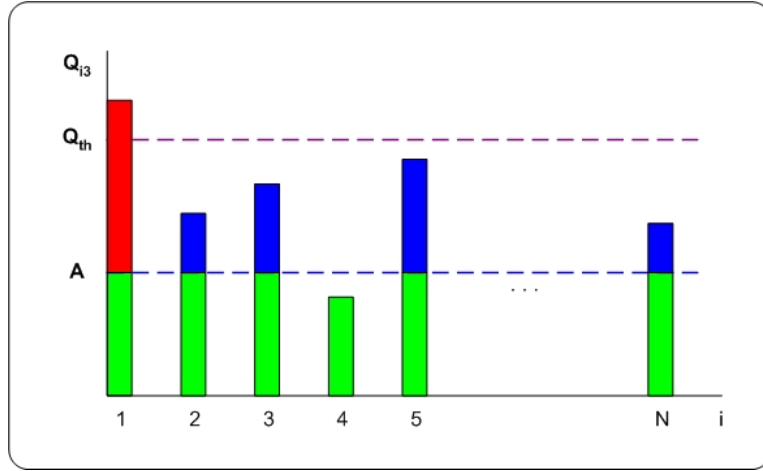


Figure A.2: Second example of EQL mechanism

Is it possible that Avg' is smaller than Q_{k3} ? The answer is “not possible”. We can see the proof as follows:

$$\begin{aligned}
 Avg' &= \frac{\sum_{j \neq k, j=1}^N Q_{j3} - R}{N-1} = \frac{\left(\sum_{j=1}^N Q_{j3} - Q_{k3} \right) - R}{N-1} = \frac{N \left(\sum_{j=1}^N Q_{j3} - R \right) - Q_{k3}}{N-1} \\
 \therefore Avg &= \frac{\left(\sum_{j=1}^N Q_{j3} - R \right)}{N} > Q_{k3} \\
 \therefore Avg' &= \frac{N \left(\sum_{j=1}^N Q_{j3} - R \right) - Q_{k3}}{N-1} > \frac{N}{N-1} \left(Q_{k3} - \frac{Q_{k3}}{N} \right) = Q_{k3} \cdot \frac{N}{N-1} \left(1 - \frac{1}{N} \right) = Q_{k3} \\
 \Rightarrow Avg' &> Q_{k3}
 \end{aligned}$$

If we still can not achieve the goal of equal occupancy, then second smallest ones out and allocate the resource to the others. This step may be repeated more time until we find the optimal solution. Obviously, by adopting EQL mechanism, the balance between the queue occupancies is achieved, and the possibility of packet dropping of emergent queues can be decreased.

The EQL mechanism described before can also be described as mathematic form as follows:

$$\begin{cases} Q_{13} - G_{13} = Q_{23} - G_{23} = \dots = Q_{N3} - G_{N3}, \\ G_{13} + G_{23} + \dots + G_{N3} = R. \end{cases} \quad (A.11)$$

If one or more of the solutions, denoted as G_{j3} , where $j \in N$, are negative. In this situation, we set the most negative one to zero, i.e. $\min_{l \in j} \{G_{l3}\} = 0$, and recalculate the grant transmission capacity in equation (A.11). Now the number of variables and equations becomes $N-1$. This operation will be repeated until that all the grant transmission capacities we solve are not less than zero.



Bibliography

- [1] G. Kramer and G. Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access Network," *IEEE Commun. Mag.*, vol. 40, pp.66-73, Feb. 2002.
- [2] H. Ueda, K. Okada, B. Ford, G. Mahony, S. Hornung, D. Faulkner, J. Abiven, S. Durel, R. Ballart, and J. Erickson, "Deployment status and common technical specifications for a B-PON system," *IEEE Commun. Mag.*, vol. 39, pp. 134-141, Dec. 2001.
- [3] Jingshown Wu, F. R. Gu, and H. W. Tsao, "Jitter performance analysis of SOCDMA-based EPON using perfect difference codes," *IEEE Journal of Lightwave Technology*, vol. 22, no. 5, pp. 1309-1319, May 2004.
- [4] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON (ePON): design and analysis of an optical access network," *Photonic Network Commun.*, vol. 3, no. 3, pp. 307-319, July 2001.
- [5] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A Dynamic Protocol for an Ethernet PON (EPON)," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 74-80, Feb. 2002.
- [6] K. Rege et al., "QoS management in trunk-and-branch switched Ethernet networks," *IEEE Commun. Mag.*, vol. 40, pp. 30-36, Dec. 2002.
- [7] G. Kramer, B. Mukherjee, S. Dixit, Y. Ye, and R. Hirth, "Supporting differentiated classes of service in Ethernet passive optical networks," *J. Opt. Networks*, vol. 1, no. 8/9, pp. 280-298, Aug. 2002.
- [8] S. Choi and J. Huh, "Dynamic bandwidth allocation algorithm for multimedia services over Ethernet PONs," *ETRI J.*, vol. 24, no. 6, pp. 465-468, Dec. 2002.
- [9] M. Ma, Y. Zhu, and T. H. Cheng, "A bandwidth guaranteed polling MAC protocol for Ethernet passive optical networks," in *Proc. IEEE INFOCOM*, San Francisco, CA,

Mar.-Apr. 2003, pp. 22-31.

- [10] C. 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 Commun.*, vol. 21, no. 9, pp. 1467-1477, Nov. 2003.
- [11] D. Liu, N. Ansari, and E. Hou, "QLP: A Joint Buffer Management and Scheduling Scheme for Input Queued Switches," *IEEE Workshop on High Performance Switching and Routing*, pp. 29-31, May 2001.
- [12] D. -S. Lee, "Generalized longest queue first: An adaptive scheduling discipline for ATM networks," *IEEE INFORCOM'97*, vol. 3, pp. 1096-1104
- [13] Mu Si, Ding Quanlong, and Ko Chi Chung, "Improving the network performance using prediction based longest queue first (PLQF) scheduling algorithm," *IEEE International Conference on ATM and High Speed Intelligent Internet Symposium*, vol. 22-25, pp. 344-348, April 2001.
- [14] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, "An Architecture for Differentiated Services," *IETF*, RFC 2475, Dec. 1998.
- [15] *Virtual Bridged Local Area Networks*, IEEE Standard 802.1Q, 1998.
- [16] IEEE standard 802.3ah task force home page [Online]. Available: <http://www.ieee802.org/3/fem>
- [17] A. Adas, "Traffic models in broadband networks," *IEEE Commun. Mag.*, vol. 35, no. 7, pp. 82-89, July 1997.
- [18] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," In *Proc. ACM SIGCOMM'95*, pp. 100-113, Cambridge, MA, Aug. 1995.
- [19] 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.

- [20] R. Jain, *The Art of Computer Systems Performance Analysis*, John Wiley and Sons Inc, 1991.



Vita

姓名：彭崇禎

學歷：

2002 ~ 2004 交通大學電信工程研究所

2000 ~ 2002 台灣科技大學電機工程系

1995 ~ 2000 台北科技大學電機工程科

E-mail: jackpeng.cm91g@nctu.edu.tw

