行政院國家科學委員會專題研究計畫　成果報告

子計畫三：光纖擷取網路技術之分析研究(3/3)

計畫類別：整合型計畫
計畫編號：NSC94-2213-E-009-018-
執行期間：94 年 08 月 01 日至 95 年 07 月 31 日
執行單位：國立交通大學電信工程學系(所)

計畫主持人：張仲儒

計畫參與人員：林立峰、唐文祥、黃琴雅、張家源、蔡俊帆

報告類型：完整報告

處理方式：本計畫可公開查詢

中　華　民　國 95 年 10 月 27 日

行政院國家科學委員會補助專題研究計畫 ■ **成果報告** □**期中進度報告**

光纖擷取網路技術之分析研究
(Optical Access Networking Technology)

計畫類別：□ 個別型計畫 ■ 整合型計畫
計畫編號：NSC 94 -2213-E-009 -018
執行期間： 94 年 8 月 1 日 至 95 年 7 月 31 日

計畫主持人：張仲儒 教授
共同主持人：
計畫參與人員： 林立峰、唐文祥、黃琴雅、張家源、蔡俊帆

成果報告類型(依經費核定清單規定繳交)：□精簡報告 ■完整報告

本成果報告包括以下應繳交之附件：
□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

處理方式： 除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢
□ 涉及專利或其他智慧財產權，□一年□二年後可公開查詢

執行單位：國立交通大學電信工程學系

中 華 民 國 95 年 10 月 29 日

■ 中、英文摘要及關鍵詞(keywords)：(此部分的頁碼編寫請用羅馬字Ⅰ、Ⅱ、Ⅲ…標在每頁下方中央)

■ 目錄：(此部分的頁碼編寫請用羅馬字Ⅰ、Ⅱ、Ⅲ…標在每頁下方中央)

■ 報告內容：(「報告內容」至「附錄」部分請以阿拉伯數字 1.2.3.…順序標在每頁下方中央)

●計畫緣由及目的

●研究方法

●結果與討論：含結論與建議

●參考文獻

※若該計畫已有論文發表者，可以 A4 紙影印，作為成果報告內容或附錄，並請註明發表刊物名稱、卷期及出版日期。

※若有與執行本計畫相關之著作、專利、技術報告、或學生畢業論文等，請在參考文獻內註明之，俾可供進一步查考。

■ 計畫成果自評：請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

■ 可供推廣之研發成果資料表：凡研究性質屬<u>應用研究</u>及<u>技術發展</u>之計畫，請依國科會提供之表格，每項研發成果填寫一份。

■ 附錄(或附件)

計畫中文摘要

隨著各設備發展廠商在光電通訊技術準化的發展，以及網際網路(Internet)的快速發展全球資訊網(World Wide Web)的風行，以及數位多媒體內容的增加，大眾對高速網路的網路頻寬需求呈倍數般爆量成長，因此電信業者提高對於寬頻光纖網路系統與光網路的建置，提供更多更快的數位網路服務。

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，乙太被動光纖網路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。針對此一研究主題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。同時，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。此外，我們也提出了一個能適用於乙太被動光纖網路上改善服務品質的動態頻寬分配方法，只要針對及時性的服務設計了已延遲以及封包被丟掉為考量的頻寬分配機制。同時，也針對非及時性的服務設計了以避免他過度被犧牲作為考量的的排成機制。

雖然光纖網路的傳輸容量大幅的提升，然而相對於一個波長提供的頻寬，訊務的流量仍是相當的小。所以為了有效的利用資源，就必須利用訊務彙集的技術，把需求頻寬小的訊務彙集到頻寬大的波長上。我們考慮動態訊務彙集的問題。而且我們的目標是讓波長頻寬的使用率達到最高，同時降低新連線呼叫拒絕率。根據要達成的目標，我們把這個問題公式化成整數線性規劃(ILP)的問題，並且為了解這個問題，我們從模擬退火演算法的精神，提出一個 STGA 演算法來求的最佳解。但是因為計算複雜度的關係，我們另外提出計算法雜度較低的 HTGA 的演算法。在 HTGA 演算法裡面主要有三個步驟：訊務彙集的動作、波長指派的動作、訊務重新被安排的動作。這些動作的主要目的是在儘可能不要改變光通道的情況下讓系統的使用率更好。

最後我們將完成系統模擬程式的撰寫，利用電腦模擬的方式來驗證所獲得研究成果的正確性。

關鍵詞：　乙太被動光纖網路、動態頻寬分配方法、公平性、服務品質、動態訊務彙集、線性規劃、退火演算法

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. So, 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. In addition, we also 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.

In the meantime, we propose a QoS-promoted dynamic bandwidth allocation (Q-DBA) method to support transmitting delay-constrained voice and video packets, and starvation-considered data packets. The goal of the Q-DBA method is to meet delay requirement of voice service when the voice dropping probability is set to zero, to meet delay requirement, and to meet dropping probability of video probability. Simultaneously the goal is also to maintain the delay of data packets. The data packets should not endure a long delay time although they do not have strict delay requirement. When the QoS requirement of real-time service can be satisfied, the performance of non-real-time service should be improved.

The transmission capacity of an optical network largely increases because of the development and application of wavelength division multiplexing (WDM) technologies. However the required bandwidth of a traffic stream is also much smaller than the bandwidth capacity of a wavelength. Thus, in order to efficiently utilize the network resources, many lower-speed traffic streams can be multiplexed onto a high-speed wavelength by traffic-grooming technique. We consider the traffic grooming problem, which the property of the traffic is dynamic and nonuniform traffic. Our objective is to effectively reduce the new call blocking rate and maximize the utilization efficiency of the wavelengths which are used by the system. Integer linear problem (ILP) methodology is applied to formulate for this problem. And, in

order to solve the ILP, we first propose a simulated annealing-based traffic grooming (STGA) algorithm to obtain the optimal solution. However, the STGA algorithm is infeasible due to its computation complexity. Alternatively, we propose a heuristic algorithm, called a heuristic-based traffic grooming (HTGA) algorithm. There are the three main operations in HTGA algorithm: operation of traffic grooming, operation of wavelength assignments, operation of traffic rearrangement. The main purpose of these operations is to achieve the better system utilization without changing the lightpath topology as could as possible.

Keywords： Ethernet Passive Optical Networks (EPON), Dynamic Bandwidth Allocation (DBA), Fairness, QoS, Traffic Grooming, Integer Linear Problem, Annealing Method

Agenda

I. Introduction and Motivation

In our proposal isuues, we consider the QoS-grauteed base mechanism in all optical acess metro network. It is due to that for the excellent performance of silica fiber, such as wide bandwidth, low transmission loss, lightweight, and immunity to interference, it is broadly deployed in long haul trunk and is gradually migrated into the access network to form the Fiber-In-The-Loop (FITL) architecture. Many efforts have been made to overcome the economic barrier for the mass deployment of FITL systems since it can improve the quality of the network and enhance the flexibility to provide the new broadband services. The rapid spread of the World Wide Web and increasing applications of digital contents have significantly increased the demand for high-speed Internet access. The networking of company LANs for intranets and extranets has also been increasing. They all need the broadband fiber access system and optical network to provide such high bandwidth and services. Also, the transmission capacity of an optical network largely increases because of the development and application of wavelength division multiplexing (WDM) technologies. However the required bandwidth of a traffic stream is also much smaller than the bandwidth capacity of a wavelength. Thus, in order to efficiently utilize the network resources, many lower-speed traffic streams can be multiplexed onto a high-speed wavelength by traffic-grooming technique. Next, we will discuss theses problem into three topics.

I.1. Fair Scheduling Schemes for Non-Real-Time Service in EPON Access Network

A passive optical network (PON) is a point-to-multipoint optical network with no active element along the path connecting sources and destinations. An Ethernet passive optical network (EPON), which uses low-cost Ethernet equipment and fiber infrastructure to carry Ethernet packets [1], is one of the candidates for the next-generation access network. An EPON consists of one optical line terminal (OLT) connecting with several optical network units (ONUs). The OLT resides in the central office (CO) and is responsible for connecting EPON to a metropolitan area network (MAN) or a wide-area network (WAN). The ONU locates at either the curb or the end-user site for providing broadband services to end-users.

In EPON, user data is first encapsulated into Ethernet frames and then exchanged between OLT and ONUs. In the downstream direction (i.e., from the OLT to ONUs), OLT broadcasts Ethernet packets to all ONUs via a passive splitter. Each ONU can easily receive its packets and ignore the others' packets by filtering the MAC addresses of incoming packets. In contrast, all ONUs have to share a common resource to transmit packets to OLT in the upstream direction (i.e., from ONU to OLT). Therefore, the network requires a multiple access scheme to prevent ONUs from packet collisions. Several multiple access schemes have been proposed for the upstream channel access in EPON, which include wavelength division multiple access (WDMA)-based schemes [1], contention-based multiple access schemes [1], optical code division multiple access (OCDMA)-based schemes [2], and time division multiple access (TDMA)-based schemes [3]. Among them, the TDMA-based scheme seems to be the most cost-effective solution for EPON because it requires only one transceiver in the OLT and is easy to implement. Hence, the TDMA-based approach is adopted herein.

Kramer, Mukherjee, and Pesavento [3] proposed a TDMA-based scheme that allocates a fixed-duration timeslot to each ONU for its upstream packet transmission. The authors further proposed an Interleaved Polling with Adaptive Cycle Time (IPACT) scheme [4] to improve the bandwidth utilization of their TDMA-base scheme. In IPACT, the timeslot is dynamically allocated to ONUs according to their traffic loading. The authors investigated performance of IPACT with several scheduling algorithms (e.g., fixed, limited, gated, constant credit, linear credit, and elastic) and found that IPACT with limited scheduling exhibits the best performance. In IPACT with limited scheduling, OLT allocates the time duration to each ONU according to its requested data volume unless an upper bound is exceeded. This approach is simple but lack of flexibility since an overloaded ONU cannot get extra resource even if the network is under-utilized. Several scheduling schemes were proposed to be used in conjunction with IPACT to support IP-based multimedia traffic with diverse quality-of-service (QoS) requirements [5]. In [6], the authors adopted IPACT with limited-service scheme to support different class of service in EPON. It adopts the gated transmission mechanism provided by the multi-point control protocol (MPCP) [7]. Other proposals [8-10] also employed a TDMA-based approach and dynamically allocate bandwidth to end users according to their QoS requirements.

Most of the aforementioned schemes aim to guarantee the bandwidth or delay requirements for real-time services. However, the performance of non-real-time services is not addressed. Hence, it would be helpful to take the fairness of both average packet delay and average packet blocking probability of non-real-time services as another performance metric in EPON. Previous studies had demonstrated that queue length proportional (QLP) scheme and longest queue first (LQF) scheme favor the fairness in packet delay and blocking probability, respectively [11-12].

I.2. QoS-Promoted Dynamic Bandwidth Allocation (Q-DBA) for Ethernet Passive Optical Networks

A high-speed broadband access network is required to support multimedia services, such as high-definition television (HDTV), video on demand (VoD), Internet telephone, and so on. Ethernet passive optical networks (EPONs), which is backward compatible with existing IEEE 802.3, have emerged as one of the most promising access network and are gaining high momentum due to the convergence of low-cost Ethernet equipment, low-cost fiber infrastructure, and the large bandwidth of the optical fiber link.

There is an upper bound of a wavelength's bandwidth, the shared bandwidth for serving upstream traffic is limited. Serval methods are proposed to efficiently use the limited bandwidth and support different requirements on the bandwidth allocation for the upstream transmission.

With the TDMA approach, the timeslot may be fixed or variable for transmitting one or more packets depending on the OLT's allocation algorithm. A fixed timeslot assignment algorithm in EPON [13] was easy to implement but it did not meet ONUs' dynamic requirements, and resulted the bandwidth under utilization. The granted bandwidth is either insufficient for a longer queue occupancy or too much for a shorter one due to the different queue occupancies of all ONUs every cycle and the bursty nature of network traffic. To overcome this problem, a polling based scheme, interleaving polling with adaptive cycle time (IPACT) [14], was proposed. The next ONU was polled before receiving packets from previous one, and the OLT dynamically assigned bandwidth to all ONUs in accordance with ONUs' demands. The bandwidth was cable of being used more efficiently in [14]. However, if all packets played the same status and served according to FCFS (first come, first serve), all packets do attain the same average delay time no matter what they were delay sensitive or not. In this way, it was not suitable for delay and jitter sensitive services.

As EPONs' technology matures, related QoS issues are becoming a key concern. The scheduling algorithms, relying on dynamic bandwidth allocation (DBA) algorithms [15], [16], [17], were proposed to support a variety of network services. Burst-polling based delta dynamic bandwidth allocation [15] and dynamic bandwidth allocation with multiple services (DBAM) algorithm [16] used a class-based traffic prediction to estimate queue's traffic arriving for enhancing QoS metrics such as average delay, but the introduced maximum window parameter of a specific class of traffic decreased the performance due to the ONU did not ask and obtain bandwidth more than the maximum window even if bandwidth was left. Traffic-class burst-polling based delta DBA (TCBPDDBA) [17] proposed a traffic class based message polling scheme. It not only reallocated surplus bandwidth to the heavily loaded ONUs to improve the under utilized problems in [16] but also provided a QoS guarantee to delay sensitive services. However, it did not individually allocate bandwidth to each class, and resulted the longest delay to non-delay sensitive services if the traffic load was heavy and the ONU did not arrange the transmission well.

To find another solution to make high bandwidth utilization after the QoS is satisfied, some people proposed ”threshold-based” mechanisms referring to queue occupancy or number of packets in the ONU [18], [19], [20]. The authors ([18]) consider the tradeoff between the bandwidth efficiency and the delay characteristics to make the bandwidth be fully utilized while the packet delay and delay variation increased. The authors ([19]) take a completed packet and the threshold into account and their method achieves a nearly optimal bandwidth utilization. However, the report message was not large enough to contain all information about packets' size, so the improvement was still limited. Dynamic credit distributed (D-CRED) based mechanism [20] proposed a dynamic queue management for the weighted fair allocation to reach a higher efficiency, but the delay variation increases with the increasing of traffic load, especially of the low priority traffic.

I.3. Computation-Efficient Algorithms for Dynamic Traffic Grooming in Metro-Access Ring Network

In recent years, synchronous optical network (SONET) ring networks have been widely deployed for the optical network infrastructure. The progresses of optical networks evolve with time; meanwhile, the carried traffic streams surge. The transmission capacity of an optical network largely increases because of the development and application of wavelength division multiplexing (WDM) technologies which allow multiple wavelengths to be delivered in a single fiber at the same time [21].

A traffic stream is normally assigned in one wavelength. However, the required bandwidth of a traffic stream, called lower-speed traffic stream, is much smaller than the bandwidth capacity of a wavelength. If every lower-speed traffic stream occupies an individual wavelength, bandwidth waste is an obvious and serious drawback. In order to efficiently utilize the network resources, many lower-speed traffic streams can be multiplexed onto a high-speed wavelength by traffic-grooming technique. An overview of the traffic-grooming technique and survey of some typical works were reported in [22].

In the design of traffic-grooming, three issues to be considered: network configuration, traffic characteristics, and cost function. For the network configuration issue, ring network architectures were mostly studied for grooming [23-30], which may be long-haul backbone networks or metro networks. In addition, many researches [31], [32] also focus on the mesh network, which their backbone is a long-haul backbone network. In the early stage of optical network, the ring network is employed in a long-haul backbone network. With the increasing of the traffic load, the mesh network has been studied by many researches because the mesh network sustains more traffic load than the ring network. However, for metro networks, ring network is preferable because the mesh network is an over-provisioning and expensive solution. Instead, many people study the ring network for the metro networks.

For the traffic characteristics, two properties are used to describe the behavior of traffics in this paper. One is bandwidth request and the other one is connection pattern. Traffics may be categorized into uniform or non-uniform ones according to their required bandwidth. The bandwidth request is the same for all the users, called uniform traffic; otherwise, it is non-uniform traffic. And, the traffics may also be categorized into static or dynamic ones according to system connection including the amount of connection, the occupied bandwidth of all connection. The system connection does not change with time for static traffic while it does for dynamic traffic. Notably, the required bandwidth of each call request does not change during its transmission time in this paper.

Besides, many objectives were proposed for the definitions of cost functions and may be categorized into five types. The first one is minimization of the required number of Line Terminating Equipment (LTE) over all the network nodes. The second one is minimization of the average hop counts. The third one is minimization the maximum number of lightpaths originating/terminating at a network node. The fourth one is maximization of traffic throughput of the optical network. The last one is minimization of the blocking probability of the new call request in dynamic traffic.

In [31], the traffic-grooming problem was considered WDM mesh network with static and uniform traffic. The new node architecture was proposed for a WDM mesh network with traffic-grooming capacity and cost function was designed to maximize the total successfully-routed lower speed traffic. A traffic-grooming problem in WDM ring was studied in [23], where the traffic is static and non-uniform. The authors presented a framework of bounds, both upper and lower bounds, on the optimal value of the amount of traffic electrically routed. In [24], the authors considered a Min-Max objective, which was desirable to minimize the cost at the node where the cost is maximum. They presented a polynomial-time traffic grooming algorithm for minimizing the maximum electronic port cost in both uni- and bi-directional ring.

The ring network was considered with static and non-uniform traffic [25]. The simulated-annealing-based heuristic algorithm for the traffic grooming problem was proposed for minimization of the total cost of electronic equipment. In [26], both nui- and bi-directional ring were considered and two kinds of traffic are considered. One was static and uniform traffic and the other was static and non-uniform traffic. For minimizing number of wavelengths, the greedy heuristic algorithm was employed for two kinds of traffic. The author also proposed another heuristic algorithm for static and non-uniform traffic according to the same objective. In [26], the ring network was considered with static and uniform traffic. The author proposed the heuristic algorithm for minimizing number of wavelengths and also considered computational complexity.

Although the ring network has been studied by many researchers, dynamic and non-uniform traffic has not considered yet. Therefore, we are motivated to study the traffic grooming problem in metro-access ring network with dynamic and non-uniform traffic. In the paper, the metro-access ring network has nodes in WDM-PON architecture connecting to end-users. We investigate the node architecture from the current equipments and propose the node architecture which is equipped in metro-access ring network and have the ability of traffic grooming. The objective of this work is to effectively reduce the new call blocking rate and maximize the utilization efficiency of the wavelength used.

II. Approach, Conclusion and Discussion

II.1. Fair Scheduling Schemes for Non-Real-Time Service in EPON Access Network

i. System Model

➢ **No-pridiction-based**

As shown in Fig. 1, an EPON access network consisting of one OLT and N ONUs is considered herein. The maximum distance between the OLT and ONUs is assumed to be L km. Each ONU can support an aggregated data rate up to RU Mbit/s for all connected users. The link rates between OLT and an ONU in upstream and downstream directions are both assumed to be RN Mbit/s. In each ONU, three independent priority queues are provided to accommodate the real-time voice service, the real-time video service, and the non-real-time data service with priorities P_1, P_2, and P_3 (i.e., $P_1 > P_2 > P_3$), respectively. In OLT, there is a scheduler responsible for scheduling the upstream packet transmissions of all ONUs. The scheduler utilizes the REPORT and GATE messages defined by MPCP [7] to exchange control messages with ONUs. Each ONU periodically reports its queue occupancy via sending a REPORT message to the scheduler. Hence, the scheduler can allocate the granted data volumes for each ONU based on the collected requested information. Finally, OLT will announce the allocation through the GATE messages and thus, each ONU can transmit data based on the allocation.

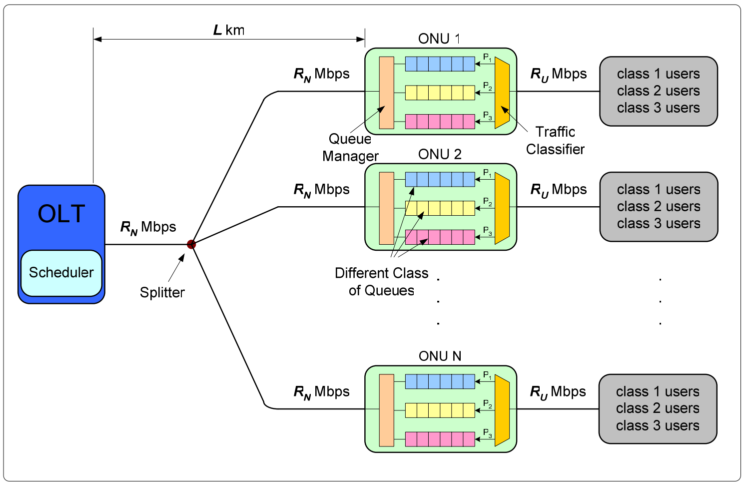


Fig. 1. EPON system architecture

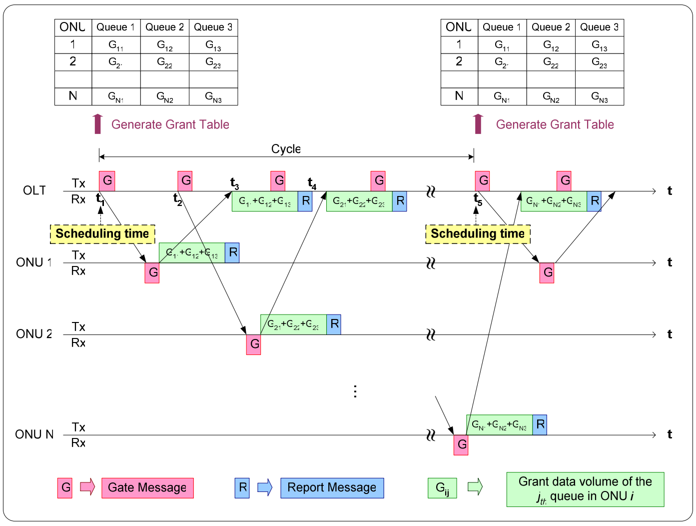


Fig. 2. IPACT-based polling procedure

The interaction between OLT and ONUs is similar to the polling mechanism defined by IPACT. As in IPACT, there is no need to synchronize the ONUs to a common reference clock. Every ONU executes the same procedure driven by the GATE messages sent from OLT. Figure 2 illustrates an example of the polling mechanism. In the example, OLT calculates the granted data volume G_{ij} ($i=1,\ldots,N$, $j=1,\ldots,3$) for the j-th queue of the i-th ONU and then generates the Grant Table at time t_1. G_{ij} is determined based on the proposed scheduling algorithm. At time t_1, the OLT sends a GATE message to ONU 1 carrying G_{11}, G_{12}, and G_{13}. The ONU 1 should immediately transmit packets from the j-th queue up to G_{1j} after receiving the GATE message, and piggybacks a REPORT message notifying OLT the remaining data volume of each queue after transmission. In the meantime, OLT further sends a GATE message to ONU 2 at time t_2 to schedule its transmission. Note that t_2 is properly chosen to ensure that ONU 2's packets will arrive at OLT right after receiving ONU 1's packets (i.e., time t_4) plus a predetermined guard time b (unit: second). As illustrated in Fig. 2, the scheduler in OLT always schedule the packet transmissions of ONUs at the beginning of the each transmission cycle (i.e., t_1, t_5) according to the information carried by all REPORT messages collected during the previous transmission cycle. The transmission cycle is the time interval between adjacent GATE messages sent to ONU 1. The length of each transmission cycle is varied with time since it depends on the granted data volumes for all N ONUs.

- **Prediction-based**

In the prediction-based system model, the EPON network still remain the same as the one for the non-prediction-based EPON network model, but the scheduler is replaced by the prediction-based scheduler, as shown in Figure1. There are still one OLT and N ONUs supported in the prediction-based EPON architecture. Also, each

ONU can support three classes of services, denoted as P_1, P_2, and P_3, by equipping three priority queues.

The proposed prediction-based scheduler architecture is shown in Figure 3. 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 the predictor to estimate 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. The principle of the predictor will be described in the later content. In the meantime, the timestamp information also passes into *Timing Function* to calculate the latest round-trip time (RTT). 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.

Before the beginning of the next cycle, the *Decision Maker* will calculate the timing and the granted data volume of each queue to be transmitted in the next cycle. The input of the decision maker 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.

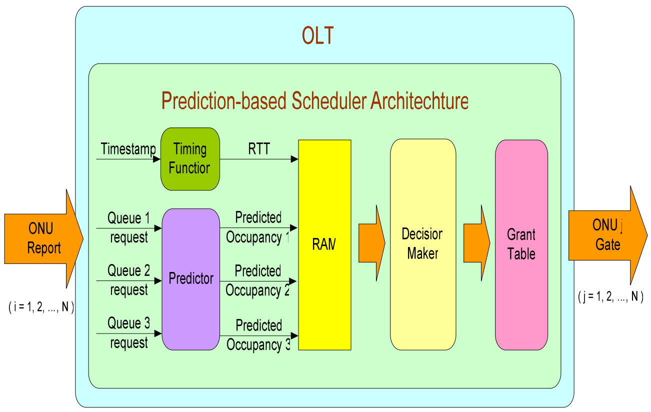


Figure 3: The prediction-based scheduler architecture

ii. **Scheduling Algorithm**

➢ **Hybrid EQL-QLP Scheme and Hybrid LQF-QLP Scheme**

The proposed scheduling algorithm is a gated-based scheme, where the scheduler only utilizes the information obtained before the scheduling time. The information received after scheduling time will be used by the next scheduling operation. The goal of the proposed scheduling algorithms is to meet the delay bound of voice service and, in addition, to simultaneously maintain the fairness of both packet delay and packet blocking probability for the non-real-time data service. The delay bound of voice service is defined as

$$\bar{d} \le d_{\max}, \tag{1}$$

where $\bar{d}$ and $d_{\max}$ are the average delay and the delay bound of voice service, respectively. The overall fairness index, F for non-real-time data service is defined as

$$F = x \cdot I_D + (1-x) \cdot I_B, \tag{2}$$

where x is a weighting factor ranging form 0 to 1; I_D and I_B are the fairness indices for average packet delay and average packet blocking probability, respectively. The I_D and I_B are defined as [13]

$$I_D = \left[\sum_{i=1}^{N} D_i\right]^2 \Bigg/ \left[N \times \sum_{i=1}^{N} \left(D_i\right)^2\right], \tag{3}$$

$$I_B = \left[\sum_{i=1}^{N} P_{B,i}\right]^2 \Bigg/ \left[N \times \sum_{i=1}^{N} \left(P_{B,i}\right)^2\right], \tag{4}$$

where D_i and $P_{B,i}$ are the average packet delay and the average packet blocking probability of ONU i, respectively. Note that I_D, I_B, and F are all in the range from 0 to 1. The parameters used in the proposed scheduling algorithms are first defined and are listed below.

Assume that the reported and the granted data volumes (in bytes) for the the j-th queue of the i-th ONU are denoted as Q_{ij} and G_{ij}, respectively, where $i = 1, \ldots, N$, and $j = 1, 2, 3$. The resource is first allocated to real-time services to guarantee their delay bound. The remaining resource is then fairly allocated to the non-real-time services in each ONU. As in IPACT, the aggregated granted data volumes of real-time services for each ONU may not exceed an upper bound $L_{\max}$, within which, the capacity is first allocated to voice services. Hence, the granted data volumes for voice and video services of ONU *i*, denoted by G_{i1} and G_{i2}, respectively, are given by

$$\begin{cases} G_{i1} = \min(L_{\max}, Q_{i1}), \\ G_{i2} = \min(L_{\max} - G_{i1}, Q_{i2}), \end{cases} \tag{5}$$

where $i = 1, 2, ..., N.$ Given the granted data volumes of real-time services, the residual available capacity that can be used by non-real-time data services, *R*, can be obtained by

$$R = (T_{\max} - N \cdot b) \times R_N - \sum_{i=1}^{N} (G_{i1} + G_{i2}). \tag{6}$$

The allocation of the residual available capacity depends on the chosen scheduling algorithm. A hybrid equal queue length - queue length proportional (hybrid EQL-QLP) scheme and a hybrid longest queue first - queue length proportional (hybrid LQF-QLP) scheme are proposed to allocate the residual available capacity to the non-real-time services of all ONUs. That is, both schemes should decide the granted data volumes of data services for ONU i, denoted by G_{i3}, according to their requested data volumes Q_{i3} and the residual available capacity *R* such that the fairness index *F* is maximized. Details of the two schemes are given below.

A. Hybrid EQL-QLP Scheme

The first proposed scheduling scheme is the hybrid EQL-QLP scheme. This scheme aims to balance the queue occupancies among ONUs such that users in different ONUs will experience the similar packet delay and packet blocking probability. In hybrid EQL-QLP scheme, G_{i3} is given by

$$G_{i3} = \begin{cases} Q_{i3}, & \sum_{j=1}^{N} Q_{j3} \le R, \quad (7) \\ R \times \dfrac{Q_{i3}}{\sum_{j=1}^{N} Q_{j3}}, & \sum_{j=1}^{N} Q_{j3} > R, \max\{Q_{13}, Q_{23}, ...Q_{N3}\} \le Q_{th}, \quad (8) \\ f_a(R, Q_{13}, Q_{23}, ...Q_{N3}), & \sum_{j=1}^{N} Q_{j3} > R, \max\{Q_{13}, Q_{23}, ...Q_{N3}\} > Q_{th}, \quad (9) \end{cases}$$

for $i = 1, 2, ..., N.$ The basic concept of hybrid EQL-QLP scheme is that all requested data volumes are granted if their total requested data volume is less than the residual available capacity *R*, as shown in Eq. (7). In case that the total requested data volume is higher than *R*, either a QLP method or an EQL method will be adopted. The QLP method is adopted if all of the requested volumes are smaller than a given queue length threshold Q_{th}, as indicated in Eq. (8). Otherwise, EQL method is adopted as given in Eq. (9). The hybrid EQL-QLP scheme is formulated as below. Define an index set $K = \{k_1, k_2, ..., k_n\}$, where $K \in \{1, 2, ..., N\}$,

$$Q_{k_j,3} > avg, \; j = 1, 2, ..., n, \tag{10}$$

and

$$avg = \left(\sum_{j=1}^{n} Q_{k_j,3} - R \right) \Big/ n. \tag{11}$$

Then the granted data volumes are given by

$$G_{i3} = f_a(R, Q_{13}, Q_{23}, \ldots, Q_{N3}) = \begin{cases} Q_{i3} - avg, & i \in K, \\ 0, & i \notin K. \end{cases} \tag{12}$$

Equation (12) indicates that OLT does not allocate resource to ONUs that have queue lengths less than avg. The allocated resource to each of the remaining ONUs is proportional to the difference between its queue length and avg.

B. Hybrid LQF-QLP Scheme

Although the EQL method is excellent in balancing the queue sizes of data services, it may result in high blocking probabilities for ONUs with long queue lengths. This problem could be alleviated by adopting a longest queue first (LQF) method to replace EQL. Hence, a hybrid LQF-QLP scheme is further proposed. The hybrid LQF-QLP scheme differs than hybrid EQL-QLP scheme when either one of the requested volumes is greater than Q_{th}. Under this condition, the LQF method is adopted and thus, G_{i3} are determined by

$$G_{i3} = \begin{cases} Q_{i3}, & \sum_{j=1}^{N} Q_{j3} \le R, & (13) \\ R \times \dfrac{Q_{i3}}{\sum_{j=1}^{N} Q_{j3}}, & \sum_{j=1}^{N} Q_{j3} > R, \max\{Q_{13}, Q_{23}, \ldots Q_{N3}\} \le Q_{th}, & (14) \\ f_b(R, Q_{13}, Q_{23}, \ldots Q_{N3}), & \sum_{j=1}^{N} Q_{j3} > R, \max\{Q_{13}, Q_{23}, \ldots Q_{N3}\} > Q_{th}. & (15) \end{cases}$$

In the hybrid LQF-QLP scheme, Eq. (15) is used to replace Eq. (9) defined in the hybrid EQL-QLP scheme when the total requested data volume is higher than R and the requested data volume of at least one ONU is higher than Q_{th}.

Define an integer permutation function $\pi : [1, N] \to [1, N]$, such that $Q_{\pi(1),3} \ge Q_{\pi(2),3} \ge \ldots \ge Q_{\pi(N),3}$. Note that π is an index set with descent order of queue length which is mapped from the original index. Then the granted data volumes of data service queues are defined by

$$\begin{aligned} G_{\pi(i),3} &= f_b(R, Q_{13}, \ldots, Q_{N3}) \\ &= \begin{cases} \min\left[\max\left(Q_{\pi(i),3} - Q_{th}, 0\right), R \right] + R' \cdot P_{\pi(i)}, & i = 1, \\ \min\left[\max\left(Q_{\pi(i),3} - Q_{th}, 0\right), R - \sum_{j=1}^{i-1} G_{\pi(j),3} \right] + R' \cdot P_{\pi(i)}, & i = 2, \ldots, N, \end{cases} \end{aligned} \tag{16}$$

where

$$R' = \max\left(R - \sum_{j=1}^{N} \max\left(Q_{j3} - Q_{th}, 0\right), 0\right), \tag{17}$$

and

$$P_{\pi(i)} = \frac{Q_{\pi(i),3} - \max\left(Q_{\pi(i),3} - Q_{th}, 0\right)}{\sum_{j=1}^{N}\left[Q_{\pi(j),3} - \max\left(Q_{\pi(j),3} - Q_{th}, 0\right)\right]}. \tag{18}$$

An amount of capacity that equals to the difference between its requested data volume and Q_{th} is first allocated to ONU that has the longest queue among the others, as shown in Eq. (16). It helps to reduce the packet blocking probabilities and is beneficial for I_B. The remaining capacity, denoted as R', is the allocated to the other ONUs based on their remaining queue occupancies. The hybrid LQF-QLP scheme may prevent from over assigning resource to greedy ONUs, which is beneficial for I_D.

- **Predictor**

The concept of the prediction method is shown in Figure 3. We see that before the beginning of Cycle (n+1), we must perform 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).

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

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

$$V_{ij}[n] = D_{ij}[n] + Q_{ij}[n] - Q_{ij}[n-1], \tag{19}$$

where i is an integer and ranged from 1 to N, and j is also an integer and 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

$$V'_{ij}[n+1] = \frac{1}{k}\sum_{m=0}^{K-1} V_{ij}[n-m]. \tag{20}$$

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

$$O_{ij}[n+1] = Q_{ij}[n] + V'_{ij}[n+1]. \tag{21}$$

When OLT receives a REPORT message, it will generate the 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.

iii. SIMULATION RESULTS

A. Simulation Environment

Simulations are conducted to verify the effectiveness of the proposed scheduling mechanism. In the simulations, the traffic models presented in [6] are adopted. The constant bit rate (CBR) traffic model is chosen to emulate a T1 link carrying voice packets. Each voice packet is generated every 125 μs and consists of 24 bytes of data, 8 bytes of UDP header, 20 bytes of IP header, and 18 bytes of Ethernet header, which result in a row data rate of 4.48 Mbps. The self-similar traffic model [14] is used to mimic the behavior of video and data services. The parameters used in the simulation are summarized in Table 1.

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 buffer size reserved for each type of service	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 μs
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 1. System parameters used in the simulation

In the simulations, the $d_{\max}$ and $T_{\max}$ are chosen according to the delay bound of voice service specified by ITU G.114 as in [6]. Figure 4 demonstrates an example illustrating the packet arrivals and departures in ONU i. The ONU i reports the amount of data volumes accumulated during the first cycle to the OLT at the beginning of the second cycle. Then, OLT sends a GATE message to the ONU i indicating the granted data volume. Generally, the requested data volume of voice service reported in second cycle can be completely allowed to transmit in third cycle. Hence, the average packet delay of voice service, $\overline{d}$, is given by

$$\begin{aligned}\bar{d} &= E[\text{packet delay in 1st Cycle}] + E[\text{packet delay in 2nd Cycle}] \\ &= 0.5 \text{ cycle time} + 1 \text{ cycle time} = 1.5 \text{ cycle time},\end{aligned} \quad (22)$$

where the voice packets arrive at ONU i in first cycle follows uniform distribution. Thus, the average packet delay would be 0.5 cycle in first cycle. Also, Eq. (22) must satisfy the delay bound defined in Eq. (1), that is

$$\bar{d} \leq 1.5 T_{\max} \leq d_{\max}. \quad (23)$$

In the simulations, $T_{\max}$ was set to be 0.72 ms such that $d_{\max} = 1.02$ ms could be achieved. As in [6], the ONUs were divided into two groups where the loading of the first group is fixed while the loading of the second group is dynamically changed. The average loads of voice and video services are 4.48 Mbps and 15 Mbps for all ONUs, respectively. For data service, the average load is fixed to 15 Mbps for 10 light-loaded ONUs and the average load of the remaining 6 heavy-loaded ONUs are changed from 15 Mbps to 80 Mbps. As a result, the average system load varies form around 600 Mbps to 1000 Mbps.

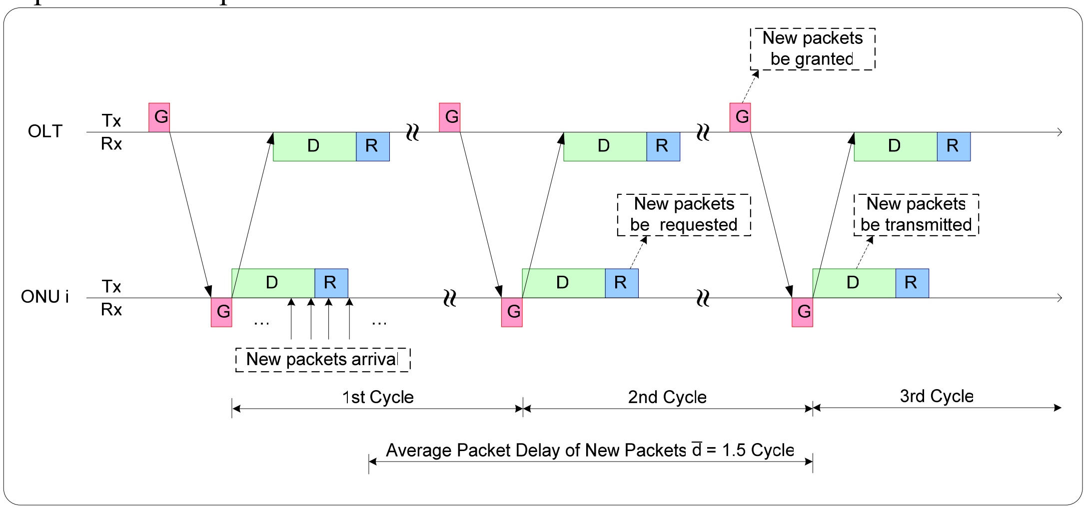


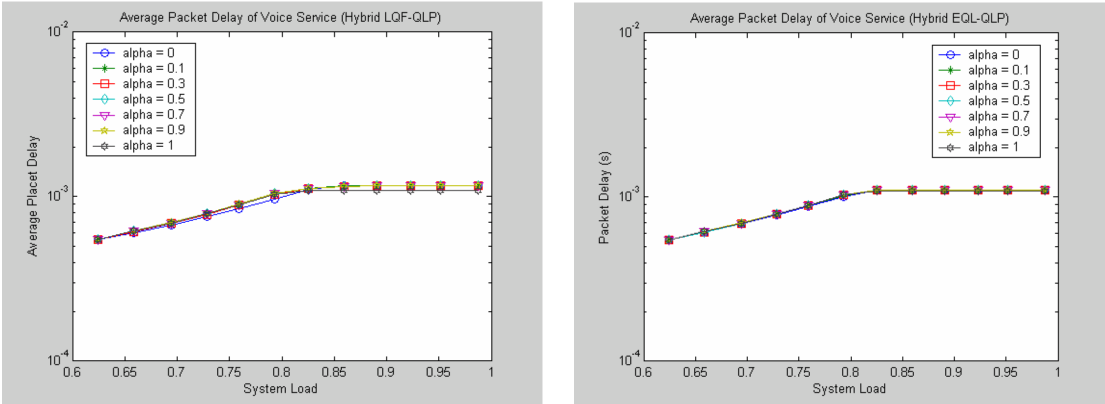
Fig. 4. The average packet delay of voice service

B. Performance of Fairness

The queue length threshold Q_{th} used throughout the simulations is given by

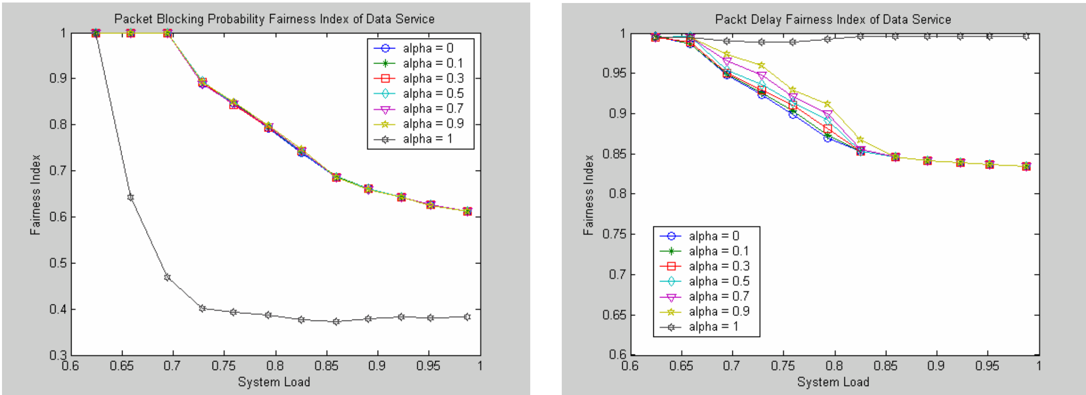
$$Q_{th} = \alpha \cdot Q, \quad (21)$$

where α is a constant ranging from 0 to 1, and Q is the buffer size reserved for each type of service in an ONU. Figures 5(a) and 6(b) show the average packet delay of voice service by adopting the two proposed scheduling algorithms. It can be verified that the delay requirement of the voice services is always guaranteed throughout the simulation for different α and system loads. In the following, the fairness of packet delay and packet blocking probability for data services were investigated.



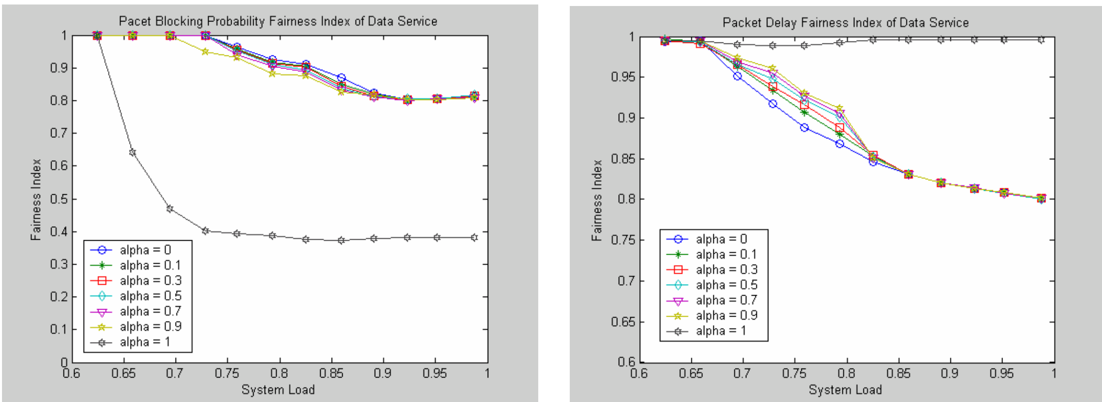
(a) Hybrid LQF-QLP (b) Hybrid QLP-QLP

Fig. 5. Average packet delay of voice service.



(a) packet blocking probability (b) packet delay

Fig. 6. Fairness index of data service by adopting hybrid EQL-QLP scheme



(a) packet blocking probability (b) packet delay

Fig. 7. Fairness index of of data service by adopting hybrid LQF-QLP scheme

Figures 6(a) and 6(b) show the fairness indices in packet blocking probability and in packet delay for data service by adopting hybrid EQL-QLP scheme, respectively. Note that $\alpha = 0$ and $\alpha = 1$ are equivalent to EQL and QLP scheme, respectively. In these two figures, we see the effect while adjusting the constant α from 0 to 1. It is found in Fig. 5(a) that the fairness index of packet blocking probability in EQL scheme is higher than that in QLP scheme. It is because EQL scheme allocates more resource to heavy-loaded ONUs which results in similar packet blocking probabilities in all ONUs. The change of fairness in packet blocking probability is not obvious when α moves from 0 to 1 because the resource assigned to each ONU is independent to α, as show in Eq. (12). In contrast, as illustrated in Fig. 5(b), QLP scheme provides a better fairness in packet delay because the resource is allocated according to the proportion of data volumes requested by each ONU.

Figures 7(a) and 7(b) show the fairness index in packet blocking probability and in packet delay for data service by adopting hybrid LQF-QLP scheme, respectively. Similarly, $\alpha = 0$ and $\alpha = 1$ are equivalent to LQF and QLP scheme, respectively. The effect of adjusting α from 0 to 1 in Fig. 7 is similar to the one shown in Fig. 6. However, there are differences between them. First, the fairness index in packet blocking probability is more sensitive to the constant α by adopting hybrid LQF-QLP scheme, due to the relationship between the resource assignment and queue length threshold, as described in Eq. (16). Second, LQF scheme favors the fairness of packet blocking probability but EQL scheme prefers the fairness of packet delay. It is because that EQL scheme tries to balance the queue occupancies among all data queues while LQF scheme always reserves resource to heavy-loaded ONUs. Therefore, the light-loaded ONUs has more opportunities to get the resource if EQL scheme is adopted. That is the reason why EQL scheme constitutes comparable packet delays and LQF scheme provides a similar packet blocking probabilities.

System Load	Overall Fairness Index $F = 0.5\, I_D + 0.5\, I_B$		
	QLP	Hybrid EQL-QLP (α = 0.9)	EQL
Low	0.99	0.99	0.99
	0.81	0.99	0.99
	0.78	0.98	0.96
↓	0.75	0.92	0.90
	0.71	0.88	0.86
↓	0.69	0.85	0.82
	0.68	0.80	0.78
	0.68	0.76	0.75
↓	0.68	0.75	0.75
	0.68	0.74	0.74
	0.68	0.73	0.73
High	0.68	0.72	0.72

(a)

System Load	Overall Fairness Index $F = 0.5\, I_D + 0.5\, I_B$		
	QLP	Hybrid LQF-QLP (α = 0.7)	LQF
Low	0.99	0.99	0.99
	0.81	0.99	0.99
	0.78	0.98	0.97
↓	0.75	0.97	0.95
	0.71	0.94	0.91
↓	0.69	0.92	0.89
	0.68	0.89	0.87
	0.68	0.86	0.85
↓	0.68	0.83	0.82
	0.68	0.81	0.80
	0.68	0.80	0.80
High	0.68	0.80	0.80

(b)

Table 2. The overall fairness index by adopting (a) hybrid EQL-QLP scheme (b) hybrid LQF-QLP scheme

Table 2 shows the overall fairness index of the two proposed hybrid schemes (i.e., which is defined in Eq. (2) with $x = 0.5$). Table 2(a) shows the overall fairness index of hybrid EQL-QLP scheme, which is compared with QLP ($\alpha = 1$) and EQL ($\alpha = 0$) schemes. It is found that hybrid EQL-QLP with $\alpha = 0.9$ results in the best overall fairness index. Table 2(b) shows the overall fairness index of hybrid LQF-QLP scheme, in which $\alpha = 0.7$ provides the best overall fairness index.

In general, it can be found in Table 2 that the overall fairness index of the proposed hybrid schemes is better than that of QLP, LQF, or EQL scheme. The reason is that the proposed hybrid schemes simultaneously consider the packet delay and packet blocking probability. It is also found in Table 2 that hybrid LQF-QLP scheme performs slightly better than hybrid EQL-QLP scheme because the increment of fairness in packet blocking probability is faster than the decrement of fairness in packet delay if hybrid LQF-QLP scheme is adopted.

C. Performance of Fairness

In the simulation, the queue length threshold Q_{th} is chosen to be $Q_{th} = 0.7$ for the Hybrid PLQF-PQLP scheme and $Q_{th} = 0.9$ for the Hybrid PEQL-PQLP scheme, respectively, according to the best performance of each algorithm. 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 8 shows the system throughput of these two proposed scheduling schemes. For each scheme, the prediction-based scheme is compared with the non-prediction-based scheme. We can see that the system throughput in prediction-based schemes is a little bit worse than the 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.

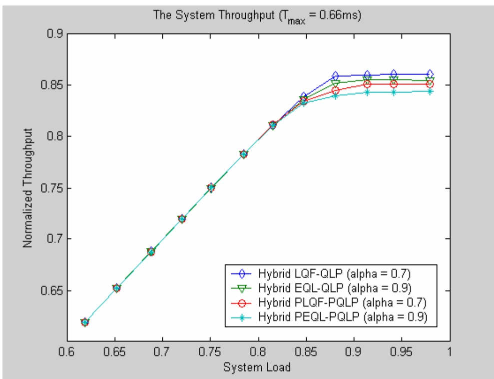


Figure 8: The system throughput (T_{max} = 0.66ms)

And then, the difference of the average packet queueing delay of voice service between the prediction-based scheme and the non-prediction-based scheme is shown in Figure 9. The result shows that the average packet queueing delay in prediction-based scheme can be reduced more, because the average packet queueing delay of the prediction-based scheme is half of the cycle, whereas the average packet queueing delay is one and half of the cycle in non-prediction-based scheme.

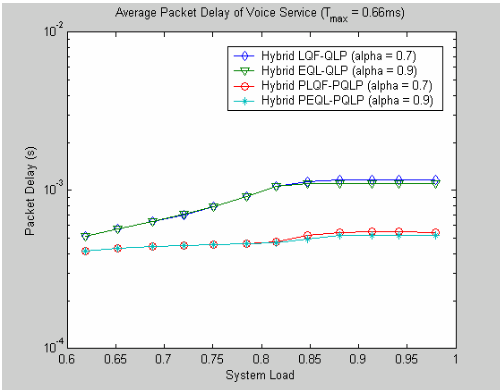


Figure 9: Average packet delay of Voice service (T_{max} = 0.66ms)

In Figure 10, the result shows that the fairness of 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 11, the results show that the fairness of packet queueing delay 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 fairness performance of the 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 an overall fairness index, we can find that the fairness performance is better in the prediction-based environment than in the non-prediction-based environment.

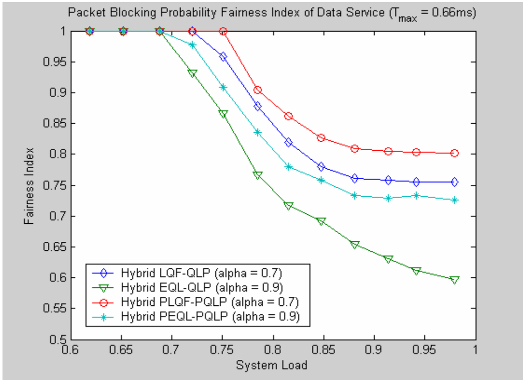


Figure 10: Packet blocking probability fairness index of data service (T_{max} = 0.66ms)

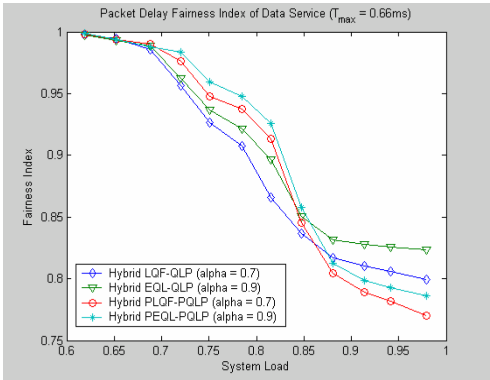


Figure 11: Packet delay fairness index of data service (T_{max} = 0.66ms)

According to the 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}) under high-loading conditions. However, the queueing delay for the voice service will also increase with T_{max} in the meanwhile. As a result, with the lower queueing delay as shown in Figure 9, the maximum cycle time can be carefully extended to improve the performance of system throughput while the increasing queueing delay for the voice service can still satisfy the delay criterion.

Now we extend the maximum cycle time T_{max} from 0.66ms to 2ms, and evaluate the performance difference compared with the T_{max} = 0.66ms. The simulation results are illustrated in Figure 12 and Figure 13. In Figure 12, we can see that for both of the proposed prediction-based scheduling algorithms, the average packet queueing delay indeed increases with the T_{max}, but still maintains lower than the specified delay bound (1.5ms). Meanwhile, as the result shown in Figure 13 where the performance of system throughput is observed, 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 a longer value than the non-prediction-based schemes and thus the percentage of the overhead, such as control messages and guard, will be smaller. Therefore, the data transmission throughput can be greatly improved under high-loading conditions. Besides, 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 effect of packet truncation error is smaller in the (P)LQF-(P)QLP scheme than in the (P)EQL-(P)QLP scheme. Furthermore, we also compare the performance with limited service, which is proposed in [6]. 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 theses 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.

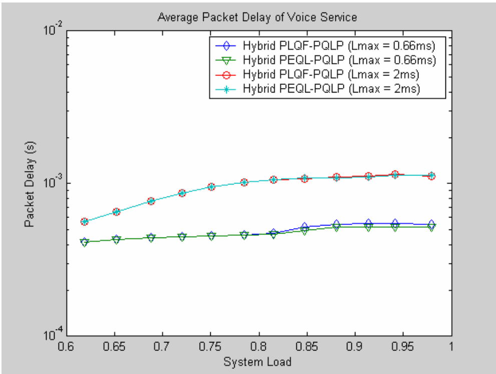


Figure 12: Average packet delay of voice service

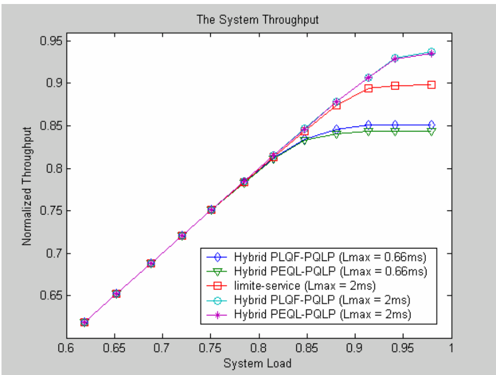


Figure 13: The System Throughput

II.2. QoS-Promoted Dynamic Bandwidth Allocation (Q-DBA) for Ethernet Passive Optical Networks

i. System Model

A. System Architecture

An EPON architecture is shown in Fig. 14. It is a point-to-multipoint configuration following the multi-point control protocol (MPCP). There are an (M) OLT (ONUs) with line rate R_u (R_E) (bps) between OLT (ONU) and each ONU (its own end users) to serve downstream (upstream) traffic. The A splitter is used to broadcast packets from OLT to all ONUs and each OUN takes its own packets. Each OUN transmits packets to OLT bytes by bytes until the end of its timeslot. The guard time is used to distinguish packets from the different ONUs. The total upstream bandwidth is shared by all ONUs' real-time and non-real-time packets. A Q-DBA scheme processes ONUs' demands and determines the timeslot as well as the volume of bandwidth for ONUs to transmit packets.

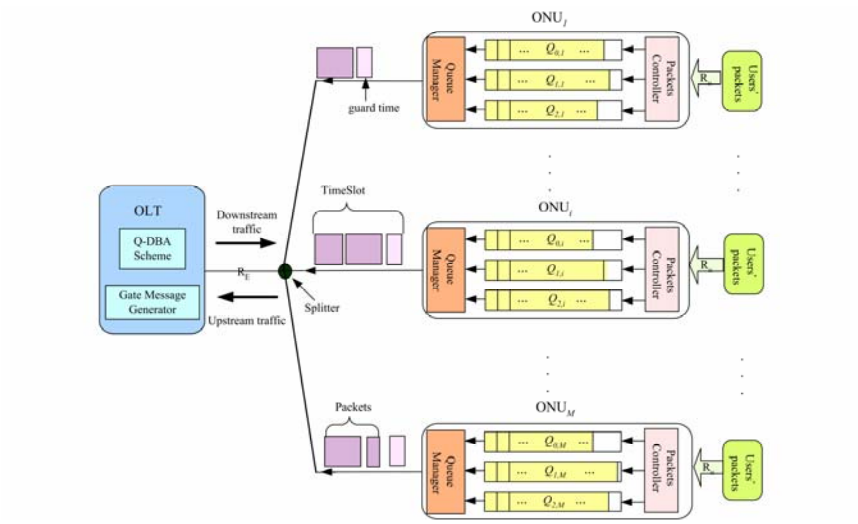


Figure 14: The system architecture of EPON

In $\mathrm{ONU}_i, 1 \le i \le M$, three classes of queues, $Q_{0,i}$, $Q_{1,i}$, and $Q_{2,i}$, are provided to store real-time voice, real-time video, and non-real-time data packets and their queue size are denoted as $|Q_{0,i}|$, $|Q_{1,i}|$, and $|Q_{2,i}|$, respectively. The packets controller puts them into the different queue depending on their types when packets arrive at the ONU, and it drops some type of packets if their queue is full. Also, the packets controller drops real-time packets as well if their delay criteria is violated. The queue manager takes charge of generating the report message and the transmission between OLT and the ONUs, transmits packets to the OLT at its granted timeslot, and receives its packets.

The report and gate messages are exchanged between OLT and ONUs. The report message can transmit eight different queue conditions of $\mathrm{ONU}_i, 1 \le i \le M$, to OLT, and the gate message can provide four different granted results to the $\mathrm{ONU}_i, 1 \le i \le M$. The queue manager in $\mathrm{ONU}_i, 1 \le i \le M$, generates the report message based on the conditions of their queues which are required for bandwidth allocation and presented in the form of the total amount of packets' size in bytes. The OLT receives all report messages, and they are processed by the Q-DBA scheme. The Q-DBA scheme, according report message and bandwidth allocation mechanism, determines the bandwidth allocation, and then, the gate message generator issues gate messages to all ONUs. When ONUs receive gate messages, they transmit packets at their assigned timeslot.

B. Source Model

Three kinds of packets: real-time voice, real-time video, and non-real-time data, are considered. Voice (data) packets are classified as the highest (lowest) priority. The two-level Markov modulated deterministic process (MMDP) is generally used to formulate ON-OFF voice traffic stream. To emulate a T_1 connection, the generation rate of a voice packet is constant bit rate (CBR) and the packet size is 24 bytes in T_1. By adding the overhead such Ethernet, UDP and IP headers in a packet, the packet results in a 70-byte frame. Video and data traffic are to emulate variable bit rate (VBR) that exhibit properties of self-similarity and long-rang-dependence (LRD) [33, 34]. The packet size is in uniform distribution and ranges from 64 to 1518 bytes. R_v (bps) is the arrival rate of video packets to ONUs, and R_d (bps) is the arrival rate of data packets.

ii. Q-DBA METHOD

Notation	Definition
$Q_{h,i}$	The hth type queue in the ith ONU, $0 \leq h \leq 2$ and $1 \leq i \leq M$.
$L_{h,i}$	The occupancy of queues $Q_{h,i}$, $0 \leq h \leq 2$ and $1 \leq i \leq M$.
$L_{dp,i}$	The total amount of video packets which will be dropped at the next cycle end if they are not transmitted (due to violating T_d^*).
$L_{d,i}$	The amount of video packets which should be transmitted at the next timeslot in order to sustain P_d^*.
$L_{w,i}$	the amount of data packets which is with a waiting time larger than T_w^*.
T_d^*	The video's delay constraint.
P_d^*	The requirement of video packet dropping probability.
T_w^*	A waiting bound.
T_{cycle}	The cycle time of EPONs.
$T_{d,n}$	The delay time of the nth packet.
$S_{n,h}$	The number of bytes of the nth packet's size in $Q_{h,i}$, $0 \leq h \leq 2$ and $1 \leq i \leq M$.
$P_{d,i}$	The dropping probability of video packets at $Q_{1,i}$ measured by ONU$_i$, $1 \leq i \leq M$.
$G_{h,i}$	The granted bandwidth for $Q_{h,i}$ of ONU$_i$, $0 \leq h \leq 2$ and $1 \leq i \leq M$.
$G_{0,i}^{'}$	The allocated bandwidth to voice packets in $Q_{0,i}$, $1 \leq i \leq M$.
$G_{1,i}^{'}$	The allocated bandwidth to video packets with the second and third priorities in $Q_{1,i}$, $1 \leq i \leq M$.
$G_{2,i}^{'}$	The allocated bandwidth to data packets with the fourth priority in $Q_{2,i}$, $1 \leq i \leq M$.
$G_{1,i}^{''}$	The allocated bandwidth to video packets with the fifth priority in $Q_{1,i}$, $1 \leq i \leq M$.
$G_{2,i}^{''}$	The allocated bandwidth to data packets with the sixth priority in $Q_{2,i}$, $1 \leq i \leq M$.
$G_{0,i}^{''}$	The proportional allocated bandwidth to voice packets in $Q_{0,i}$ with $G_{1,i}^{'''}$ at the same time under the six priorities have been allocated, $1 \leq i \leq M$.
$G_{1,i}^{'''}$	The proportional allocated bandwidth to video packets in $Q_{1,i}$ with $G_{0,i}^{''}$ at the same time under the six priorities have been allocated, $1 \leq i \leq M$.
$G_{2,i}^{'''}$	The proportional allocated bandwidth to video packets in $Q_{2,i}$ under $G_{0,i}^{''}$ and $G_{1,i}^{'''}$ are allocated, $1 \leq i \leq M$.
B	The total bandwidth of the fiber link.

Table 3. THE NOTATIONS OF Q-DBA METHOD

Table 3 defines the notations of QoS-guaranteed dynamic bandwidth allocation (Q-DBA) method in this paper. The Q-DBA assumes that $\mathrm{ONU}_i, 1 \leq i \leq M$, sends the report message including six kinds of information of queues, $L_{0,i}$, $L_{1,i}$, $L_{2,i}$, $L_{dp,i}$, $L_{d,i}$, and $L_{w,i}$, where $L_{dp,i}$ and $L_{d,i}$, and $L_{w,i}$ are used to ensure the QoS guarantee, just like p_i [35] and δ [36] are used according to the conception of the "earliest deadline first" (EDF). For video packets in $Q_{1,i}, 1 \leq i \leq M$, at the present time, the $n=1$ means the first packet in $Q_{1,i}$, and the packet will be transmitted firstly according the FCFS service principle. Denote x to be the xth packet with the least delay time which will violate the delay criterion at the next cycle end. It can be calculated as follows and any video packet queued before it will be dropped at the next cycle end. Then the $L_{dp,i}$ can be obtained as follows.

$$x=\arg\min_{n}\{T_{d,n}+T_{cycle},\forall n,\text{and }T_{d,n}+T_{cycle}>T_d^*\},\text{ and}L_{dp,i}=\sum_{n=1}^{x}S_{n,1}. \tag{22}$$

A moving time window of observation, introduced to calculate the dropping probability, contains the latest N output video packets which have been dropped and transmitted, or are going to be dropped or transmitted at the next cycle in ONU_i. Among the N video packets, assume there are N_d video packets which have been dropped so far, and there are x video packets which are waiting in the queue, $Q_{1,i}$ and will be dropped if they are not transmitted at the next timeslot. A number of packets among the x packets, denoted it by y, must be transmitted otherwise P_d^* will be violated. The y and $L_{d,i}$ can be obtained as follows.

$$y=(N_d+x-\lceil N\times P_d^*\rceil)^+,\text{ and}L_{d,i}=\sum_{n=1}^{y}S_{n,1}, \tag{23}$$

where $(a)^+=a$ if $a\geq 0$, $(a)^+=0$ if $a<0$; and $\lceil b\rceil$ denotes the smallest integer greater than b. The $L_{w,i}$ is used to tell OLT how much bandwidth is required for data packets in $Q_{2,i}$. A number of packets with a waiting time larger than T_w^*, denoted by k ,had better be served, and it is given and the $L_{w,i}$ can be obtained as follows

$$k=\arg\min_{n}\{T_{w,n}-T_w^*,\forall n,\text{and }T_{w,n}-T_w^*>0\},\text{ and}L_{w,i}=\sum_{n=1}^{k}S_{n,2}. \tag{24}$$

The OLT determines the bandwidth allocation when receiving all *report* messages.

The allocated bandwidths are individually sent to each ONU by the *gate* message which includes $G_{0,i}$, $G_{1,i}$, and $G_{2,i}$. The gate message tells ONU_i the amount bandwidth which it can transmit. As soon as ONUs receive their gate message, they prepare to transmit the packets in themselves time slot.

The Q-DBA method allocates bandwidth in unit of bytes to all ONUs from the service of the highest priority to that of the lowest one successively until all the bandwidths are used up. In order to guarantee QoS requirements (the Q-DBA classifies the packets of all services into six priorities, voice packets, video packets with problem of delay constraint, video packets with problem of dropping probability constraint, data packets considering waiting bound, video packets, and data packets from the highest to the lowest in order) and they are defined as $L_{0,i}$, $L_{dp,i}$, $L_{d,i}$, $L_{w,i}$, $L_{1,i}$, and $L_{2,i}$ in order. The voice packets are first allocated the bandwidth and the video packets which will be dropped if they are not transmitted at the next timeslot are secondly allocated. Next, the Q-DBA continues to allocate the bandwidth to the data packets whose waiting times exceed the waiting bound. Data packets do not be dropped but the starvation, means the data packets wait for a long time in the queue, may occur. To avoid the starvation, the Q-DBA raises the priority of data packets (whose waiting bound are violated) higher than the video packets with a non-violating delay time. It is due to the video packets which will be dropped have already been processed, and the QoS requirement of residual video packets is satisfied. Also, the Q-DBA allocates the bandwidth to the unallocated video and data packets in order. Finally, the voice and video packets proportionally share the residual bandwidth based on their queue occupancy to use up the bandwidth and guarantee QoS in further.

The Q-DBA method is described in detail as follows:

Step 1: [*Bandwidth Allocation to Voice*]

The allocated bandwidth to voice, $G_{0,i}^{'}$, is given by

$$G_{0,i}^{'} = \begin{cases} L_{0,i}, & \text{if } \sum_{i=1}^{M} L_{0,i} \le B, \\ B \times \dfrac{L_{0,i}}{\sum_{i=1}^{M} L_{0,i}}, & \text{elsewhere.} \end{cases} \tag{25}$$

Step 2: [*Bandwidth Allocation to Voice Packets with the second and the third priorities*]

After $G_{0,i}^{'}$ allocation, the allocated bandwidth to video packets with the second and the third priorities, $G_{1,i}^{'}$, is given by

$$G_{1,i}^{'} = \begin{cases} L_{dp,i}, & \text{if } B - \sum_{i=1}^{M} G_{0,i}^{'} \geq \sum_{i=1}^{M} L_{dp,i}, \\ L_{d,i} + (B - \sum_{i=1}^{M} [G_{0,i}^{'} + L_{d,i}])^{+} \times \dfrac{L_{dp,i} - L_{d,i}}{\sum_{i=1}^{M} (L_{dp,i} - L_{d,i})}, & \text{if } \sum_{i=1}^{M} L_{d,i} < B - \sum_{i=1}^{M} G_{0,i}^{'} < \sum_{i=1}^{M} L_{dp,i}, \\ (B - \sum_{i=1}^{M} G_{0,i}^{'})^{+} \times \dfrac{L_{d,i}}{\sum_{i=1}^{M} L_{d,i}}, & \text{if } B - \sum_{i=1}^{M} G_{0,i}^{'} \leq \sum_{i=1}^{M} L_{d,i}. \end{cases} \tag{26}$$

Step 3 : [*Bandwidth Allocation to Data Packets with the fourth priority*]

Next, the allocated bandwidth to data packets considering waiting bound, $G_{2,i}^{'}$, is given by

$$G_{2,i}^{'} = \begin{cases} L_{w,i}, & \text{if } B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'}] > \sum_{i=1}^{M} L_{w,i}, \\ (B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'}])^{+} \times \dfrac{L_{w,i}}{\sum_{i=1}^{M} L_{w,i}}, & \text{elsewhere.} \end{cases} \tag{27}$$

Step 4 : [*Bandwidth Allocation to video packets with the fifth priority*]

Based on the amount of unallocated video packets, $L_{1,i} - L_{dp,i}$, and the residual bandwidth of the fiber link, the $G_{1,i}^{''}$ is given by

$$G_{1,i}^{''} = \begin{cases} L_{1,i} - G_{1,i}^{'}, & \text{if } B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'} + G_{2,i}^{'}] > \sum_{i=1}^{M} L_{1,i} - G_{1,i}^{'}, \\ (B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'} + G_{2,i}^{'}])^{+} \times \dfrac{L_{1,i} - G_{1,i}^{'}}{\sum_{i=1}^{M} (L_{1,i} - G_{1,i}^{'})}, & \text{elsewhere.} \end{cases} \tag{28}$$

Step 5 : [*Bandwidth Allocation to Data Packets with the sixth priority*]

Considering the amount of unallocated data packets, $L_{2,i} - L_{w,i}$, and the residual bandwidth, the $G_{2,i}^{''}$ is given by

$$G_{2,i}^{''} = \begin{cases} L_{2,i} - G_{2,i}^{'}, & \text{if } B - \sum_{i=1}^{M} (G_{0,i}^{'} + G_{1,i}^{'} + G_{2,i}^{'} + G_{1,i}^{''}) > \sum_{i=1}^{M} (L_{2,i} - G_{2,i}^{'}), \\ (B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'} + G_{1,i}^{''} + G_{2,i}^{'}])^{+} \times \dfrac{L_{2,i} - G_{2,i}^{'}}{\sum_{i=1}^{M} (L_{2,i} - G_{2,i}^{'})}, & \text{elsewhere.} \end{cases} \tag{29}$$

Step 6 : [*Residual Bandwidth Allocation*]

Finally, according to $L_{0,i}$, $L_{1,i}$, and the residual bandwidth, the $G_{0,i}^{''}$ and $G_{1,i}^{'''}$ are given by

$$\begin{cases} G_{0,i}^{''} = (B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'} + G_{1,i}^{''} + G_{2,i}^{'} + G_{2,i}^{''}])^{+} \times \dfrac{L_{0,i}}{\sum_{i=1}^{M} (L_{0,i} + L_{1,i})}, \\ \\ G_{1,i}^{'''} = (B - \sum_{i=1}^{M} [G_{0,i}^{'} + G_{1,i}^{'} + G_{1,i}^{''} + G_{2,i}^{'} + G_{2,i}^{''}])^{+} \times \dfrac{L_{1,i}}{\sum_{i=1}^{M} (L_{0,i} + L_{1,i})}. \end{cases} \tag{30}$$

Step 7 : [*Gate Message Generation*]

The granted bandwidths , $G_{0,i}$, $G_{1,i}$, and $G_{2,i}$ are given by

$$\begin{cases} G_{0,i} = G_{0,i}^{'} + G_{0,i}^{''}, \\ G_{1,i} = G_{1,i}^{'} + G_{1,i}^{''} + G_{1,i}^{'''}, \\ G_{2,i} = G_{2,i}^{'} + G_{2,i}^{''} + G_{2,i}^{'''}. \end{cases} \tag{31}$$

The $G_{0,i}$, $G_{1,i}$, and $G_{2,i}$ are included in the *gate* message which is sent by the OLT to the ONUs. When the ONUs receive the gate message, they transmit packets combining OLT's grants and queue conditions at that time. If there is some bandwidth left at some queues, the left can be reallocated to other queues in the same ONU. It not only decreases delay and dropping probability but also uses bandwidth more efficiently.

iii. Simulation Results and Discussions

A. Simulation Environment

An event-driven packet-based simulation is developed the performance of the proposed Q-DBA and the DBAM [16]. We consider an EPON architecture with 32 OUNs connected in a tree topology. The length of the fiber between the OLT (each ONU) and the splitter is 20 (5) km, and the line rate is 1 Gbps (100 Mbps). The cycle time is set to 0.72 ms. Each ONU supports three priority queues with the same buffering space of size 1 Mbytes. The guard time separating two consecutive transmission windows is set to 1 μs.

Three kinds of traffic, voice, video, and data, are considered in the system. The voice traffic is transmitted with the highest priority, and is generated by a two-level MMDP. To emulate T_1 connection, in a ONU, there are 24 channels in a T_1 link. The ONU aggregates the traffic of each channel. During ON state, the generate rate is decided by the number of channels which are at ON state. The mean durations of talk spurts and silence periods are assumed to be exponentially distributed with $1/\alpha = 1$ sec. and $1/\beta = 1.35$ sec., respectively. The packet size is fixed to 70 bytes, and the generation rate is constant bit rate (CBR), during ON state (talk spurts). On the other hand, the video and data packets are modeled by ON/OFF Parato-distributed model, to generate the highly bursty video and data packets, and packet sizes are uniformly distributed between 64 and 1518 bytes. The voice delay bound is 1.5 ms as one way propagation delay based on the specification. International Telecommunication Standardization Sector (ITU-T) Recommendation G.114. The video delay bound is defined 10 according to the ITU-TS [37]. The requirement (delay bound) of video dropping probability (data packets) is defined 1% (500 ms).

B. Fairness Index

The fairness index is used to see the fairness of average delay, blocking probability, and dropping probability within all ONUs. If the fairness index equals to 1, it means all ONUs have almost the same simulation results of data packets; if the fairness index equals to 0, it means the difference among all ONUs is too large. The overall fairness index, F_1 (F_2) for video (data) packets is defined the same as in [38].

C. Simulation Results

We show the performance of Q-DBA method. The traffic arrival rates are set as follows:

Type	Case one	Case two
Voice service	4.48Mbps × 32(iid)	4.48Mbps × 32(iid)
Video service	0.55Mbps × 32(iid) ~ 15.75Mbps × 32 (iid)	0.55Mbps × 20 (iid) +0.41Mbps × 12 (iid) ~ 15.75Mbps × 20 (iid) +11.31Mbps × 12 (iid)
Data service	0.28Mbps × 32 (iid) ~ 7.27Mbps × 32 (iid)	0.28Mbps × 20 (iid) +0.41Mbps × 12 (iid) ~ 7.87Mbps × 20 (iid) +11.31Mbps × 12 (iid)

Fig. 15 illustrates the average video dropping probability versus system load in EPON. The dropping probability in Q-DBA is zero due to the video packets with the problem of delay requirement are served earlier by raising their priorities when the system load is less than 0.8, but, the dropping probability exceeds the video dropping probability requirement due to the limited capacity of fiber link when the system load is larger than 0.8. Unfortunately, it also can be found that the dropping probability in DBAM cannot be guaranteed due to the maximum window which is used to allocate the bandwidth of video packets cannot totally support the burst arrival. Besides, when the system load is 0.6, both in case one and case two, the performances in Q-DBA are improved 100%.

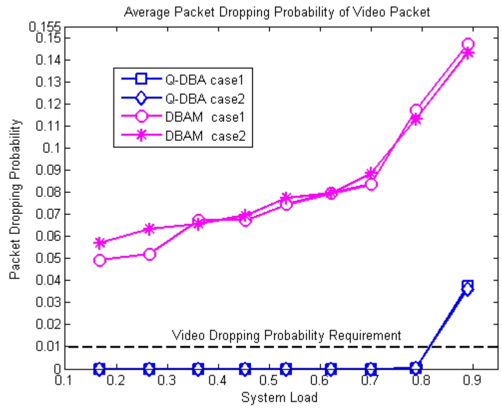


Figure 15. Average video dropping probability versus system load in EPON

Fig. 16 illustrates the blocking probability of data packets versus system load in EPON. The blocking probabilities both in Q-DBA and DBAM are zero when the traffic load is less than 0.8. However, when the system load is larger than 0.8, the blocking probability in DBAM increases greatly than that in Q-DBA. It is due to that the data burst arrival cannot be served within a short time in DBAM and the allocated bandwidth is not limited and the data packets which reach the waiting bound would be sent early in Q-DBA. It is shown that when the system load is 0.8, the blocking probability in Q-DBA is improved 100% whether in case one or case two.

Fig. 17 illustrates the starvation ratio of data packets versus traffic load in EPON. It is shown that when the system load is less than 0.7, the starvation in both Q-DBA and DBAM do not occur, but, when the system load is larger than 0.9, the starvation occurs due to that the burst arrival of data packets may not be instantly supported as both video and data packets arrival rate increases. However, the starvation ratio is still 60% improved by Q-DBA in case one, and it also 60% improved by Q-DBA in case two due to raising the priority of data packets which will reach their waiting bound in Q-DBA. Therefore, the starvation in DBAM happens earlier than that in Q-DBA.

Fig. 18 shows the average voice delay time versus system load in EPON. It can be seen that the average voice delay in DBAM increases with the increasing of system load and one in Q-DBA does not increase obviously. It is due to the Q-DBA allocates total bandwidth to all ONUs by the step of residual bandwidth allocation, and the voice packets have sufficient bandwidth to transmit the reported voice packets and new arrival voice packets. However, when system load is larger than 0.8, the average voice delay in DBAM is less than that in Q-DBA. Since the report message in DBAM includes prediction, the requirement can be satisfied as much as possible. Under this circumstances, the voice delay in Q-DBA is large than DBAM, but the voice delay requirement and the dropping probability are still satisfied.

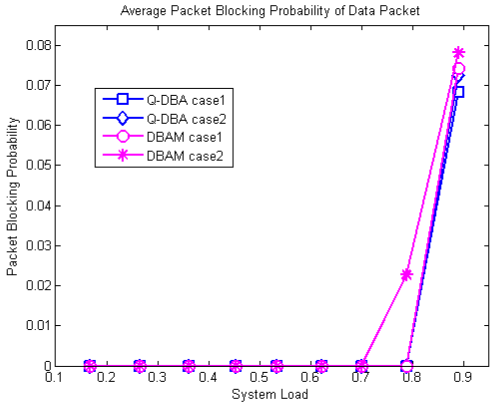


Figure 16. Average data data blocking probability versus system load in EPON

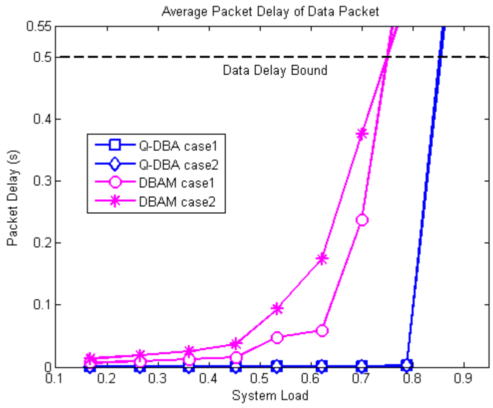


Figure 17. Average data starvation ratio versus system load in EPON

Fig. 19 shows the average video delay time versus system load in EPON. The average video delay in Q-DBA is far from the video delay requirement when the system load is below 0.8, whereas, one is close to the video delay requirement when the system load exceeds 0.8. However, the average video delay in DBAM increases almost smoothly with the increasing of system load due to the maximum window and the prediction. Furthermore, the average video delay is in the case one (two) of Q-DBA is improved 78% (77%) due to the video packets with a higher priority than data packets. Besides, the video packet with the problem of delay requirement could be transmitted with a higher priority. Thus, the delay can be decreased and the dropping probability can be also satisfied even if the burst arrival occurs.

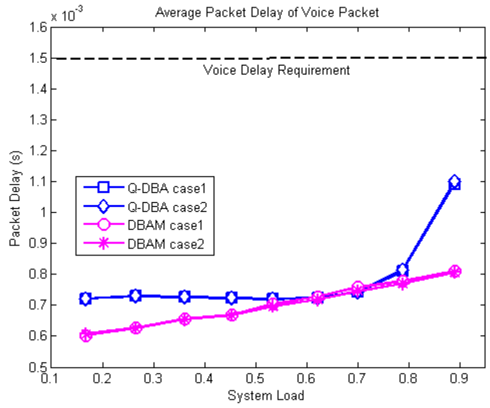


Figure 18 Average voice delay time versus system load in EPON

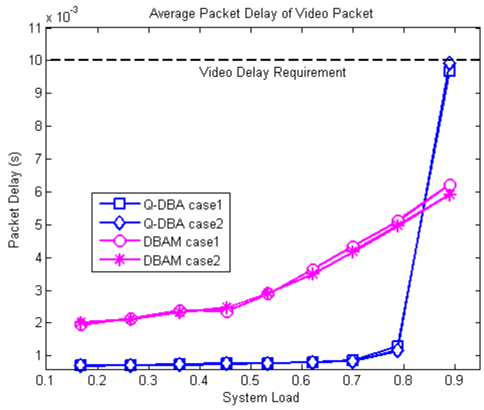


Figure 19 Average video delay time versus system load in EPON

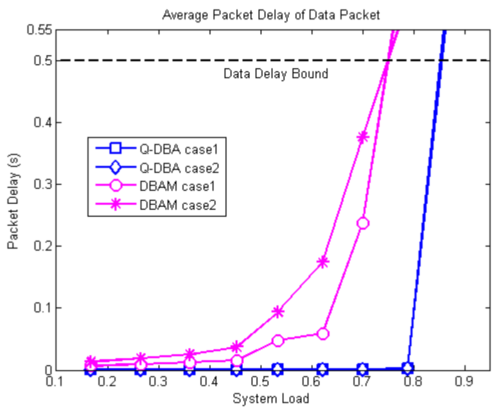


Figure 20 Average data delay time versus system load in EPON

Fig. 20 shows the average packet delay time versus system load in EPON. The average data delay in case one is less than that in case two due to the less data arrival rate. Since the Q-DBA allocates bandwidth based on the requirement of queue occupancy and the burst arrival can be transmitted more easily, and thus the delay of data packets can be small. Besides, due to the Q-DBA considers the condition of waiting bound, the data packets dose not violate the delay bound in Q-DBA as early as that in DBAM does. Also, Q-DBA improves the average data delay of the case one (two) by 98% (99%)when the system load is 0.6.

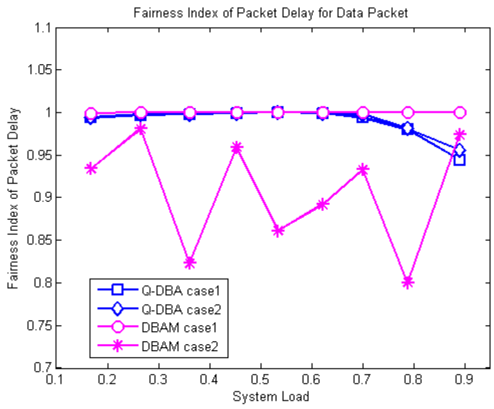


Figure 21 Fairness index of average data delay versus system load in EPON

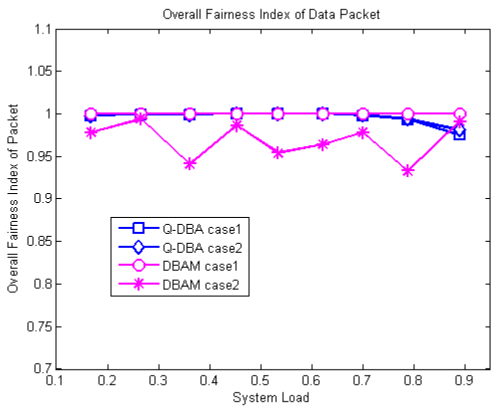


Figure 22 Overall fairness index of data packets versus system load in EPON

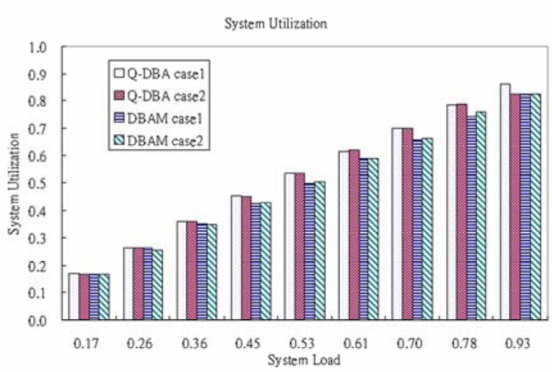


Figure 23 System utilization versus system load in EPON

Fig. 21 illustrates the fairness index of average data delay versus system load in EPON. It can be found that the fairness index of average data delay in Q-DBA and DBAM is bear to 1. It is due to the Q-DBA consider all ONUs' condition to allocated bandwidth. Furthermore, in DBAM, all ONUs have their maximum window to transmit packets according to the SLA. In this way, the fairness index of average data delay whether in Q-DBA or in DBAM is close to zero. It also can be found that the fairness index of average data delay in case two of DBAM varies greatly because the packet arrival rates are not the same, and the maximum window cannot suitably support the different traffic load.

Fig. 22 illustrates the overall fairness index of data packets versus system load in EPON. The fairness index of data packets in both Q-DBA and DBAM is close to 1. due to the DBAM uses a maximum window to allocate packets according to the SLA, and the Q-DBA allocates packets based on the requirement of all ONU. The overall fairness index of data packets in Q-DBA decreases with the increasing of traffic load when the traffic load is larger than 0.7. Since the data packets in Q-DBA are allocated in two different priority, the overall fairness is little small than 1 when the traffic load is high. In DBAM, the overall fairness index in case one is independent of different system loads. Although the DBAM in case one has a maximum window to allocate packets, the burst in high priority cannot be always processed quickly to result packets' blocking and long delay time.

Fig. 23 shows the system utilization versus system load in EPON. It can be found that the bandwidth utilization in Q-DBA is better than that in DBAM. It is because the bandwidth in Q-DBA is allocated step by step to difference class rather than set a maximum window to each class in advance. In addition, because the dropping probability of video packets is high, and the maximum window does not always meet the actual traffic condition, the bandwidth utilization in DBAM is limited. It also can be found that when system load is in 0.6, the case one of Q-DBA improves the system utilization 4.4%, and the case two improves 7.3%.

II.3. Computation-Efficient Algorithms for Dynamic Traffic Grooming in Metro-Access Ring Network

i. SYSTEM MODEL

A. Network Architecture

The network topology is a WDM metro access ring architecture consisting of of N nodes, labeled as 1, 2, …, N counterclockwise. Each node provides several PONs with tree topology to building users (BUs). Only one fiber exists between node i and node i+1 and is labeled as i, $1 \le i \le N$-1, and the fiber between node N and node 1 is labeled as N. Each fiber contains W wavelengths, carrying traffic streams in counterclockwise directions. The capacity C of a wavelength is optical-carrier (OC) 192.

B. Node Function

Each node is equipped with an optical add-drop multiplexer (OADM), an optical line terminal (OLT), m tunable transmitter/receiver pairs, and a fixed transmitter/ receiver pair, shown in Fig. 24. A tunable transmitter (receiver) possesses the capability of transmitting (receiving) any of W wavelengths, and a fixed transmitter (receiver) possesses the capability of transmitting (receiving) a fixed wavelength. Multiplexer aggregates the wavelengths into a single fiber while demultiplexer dispatches wavelengths from the fiber, and a wavelength whose traffic streams are required being dropped out of the node would be disaggregated by the demultiplexer.

An OADM provides three functions. First, it can optically bypass some wavelengths from the incoming link directly to its corresponding outgoing link if traffic streams are not necessary to be dropped in this node. Second, it can optically drop some wavelengths from the incoming link to tunable receivers which convert the traffic streams into electronic form. Finally, it multiplexes some wavelengths to the outgoing link. An OLT performs three electrical operates. First, it multiplexes low-speed streams from the BUs to a high-speed one which is sent to a tunable transmitter. Second, it extracts low-speed streams out of a high-speed stream. At last, it transmits the low-speed streams to the BUs by a fixed transmitter and receives the streams from the BUs by a fixed receiver.

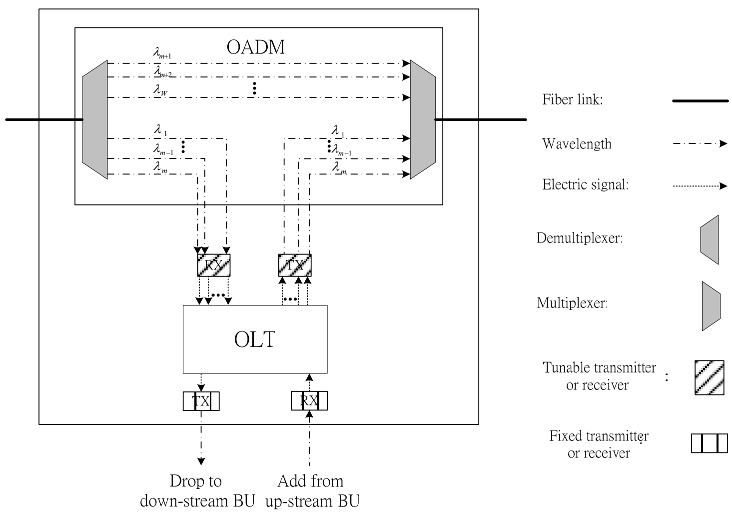


Figure 24. Node Architecture

C. Source Model

The traffic source model is a dynamic and non-uniform. The capacity of a wavelength is OC-192. The bandwidth request of a new call, R, is categorized into four types: OC-1 (type-1), OC-3 (type-2), OC-12 (type-3), and OC-48 (type-4). The probability of the type of the new call request is pk and $\sum_{k=1}^{4} p_k = 1$. Arrival process of the new call requests from each node is Poisson with mean arrival rate λ (1/sec). The traffic model is traffic streams from any source and destination (s, d) arrive and leave according to Poisson process. The service time is exponentially distributed with mean service time 1/μk (sec) for type k traffic. The destination of a traffic stream uniformly locates among all the nodes. As a result, the traffic streams will arrive one by one and the arrival time of next new call request is not known in advance.

ii. PROBLEM FORMULATION

A. Notations

Table 4 (5) (6) defines the notations (the sets used to define the problem) (the variables of the problem formulation).

Notation	Definition
(*i, j, k*)	A lightpath which traverses from node *i* to node *j* through wavelength *k* without optical-electric conversion. $i, j \in \{1, 2..., N\}$, $i \neq j$ and $1 \leq k \leq W$.
(*s, d*)	The node pair represents source node *s* and destination node *d* for an end-to-end traffic which traverses through a single or multiple lightpaths. $s, d \in \{1, 2..., N\}$ and $s \neq d$.
l	The fiber link between node *l* and node *l*+1. $l \in \{1, 2, ..., N\}$. $l = N$, if node *N* to node *1*.

Table 4: The notations of mathematical formulation

Set	Definition
$B(l)$	$B(l) = \{(i, j, k) \mid$ lightpath (i, j, k) passes through fiber link l, $\forall i, j.\}$
$L'(s, d)$	$L'(s, d) = \{(i, j, k) \mid$ lightpath (i, j, k) forms a complete path s to d, where $i = s$ is for the first lightpath, $j = d$ is the last one, and k could be different for various one.$\}$

Table 5: The sets of mathematical formulation

Variable	Definition
$t_{(i,j,k)}(s, d)$	The total traffic load in unit of OC-1 from (s, d) carried by the lightpath (i, j, k) : $t_{(i,j,k)}(s, d) \in \{0, 1, 2, \cdots, C\}, \forall (s, d), (i, j, k)$.
$t_{i,j}$	The total traffic load carried by the lightpaths belonging to $L(i, j)$ before the acceptance of a new call request or the release of an on-going call : $t_{i,j} = \sum_{(s,d)} \sum_{(i,j,k) \in L'(s,d)} t_{(i,j,k)}(s, d)$.
$t'_{i,j}$	The total traffic load carried by the lightpaths belonging to $L(i, j)$ after the acceptance of a new call request or the release of an on-going call. i) If a new call with required traffic load R is accepted, $t'_{i,j} = t_{i,j} + R$. ii) If an on-going call with required traffic load R releases, $t'_{i,j} = t_{i,j} - R$.
$c_{(i,j,k)}$	The lightpath (i, j, k) indicator : $c_{(i,j,k)} = \begin{cases} 1 & \text{a lightpath } (i, j, k) \text{ is actively used,} \\ 0 & \text{otherwise.} \end{cases}$
$b_{i,j}$	The number of wavelengths in $L(i, j)$ before the acceptance of a new call request or the release of an on-going call, $b_{i,j} \in \{0, 1, \cdots, m\}$.
$b'_{i,j}$	The number of wavelengths in $L(i, j)$ after the accpetance of a new call request or the release of an on-going call and the relationship between number of wavelengths and lightpath indicator can be described by the equation. $\sum_{k=1}^{W} c_{(i,j,k)} = b'_{i,j}, \forall (i, j)$, where $b'_{i,j} \in \{0, 1, \cdots, m\}$.
$[x_{ij}]$	The N by N matrixes of which the (i, j)th element is x_{ij}, and x is the general expression $x_{ij} \in \{t'_{i,j}, t_{i,j}, b'_{i,j}, b_{i,j}\}$

Table 6: The variables used in mathematical formulation

B. Generalized Problem Formulation

The dynamic traffic grooming is formulated as an integer linear programming problem (ILP), where the goal is to maximize the utilization efficiency of the wavelength used. This ILP problem is expressed as

$$Maximize \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{t'_{i,j}}{C * b'_{i,j}}, \tag{32}$$

subject to three constraints :

$$t'_{i,j} \leq b'_{i,j} * C \text{, for } \forall (i,j), \tag{33}$$

$$\sum_{(i,j) \in B(l)} b'_{i,j} \leq W \text{, for } \forall l, \tag{34}$$

$$\sum_{(i,j) \in B(l)} c_{(i,j,k)} \leq 1 \text{, for } \forall l,k, \tag{35}$$

$$\sum_{j} b'_{i,j} \leq m \text{, for } \forall i, \tag{36}$$

$$\sum_{i} b'_{i,j} \leq m \text{, for } \forall j, \tag{37}$$

Eq. (33) is the traffic constraint and indicates that the total traffic load from (i, j) must be less than the total capacity. Eq. (34) and Eq. (35) are the wavelength constraints. Eq. (34) the former expresses the bound imposed by the number of wavelengths available while Eq. (35) ensures that no wavelength clash is allowed. Eq. (36) and (37) are the transceiver constraints and they ensure that the number of lightpaths from node i to node j is less than or equal to the number of transmitters and receivers, respectively.

Since the ILP problem is NP-complete [23] and some time-consuming optimization algorithms are not feasible to deal with dynamic traffic model, we propose a SA-based method for optimal solution and a heuristic-based method for suboptimal solution to solve this problem.

iii. COMPUTATIONAL-EFFICIENT ALGORITHM FOR DYNAMIC TRAFFIC GROOMING

A. Simulated Annealing-based Traffic Grooming Algorithm (STGA)

SA algorithm [22] is applied to obtain the optimal solution of the dynamic traffic grooming problem. A state of the Markov chain is defined as the matrix $[b^{'}_{i,j}]$ which is the distribution of the lightapths in the network for adapting SA algorithm, and its goal is to find an optimal matrix $[b^{'}_{i,j}]$ having the maximum value of the object function. According to SA algorithm, we propose simulated annealing-based traffic grooming algorithm (STGA). Fig. 25 shows the flowchart of STGA, which consists of three main processes: *initialization*, *selection and comparison*, and *annealing*.

- *Initialization*

First of all, initial temperature, T_{max}, final temperature, T_{min}, decrement rate of temperature, γ, the maximum number of iteration at a temperature level, L_{max}, should be set to appropriate values. Second, the initial matrix $[b^{'}_{i,j}]_0$ is selected randomly and it must satisfy the constraints. Besides, the STGA is insensitive to the initial state, so choosing a reasonable $[b^{'}_{i,j}]_0$ in this case is enough for guaranteeing convergence of the STGA.

- *Selection and Comparison*

If $[b^{'}_{i,j}]_p = [b^{'}_{i,j}]_0$, p denotes the present state, the next state $[b^{'}_{i,j}]_q$ will be generated from the neighborhood of $[b^{'}_{i,j}]_0$. If the next state $[b^{'}_{i,j}]_q$ is unreasonable, then it will be put in the taboo list to prevent to be chosen again in following iterations, and this makes the STGA more efficient. If the next state $[b^{'}_{i,j}]_q$ is reasonable, the SA algorithm will determine whether it accepts $[b^{'}_{i,j}]_q$ or not according to the acceptance function as below. The $[b^{'}_{i,j}]_q$ will be absolutely accepted if U_q is larger than U_p, or it will be accepted by probability. The acceptance function is

$$f_T(U_q - U_p) = \begin{cases} \exp(\dfrac{U_q - U_p}{T}) & \text{if } U_q - U_p < 0 \\ 1 & \text{if } U_q - U_p \geq 0 \end{cases}, \tag{38}$$

where

$$U_k = \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \frac{t'_{i,j}}{C * b'_{i,j}},\ b'_{i,j} \text{ is an entry of } \left[b'_{i,j} \right]_k,\ k = p \text{ or } q.$$

- *Annealing*

If the number of iteration at a temperature level reaches L_{max}, then the temperature will decrease by the decrement rate γ until the temperature reaches the final temperature T_{min}.

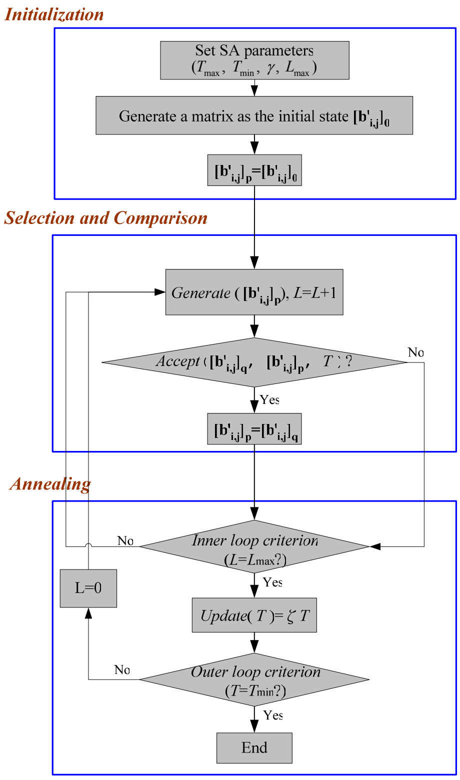


Figure 25. Flowchart of the Simulated Annealing-based Traffic Grooming algorithm

B. Heuristic-based Traffic Grooming Algorithm (HTGA)

The STGA describes the dynamic model as a static model to obtain optimal solution according to the snapshot approach. However, the computational complexity remains unsolved. Thus, we proposed a heuristic algorithm, called a heuristic-based traffic grooming algorithm (HTGA), to reduce the complexity. The heuristic ideas of HTGA states are described that traffic grooming or wavelength assigned into a new lightpath is performed for new call requests, and traffic grooming and re-arrangement are performed when some lightpaths are under-utilized. Therefore the HTGA will solve the ILP problem by solving the three subproblems : *(a) traffic grooming* *subproblem, (b) wavelength assignments subproblem, (c) traffic rearrangement subproblem*. As shown in Fig. 4, the HTGA consists of three block diagram, *traffic grooming* (TG) block, *wavelength assignment* (WA) block, and *traffic rearrangement* (TR) block to solve the (a), (b), (c) subproblems, respectively. The three subproblems and their solution are described in the following.

(a) Traffic grooming subproblem

Since the utilization efficiency of the used wavelength increases as long as a traffic can be groomed into the lightpaths suitably, we use TG block to looks for a path from the present lightpaths, and considers the lowest residual capacity first.

(b) Wavelength assignment subproblem

This subproblem is considered how to assign a wavelength to a lightpath suitably. Many wavelength-assignment approaches were discussed in [25] and were found to perform similarly. Therefore, according to the results in [25], one simple approach, first-fit (FF) [26], is selected to solve *wavelength assignment subproblem*. And, the FF is used in the WA block.

(c) Traffic rearrangement subproblem

This subproblem is considered how to rearrange a portion of lightpaths topology suitably. In order to increase the performance of the network, we still consider whether the present wavelengths are able to carry the traffic which is carried by a lower utilized wavelength. The TR block is used to find the path and to groom the traffic in it.

The flowchart of the HTGA is shown in Fig. 26. There are two trigger events in HTGA; event 1 indicates new call arrival, and includes TG block and WA block of block diagram, event 2 means type-4 call departure and includes TR block. Because capacity of type-4 call is larger than the other types, the wavelength utilization may decrease largely when type-4 calls leave.

When event 1 occurs, the HTGA will work as follows.

- **Step 1.1:** *Check (i, j,k) of the current lightpaths.*

  If (i, j,k) of a lightpath which satisfies (s, d) of new call request exists and has enough capacity, the HTGA grooms the new call into this lightpath, which is called direct lightpath and with the largest utilization of lightpath, and go to Step 1.4. Otherwise, go to Step 1.2.

- **Step 1.2:** *Check complete paths.*

  The HTGA finds whether the complete path exists and also has enough capacity. If it exists, the HTGA grooms the new call to this path and go to Step 1.4. Otherwise, go to Step 1.3.

- **Step 1.3:** *Check free wavelengths*

  If there are unused wavelengths from (*s*, *d*), the HTGA will establish the new lightpath for the new call and go to Step 1.4. Otherwise, the new call is blocked and the algorithm ends.
- **Step 1.4:** *Sort the utilization*

  The algorithm will decreasingly sort the utilization efficiency of the current lightpaths according to the value of utilization of lightpaths, and then ends.

When event 2 occurs, the HTGA will work as follows.

- **Step 2.1:** Check the lightpath of released call

  If utilization efficiency of the lightpaths of departure call is small than θ, the HTGA will reroute lightpaths and groom traffic onto the other lightpath. Otherwise, the algorithm ends.

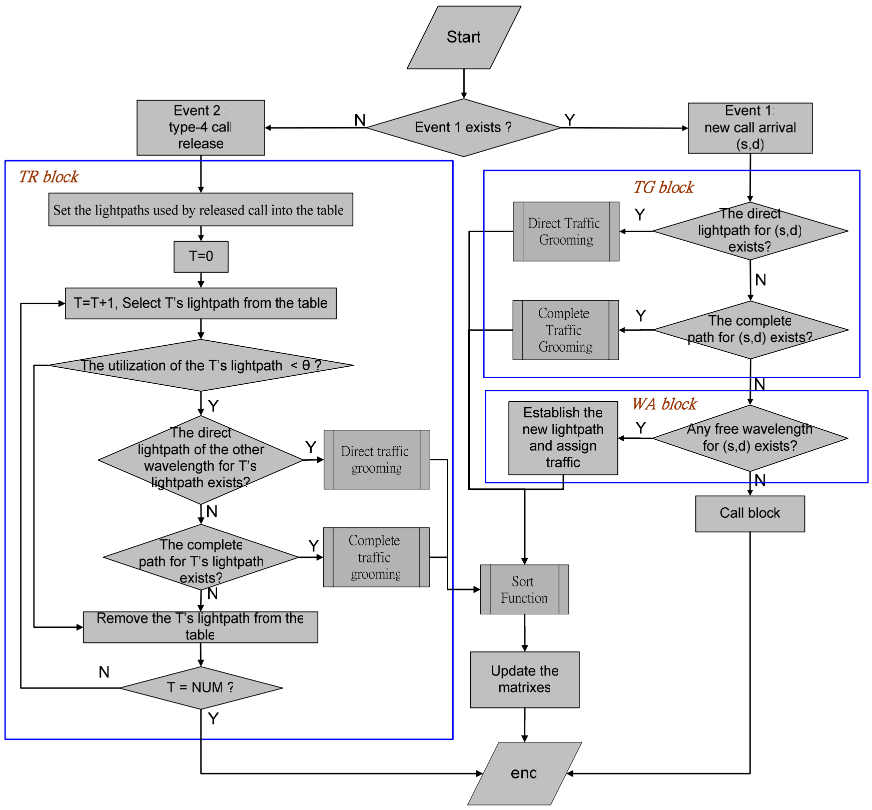


Figure 26. Flowchart of the HTGA algorithm

iv. SIMULATION RESULTS

Simulation parameters are introduced first. Then, we evaluate the performance of HTGA and STGA in a fixed network configuration. In order to reveal the performance of the algorithms with respects to various parameters, three kinds of scenario cases were simulated; ***Scenario 1*** is that the number of nodes N is varied, but the other parameters are the same, ***Scenario 2*** is that the mean service time $1/\mu_k$ is varied, but the other parameters are always the same, ***Scenario 3*** is that the ratio of the probability among difference type p_k is varied, but the other parameters are always the same.

1) Simulation Parameters

The value of the parameters are listed in Table 7, respectively. In table 8, the parameters of the STGA are computed according to [27].

Parameter	**Values**
N	16, 24,32
W	20
m	10
$1/\mu_k$(sec)	60, 180
(p_1, p_2, p_3, p_4)	(0.4, 0.3, 0.2, 0.1), (0.1, 0.2, 0.3, 0.4).
θ	0.5

Table 7: The value of the parameters used in the simulation

Different simulation environment defined by **(***N, W, m***)**	The corresponding STGA parameters for difference environments **(**$T_{max}, T_{min}, L_{max}, \gamma$**)**
(16, 20, 10)	(700,10,16000,0.9)
(24, 20, 10)	(1000,10,24000,0.9)
(32, 20, 10)	(1400,10,32000,0.9)

Table 8: The value of the parameters used in STGA

2) Performance Evaluation in Fixed Network Configuration

Fig. 27 shows the utilization efficiency of the used wavelengths of the two algorithms. The utilization efficiency of the used wavelengths in the HTGA is only slightly less 10% than the STGA. The STGA has better performance over the HTGA because it always finds the optimal allocation for every arrival call. Conversely, the HTGA just arranges every arrival call from the less residual capacity without changing the present lightpaths as could as possible, and it only rearranges traffic when the calls with large bandwidth request leaves the network.

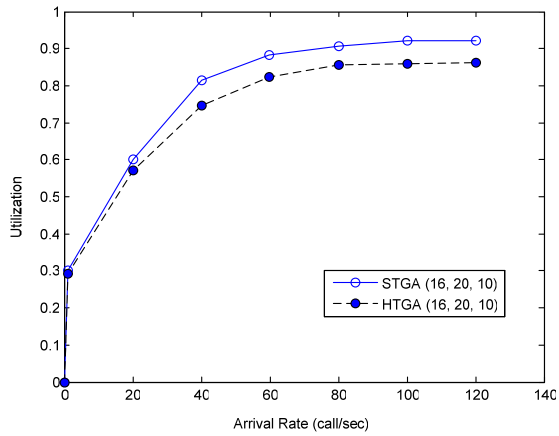


Figure 27. The utilization efficiency of the used wavelengths in two algorithms

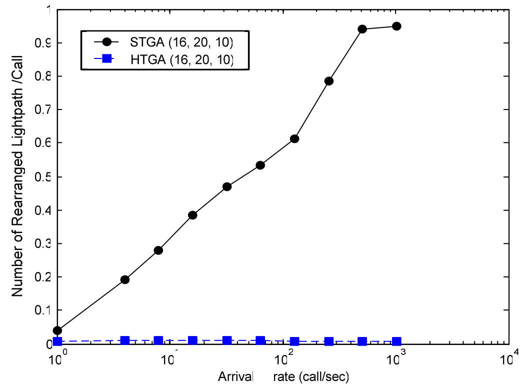


Figure 28. Number of rearranged lightpaths per call in two algorithms

Fig. 28 illustrates two curves of the number of rearranged lightpaths for the two algorithms. As shown in the figure, the number of rearranged lightpaths in the STGA is larger than that in the HTGA. The reason is that, when a call is allowed to enter the network or departures from the network, the STGA rearranges all serving calls whereas the HTGA just rearranges parts of serving calls. However, the network has to take more time to establish the varied lightpaths if lightpaths are rearranged frequently. Lower number of rearranged lightpaths is preferable from the system's point of view. In this aspect, the HTGA is more practical than the STGA obviously.

The overall computation complexity of STGA is $O(ln(N*W)*\chi*N^3*W)$ and the complexity of the HTGA is $O(N^2*W)$, where χ is the number of current served call. The STGA requires very large computation power although it can attain optimal solution. However, the HTGA can attain acceptable system utilization close to the STGA while its computation complexity is extremely low. Moreover, the HTGA only has to change a small portion of lightpaths, compared to the STGA. Thus, the HTGA effectively reduces the computation complexity without severely suffering the system utilization, making it more practical and implementation-feasible for the current optical equipments.

3) Performance Evaluation in Various Network Configurations

Fig. 29 shows the results of the utilization efficiency of the used wavelengths for ***Scenario 1***. When number of nodes increases, the HTGA can still attain high system utilization. In addition, the utilization curves of two algorithms reach the saturation state at lower arrival rate as number of nodes increases. The reason is that the optical network accommodates more service calls at the saturation state when the network size enlarges.

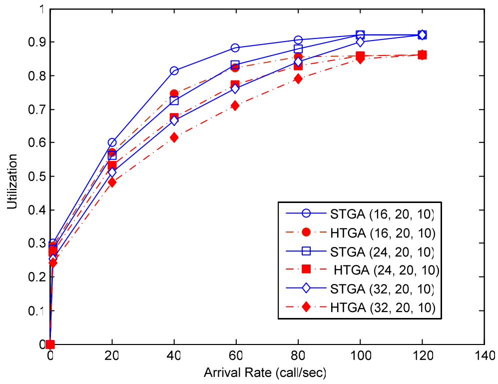


Figure 29. The utilization efficiency of the used wavelengths in ***Scenario 1***

The result of the ***Scenario 2*** is shown in Fig. 30. As mean service time increased, the curve of utilization efficiency reaches the saturation state at a lower arrival rate. Moreover, as mean service time becomes larger, the utilization gap between the two algorithms widens. The reason is that the number of serving call in the network becomes large with the same arrival rate and also the chances of rearranged traffic decrease due to the lower trigger probability of event 2.

Fig. 31 shows the utilization efficiency of the used wavelengths of two algorithms. The arrival rate to reach saturation in the small-call case is higher than that in the large-call case; moreover, the utilization gap between the two algorithms in the large-call case is smaller than that in the small-call case. The large-call case accommodates more type-4 calls so that it has a larger total traffic load and higher chances to trig the rearrangement event. As a result, the system performance in the large-call case is better.

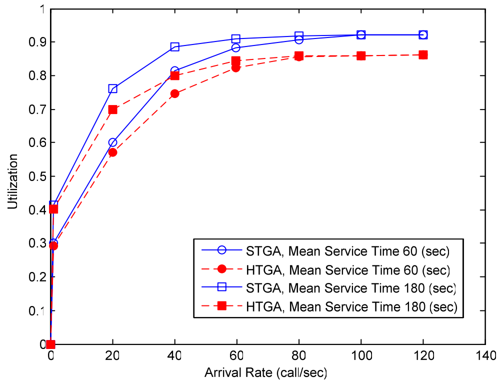


Figure 30. The utilization efficiency of the used wavelengths in ***Scenario 2***

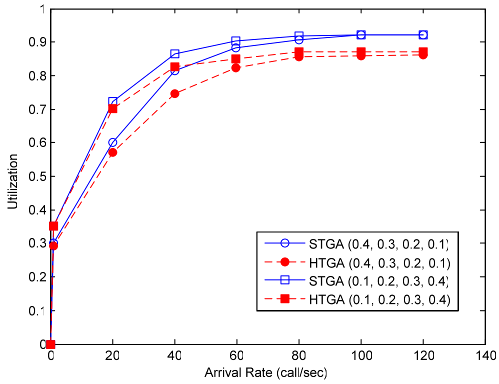


Figure 31. The utilization efficiency of the used wavelengths in ***Scenario 3***

III. Reference

[1] G.. Kramer and G. Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Acces Network," *IEEE Commun Mag.*, vol. 40, pp.66-73, Feb. 2002.

[2] J. 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.

[3] 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.

[4] 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.

[5] K. Rege et al., "QoS management in trunk-and-branch switched Ethernet networks," *IEEE Commun. Mag.*, vol. 40, pp. 30-36, Dec. 2002.

[6] 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.

[7] IEEE standard 802.3ah task force home page [Online]. Available: http://www.ieee802.org/3/fem

[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," *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 Queueed 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*, vol. 3, pp. 1096-1104.

[13] G. Kramer,B. Mukherjee, and G. Pesavento, "Ethernet PON (EPON): design and analysis of an optical access network," Phontic Network Commun., vol. 3, no.3, pp.307 319, July 2001,

[14] "IPACT: A dynamic protocol for an Ethernet PON (EPON)," IEEE Commun.

Mag., vol. 40, no2, pp. 74–80, Feb. 2002.
[15] Y. Yang, J. Nho and B. Ahn, “An enhanced burst-polling based delta dynamic bandwidth allocation scheme for QoS over E-PONs,” NRBC’04, pp.31–36, Oct. 2004.
[16] Y. Luo and N. Ansari, “Bandwidth allocation for multiservice access on EPONs,” IEEE Opt. Commun., vol. 43, pp. s16–s21, Feb. 2005.
[17] Y. Yang, J. Nho, N. P. Mahalik, K. Kim and B. Ahn, “QoS provisioning in the EPON systems with traffic-class burst-polling based delta DBA,” IEICE Trans. Commun., vol. 89, pp.419–426, Feb. 2006.
[18] D. Nikolova, B. V.Houdt and C. Blondia, “Dynamic bandwidth allocation algorithms in EPON: a simulation study,” Proc. of SPIE vol. 5285, pp. 369–380, 2003,
[19] “Dynamic bandwidth allocation algorithms for Ethernet Passive Optical Networks with threshold reporting,” Telecommun. Sys., vol. 28, pp. 31–52, 2005.
[20] H. Miyoshi, T. Inoue and K. Yamashita, “QoS-aware dynamic bandwidth allocation scheme in gigabit-Ethernet Passive Optical Networks,” IEEE International Commun. Conference, vol. 1, pp. 90–94, June 2004.
[21] B. Mukherjee, S. Ramamurthy, D. Banerjee, and A. Mukherjee, “Some principles for designing a wide-area optical network,” Proc. of IEEE INFOCOM ‘94, vol. 1, pp. 110 – 119, June 1994.
[22] E. Modiano and P. J. Lin, “Traffic grooming in WDM networks,” IEEE Comm. Mag., vol. 39, no. 7, pp. 124 – 129, July 2001.
[23] R. Dutta, and G.. Rouskas, “On optimal traffic grooming in WDM rings,” IEEE Journal on Selected Areas in Communications, vol. 20, no. 1, pp. 110 – 121, Jan. 2002.
[24] B. Chen, G. Rouskas, and R. Dutta “Traffic grooming in WDM ring networks with the min-max objective,” Proceedings of Networking 2004, pp. 174 – 185, May. 2004.
[25] E. H. Modinao, and A. L. Chiu “Traffic grooming algorithms for reducing electronic multiplexing costs in unidirectional WDM ring network,” Journal of Lightwave Technology , vol.18, no. 1, pp. 2-12, Jan. 2000.
[26] X. Zhang, and C. Qiao “An effective and comprehensive solution to traffic grooming and wavelength assignment in SONET/WDM rings,” IEEE/ACM Transactions Journal on Networking, vol. 8, no. 5, pp. 608-617, Oct. 2000.
[27] Moon-Gil Yoon “Traffic grooming and light-path routing in WDM ring networks with hop-count constraint,” IEEE International Conference on Communications, vol. 3, pp. 731-737, June 2001.

[28] V.R. Konda, and T.Y. Chow "Algorithm for traffic grooming in optical networks to minimize the number of transceivers," IEEE Workshop on High Performance Switching and Routing, pp. 218-221, May 2001.
[29] R.B. Billah, B. Wang, and A.S. Awwal "Efficient traffic grooming in synchronous optical network/wavelength-division multiplexing bidirectional line-switched ring network," Optical Engineering, vol. 43, no. 5, pp. 1101-1114, May 2004.
[30] J. Wang, W. Cho, V. R. Vemuri, and B. Mukerjee "Improved approaches for cost-effective traffic grooming in WDM ring networks: ILP formulations and single-hop and multiple connections," Journal of Lightwave Technology, vol. 19, no.11, pp. 1645 – 1653, November. 2001.
[31] K. Zhu, and B. Mukherjee, "Traffic grooming in an optical WDM mesh network," IEEE Journal on Selected Areas in Communications, vol. 20, no. 1, pp. 122 – 133, Jan. 2002.
[32] P. Prathombutr, J. Stach, and E.K. Park "An algorithm for traffic grooming in WDM optical mesh networks with multiple objectives," ICCCN 2003, pp. 405-411, Oct. 2003.
[33] A. Adas, "Traffic models in broadband networks," *IEEE Commun. Mag.*, vol. 35, no.7, pp.82-89, July 1997.
[34] W. Willinger, M. Taqqu, R. Sherman, and D. Wilson, "Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level," *Proc. ACM SIGCOMM 1995*, pp.100-113, Aug.1995.
[35] D.Liu and Y.H. Lee, "An efficient scheduling discipline for packet switching networks unsing earliest deadline first round robin," *Telecommunication Sys*, vol. 28, no. 3.
[36] W. T. Chen, R. R. Lee, and H. J. Lin, "A QoS-guaranteed and delay-minimized cell multiplexing method in ATM networks," *IEEE International Commun. Conference*, vol. 3, pp.1723-1727, June. 1998.
[37] Y. Zhu, "A survey of network requirements to support current and future data streams," *CS 7001 Introduction to Graduate Studies*, vol. 1, pp.1-7, June. 2000.
[38] R. Jain, *The art of computer systems performance analysis: Techniques for experimental design, measurement, simulation, and modeling*. New York: John Wiley Sons, INC., 1991.
[39] Chun-Chen Peng（彭崇禎）, "Uplink Scheduling Schemes in EPON Access Network（乙太被動光纖擷取網路之上鏈路排程機制）," M.S. thesis（碩士論文）, Department of Communication Engineering, National Chiao Tung University, Taiwan, 2004.
[40] Chen-Fong Kuo（郭朕逢）, "Computation-Efficient Algorithms for Dynamic

Traffic Grooming in Metro-Access Ring Network（都會擷取環形光網路之低運算複雜度動態訊務彙集機制）," M.S. thesis（碩士論文）, Department of Communication Engineering, National Chiao Tung University, Taiwan, 2005.

[41] Chin-Ya Huang（黃琴雅）, "QoS-Promoted Dynamic Bandwidth Allocation (Q-DBA) for Ethernet Passive Optical Networks（以太被動光纖網路下改善服務品質的動態頻寬分配）," M.S. thesis（碩士論文）, Department of Communication Engineering, National Chiao Tung University, Taiwan, 2006.

IV. Comment

◆ Fair Scheduling Schemes for Non-Real-Time Service in EPON Access Network

We first proposed a scheduling method that is suitable for EPON access network, and 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, Hybrid EQL-QLP and Hybrid LQF-QLP schemes, to obtain the overall fairness which combines the fairness of packet delay with fairness of packet blocking probability. 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 and 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.

We also 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.

◆ QoS-Promoted Dynamic Bandwidth Allocation (Q-DBA) for Ethernet Passive Optical Networks

In this field, we proposed a Q-DBA method to allocate bandwidth with three classes of packets, voice, video, and data based on the priorities from the highest to the lowest step by step. Thus, Q-DBA sets the voice packets to be the first priority, and the video packets' priority is changed from the fifth priority into the third priority when the video packets' delay criterion will be violated at the end of next cycle. To sustain the video dropping probability, the priority of video packets can be raised to the second priority in further. Although data packets do not have any delay constraint, they still should not be sacrificed. The Q-DBA also raises the priority of data packets, which is from the sixth priority to the fourth priority when the waiting time of data packet exceeds the waiting bound. It can avoid the data packets from being in the starvation condition. Furthermore, the bandwidth of the fiber link is totally allocated to make the system fully utilized.

The simulation results show that the Q-DBA has a better performance than DBAM and the QoS of voice packets are also fully guaranteed. The delay time of video packets is almost below the video delay criterion and the dropping probability of video packets is also guaranteed. Besides, the data packets are transmitted without being in the starvation condition. Above that, the Q-DBA can support most of the system load without violating the QoS requirements. In addition, with the development of WDM, the proposed Q-DBA method may be adapted to manage the resource in the EPONs, and customers can get the most benefits.

◆ Computation-Efficient Algorithms for Dynamic Traffic Grooming in Metro-Access Ring Network

In this field, we want to solve the problem of dynamic traffic grooming and wavelength assignment in metro-access ring network. The goal is to effectively maximize the utilization efficiency of the wavelength used and reduce the new call blocking rate. We present the ILP formulation for the problem. In order to solve this problem, we propose STGA algorithm to obtain the optimal solution. However, the STGA algorithm is infeasible due to its computation complexity. Alternatively, we propose a heuristic algorithm-based HTGA algorithm. Although the HTGA algorithm is a suboptimal solution, the computation complexity of the HTGA is much lower than that of the STGA.

From the simulation comparisons between the two algorithms in the same environment, we can summarize the pros and cons of the STGA and HTGA algorithm. For STGA algorithm, the advantage is that it can attain an optimal solution while the disadvantage is that the computation complexity and the number of rearranged lightpaths are large. On the other hand, for HTGA algorithm, the advantage is that the computation complexity and the number of rearranged lightpaths are small while the disadvantage is that it just can obtain sub-optimal solution. However, the disadvantage of the HTGA algorithm is accepted because the gap of the throughput utilization is just less than 10%. Therefore, the HTGA algorithm is a feasible and attractive algorithm.

Besides, another advantage of the HTGA can be observed in various environments. Simulation results show that the throughput gap between the two algorithms is no more than 10% in all the three Scenarios. Thus, we can conclude that the performance of the HTGA algorithm does not deteriorate substantially in various network configurations, including different network size, mean service time, and probability of new call request type. Therefore, the results demonstrate once again the superiority of the HTGA algorithm.

According to the research results during latest three years, we believe that there is close relationship between the research issues and the original projects. The research results have carried out the project taget we expected. In the field of the design of dynamic bandwidth allocation on EPON, we proposed two schemes, Hybrid EQL-QLP and Hybrid LQF-QLP schemes, and a predictor to improve the fairness of non-real time traffic. From the simulation result, we can find that our algorithm is better. Also, we design a QoS-Promoted dymanic bandwidth allocation (Q-DBA) to guarantee the requiredment of the different traffic classes. Also, simulation results prove our method is better than DBAM. Therefor, we can provide an over all solution to deal with the problem of the dymanic bandwidth allocation on EPON with considering the fairness of non-real time traffic and guaranteeing the QoS.

In this field of traffic grooming, we discuss the dymanic traffic grooming under the nonuniform traffic. This problem has not discussed before. Integer linear problem (ILP) methodology is applied to formulate for this problem. Also, we modified a smart method, annealing method, to solve ILP, annealing-based traffic grooming. However, the complexity is very much. We also design another method, heuristic-based traffic grooming. The scheme provides the better performance than the front scheme. For the access-metro ring network, if you want to achieve the highest bandwidth ultilization, you can use the annealing-based traffic grooming scheme. However, if you want to reduce the dropping and also improve the bandwidth ultilization, you should use heuristic-based traffic grooming scheme. Therefor, we provide two schemes to solve the problem of the dymanic traffic grooming under the nonuniform traffic. Also, you can use one of them under the different requirements.