A Bandwidth Guaranteed Polling MAC Protocol for Ethernet Passive Optical Networks

Maode Ma, Yongqing Zhu, Tee Hiang Cheng
School of Electrical & Electronic Engineering
Nanyang Technological University
Singapore 639798

Abstract-While the backbone networks have experienced substantial changes in the last decade; the access networks have not changed much. Recently, passive optical networks (PONs) seem to be ready for commercial deployment in access networks due to the maturity of a number of enabling technologies. long distance and reduced maintenance. Among PON technologies, the Ethernet PON (EPON) presently being standardized by the IEEE 802.3ah Ethernet in the First Miles (EFM) Task Force is most attractive because of its high speed, low cost, familiarity, interoperability and low overhead. In this paper, we propose a novel Bandwidth Guarantee Polling (BGP) scheme that will allow the upstream bandwidth to be shared based on the service level agreement (SLA) between each subscriber and the operator. It is able to provide bandwidth guarantee for premium subscribers according to the SLAs while providing best-effort service to other subscribers. The analytical and simulation results prove that the proposed scheme does best in what it is designed to do compared to another well-known scheme that has not considered providing differentiated services. With business customers preferring premium services with guaranteed bandwidth and residential users preferring low-cost best effort services, our scheme could benefit both groups of subscribers as well as the operators.

Keyword--Ethernet Passive Optical Network (EPON), Medium Access Control (MAC) Protocol, Service Level Agreement (SLA), Access Network, Polling Systems

I. INTRODUCTION

While access network has experienced little changes in recent years, long-haul networks have been changed dramatically due to the emergence of wavelength division multiplexing (WDM) technologies. During the same period, local area networks (LANs) have scaled up in speed from 10Mb/s to 100Mb/s, and subsequently to 1Gb/s. Even 10Gb/s Ethernet products have started to emerge. As a result, there is a growing gulf between high-capacity LANs and backbone networks with the low-speed access networks being the bottleneck. There is an urgent need for a technology to develop access networks that are inexpensive, simple, scalable, and capable of delivering integrated voice, data, and video services to the subscribers. Passive optical network (PON) is such a technology considered as an attractive solution for access networks.

A PON is a point-to-multipoint optical network employing passive optical components, such as couplers and splitters with no active elements. A PON could be used to implement a fiber-to-the-home (FTTH), fiber-to-the-building (FTTB) or fiber-to-the-curb (FTTC) subscriber access network. PON could adopt a tree, tree-and-branches, ring or bus topology with tree topology being the most common. All transmissions in a tree-based PON are performed between an optical line terminal (OLT) and optical network units (ONUs). The OLT resides in the local exchange (central office), connecting the optical access network to the backbone. The ONUs may locate at the curbs (FTTC solution) or subscriber premises (FTTH and FTTB solutions) to provide an integrated broadband service to the traffic with combination of data, voice, and video. In the downstream direction (from OLT to ONUs), the traffic goes from one point to multiple points. In the upstream direction (from ONUs to OLT), the traffics from multiple points could only reach one point (OLT) [1-2]. The most cost-effective PON architecture is one that uses 1550nm wavelength to broadcast downstream traffic and 1310nm wavelength for upstream transmission.

There are two prevalent datalink layer protocols for above PON architecture, namely asynchronous transfer mode (ATM) PON and Ethernet PON (EPON). APON was developed and defined initiatively as a PON-based optical access network that uses ATM as its layer 2 protocol. There are a numbers of proposals for APONs to share upstream bandwidth [3-5]. However, the expense and the complexity of an APON have made it to a declining interest of industry. An EPON is a PON that carries all data encapsulated in Ethernet frames and is backward compatible with existing IEEE 802.3 Ethernet standards as well as other relevant IEEE 802 standards. Since Ethernet is an inexpensive technology that is ubiquitous and interoperable with a variety of legacy equipments, it becomes a perfect choice of PON for delivering IP packets and supporting multimedia traffics efficiently. An EPON cannot be simply considered as either shared medium or a point-to-point network. In the downstream direction, Ethernet frames transmitted by the OLT pass through a 1:N passive splitter to reach each ONU. The packets broadcasted by the OLT will be filtered by their destination ONUs based on the MAC addresses. In the upstream direction, due to the directional property of a passive combiner (optical splitter), packets from any ONU can only reach the OLT, not other ONUs. As in the upstream, the ONUs have to share the trunk fiber channel, a medium access control protocol is required to arbitrate ONUs' transmission to prevent data from being corrupted due to multiple ONU transmitting at the same time.

It is possible to use WDM technique for multiple ONUs to share the upstream channel. By the WDM scheme, each ONU could transmit at a unique wavelength and not interfering with

other ONUs' transmission. However, this requires each ONU to use a fixed transmitter operating at a different wavelength that will give rise to a lot of inventory and operation and maintenance problems because different types of transceivers have to be stored and installed. Alternatively, a tunable transmitter or a transmitter array can be used for each ONU. This is also not attractive because of the high cost of tunable transceiver and transceiver array. Similarly, either a tunable receiver or a receiver array is required at the OLT to receive the information from different ONUs if each ONU uses a different wavelength for transmission. This makes the WDM scheme a costly proposition.

The conventional contention-based carrier-sense multiple access protocol is difficult to implement in EPON because ONUs cannot easily detect a collision at the OLT due to the directional property of optical splitter/combiner. Although the OLT could detect a collision and inform ONUs by sending a contention signal [6], the long propagation delay in EPON will greatly reduce its efficiency. Furthermore, the bandwidth apportion to each ONU cannot be controlled and guaranteed and makes it very difficult for any form of quality of service (QoS) to be supported. Considering the above factors, a scheduling or reservation-based time-division multiple access (TDMA) scheme on a common wavelength is more attractive for upstream traffic. There is no contention and the bandwidth to each ONU can be allocated based on the real-need and/or based on the SLAs. Moreover, with TDMA, all ONUs use the same wavelength and hence the same type of transceiver, which makes the equipment more easy and cost effective to design, manufacture, operate and maintain.

With the TDMA approach, each ONU is allocated a fixed or variable size time window for transmission of one or more Ethernet frames. An ONU should buffer frames received from the users until they are transmitted in the assigned time window. When the assigned time window arrives, the ONU will burst out frames at full channel speed. There are various allocation schemes for the time window. Interleaved Polling with Adaptive Cycle Time (IPACT) proposed in [7] is a very well designed scheme for dynamic bandwidth allocation in EPON because it minimizes unused bandwidth by using carefully time polling messages which are interleaved or piggybacked on downstream Ethernet traffic. In IPACT, the OLT distributes variable size time windows whose sizes depend on the amount of frames buffered at the respective ONUs, as reported by the respective ONUs using control messages. Each ONU executes the same procedure driven by a grant message received from the OLT. It is an OLT-based centralized dynamic medium access arbitration scheme; thus, it is easy to adaptively change the scheduling at run-time based on the actual network condition. Also, it takes into consideration the large propagation delay of an EPON that could span as far as 20km. IPACT is an attractive proposition because, firstly, no synchronization is needed among ONUs and, secondly, polling message is interleaved with frame transmission so that the overall overhead arisen from the propagation delay is reduced and the efficiency is higher. Unfortunately, IPACT does not explicitly consider the fact that different ONUs have different bandwidth requirements due to the differences in subscribers' service level agreements (SLAs). In addition, IPACT's dynamic bandwidth allocation scheme also introduces burstiness in the transmission of traffic from each ONU to the OLT.

In this paper, we propose a novel Bandwidth Guarantee Polling (BGP) MAC protocol for the EPON. Our proposed BGP could be used to control the sharing of the upstream data channel. The proposed scheme explicitly incorporates the SLAs [8-9] into medium access control (MAC) protocol design to achieve bandwidth guarantee for subscribers that subscribe a premium bandwidth-guaranteed service while providing best-effort service to subscribers that are more price-sensitive and prefer a lower-cost best-effort service.

The remainder of this paper is organized as follows. Section 2 specifies the EPON network architecture and the service. Section 3 presents our BGP scheme in detail. Section 4 presents the analysis we have developed to evaluate the performance of the proposed protocol. Section 5 presents and compares the results got from simulation experiments and analytical calculations. Finally, section 6 concludes the paper.

II. Network Architecture and Service

In this paper, we consider a similar tree-based topology for the EPON as that in [7]. In this architecture, the only OLT connects to *N* (Maximum 64) ONUs through a 1:N passive splitter. The OLT resides in the local exchange (central office), connecting the optical access network to an IP, ATM, or SONET backbone. The ONUs reside on subscriber premises in the FTTH and FTTB scenarios to provide data, voice, and video services to end users. All transmissions in the EPON are performed between OLT and ONUs. There is no direct traffic among ONUs. The system operates in point-to-multipoint way in the downstream direction. It has multipoint-to-point transmission in the upstream direction. In the downstream, packets are broadcasted by the OLT and extracted by their destination ONUs based on the MAC addresses. Figure 1 shows the downstream traffic. In the upstream, the transmission medium is shared by *N* ONUs and a medium access control (MAC) protocol is required to arbitrate among ONUs that have data to transmit. Figure 2 exemplifies the upstream transmission.

The ONUs in the network are divided into two disjointed ONU sets. One set contains a number of bandwidth-guaranteed ONUs while another contains ONUs that do not require bandwidth-guaranteed service. In the implementation

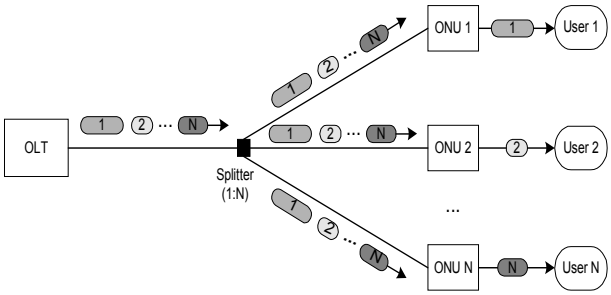

Figure 1. Downstream traffic in tree-based EPON

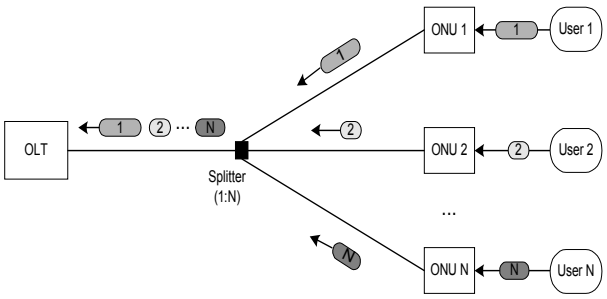

Figure 2. Upstream traffic in tree-based EPON

and deployment, a subscriber premise that supports both types of services can be served with both types of ONUs, which could be two separate pieces of equipment or implemented as two separate blades plagued to the same module, rack or chaise. In the later, only one branch of the PON needs to be terminated at the subscriber premise because the optical signal power could be split internally within the equipment for the two ONU blades in the same equipment.

The system is designed to ensure that bandwidth-guaranteed ONUs are granted sufficient transmission time agreed upon between the subscriber and the service provider in the SLA while making use of any unused bandwidth to provide a best-effort service to non-bandwidth-guaranteed ONUs. In the next section, we propose and present a novel MAC protocol to provide both bandwidth guaranteed service to bandwidth-demand users and best-effort service to non bandwidth-demand users based on the architecture we presented in this section.

III. Bandwidth Guaranteed Polling Scheme

The major challenge of design a MAC protocol to provide bandwidth guaranteed service is how to explicitly incorporate the bandwidth requirements from users into the design of the scheduling scheme so that all of the users can share the transmission link with almost same bandwidth as their SLA. As mentioned in the first section, the TDM scheme is the most pragmatic way to share the transmission link. Taking into the consideration of the various characteristics of the EPON, we believe that polling scheme has potential and advantages to dynamically schedule upstream traffic generated by ONUs. Hence, we have the ideas to combine both to approximate a TDM system with dynamic scheduling ability. The advantages of this approach are obvious. First, the difficult synchronization problem that is the major shortcoming of a multipoint TDM system has been eliminated with the transmit-when-polled mechanism. Second, inflexibility and poor efficiency of TDM scheme is overcome with dynamic scheduling to allocate bandwidth to ONUs that require it at any point in time. Most importantly, the SLAs and instantaneous bandwidth requirements of OUNs can be explicitly incorporated into the scheduling mechanism to ensure that the SLAs are met while any excess bandwidth could be utilized to the fullest for non-bandwidth-guaranteed services. Our proposed protocol is aptly named as Bandwidth-guaranteed Polling (BGP) because it is a roll-call polling system in which the OLT polls ONUs one after another in the adaptive order to allow ONUs to send data to the OLT.

In our BGP scheme, the OLT is the central controller that polls ONUs by sending polling messages regularly to each ONU to grant transmission windows. After receiving the polling message, ONUs start to send data to the OLT. The ONUs are divided into two disjoint groups as bandwidth-guaranteed ONUs and non-bandwidth-guaranteed ONUs. The total upstream bandwidth is divided into equivalent bandwidth units. The OLT maintains an Entry Table that keeps the sequence of entries being polled. Each entry possesses one bandwidth unit that will either be allocated to a bandwidth-guaranteed ONU or be dynamically assigned to a non-bandwidth-guaranteed ONU. There are two sections of each entry. The first section of an entry stores the ID number of the ONU, which occupies the entry. The second section of an entry keeps the propagation delay of the ONU to the OLT. The total number of the entries in the Table is the maximum number of the bandwidth units of the uplink. The OLT also maintains a non-bandwidth-guaranteed ONUs' List that determines the polling sequence of non-bandwidth-guaranteed ONUs. The non-bandwidth-guaranteed ONUs' List has the similar structure as the Entry Table. Each element of it has two sections too. The first section of an element keeps the ID number of the non-bandwidth-guaranteed ONU. The second section keeps the propagation delay of the ONU to the OLT. The total number of the elements in the List is not fixed.

Each bandwidth-guaranteed ONU can be distributed one or more bandwidth units (entries) according to the SLA on their bandwidth requirements, which is corresponding to the payment of end-users. Bandwidth-guaranteed ONUs with more entries will take more bandwidth of the upstream link. And they will be polled more than one time by the OLT in a round of polling. If one bandwidth-guaranteed ONU occupies more than one entry, they will be evenly distributed among all entries in the Entry Table by some algorithm. Best-effort service without bandwidth guarantee will be provided to the non-bandwidth-guaranteed ONUs. Entries that are not occupied by bandwidth-guaranteed ONUs can be assigned to non-bandwidth-guaranteed ONUs dynamically. The OLT polls non-bandwidth-guaranteed ONUs in these available entries in the order presented in the List. In addition, the superfluous transmission window in one entry that is occupied yet consumed by an ONU could be assigned to a non-bandwidth-guaranteed ONU dynamically. In this case, there should be a predefined threshold to determine whether to use the superfluous transmission window or give it up. If the entire transmission window consumed by an ONU in one entry is less than the threshold, then the superfluous could be used to poll a non-bandwidth-guaranteed ONU; otherwise, it has be given up. The maximum transmission window size in terms of number of packets for each entry is a system parameter and can be adjusted according to the network environment.

The OLT polls ONUs one by one in the order of the entry sequence in the Entry Table by a pointer indicating the current entry. If one entry in the Entry Table is not allocated to a bandwidth-guaranteed ONU, it can be used to poll a non-

bandwidth-guaranteed ONU indicated by another pointer as current being served ONU in the non-bandwidth-guaranteed ONUs' List. If one entry in the Entry Table has allocated to a bandwidth-guaranteed ONU or it has been used to poll a non-bandwidth-guaranteed ONU and there is a superfluous transmission window left, it can be used to poll another non-bandwidth-guaranteed ONU.

The proposed BGP scheme has two parts. The first one of them is the scheduling algorithm to presents the polling scheme. The other is an Even Distribution Algorithm (EDA) to evenly distribute multiple entries of the same bandwidth-guaranteed ONUs among all of the entries in the Table.

A. BGP Scheduling

1. Initially, the OLT initializes the Entry Table of bandwidth-guaranteed ONUs and the non-bandwidth-guaranteed ONUs' List according to the SLA on the bandwidth requirements from the ONUs in the network and other system parameters such as propagation delay for individual ONU and maximum transmission window size W_{max}. In the initialization, the EDA is also employed to evenly distribute multiple entries of bandwidth-guaranteed ONUs among all in the Table.

2. The OLT starts to poll bandwidth-guaranteed ONUs in the order as determined in the Entry Table, or poll non-bandwidth-guaranteed ONUs dynamically according to the sequence of the List, by sending Grant message with appropriate window size *G* through downstream channel.

3. On receiving Grant message, the ONU (including bandwidth-guaranteed and non- bandwidth-guaranteed) will:

(1) Get the granted window size *G*;

Get the buffer length *L* (number of packets waiting in the buffer);

(2) Decide the number of packets for transmission:

If *L* is less than *G*, the ONU can send all packets in the buffer;

If *L* is larger than *G*, the ONU can only send packets up to *G*;

(3) Send a Reply message to the OLT to indicate the number of packets in transmission *B*;

(4) Send the data in the buffer of the evaluated amount *B*.

4. The OLT keeps on receiving packets transmitted from ONUs. On receiving a Reply message from ONU, the OLT will:

(1) First get the granted window size *G* and compare it with W_{max};

(2) If *G* equals to W_{max}, this Reply message is coming from an ONU that is assigned a whole entry. Get the number of packets in transmission *B*:

If *B* equals to zero, poll next entry in the Entry Table immediately;

If *B* is between zero and the threshold, the superfluous transmission window in this entry can be used to poll a non-bandwidth-guaranteed ONU, send a Grant message with granted window size $G=(W_{max} - B)$ to next non-bandwidth-guaranteed ONU in the List;

If *B* is between the threshold and W_{max}, give up the superfluous transmission window in this entry, poll next entry in the Entry Table after the time of this entry passes by;

(3) If *G* is less than W_{max}, the Reply message is coming from a non-bandwidth-guaranteed ONU that is assigned the superfluous transmission window in an entry which is not fully utilized by an ONU. Poll next entry in the Entry Table immediately.

In the BGP scheduling, the propagation delay of the polling (Grant) message to the next ONU can be interleaved with the current data transmission except the number of packets in current transmission *B* is or almost is 0. The interleaving can be implemented if the OLT can send out the polling (Grand) message at a proper time so that the message can arrive at the desired destination ONU to activate its data transmission within the time period of current data transmission. By this interleaving, the total average delay experienced by a packet can be reduced.

B. Even Distribution Algorithm (EDA)

1. Fill the Entry Table starting from the bandwidth-guaranteed ONUs with maximum entries to the one with minimum entries.

2. For each bandwidth-guaranteed ONU_i with K_i entries, determines:

(1) First entry number $E_i[1]$: Let entry *i* to be the first entry of ONU_i, $E_i[1]=i$. If entry *i* is occupied by other ONU, check its neighbor entries $E_i[1]= i \pm n$, $n=1,2,3,\ldots$until the nearest entry that is not occupied by other bandwidth-guaranteed ONUs, and fill this entry with the ONU ID *i*.

(2) Other entry number $E_i[j]$ ($j=2,3,\ldots K_i$): the j^{th} entry number of this bandwidth-guaranteed ONU should be $E_i[j]=mod((E_i[1]+int((j-1)*K/K_i)), K)$, where *K* is the total entry number of the Entry Table. If j^{th} entry is occupied by other ONU, check its neighbor entries $E_i[j]=mod((E_i[1]+int((j-1)*K/K_i) \pm n), K)$, $n=1,2,3,\ldots$ until the nearest entry that is not occupied by other bandwidth-guaranteed ONUs, and fill this entry with the ONU ID *i*.

3. All entries that are not occupied by bandwidth-guaranteed ONUs will be assigned to non-bandwidth-guaranteed ONUs dynamically.

According to the EDA, all entries of the same bandwidth-guaranteed ONU should be separated from the same distance in the Entry Table, so that bandwidth of this ONU can be evenly distributed on the upstream channel when they are polled based on the Entry Table. It will benefit to the bandwidth-guaranteed ONUs with high traffic since the packets in the buffer can be allowed to transmit after the same intervals. Thus there will be no burst, caused by the scheduling, introduced to the data transmission.

IV. Performance Analysis

In this section, we will evaluate the performance of the proposed MAC protocol by setting up a simple mathematical model. The major metric of this analysis is the average delay experienced by a packet from the time of its arriving at an ONU to the time of its arriving at the OLT.

The point-to-point average delay obviously consists of three parts: access delay, transmission delay and propagation delay. Access delay is the time from the packet arrives at an ONU to the time it starts to be transmitted over the upstream channel. The access delay A experienced by a packet can be divided into two parts: $A=W_1+W_2$, where W_1 is the waiting delay while other ONUs are being polled, the ONU, which contains this packet, is idle and waiting for its turn to be polled; W_2 is the waiting delay while the ONU is being polled and transmitting earlier packets in the buffer. The transmission delay of a packet is the time period the ONU to send out the whole frame. The propagation delay is the time period for the packet traveling from the ONU to the OLT in the upstream channel. Based on above considerations, we can get the expression of the average message delay as

$$E(D) = E(A) + \bar{t} + \bar{\tau}. \quad (1)$$

where, $E(A) = E(W_1) + E(W_2)$ (2)

$\bar{t}$ is the average transmission time for a packet, and $\bar{\tau}$ is the average propagation delay for each ONU to the OLT.

In order to evaluate the average waiting time in a polling system, we need to consider another time period t_c, which is the time needed to complete one round of polling to all of the ONUs in the network.

The scan time t_c is equal to walk time L, which is the total time delay for transmission of polling messages, plus total data transmission time t. In our BGP scheme, an entire polling will poll both bandwidth-guaranteed and non-bandwidth-guaranteed ONUs presented in all the entries of the Entry Table. Data transmission time for each entry equals to the average transmission time for one packet times the maximum transmission window. To poll all the entries in the Entry Table, it will take time t for transmission of all the packets. We have $t = N * W_{\max} * \bar{t}$, where N is the total number of entries in the Entry Table, W_{max} is the maximum transmission window size, and $\bar{t}$ is the average transmission time for a packet. The walk time in a poll equals to the sum of polling message transmission time, round-trip propagation time and synchronization time. We consider that the synchronization time and polling message transmission time can be ignored. The walk time can be divided into: $L = L_{up} + L_{down}$, where L_{up} is the upstream walk time, the propagation time experienced for a packet to be transmitted from an ONU to the OLT, L_{down} is the downstream walk time, the propagation time for a polling (Grant) message traveling from the OLT to ONUs. As one entry may be used to poll one or two ONUs, L_{up} will range from $N*\bar{\tau}$ to $2N*\bar{\tau}$, where $\bar{\tau}$ is the average propagation delay from an ONU to the OLT, in terms of different traffic load to the whole network, ρ and threshold of the transmission window T. The polling message propagation time can be ignored when the previous ONU sends packets with full transmission window. Let P to be the number of entries that could be used to poll a second ONU, which is a random variable that varies in terms of different ρ. Then $L_{up} = N*\bar{\tau} + P*\bar{\tau}$, $L_{down} = 2P*\bar{\tau}$. The average walk time will be $L = N*\bar{\tau} + 3P*\bar{\tau}$, then the average scan time will be

$$\bar{t_c} = L + t = N*\bar{\tau} + 3P*\bar{\tau} + N*W_{\max}*\bar{t}. \quad (3)$$

To determine the value of P, the concept of effective traffic load is introduced as follows. As the full channel capacity μ is divided into N unit capacity in our BGP model and each ONU is assigned different unit capacity according to the number of entries it occupies, the effective service rate for each ONU can be expressed as: $n\mu / N$, where n is the number of entries distributed to this ONU. A specific ONU's effective traffic load can be defined as: $\rho_e = \frac{\lambda_i}{n\mu/N} = \frac{\rho\mu/M}{n\mu/N} = \frac{N}{nM} * \rho$, where M is the total number of ONUs in the network. P can be expressed as: $P = \sum_{n=1}^{N} n * m_n * p_n$, where m_n is the number of bandwidth guaranteed ONUs that have n entries, p_n is the probability that one entry can be used to poll two ONUs, defined as $p_n = \begin{cases} \alpha, & \text{if } \rho_e < \rho_s \\ 0, & \text{if } \rho_e \geq \rho_s \end{cases}$. ρ_s is the saturate traffic load defined as the point of effective traffic load when a specific ONU starts to loss packets due to its finite buffer, determined by the total traffic load ρ and threshold of the transmission window T. Let's assume that p_n is uniformly distributed with value α, $0 \leq \alpha \leq 1$, then the average value of P is: $\bar{P} = \sum_{n=1}^{N} n * m_n * \overline{p_n}$, where the mean value of p_n is 0.5 when $\rho_e < \rho_s$.

With the average scan time $\bar{t_c}$, it is possible for us to derive the average waiting delay for an ONU to be polled, $E(W_1)$. We derive it for bandwidth-guaranteed and non- bandwidth-guaranteed ONUs as follows.

During the scan time $\bar{t_c}$, the number of packets arrived at each ONU is: $\lambda_i * \bar{t_c}$. To a specific bandwidth-guaranteed ONU with n entries, if $\lambda_i * \bar{t_c} < n * W_{\max}$, we regard that it could send out all its waiting packets when it is polled. Because this ONU is polled every $(\bar{t_c} / n)$ time, the packets arrived during this time period is $\lambda_i * \bar{t_c} / n$. The average transmission time for all these packets is

$\bar{t} * \lambda_i * \overline{t_c} / n = \rho_i * \overline{t_c} / n$. Then the remaining part of the average polling cycle ($\overline{t_c} / n$), during which this ONU is idle, is: $(1-\rho_i) * \overline{t_c} / n$. It is also the average waiting time for this ONU to be polled. At steady state, the average waiting delay that a packet experienced is

$$E(W_1) = \frac{1}{2}(1-\rho_i) * \overline{t_c} / n . \quad (4)$$

If $\lambda_i * \overline{t_c} > n * W_{\max}$, then this ONU can only send out packets up to the W_{max} when it is polled. The average transmission time for these packets is: $W_{\max} * \bar{t}$. The average time this ONU is waiting for a poll is $\overline{t_c} / n - W_{\max} * \bar{t}$. At steady state, the average waiting delay that a packet experienced is

$$E(W_1) = \frac{1}{2}(\overline{t_c} / n - W_{\max} * \bar{t}) . \quad (5)$$

Since there are m_n bandwidth-guaranteed ONUs that occupy n entries, the m' non-bandwidth-guaranteed ONUs can only be assigned the left ($N - \sum_{n=1}^{N} n * m_n$) entire entries by the OLT. In addition, the superfluous transmission window in one entry that is not fully consumed by a bandwidth-guaranteed ONU could be assigned to a non-bandwidth-guaranteed ONU dynamically. As mentioned before, there are $\overline{P}$ on average superfluous windows for non-bandwidth-guaranteed ONUs. So the average number of entries that one non-bandwidth-guaranteed ONU is assigned can be approximate to $\bar{n} = \frac{1}{m'} * \left[(N - \sum_{n=1}^{N} n * m_n) + \overline{P} \right]$.

The transmission window size for an entire entry is W_{max}, while the superfluous transmission windows have random sizes (less than W_{max}) due to the first ONU's consumption. If the packets arrived at a bandwidth-guaranteed ONU with n entries during ($\overline{t_c} / n$) time is less than the threshold T, then the superfluous window ($W_{\max} - \lambda_i * \overline{t_c} / n$) can be used to poll a non-bandwidth-guaranteed ONU. As the first ONU can take window size up to threshold T, the average window size for first ONU is $T/2$, the average superfluous window size for a non-bandwidth-guaranteed ONU is ($W_{\max} - T/2$). All window size for non-bandwidth-guaranteed ONUs is

$$W_{non} = (N - \sum_{n=1}^{N} n * m_n) * W_{\max} + \overline{P} * (W_{\max} - T/2) \quad (6)$$

The average window size for non-bandwidth-guaranteed ONUs is

$$\overline{W_{non}} = \frac{1}{m'} * W_{non} = \frac{1}{m'} * \left[(N - \sum_{n=1}^{N} n * m_n) * W_{\max} + \overline{P} * (W_{\max} - T/2) \right] \quad (7)$$

To a specific non-bandwidth-guaranteed ONU, if $\lambda_i * \overline{t_c} < \overline{W_{non}}$, then we can regard this ONU will send out all its waiting packets when it is polled. Because this ONU is polled every ($\overline{t_c} / \bar{n}$) time, the packets arrived during this period of time is $\lambda_i * \overline{t_c} / \bar{n}$. The average transmission time for all these packets is $\bar{t} * \lambda_i * \overline{t_c} / \bar{n} = \rho_i * \overline{t_c} / \bar{n}$. Then the remaining part of the average polling cycle ($\overline{t_c} / \bar{n}$), during which this ONU is idle, is $(1-\rho_i) * \overline{t_c} / \bar{n}$. It is also the average time the ONU is waiting for a poll. At steady state, the average waiting delay that a packet experienced is

$$E(W_1) = \frac{1}{2}(1-\rho_i) * \overline{t_c} / \bar{n} . \quad (8)$$

If $\lambda_i * \overline{t_c} > \overline{W_{non}}$, the ONU can only send out packets up to the $\overline{W_{non}} / \bar{n}$ when it is polled. The average transmission time for these packets is $\overline{W_{non}} * \bar{t} / \bar{n}$. The average time for this ONU waiting for a poll is $(\overline{t_c} - \overline{W_{non}} * \bar{t}) / \bar{n}$. At steady state, the average waiting delay that a packet experienced is

$$E(W_1) = \frac{1}{2}(\overline{t_c} - \overline{W_{non}} * \bar{t}) / \bar{n} . \quad (9)$$

We can evaluate $E(W_2)$ of all kinds of ONUs according to different traffic loads. We notice that each ONU could experience three stages of traffic: light, moderate and heavy traffic load, each of which is referred to the effective transmission capacity of the ONU. And different kinds of ONUs could have different values for light, moderate and heavy traffic loads.

When the traffic load ρ is low, that means $\lambda_i \overline{t_c} < n\, W_{max}$ for the bandwidth-guaranteed ONUs or $\lambda_i \overline{t_c} < \overline{W_{non}}$ for the non-bandwidth-guaranteed ONUs. The whole system can be approximated to multiple M/G/1 queues, each of which corresponds to one kind of ONUs. We define $\rho' = \frac{m_n}{M} \rho$ as the traffic load for each kind of ONUs. Then the average waiting delay for the ONU (no matter bandwidth-guaranteed or non-bandwidth-guaranteed ONU) can be expressed based on the assumption that all new arrival data can be transmitted during the same polling cycle they arrive:

$$E(W_2) = \frac{\lambda E(\tau^2)}{2(1-\rho)} = \frac{\rho'}{2(1-\rho')} \bar{t} \quad (10)$$

As the load ρ increases, $\lambda_i \overline{t_c}$ of bandwidth-guaranteed ONUs approaches to $n\, W_{max}$. Bandwidth-guaranteed ONUs experience a moderate traffic load. Now we can assume $\lambda_i \overline{t_c} \geq w\, n\, W_{max}$ where w (around 95%) is the weight of effective transmission capacity, and there is no packet loss for the bandwidth-guaranteed ONUs. Since the new arrival data can't be sent out during the same polling cycle they arrive, we

assume they will be transmitted after another k cycles of polling, where k is the random variable and $\bar{k}$ is the average value. The $E(W_2)$ for the bandwidth-guaranteed ONU can be expressed as:

$$E(W_2) = \frac{\bar{k}}{n}\ \bar{t_c} + \frac{\rho'}{2(1-\rho')}\ \bar{t} \qquad (11)$$

If we regard $\bar{n}$ as the number of entries of a non-bandwidth-guaranteed ONU occupies, then the analysis will be similar to that of the bandwidth-guaranteed ONUs. For non-bandwidth-guaranteed ONUs, when $\lambda_i \bar{t_c} \geq w\ \overline{W_{non}}$, they experience moderate traffic load. And the $E(W_2)$ for non-bandwidth-guaranteed ONU can be expressed as:

$$E(W_2) = \frac{\bar{k}}{n}\ \bar{t_c} + \frac{\rho'}{2(1-\rho')}\ \bar{t} \qquad (12)$$

When the load ρ is heavy, that means $\lambda_i \bar{t_c} \geq n\ W_{max}$. For the bandwidth-guaranteed ONUs, there will be data loss from the ONUs due to the ONU's finite buffer filled with waiting data packets. Let B to be the ONU's buffer size, the new arrival data should be sent out after waiting for enough polling cycles to transmit the previous (B-1) data. So the average waiting delay for bandwidth-guaranteed ONUs can be expressed as:

$$E(W_2) = int\left(\frac{B}{nW_{max}}\right)\ \bar{t_c} + \frac{\rho'}{2(1-\rho')}\ \bar{t} \qquad (13)$$

For non-bandwidth-guaranteed ONUs with heavy traffic load, $\lambda_i \bar{t_c} \geq \overline{W_{non}}$, there will also be packet loss from the ONUs. Then the waiting delay for non-bandwidth-guaranteed ONU is expressed as:

$$E(W_2) = int\left(\frac{B}{\overline{W_{non}}}\right)\ \bar{t_c} + \frac{\rho'}{2(1-\rho')}\ \bar{t} \qquad (14)$$

The $E(W_1)$ in Formula (2) can be substituted by Formula (4) or Formula (5) when bandwidth-guaranteed ONUs are concerned. And Formula (8) or Formula (9) can be used when non-bandwidth-guaranteed ONUs are concerned. The $E(W_2)$ in Formula (2) can be substituted by Formula (10), Formula (11) or Formula (13) when bandwidth-guaranteed ONUs are concerned. And Formula (10), Formula (12) or Formula (14) can be used when non-bandwidth-guaranteed ONUs are concerned. With $E(W_1)$ and $E(W_2)$, it is easy to get *E(A)*. Then, at last, with network parameters, from Formula (1), the point-to-point delay *E(D)* could be evaluated.

V. PERFORMANCE RESULTS

In this section, in order to evaluate the performance of the proposed MAC protocol, we present and compare the results got from a set of simulation experiments and analytical calculation. Extensive discrete event simulation experiments have been conducted with system parameters close to the real world. We present the experiment design as well as the experiment results in the following subsections. The calculation takes the same network parameters and scenario as those in the simulation experiments based on the analysis we have in the previous chapter.

A. Experiments Design

In the experiments, we design our simulation model with reference of the real EPON architecture and system parameters.

We built up the system model as described in Figure 1 & 2. The total number of the ONUs in the EPON is 64. The 64 ONUs are divided into two groups: 20 bandwidth-guaranteed ONUs and 44 non-bandwidth-guaranteed ONUs. The IDs of bandwidth-guaranteed and non-bandwidth-guaranteed ONUs are assumed to be continuous respectively. ONU_1~ONU_{20} are bandwidth guaranteed and ONU_{21}~ONU_{64} are bandwidth non-guaranteed. This assumption makes non-bandwidth-guaranteed ONUs' List simple to implement. It composes of ONUs with continuous IDs that start from ONU_{21} and end at ONU_{64}. Among those bandwidth-guaranteed ONUs, ONU_5 has 20 entries; ONU_8, ONU_{12}, ONU_{17} have 10 entries; ONU_1, ONU_3, ONU_6, ONU_{10}, ONU_{15}, ONU_{18} have 4 entries; all of the 10 remaining ONUs have only 1 entry. As the bandwidth-guaranteed ONUs occupy 84 entries totally, the remaining 16 entries could be assigned to non-bandwidth-guaranteed ONUs. The Entry Table is filled by EDA algorithm. According to EDA, we first allocate entries to ONU_5 that has the maximum number of entries. Then ONU_8, ONU_{12} and ONU_{17} are allocated. Following them, ONU_1, ONU_3, ONU_6, ONU_{10}, ONU_{15} and ONU_{18} will be allocated. Finally, the bandwidth-guaranteed ONUs requiring only one entry will be assigned entries. Let's see an example of ONU_8. The first entry of ONU_8 is entry 8 (number of the ONU ID) based on EDA. Since ONU_8 occupies 10 entries in all 100 entries, the distance between two neighboring entries of ONU_8 should be 100/10 = 10. Other entries of ONU_8 should be entry 18, 28, 38, 48, 58, 68, 78, 88, and 98. The final result of entry assignment is shown in the Table 1.

TABLE I. ENTRY TABLE FOR THE SIMULATION MODEL

1 1	2 12	3 3	4 2	5 5	6 6	7 17	8 8	9 4	10 5
11 10	12 12	13 7	14 9	15 5	16 15	17 17	18 8	19 18	20 5
21 11	22 12	23 13	24 14	25 5	26 1	27 17	28 8	29 3	30 5
31 6	32 12	33 16	34 19	35 5	36 10	37 17	38 8	39 20	40 5
41 15	42 12	43	44 18	45 5	46	47 17	48 8	49	50 5
51 1	52 12	53 3	54	55 5	56 6	57 17	58 8	59	60 5
61 10	62 12	63	64	65 5	66 15	67 17	68 8	69 18	70 5
71	72 12	73	74	75 5	76 1	77 17	78 8	79 3	80 5
81 6	82 12	83	84	85 5	86 10	87 17	88 8	89	90 5
91 15	92 12	93	94 18	95 5	96	97 17	98 8	99	100 5

According to IEEE 802.3, the Ethernet frame can be variable from 64 bytes to 1518 bytes. In the model, we design packet with fixed size as 500 bytes or 4000 bits. Each ONU has its own source generating the fixed size packets, with exponential distributed inter-arrival time, ranging from 0.256ms to 2.56ms making total traffic load of the entire network changing from 0.1 to 1.0 accordingly.

In the simulation model, the OLT generates Grant messages to poll ONUs with granted window size; ONUs produce Reply messages to inform the OLT about the number of the packets in transmission. The Grant message is formatted to contain the ONU ID being polled and granted transmission window size. The Reply message contains its own ID, the number of the packets in transmission and transmission window size being granted by the OLT. The ONU ID occupies 1 byte. The granted window size and the number of the packets in transmission need 2 bytes respectively. Every ONU has a round-trip propagation delay from the OLT to the ONU. The round-trip propagation delay for each ONU is assumed to be randomly (uniformly) over the interval [50 μ s, 100 μ s]. These values correspond to distances between the OLT and ONUs ranging from 10 to 20 km.

In this study, we consider the transmission speeds of the EPON is 1Gb/s, the upstream link rate is divided into 100 bandwidth units, each with 10Mb/s. It makes the Entry Table have 100 entries. The maximum transmission window size W_{max} is set to 15000 bytes or 30 packets. The threshold T is set to $2/3W_{max}$, equal to 20 packets. Every ONU has a finite memory buffer of size, which is set to 10 Mbytes or 20000 packets.

B. Simulation Results

In this section, we present and explain a series of simulation results. In addition to the performance of different kinds of ONUs in the BGP scheme, we also present the performance of the ONUs in the IPACT scheme for comparison. The simulation design for the IPACT model is same as that for the BGP scheme, with the exception that all ONUs in IPACT are treated in the same way as in [7], no difference on bandwidth requirements from the ONUs.

Figure 3 shows the relationship between the average packet delay for different kinds of ONUs' and the entire network load. It shows that, in the BGP scheme, the more entries a bandwidth-guanteed ONU occupies, the less average delay it suffers. The average delays for the bandwidth-uaranteed ONUs with 20 and 10 entries are kept at the lowest values. They are almost unchanged despite of the increase of the traffic load. The curve of average delay for the bandwidth-uaranteed ONUs with 4 entries is kept smooth and low when the load is less than 0.8, after which it increases to a new balanced point when the load reaches 1.0. The bandwidth-guaranteed ONUs with only one entry experience much more average delay than other bandwidth-guaranteed ONUs. The curve increases in the very beginning with the network load of 0.1. It quickly becomes smooth and at quite high value after the network load reaches 0.3. The delay for non-bandwidth-guaranteed ONUs is more than those of bandwidth-

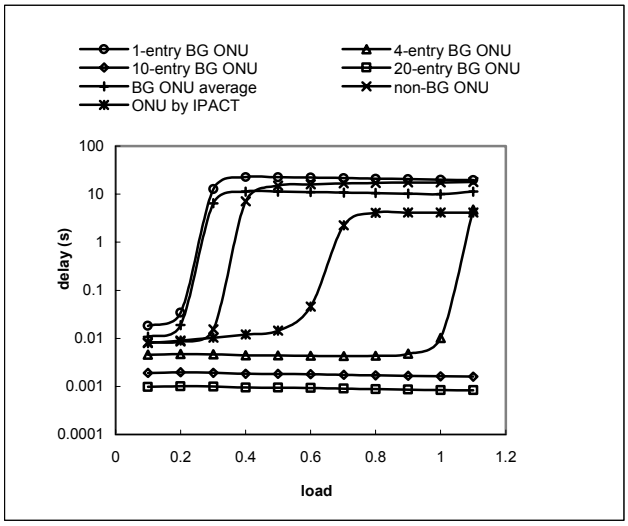

Figure 3. Average delay for different kinds of ONUs

guaranteed ONUs with multiple entries, while less than that of bandwidth-guaranteed ONUs with only one entry. The curve gradually tends to the curve of the bandwidth-guaranteed ONUs with one entry. At last, the two curves converge to a same value. The average delay for the ONUs in the IPACT scheme is very similar to that of the non-bandwidth-guaranteed ONUs in the BGP scheme when the network load is lower than 0.2. However, it increases slowly before the load of 0.6, after which it promptly climbs up. The average delay reaches a balanced point after the load of 0.7.

Figure 4 shows the relationship between the throughput of different kinds of ONUs and the entire network load. All throughputs increase as the network load rises. In the BGP scheme, as the number of non-bandwidth-guaranteed ONUs is larger than that of bandwidth-guaranteed ONUs, the throughput of the former group preponderates over that of the latter one when the traffic is light. With the increase of the network load, the throughput of the latter group exceeds that of the former group at the load of 1.0. The throughput of all ONUs in the IPACT scheme is very similar to that of all ONUs in the BGP scheme when the load is very low. After the load of 0.2, it increases more rapidly than the throughput of all ONUs in the BGP scheme.

Figure 5 shows the relationship between queue length of different kinds of ONUs and the entire network load. The curve shapes of this figure are very similar to the one of average delay vs. network load. In the BGP scheme, the more entries a bandwidth-guaranteed ONU occupies, the shorter queue length existing in its buffer. The queue lengths of the bandwidth-guaranteed ONU with 20 and 10 entries are kept at the lowest values. They increase slowly as the network load increases. The curve of the queue lengths for bandwidth-guaranteed ONUs with 4 entries goes up slowly when the network traffic load is less than 0.8. There is a sharp increase after the load of 0.9. Finally it reaches the bound of buffer size of 20000 packets when the load exceeds 1.0. When traffic is light, non-bandwidth-guaranteed ONUs have shorter queue lengths than bandwidth-guaranteed ONUs with only one

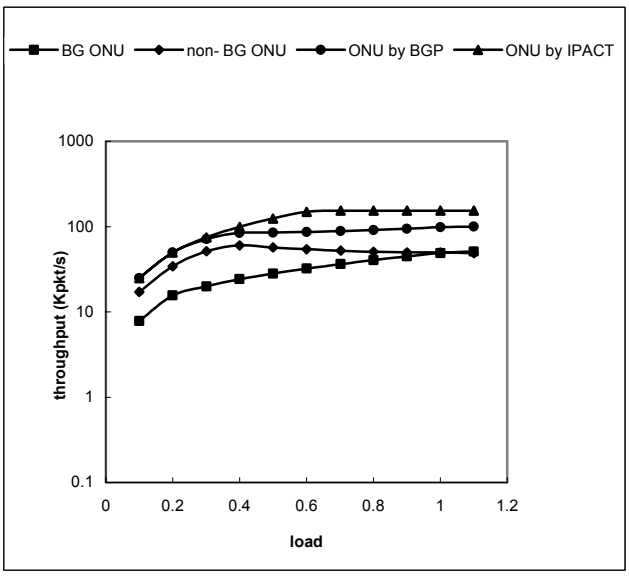

Figure 4. Throughput for different kinds of ONUs

entry. After the network load of 0.4, it catches that of the latter one. At last, queues of both kinds of ONUs come to the limited length of 20000 packets. Among all kinds of ONUs, the bandwidth-guaranteed ONUs with one entry are the earliest ones to reach the upper bound buffer size, when load is only 0.3.

The curve of queue length of ONUs in the IPACT scheme is very similar to that of the non-bandwidth-guaranteed ONUs in the BGP scheme when the network load is lower than 0.2. It increases slowly before the load of 0.6. After that it promptly climbs to the limited buffer of 20000 packets at load of 0.7.

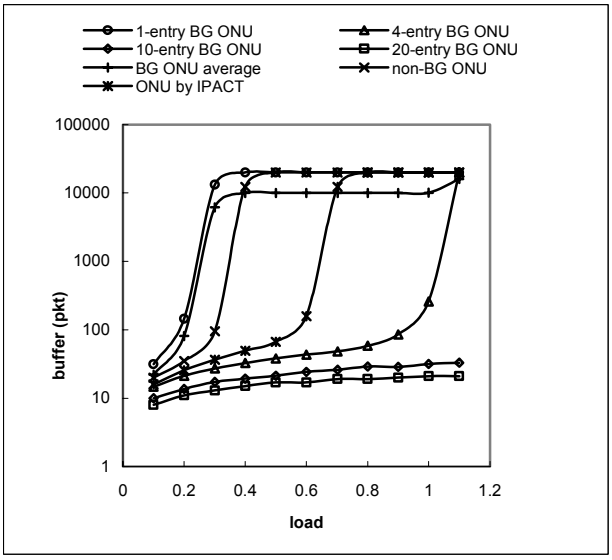

Figure 5. Average queue length for different kinds of ONUs

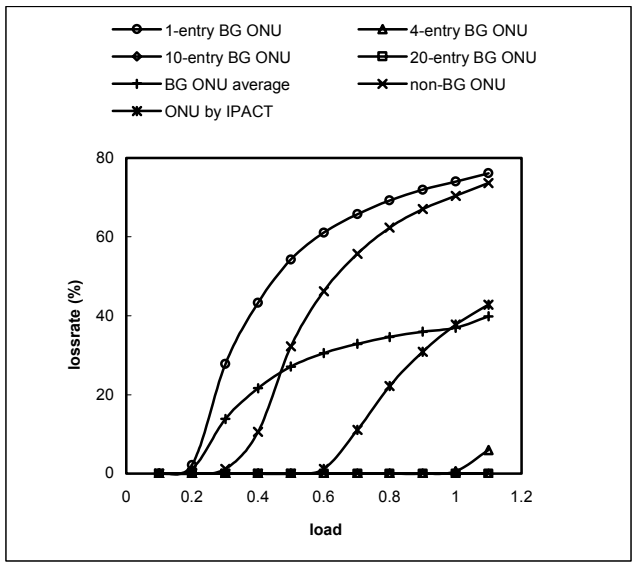

Figure 6. Average loss rate for different kinds of ONUs

Figure 6 shows the relationship between the average loss rate of different kinds of ONUs and the entire network load. In the BGP scheme, the bandwidth-guaranteed ONUs with 20 and 10 entries have no data loss despite the increase of the network load. The bandwidth-guaranteed ONUs with 4 entries experiences data loss after the network load exceeds 0.9. And it rises continuously. The bandwidth-guaranteed ONUs with only one entry begin to drop packets after the load of 0.2, earlier than the non-bandwidth-guaranteed ONUs, for which loss starts to occur at the load of 0.3. The average loss rate of the non-bandwidth-guaranteed ONUs is less than that of the bandwidth-garanteed ONUs with one entry. But it gradually approaches to the later ones. The two loss rates almost become the same value at the load of 1.1.

The ONUs, in the IPACT scheme, have no data loss before the network load of 0.6. After that, they begin to drop data. The loss rate increases continuously and locates between the curves of the bandwidth-guaranteed ONUs with 4 entries and the non-bandwidth-guaranteed ONUs in the BGP scheme.

C. Calculation Comparison

According to the analysis in the previous chapter and network parameters in the experiment design, we can get a group of outcomes on the point-to-point average delay, *E(D)*, by Formulas (4), (5), (10), (11) and (13) for the bandwidth guaranteed ONUs and Formulas (8), (9), (10), (12) and (14) for the bandwidth non-guaranteed ONUs.

Figure 7 compares the analytical outcomes with the simulation results of different kinds of ONUs. We calculate the point-to-point average delay, *E(D)*, for different kinds of ONUs with entire network traffic load changing from 0.1 to 1.0. The analytical results match well with the simulation results, especially, for the bandwidth-guaranteed ONUs with

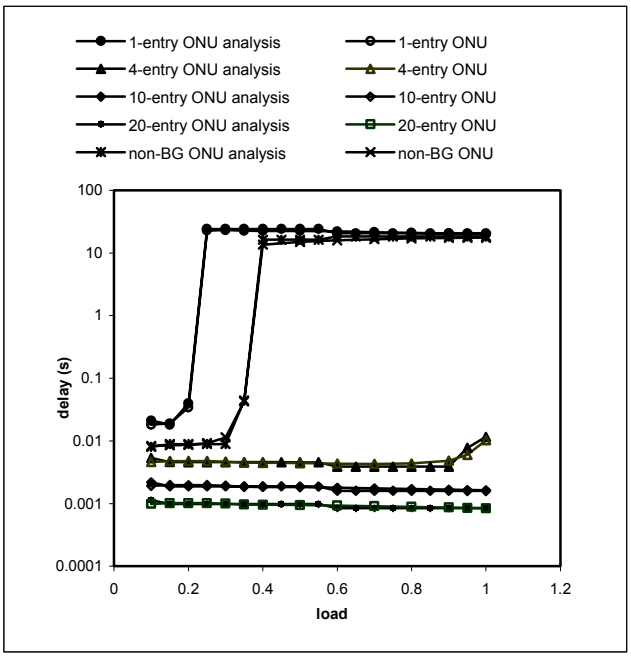

Figure 7. Average buffer length for different kinds of ONUs

1, 10 or 20 entries. There are some discrepancies existing between the analytical and simulation results for the non-bandwidth-guaranteed ONUs simply because some approximations in the analysis have been made to get the value of $\overline{P}$, $\overline{n}$ and $\overline{W_{non}}$, etc. Our analytical model is reasonably well, since it proves the efficiency of the proposed BGP protocol, testifies the rationality of the simulation results, and finally verifies the accuracy of the analytical model itself.

Through the experimental simulations and performance analysis, we have the following conclusions. In the BGP scheme, the more entries a bandwidth-guaranteed ONU occupies, the better performance it can behave in terms of less average delay, higher throughput, shorter queue length and lower loss rate. A bandwidth-guaranteed ONU with multiple entries has better performance than an ONU in IPACT scheme. As a sacrifice, the performance of a non-bandwidth-guaranteed ONU is worse than that of an ONU in IPACT scheme.

VI. Conclusion

In this paper, we have proposed a novel MAC protocol for the EPON to provide efficient network service to different kinds of OUNs with diverse bandwidth requirements. The performance evaluation shows that, although the low-demand users are served on the best-effort basis, the proposed BGP scheme can guarantee bandwidth for high-demand users and could provide QoS service to them. The bandwidth-guaranteed ONUs with multiple bandwidth units have better performance than an ONUs, in the IPACT scheme, which are not explicitly differentiated on bandwidth requirements. The major contribution of this research is that a novel MAC protocol, BGP scheme, has been proposed and proved by a reasonable mathematical model and extensive simulation experiments. The significance of the GBP scheme is that it can provide differentiated service to different users with various bandwidth requirements. It is also a potential MAC protocol to provide QoS service to different traffics. The BGP protocol has also been proved to be a powerful scheme that incorporates the SLA into MAC protocol design. Both the service provider and subscribers can benefit greatly from it.

References

[1] Glen Kramer, Gerry Pesavento, "Ethernet Passive Optical Network (EPON): Building a Next-Generation Optical Access Network", *IEEE Communications Magazine*, Vol. 40, Issue 2, February 2002, pp.66-73

[2] Howard Frazier, Gerry Pesavento, "Ethernet Takes on the First Mile", *IT Professional*, Vol. 3, Issue 4, July-Aug. 2001, pp.17-22

[3] J.D. Angelopoulos, I.S. Venieris, and G.I. Stassinopoulos, "A TDMA Based Access Control Scheme for APONs", *Journal of Lightwave Technology*, Vol. 11 No. 5/6, May/June 1993, pp. 1095-1103

[4] J.D. Angelopoulos, I.S. Venieris, and E.N. Protonotarios, "A Transparent MAC Method for Bandwidth Sharing and CDV Control at the ATM Layer of Passive Optical Networks", Journal of Lightwave Technology, Vol. 14 No. 12, Dec. 1996, pp. 2625-2635

[5] Youngjin M, Changhwan O, and Kiseon K, "A MAC Scheme Based on the Cell Arrival Timing Information for Multimedia Services over ATM-Based PON", *IEICE Transaction Communications*, Vol. E82-B, No. 10, Oct.1999, pp. 1702-1705

[6] Chang-Joon Chae, Elaine Wong, and Rodney S. Tucker, "Optical CSMA/CD Media Access Scheme for Ethernet over Passive Optical Network", *IEEE Photonics Technology Letters*, Vol. 14, No. 5, May 2002, pp.711-713

[7] Glen Kramer, Biswanath Mukherjee, and Gerry Pesavento, "IPACT: A Dynamic Protocol for an Ethernet PON (EPON)", *IEEE Communications Magazine*, Vol. 40, Issue 2, February 2002, pp.74-80

[8] Jong-Tae Park, Jong-Wook Baek, and J. Won-Ki Hong, "Management of Service Level Agreements for Multimedia Internet Service Using an Utility Model", *IEEE Communications Magazine*, Vol. 39 Issue 5, May 2001, pp.100-106

[9] Eric Bouillet, Debasis Mitra, and K.G. Ramakrishnan, "The Structure and Management of Service Level Agreements in Networks", *IEEE Journal on Selected Areas in Communications*, Vol. 20 No. 4, May 2002, pp. 691-699.