Analysis and Derivation of Mean Packet Delay for Gated Service in EPONs

Sailesh Bharati[1] and **Poompat Saengudomlert**[2], Non-members

ABSTRACT

Closed form mathematical expressions of network parameters such as the mean packet delay are useful for evaluating the communication network performance in the design process. This paper provides a simple derivation of closed form expression of the mean packet delay for the gated service dynamic bandwidth allocation algorithm (DBAA) in an Ethernet Passive Optical Network (EPON). Using the queuing analysis framework of a multi-user cyclic polling system with reservation, we derive the mean packet delay expression by modifying the expression for the reservation time component of the total packet delay. The derivation can also be extended for a different polling algorithm in which the gating time is at the beginning of each data interval. We verify our analysis with results from simulation experiments.

Keywords: Cyclic Polling System With Reservation, Dynamic Bandwidth Allocation, EPON, Packet Delay Analysis

1. INTRODUCTION

An Ethernet Passive Optical Network (EPON) is an inexpensive, high capacity, easy-to-upgrade and long operative access network [1]. It removes the capacity bottleneck between a high capacity user or a local area network (LAN) and a backbone network. In its simple architecture, an EPON consists of an Optical Line Terminal (OLT) at a local exchange or a central office (CO) and multiple Optical Network Units (ONUs) at customers' premises as shown in Fig. 1.

In an EPON, a single fiber connects the OLT to a passive $1 \times N/N \times 1$ optical splitter/combiner which divides/combines the signal from/to the OLT as shown in Fig. 1. Wavelength division multiplexing (WDM) is used to separate upstream (ONU-to-OLT) and downstream (OLT-to-ONU) transmissions. While upstream packets are only received by the OLT, downstream packets are broadcast to all ONUs. To avoid collisions among upstream packets from different ONUs, scheduling based on time division multiple access (TDMA) is used by the OLT.

Manuscript received on August 1, 2009 ; revised on January 10, 2010.

[1,2] The authors are with The Asian Institute of Technology, Thailand, and can be reached at , E-mail: sa_ilesh@yahoo.com and poompats@ait.ac.th.

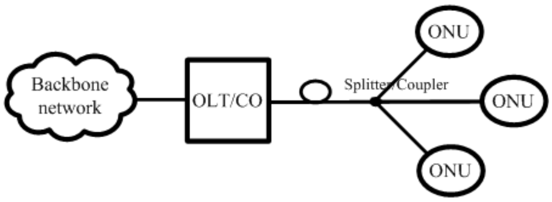


Fig.1: *A simple tree topology for an EPON.*

In an EPON, the Multi-Point Control Protocol (MPCP) [2] is a signaling protocol that facilitates the OLT's allocation of non-overlapping transmission windows (TWs) to ONUs. This process of allocating TWs to ONUs is known as a bandwidth allocation algorithm (BAA). A BAA is considered to be a dynamic BAA (DBAA) if TWs are allocated dynamically on each cycle of allocation as per ONUs' requests, traffic queues, and so on. If the allocation is static in all cycles, then it is regarded as a static BAA.

MPCP uses two 64-byte messages called GATE and REPORT messages. A GATE message is used by the OLT to inform an ONU about the length and the start time of the allocated TW. On the other hand, an ONU informs the OLT about its TW requirement via a REPORT message. Such message exchanges among the OLT and ONUs are generally referred to as polling. MPCP messages are also used to synchronize the clocks of the OLT and ONUs.

Interleaved Polling with Adaptive Cycle Time (IPACT) [3] is a polling scheme in which ONUs gain access to the upstream channel sequentially in a cyclic manner. In this scheme, the OLT transmits a GATE message to the next ONU without waiting for transmissions from previously polled ONUs to arrive. ONUs, on the other hand, starts transmitting its packets at the start time of its TW, which is embedded in the GATE message. At the end of its TW, an ONU transmits a REPORT message reporting its queue length to the OLT. The length and start time of the TW that it gets in the next cycle depends on its queue length and the BAA that the OLT uses.

Several BAAs have been proposed based on IPACT and MPCP. These schemes vary from simple to complex, static to dynamic. Each of these schemes focuses on some definite objectives. Fairness among users, high channel utilization, low mean packet delay, delivering Differentiated Services (DiffServ, i.e., class-based services to various traffic classes) for quality of service (QoS) and bandwidth guarantee (BG) services as per service level agreements (SLAs) are some important objectives of BAAs. These algorithms can be classified into at least six different service types [3, 4], namely fixed, gated, limited, constant credit, linear credit and elastic. Among these six schemes the only static scheme is the fixed BAA and the rest are all DBAAs.

While a large number of DBAAs have been proposed together with performance evaluation based on computer simulations, few analytical results are available for DBAAs in EPONs [5–8]. In [5], the authors model IPACT mathematically under the gated service and develop a closed form expression for the mean granted TW size. In [6], the authors analyze and derive an expression for the mean packet delay for the gated service with one ONU but could not extend for multiple ONUs accurately. In [7], the authors derive a closed form expression of the mean packet delay for the gated service with the gating time at the beginning of each TW, which is different from the actual gated service of IPACT whose gating time is at the end of each TW. In addition, the analytical results in [7] are not verified with any simulation or experiment.

In [8], the authors use the mean value analysis approach to derive the mean queue length in the gated service. Later, Little's theorem [9] is used to find the expression for the mean packet delay. The derived expression was also verified using simulation. This paper considers mathematical analysis of the performance of an EPON using the gated service. We use queuing theory for the analysis and verify the results by simulation experiments.

This paper is structured as follows. Section 2 discusses relevant queuing analysis of various types of polling system based on the gating time, and points out basic differences between the traditional polling systems and an EPON. Section 3 discusses the system model and various assumptions for the analysis that follows. In section 4, we derive a closed form expression of the mean packet delay for the gated service. We validate our analysis with simulation results in section 5. Section 6 provides a summary of our contribution.

2. QUEUING ANALYSIS OF POLLING SYSTEMS

In the traditional cyclic polling system with reservation [9], each time slot used by a single user consists of two intervals, which are a reservation interval followed by a data interval. In a reservation interval, the corresponding user transmits a control message to take over or reserve the channel for the data interval that follows. The choice of packets to be transmitted in a particular data interval differentiates the system types among gated, exhaustive, and partially gated systems [9] as described below. Note that the gated system should not be confused with the gated service for DBAA in an EPON.

- **Exhaustive system**: In this system, a reservation is made for those packets which arrived before the end of the data interval.
- **Partially gated**: In this system, a reservation is made for those packets which arrived before the end of the reservation interval.
- **Gated system**: In this system, a reservation is made for those packets which arrived before the beginning of the reservation interval.

All packets wait in their queues before being transmitted. We refer to the waiting time of a packet in a queue as the packet delay, which is a random variable denoted by W, and denote its mean by $\overline{W}$. The packet delay can be divided into three components as mentioned below.

- **Residual Time** (R): is the remaining time until the service time of a packet or reservation that is currently being served is complete.
- **Time spend in queue** (Q): is the time for the transmissions of all packets that are currently in the queue ahead of the packet of interest.
- **Reservation time** (Y): is the total time of reservation slots for the packet. It can be measured in terms of the number of reservation intervals that the packet experiences.

Each of this component can be derived individually. The derived components can be added to get the final expression for the mean packet delay as $\overline{W} = \overline{Q} + \overline{Y} + \overline{R}$ [9].

2.1 N-User $M/G/1$ System with Reservation

Consider a cyclic polling system in which time slots are allocated to N users in a round robin fashion such that, in a cycle, slots are allocated to user 1, user 2, and so on up to user N. Consider all users to be symmetric in terms of the statistics of packet arrivals and service times. Let the service time of each packet be random with mean $\overline{X}$ and second moment $\overline{X^2}$. Let each user's reservation time be random with mean $\overline{V}$, second moment $\overline{V^2}$ and variance ${\sigma_v}^2$. All service times and reservation times are independent. Packets from all users arrive according to a Poisson process of rate λ, i.e., λ/N is the arrival rate from a single user. Let $\rho = \lambda\overline{X}$ denote the total traffic load. This cyclic polling system can be viewed as an $M/G/1$ queue with reservation and analysis of such $M/G/1$ queue in a gated system with reservation yields [9]

$$\overline{Q} = \rho\overline{W}, \tag{1}$$

$$\overline{Y} = \frac{(N+1)}{2}\overline{V}, \tag{2}$$

$$\overline{R} = \lambda\frac{\overline{X^2}}{2} + \frac{(1-\rho)\overline{V^2}}{2\overline{V}}, \tag{3}$$

$$\overline{W} = \lambda\frac{\overline{X^2}}{2(1-\rho)} + \frac{(N+2-\rho)}{2(1-\rho)}\overline{V} + \frac{{\sigma_v}^2}{2\overline{V}}. \tag{4}$$

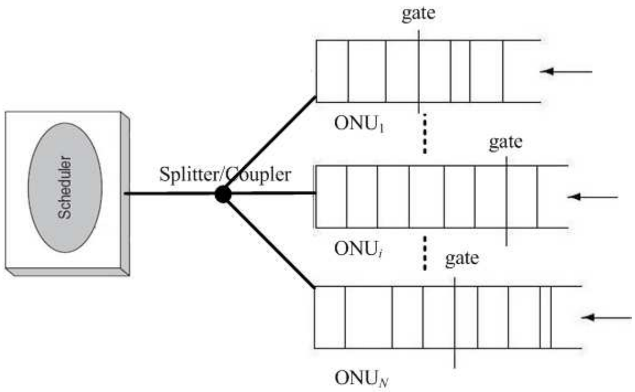


Fig.2: *EPON model with single-stage buffer.*

Under the above conditions, the expressions for $\overline{R}$ and $\overline{Q}$ are the same for all three systems, i.e., gated, partially gated and exhaustive.

2.2 Additional Considerations for EPON

The IPACT algorithm for EPON can be viewed as polling considered in the previous section. In this polling system, each ONU sends to the OLT a REPORT message, which can be considered as a reservation request for a TW for the next scheduling cycle. The time epoch when a REPORT message is sent is known as the gating time. The requested TW is equal to the ONU's queue size at the gating time. Unlike the polling system discussed in the previous section, a reservation by a REPORT message is done after (instead of before) the data interval. As a result, the above analysis cannot be directly applied to EPON.

Although the residual time ($\overline{R}$) and the time spent in a queue ($\overline{Q}$) are the same between an N-user $M/G/1$ system with reservation and an EPON, the reservation time ($\overline{Y}$) is different due to different ways of using the gating time. We shall show that, in an EPON, a packet will typically experience longer reservation time compared with the traditional N-user $M/G/1$ system with reservation. Details of this of extra reservation time will be discussed in section 4.

3. SYSTEM MODEL

Consider an EPON with single-stage buffers at ONUs, as shown in Fig. 2. It consists of N ONUs that are identical in terms of the statistics of packet arrivals and service times. We focus on upstream transmissions, which are more challenging than downstream transmissions since transmitted packets from different ONUs could potentially collide.

Each ONU is connected to the OLT via a common fiber link between the splitter/coupler and the OLT. As in [3], the scheduler in the OLT performs cyclic inter-ONU scheduling, such that the OLT starts a polling cycle by serving ONU_1 followed by serving ONU_2 and so on up to ONU_N. The scheduling order of serving ONUs is the same for every polling cycle. Each ONU queue uses a first-in-first-out (FIFO) scheme to select packets for transmissions in its TW. Assume that each ONU has a buffer large enough so that there is no packet drop.

When a packet arrives at an ONU, the ONU stores that packet in its buffer. Only the packets which were reported in the last reservation are eligible to be transmitted in the ONU's current TW. As soon as these packets have been transmitted, i.e., at the end of the data interval or equivalently at the gating time, the ONU transmits a REPORT message informing the OLT about its remaining queue size at the gating time. As indicated in Fig. 2, the gating process can be viewed as setting up a gate to allow only packets ahead of the gate to be transmitted in the next TW. A REPORT message and the guard time to set up or turn on/off the hardware of adjacent ONUs form a reservation interval of an ONU [10, 11].

Packet arrivals to each ONU's queue form a Poisson process with rate λ/N. The packet service times are random with the first and second moments equal to $\overline{X}$ and $\overline{X^2}$. The reservation times are random with the first and second moments equal to $\overline{V}$ and $\overline{V^2}$. All service and reservation times are independent. Denote the overall traffic load by $\rho = \lambda\overline{X}$. The mean reservation time is the sum of the mean guard time $\overline{t_g}$ and the time to transmit a REPORT message, i.e.,

$$\overline{V} = \overline{t_g} + 8L_{REPORT}/C_{UPSTREAM}, \tag{5}$$

where L_{REPORT} is the size of a REPORT message (in byte), and $C_{UPSTREAM}$ is the upstream transmission capacity of a fiber (in bps).

In an EPON, the OLT may or may not allocate a TW equal to what was requested by an ONU. In this paper, we shall focus on the gated service [3, 4] in which the OLT allocates TW to an ONU that is equal to what it requested in its last REPORT message.

4. MEAN PACKET DELAY OF GATED SERVICE IN EPON

In this section, we shall derive a closed form expression of the mean packet delay in gated service for DBAA in an EPON. As discussed in section 2, because of the nature of gating in EPON, we focus on the gated system for reporting packets in queues. Since multiple ONUs make reservations for TWs with the OLT in a cyclic manner, our system can be modeled as a multiuser $M/G/1$ queueing system with reservation. We decompose the mean packet delay into three components as mentioned in section 2. The delay components $\overline{Q}$ in (1) and $\overline{R}$ in (3) are still applicable. However, the reservation time component $\overline{Y}$ needs to be modified for EPON's gated service.

The gated service in EPON can be considered as the gated system with the gating time after the data interval. Compared to the gated system discussed in section 2 with the gating time before the data interval, a packet in an EPON experiences additional reservation. An expression for this additional reservation time is derived in the following theorem.

Theorem 1: The additional mean reservation delay, denoted by $\triangle\overline{Y}$, for the gated service in an EPON compared to the gated system based on a multiuser $M/G/1$ queue with reservation is equal to $(N-1)\overline{V}$.

Proof: Due to the different natures of gating, in EPON's gated service, a newly arrived packet will not be served in the following data interval of its ONU as in the traditional gated system. In the next reservation interval of its ONU, the packet will be reported to the OLT via a REPORT message. Then, this packet will be served in its ONU's data interval in the next cycle.

Fig. 3 shows the basic difference between an N-user $M/G/1$ queuing system with reservation and EPON's gated service in six possible cases. The time period indicated with "b" indicates a time interval during which a packet of interest arrives. In this time interval, a packet arrives while the OLT is serving its ONU or other ONU either with a data interval or a reservation interval. In addition, ONU which is idle i.e., not being serve by the OLT, might have been served or waiting for its turn in the current service cycle. Hence, during the time interval "b", a packet can arrive in one of six possible types of time intervals. Detailed explanations of a reservation delay experienced by a packet in all six cases are given below and illustrated in Fig. 3.

Case 1: When a packet arrives in its ONU's data interval, in the traditional gated system, it will be reported in the next reservation interval and will be served in the data interval that follows. Hence, a packet experiences reservation delay of $N\overline{V}$ in the gated system. In EPON's gated service, the packet will be reported in the next reservation interval following the packet arrival. It will be served during a data interval of its owner in the next cycle. This results in a reservation delay of $N\overline{V}$. It follows that $\triangle\overline{Y}=0$ in this case.

Case 2: When a packet arrives in its ONU's reservation interval, in the traditional gated system, a packet experiences reservation delay of $N\overline{V}$ as in case 1. On the other hand, in EPON's gated service, the packet will be reported in the reservation interval in the next cycle and is served in the data interval of the next cycle after the report. This results in a reservation delay of $(2N-1)\overline{V}$ in EPON's gated service. It follows that $\triangle\overline{Y}=(N-1)\overline{V}$ in this case.

Cases 3 and 4: Consider a packet belonging to ONU_i arriving in the data interval of ONU_j. In the traditional gated system, the packet will be reported in the next reservation interval of ONU_i and will be served in the following data interval in the same cycle. Hence, the reservation delay is the sum of reservation intervals of all ONUs that are after ONU_j and up to (including) ONU_i in the scheduling order. This sum will be $(N-j+i)\overline{V}$ if $j>i$ and $(i-j)\overline{V}$ if $i>j$.

However, EPON's gated service, the packet will be reported to the OLT after reservation intervals of ONU_j up to ONU_i in the scheduling order. Then after $(N-1)$ reservation intervals after that of ONU_i, the packet of interest will be transmitted. Hence, the total mean reservation time experienced by the packet is $(2N-j+i)\overline{V}$ if $j>i$ and $(N-j+i)\overline{V}$ if $i>j$. It follows that $\triangle\overline{Y}=N\overline{V}$ in these cases.

Cases 5 and 6: Consider a packet belonging to ONU_i arriving in the reservation interval of ONU_j. In the traditional gated system, the packet experiences the reservation delay of $(N-j+i)\overline{V}$ if $j>i$ and $(i-j)\overline{V}$ if $i>j$ as in cases 3 and 4 respectively. However, in EPON's gated service, the packet will experience one fewer reservation interval than that of cases 3 and 4. Hence, the total mean reservation time experienced by the packet is $(2N-j+i-1)\overline{V}$ if $j>i$ and $(N-j+i-1)\overline{V}$ if $i>j$. It follows that $\triangle\overline{Y}=(N-1)\overline{V}$ in these cases.

Table 1 shows a summary of the additional reservation time experienced by a packet in EPON's gated service compared to the traditional gated system based on an N-user $M/G/1$ queue with reservation. From Table 1, any packet that arrives during a data interval of another ONU experiences additional reservation time $N\overline{V}$. If its arrival is during a reservation interval of another ONU, the packet experiences additional reservation time $(N-1)\overline{V}$. Finally, any packet that arrives during a reservation interval of its own ONU experiences additional reservation time $(N-1)\overline{V}$.

Since a packet arrives with probability ρ and $1-\rho$ in data and reservation intervals respectively, the mean increase in reservation time $\triangle\overline{Y}$ is

$$\triangle\overline{Y}=\rho\frac{N-1}{N}N\overline{V}+(1-\rho)(N-1)\overline{V}=(N-1)\overline{V}, \quad (6)$$

which concludes the proof. ■

From (2) and (6), the mean reservation time for an EPON's gated service is given as

$$\overline{Y}=\frac{N+1}{2}\overline{V}+\triangle\overline{Y}=\frac{3N-1}{2}\overline{V}. \quad (7)$$

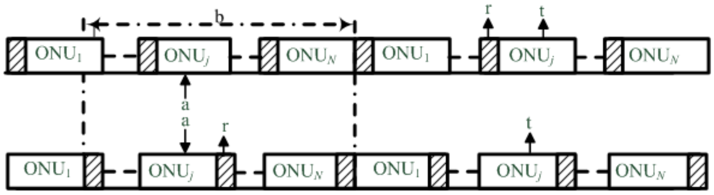
Adding (1), (3) and (7), we can find the mean packet delay experienced by an arbitrary packet in the gated service in an EPON as stated below.

Theorem 2: The mean packet delay for an EPON with the gated service is

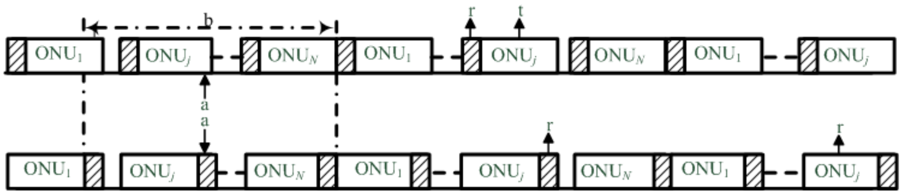
$$\overline{W}=\frac{\lambda\overline{X^2}+(3N-\rho)\overline{V}+(1-\rho)\sigma_v{}^2/\overline{V}}{2(1-\rho)}. \quad (8)$$

4.1 Comparison with Existing Result

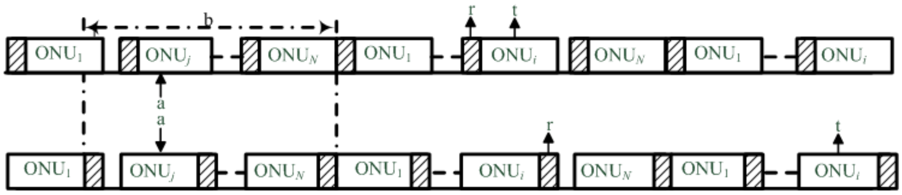
The mean packet delay in (8) is for a standard EPON with the gating time after the data interval.



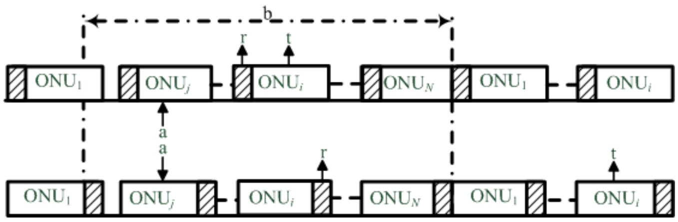
(a) Case 1: Packet arrival in its ONU's data interval



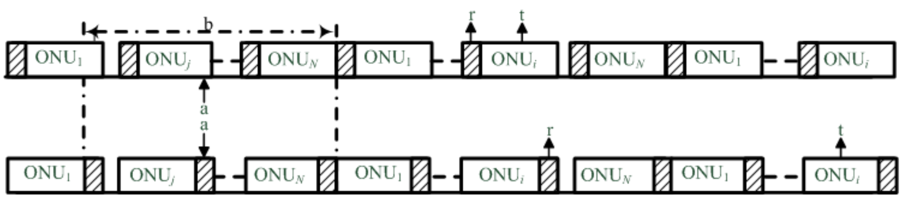
(b) Case 2: Packet arrival in its ONU's reservation interval



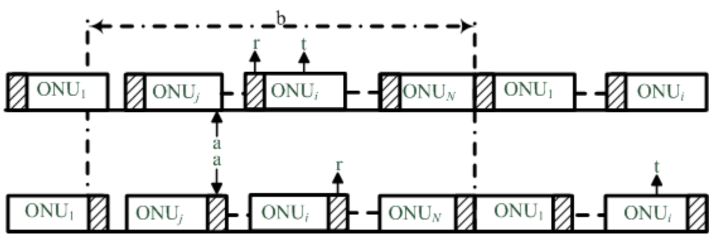
(c) Case 3: Packet of ONU_i arriving in data interval of ONU_j $(j > i)$



(d) Case 4: Packet of ONU_i arriving in data interval of ONU_j $(i > j)$



(e) Case 5: Packet of ONU_i arriving in reservation interval of ONU_j $(j > i)$



(f) Case 6: Packet of ONU_i arriving in reservation interval of ONU_j $(i > j)$

Fig.3: *Comparison of the gated system based on N-user $M/G/1$ queue with reservation (top) and the gated service in EPON (bottom) (a: arrival time, r: reported time and t: transmission time)*

Table 1: *Mean reservation delay difference between the N-user $M/G/1$ queue with reservation and EPON for the gated service in EPON*

Case	Packet arrival during:	Reservation delay	$\triangle Y$
	Own data interval		
1	N-user $M/G/1$ queue with reservation EPON	$N\overline{V}$ $N\overline{V}$	0
	Own reservation interval		
2	N-user $M/G/1$ queue with reservation EPON	$N\overline{V}$ $(2N-1)\overline{V}$	$(N-1)\overline{V}$
	Data interval of ONU_j with packet belonging to $\text{ONU}_i (j > i)$		
3	N-user $M/G/1$ queue with reservation EPON	$(N-j+i)\overline{V}$ $(2N-j+i)\overline{V}$	$N\overline{V}$
	Data interval of ONU_j with packet belonging to $\text{ONU}_i (i > j)$		
4	N-user $M/G/1$ queue with reservation EPON	$(i-j)\overline{V}$ $(N-j+i)\overline{V}$	$N\overline{V}$
	Reservation interval ONU_j with packet belonging to $\text{ONU}_i (j > i)$		
5	N-user $M/G/1$ queue with reservation EPON	$(N-j+i)\overline{V}$ $(2N-j+i-1)\overline{V}$	$(N-1)\overline{V}$
	Reservation interval ONU_j with packet belonging to $\text{ONU}_i (i > j)$		
6	N-user $M/G/1$ queue with reservation EPON	$(i-j)\overline{V}$ $(N-j+i-1)\overline{V}$	$(N-1)\overline{V}$

In [7], the authors assume the gating time before the data interval and derive the expression for the mean packet delay as

$$\overline{W} = \frac{\lambda\overline{X^2} + (3N+\rho)\overline{V} + (1-\rho)\sigma_v{}^2/\overline{V}}{2(1-\rho)}. \quad (9)$$

We now show that the delay expression in (8) can be modified to yield the expression in (9). To do so, we update the mean reservation time $\overline{Y}$. The consequence of having the gating time before the data interval is additional delay for packets that arrive during the data intervals of their ONUs. These packet arrivals occur with probability ρ/N and face additional reservation time $N\overline{V}$, resulting in a further increase of $\overline{Y}$ by $\rho\overline{V}$, i.e., $\triangle\overline{Y} = (N-1+\rho)\overline{V}$ compared to the traditional N-user $M/G/1$ queue with reservation in section 2. Hence, the mean reservation time for this gating time is

$$\overline{Y} = \frac{N+1}{2}\overline{V} + (N-1+\rho)\overline{V} = \frac{3N+2\rho-1}{2}\overline{V}. \quad (10)$$

Adding (1), (3) and (10) yields the mean packet delay expression in (9).

Despite the same result, it is worth pointing out that our derivation is much simpler compared to the method used in [7], which requires the use of pseudo conservation law [12] and probability generating functions.

Note that (8) and (9) indicates that the stability condition for the gated service is $\rho < 1$. Hence the system load ρ must be maintained below 1 for stable operations of a system. As ρ approaches 1, which is not desirable, the mean packet delay approaches infinity.

5. SIMULATION RESULTS

In this section, we present results from simulation experiments and compare them with analytical results obtained in section 4. The default network parameters used in simulations are given in Table 2.

Table 2: *Default network parameters*

Parameter	Value
N: Number of ONU	8, 16, 32
$C_{UPSTREAM}$: Line rate of EPON	1 Gbps
Inter frame gap	12 bytes
Buffer size of a queue	∞
$\bar{t}_g$: Mean guard time (constant)	1 μs
L_{REPORT}: Size of REPORT message	64 bytes

We consider constant and symmetric guard time such that it is same for all ONUs and is with zero variance. From (5), the reservation interval has mean $\overline{V} = 1512$ ns and variance $\sigma_v{}^2 = 0$.

For simulation, we considers the gated service with the REPORT message at the end and at the beginning of its TW. MATLAB software is used to implement the simulation model. The mean packet delay and the mean cycle time are used to validate the results obtained from analysis. In what follows, each data point in a plot is obtained after averaging the corresponding parameter over a total of 100,000 packets.

The packet payload sizes vary from 64 to 1518 bytes with the distribution based on [2, 7] as follow: 64 bytes (47%), 300 bytes (5%), 594 bytes (15%), 1300 bytes (5%), and 1518 bytes (28%). Assuming the inter-frame gap of 12 bytes, the corresponding service times for these packet sizes are 0.608 μs, 2.496 μs, 4.848 μs, 10.496 μs, and 12.240 μs, with the mean $\overline{X} = 5.090$ μs and the second moment $\overline{X^2} = 51.468$ $(\mu\text{s})^2$. Assume that packet arrivals to each ONU form a Poisson process with rate λ/N. We vary the total traffic load $\rho = \lambda\overline{X}$ from 0.1 to 0.95.

Fig. 4 compares the mean packet delay obtained from simulations and theorem 2 for the gated service in EPON. Fig. 5 compares the mean packet delay obtained from simulations and (9) for the gated service with a reservation interval before data interval in EPON. In addition, Fig. 6 compare the mean cycle time in gated service in EPON.

For the analytical mean cycle time, denoted by $\overline{C}$, we use the known expression in [7, 12] for the gated service, as given below.

$$\overline{C} = \frac{N\overline{V}}{(1-\rho)} \tag{11}$$

In Fig. 4−6, we observe a reasonable amount of consistency between simulation and analytical results, which verify the analytical results in section 4.
.

In general, the mean packet delay and the mean cycle time increases with the traffic load. From fig. 4 and 5, if we compare the mean packet delay in the two gated services, a packet experiences more delay in the gated service with the reservation before the data interval. This is because of those packets which arrive in a data interval of their own ONU, i.e., after the gating time and before the end of data interval of their own ONU. Putting the reservation before the data interval increases the mean reservation time by $\rho\overline{V}$ overall, causing an increase of $\rho\overline{V}/(1-\rho)$ in the mean packet delay.

6. CONCLUSIONS

We derived a closed form expression of the mean packet delay for an EPON with the gated service as the DBAA. The derivation is based on modeling an EPON as a multi-user $M/G/1$ queue with reservation. Since an EPON differs from the traditional $M/G/1$ queue with reservation because (i) it is not a broadcast system for upstream transmissions, and

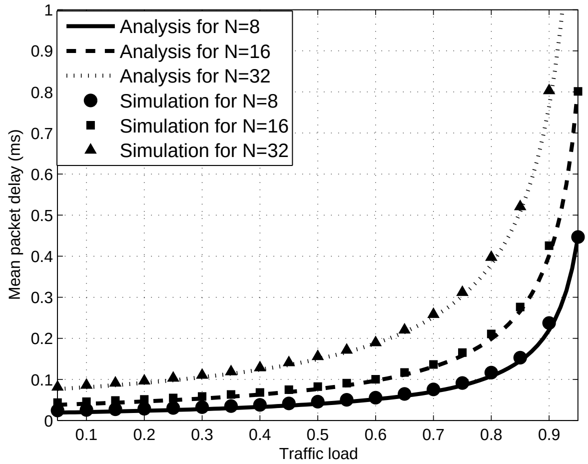


Fig.4: *Mean packet delay for the gated service in EPON with reservation after data interval.*

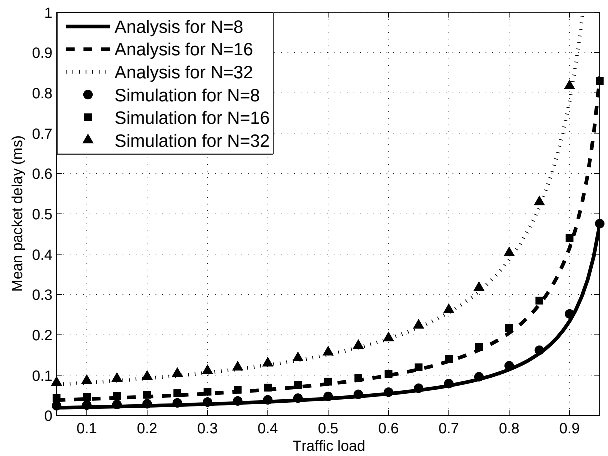


Fig.5: *Mean packet delay for the gated service in EPON with reservation before data interval.*

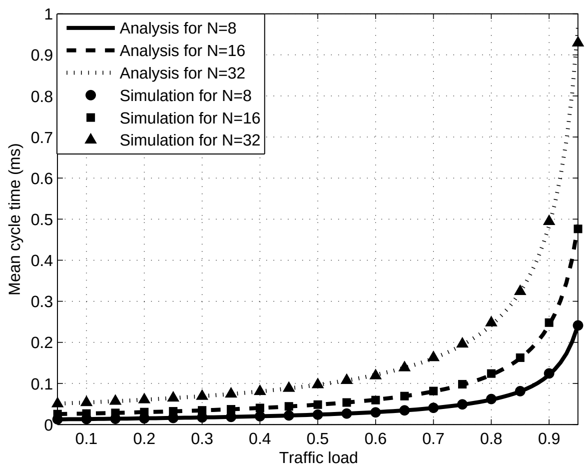


Fig.6: *Mean cycle delay for the gated service.*

(*ii*) the reservation interval is after the data interval in an allocated TW, we modified the mean packet delay analysis to take into account these differences. In addition, we further modified the analysis to handle a variation in a DBAA in which the gating time is at the beginning of a TW and obtain the same mean packet delay expression as in [7]. Compared to the method in [7], our work provides an alternative derivation that is much simpler and able to model an EPON more closely. The analytical expression of the mean packet delay was later verified with results from simulation experiments.

7. ACKNOWLEDGEMENT

We would like to thank Asian Institute of Technology (AIT) and the Government of Finland for their financial supports to conduct this research. We would also like to thank Assoc. Prof. Tapio Erke, AIT, and Assoc. Prof. Teerapat Sanguankotchakorn, AIT, for their helpful comments related to this work.

References

[1] G. Kramer and G. Pesavento, "Ethernet passive optical network (EPON): Building a next-generation optical access network", *IEEE Communications Magazine*, vol. 40, no. 2, pp. 66-73, February 2002.

[2] G. Kramer et al., "On supporting differentiated classes of service in Ethernet passive optical networks", *Journal of Optical Networking*, vol. 1, nos. 8 & 9, pp. 280-298, 2002.

[3] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: a dynamic protocol for an Ethernet PON (EPON)", *IEEE Communications Magazine*, vol. 40, no. 2, pp. 74-80, February 2002.

[4] J. Zheng and H.T. Mouftah, "Media access control for Ethernet Passive Optical Networks: an overview", *IEEE Communications Magazine*, vol. 2, no. 2, pp. 145-150, February 2005.

[5] S. Bhatia, D. Garbuzov, and R. Bartos, "Analysis of the Gated IPACT Scheme for EPONs",*Proceedings of the IEEE International Conference on Communications*, vol. 6, pp. 2693-2698, 2006.

[6] F. Aurzada et al., "Delay analysis of Ethernet Passive Optical Networks with gated service", *Journal of Optical Networking*, vol. 7, pp. 25-41, 2008.

[7] C.G Park, D.H. Han, and B. Kim, "Packet delay analysis of dynamic bandwidth allocation scheme in an Ethernet PON", *Lecture Notes in Computer Science (LNCS)*, vol. 3420, pp. 161-168, 2005.

[8] M. Thanh Ngo, A. Gravey, and D. Bhadauria, "A mean value analysis approach for evaluating the performance of EPON with Gated IPACT", *Proceedings of International Conference on Optical Network Design and Modeling*, pp. 1-6, March 2008.

[9] D.P. Bertsekas, and R.G. Gallager,*Data Networks, 2^{nd} Edition*, Prentice-Hall International, 1992.

[10] G. Kramer, *Ethernet Passive Optical Networks*, Mc-Graw Hill, 2005.

[11] C.M. Assi, Y. Ye, S. Dixit, and M.A. Ali,"Dynamic bandwidth allocation for quality-of-service over EPON", *IEEE Journal in Selected Areas in Communications*, vol. 21, no. 9, pp. 1467-1477, 2003.

[12] O.J. Boxma and W.P. Groenendijk, "Pseudo-conservation laws in cyclic-service systems," *Journal of Applied Probability*, vol. 24, no. 4, pp. 949-964, 1987.

Sailesh Bharati received the B.Eng degree in electronics and communication engineering from Tribhuvan University, Nepal in 2005 with university fellowship. He then received M.Eng degree in information and communication technologies at Asian Institute of Technology, Thailand in 2006 with scholarship from Asian Institute of Technology (AIT) and The Government of Finland for. His research interest includes performance evaluation, mathematical modeling and analysis, queuing theory, and resource allocation problems in communication networks.

Poompat Saengudomlert received the scholarship from His Majesty the King of Thailand to obtain the B.S.E. degree in electrical engineering from Princeton University, USA, in 1996. He then received the M.S and Ph.D degrees, both in electrical engineering and computer science, from Massachusetts Institute of Technology (MIT), in 1998 and 2002 respectively. Currently, he is an Assistant Professor at the Asian Institute of Technology (AIT), Thailand. His research interest includes communication theory, data networks, and resource allocation problems.