Delay Analysis for Ethernet Long-Reach Passive Optical Networks

Mohammad S. Kiaei, Kerim Fouli, Michael Scheutzow, Martin Maier, Martin Reisslein, Chadi Assi

***Abstract*—Designing low latency polling schemes is one of the most important parts for passive optical networks (PONs), particularly for long-reach PONs (LR-PON) which suffer from long propagation delays. Sophisticated and efficient bandwidth allocation mechanisms are required to cope with the imposed transmission delay in LR-PONs. In this work, we evaluate three dynamic bandwidth allocation methods in terms of transmission delay. Namely, we consider conventional or interleaved polling for traditional PON and two recently introduced scheduling paradigms for next generation LR-PON, i.e., multi-thread polling (MT-P) and real-time polling (RT-P). We examine various flavors of each scheduling method and investigate their shortcomings and advantages in a LR-PON setting. Furthermore, we provide an analytical framework for obtaining packet delay in an enhanced version of RT-P method. The simulation results highly match the analysis for this framework. Also, our results indicate that RT-P method significantly reduces frame delay in LR-PONs compared to MT-P and conventional polling methods.**

I. Introduction

LR-PONs are poised to be one of the major trends in the evolution of access-metro optical networks. They essentially have the same topology as PONs and are characterized by a longer distance between the optical line terminal (OLT) and the optical network units (ONUs) as well as a larger number of ONUs. Although both legacy PON and LR-PON use one upstream and one downstream channel, the maximum reach of standardized PONs is 20km whereas LR-PONs are expected to span lengths of 100km. The shift from PONs to LR-PONs translates into longer propagation delays and round-trip times (RTTs) from the OLT to the ONUs. This stipulates more sophisticated and efficient grant scheduling methods while taking care of the imposed transmission delay.

The problem of scheduling and bandwidth allocation in PONs has been widely investigated during past few years. Several dynamic bandwidth allocation (DBA) methods have been introduced for various PON architectures [1]–[4]. Most of previous studies investigate DBA methods through simulations; while, a few studies employ formal mathematical analysis to provide better insights into characteristics of DBA approaches [5]–[8]. An analysis has been provided in [5] for evaluating the collision probability of the messages sent by ONUs to the OLT during registration which establishes the first communication between new ONUs and the OLT. The analysis provides contention window sizes for an efficient registration process. Lannoo et al. [6] derived the mean cycle length and the approximative mean delay by using a Markov chain model for the cycle length in a multi-ONU EPON with reporting at the end of the upstream transmission. This analysis was extended in a subsequent work by Aurzada *et al.* [7], where an exact closed-form expression was derived for the delay in a gated-service single-ONU EPON with reporting at the end. They also derived a Markov chain model for the cycle length, the exact delay in a single-ONU EPON, and an approximation of the delay in a multi-ONU EPON for gated-service EPONs with reporting at the beginning of the upstream transmission. This work was followed by [8] which provides a probabilistic analysis for the throughput and packet delay of potential next generation PONs (NG-PONs).

Most of studies on delay analysis for PONs have focused on ordinary PONs which normally span less than 20km. A formal mathematical analysis for LR-PONs is lacking in the existing literature. A few polling methods have been introduced in recent years which can be used for LR-PONs. The first candidate to address the delay problem in LR-PONs is multi-thread polling (MT-P) [9] which creates multiple interleaved polling cycle instances (i.e. "threads"). Each thread is a complete polling cycle process where all ONUs are polled once. The second polling candidate for LR-PONs is real-time polling (RT-P) [10] which uses an additional separate reporting channel that allows ONUs to report increases in their queue size in real-time. These polling schemes will be further explained in Section II.

To date, no research has been reported on investigating the effect of different polling and system parameters on frame delay in LR-PON. In this work, we will design a framework for quantifying transmission delay and exploring the roles played by various system parameters in a LR-PON. We evaluate the advantages of recently introduced scheduling paradigms for next generation LR-PON and compare them with interleaved polling with adaptive cycle time (IPACT) [1] which is a pioneer algorithm for traditional PONs. We examine various flavors of each scheduling method and investigate their shortcomings and advantages in a LR-PON setting. Furthermore, we provide an analytical framework for obtaining packet delay in an enhanced version of RT-P method. The rest of this paper is structured as follows. In Section II, we address the problem statement and describe the DBA methods which are considered in this work. Delay analysis framework for an enhanced version of RT-P method is presented in Section III followed by the numerical results in Section IV. We conclude our work in Section V.

II. Problem Statement and DBA Description

Fig. 1 illustrates the main differences between the existing PON and next generation LR-PON topologies. Assuming a maximum range of 100km for LR-PON, the maximum RTT will increase to RTT_{max} = 1ms. LR-PON can divide ONUs into subsets and assign separate wavelength to each subset using wavelength division multiplexing (WDM) [11]. This can be realized by means of optical add/drop multiplexers (OADMs) and power splitter/combiner (PSC) at the remote nodes.

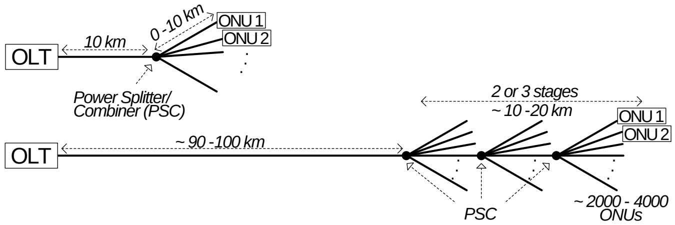


Fig. 1. Differences between PON (top) and LR-PON (bottom) topologies.

The average frame delay is defined as the average time elapsed between a frame's arrival at the ONU and its arrival at the OLT. If T_P is the propagation delay between a given ONU and the OLT, a frame generated by the ONU must experience a minimal delay of $3T_P$, in addition to its transmission time [7]. This is due to the fact that, upon the frame's generation, a REPORT must be transmitted to the OLT and a GATE must be broadcast back to the ONU before actual transmission of the frame is possible.

A. Conventional Polling

Fig. 2(a) shows an example of conventional polling (IPACT) in a PON with 2 ONUs. On the downstream channel, the OLT transmits a GATE message such that its arrival at the ONU coincides with the beginning of the granted transmission window of that ONU. It is important to note that the OLT interleaves its GATE transmissions with the reception of data so as to increase the utilization of the upstream channel. This process is called interleaved polling, and may reduce the time τ_c required by each polling cycle on the upstream channel through reducing idle time (τ_c denotes the duration of the polling cycle).

Conventional polling imposes a fateful second constraint on the downstream transmission of GATE messages: they may not be transmitted before the arrival of the REPORT message to the OLT. Since the OLT cannot generate a GATE before receiving a REPORT, τ_c in conventional polling cannot fall below RTT_{max}. This feature is the crucial limitation of conventional polling: it places an upper-bound on the frequency of REPORTs and data transmissions emanating from any ONU. As a result, the average frame delay increases. To illustrate this, consider the worst-case scenario where a frame is generated an infinitesimal time ϵ after the transmission of a REPORT message. In that case the frame experiences a delay equal to the cycle time before its bandwidth is reported to the OLT. The fact that τ_c is lower-bound by RTT_{max} hence compounds the delay and capacity problem in LR-PONs. Consequently, novel bandwidth allocation methods must be explored.

B. Multi-Thread Polling

The first candidate to alleviate the delay problem of conventional polling in LR-PONs is multi-thread polling (MT-P) [9]. MT-P creates $N \geq 1$ multiple interleaved polling cycle instances (i.e. "threads"). This is illustrated in Fig. 2(b), where the REPORTs, GATEs, and data transmissions are identified by a thread number. Each thread is a complete polling cycle process where all ONUs are polled once. The duration of each thread is initialized to $RTT_{\max}/N$, hence giving the ONUs N reporting and data transmission opportunities within an $RTT_{\max}$ period. This has the effect of reducing the waiting time of frames prior to being reported and transmitted. In MT-P, fairness is achieved by allocating the excess bandwidth in each thread to over loaded ONUs. Therefore, the bandwidth allocation in each thread is performed after collecting the REPORTs from all ONUs; thus leading to "offline scheduling" framework which results to idle time between subsequent threads. In Section IV, we consider "just-in-time" (JIT) scheduling framework [12] for MT-P. This means that the OLT collects REPORTs from various threads until channel becomes available. Then, the OLT sends the grant to next ONU using round-robbin ordering.

C. Real-Time Polling

The second polling candidate for LR-PONs is real-time polling (RT-P) [10], illustrated in Fig. 2(c). RT-P uses an additional separate reporting channel that allows ONUs to report increases in their queue size in real-time. By using a separate control channel, reporting can be done independent of the upstream data traffic taking place on the legacy upstream data wavelength channel, i.e., upstream data and control transmissions are decoupled. Each ONU has the opportunity to report a queue increment of Δ_Q bytes every reporting period T_s by transmitting a Queue Increment Report (QIR) to the OLT. In addition, REPORTs continue to be issued at the end of each granted data transmission window on the conventional upstream wavelength channel. RT-P can be realized by using optical coding (OC). OC-enabled ONUs apply remote encoding, reflection, and ON-OFF modulation of a pulse stream generated at the OLT. We assume that the applied codes are orthogonal and the operation is performed in an interference-free environment. The minimum queue increment size is equivalent to the minimum frame size, leading to $T_s = 5\mu$s in 1G-EPON. This means the pulse train is generated at 200 KHz which is practical using off-the-shelf switches. Smaller increment sizes do not yield higher performance as the Ethernet frames are not fragmented to lower granularity [10]. These OC enhancements allow the OLT to receive out-of-band queue status updates every $T_s \geq 5\mu$s which is much faster than conventional PONs, thus enabling it to determine the size of grants more accurately [13].

There are two crucial differences between QIRs and REPORT messages. First, QIRs are pulses that can be transmitted instantaneously and at very high frequencies. Second, QIRs carry incremental information about the queue whereas REPORTs carry its exact size. Therefore, the OLT needs to keep track of the queue size by adding up QIRs until the arrival of

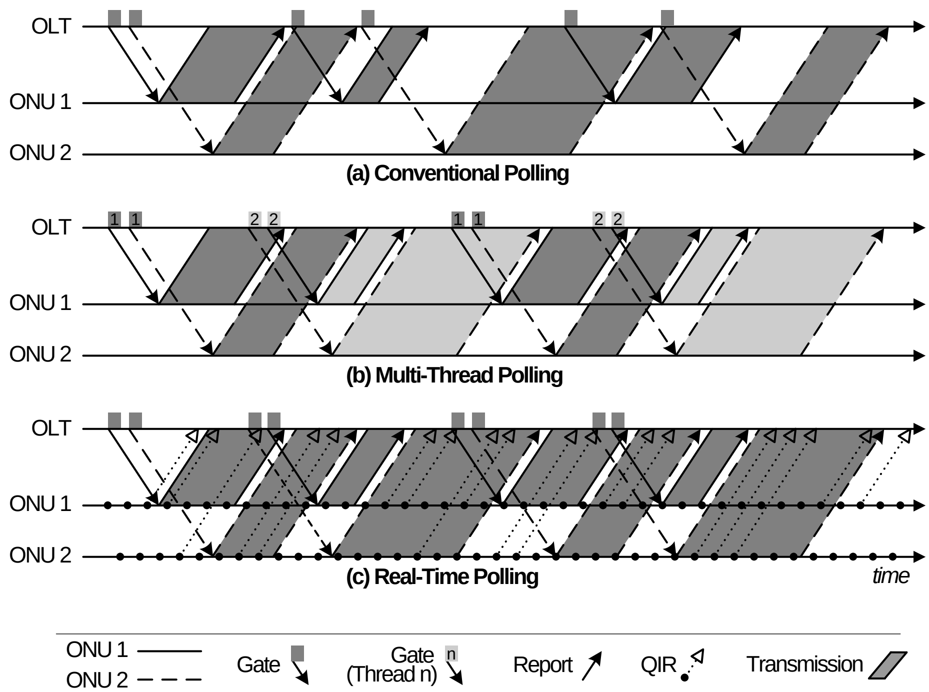


Fig. 2. Illustrations of the different LR-PON polling paradigms.

the next REPORT. QIRs enable ONUs to reduce the waiting time of frames prior to their reporting to $\tau_R < T_s$. In addition, they enable the OLT to increase the granted transmission windows on shorter notice, hence reducing the total frame delay significantly. Unlike thread-specific REPORTs in MT-P, the generation of QIRs in RT-P does not trigger GATE transmission.

D. Discussion

MT-P succeeds in reducing frame delays by giving ONUs multiple opportunities to report on their queue size and to transmit data within the RTT time-range. RT-P introduces flexibility in its polling cycle time; since QIRs are sent in real-time and on a different channel than data, τ_c is not lower-bound by RTT_{max}, as is the case in conventional polling. In fact, RT-P can implement a cycle time τ_c that is in the order of a single thread duration in MT-P ($RTT_{\max}/N$) or lower, while avoiding the complexity of MT-P.

Indeed, MT-P introduces significant complexity to the bandwidth allocation process, such as elaborate thread duration management [9] and the tagging of GATEs and REPORTs with a thread number. However, MTP does not require changes in physical-layer, whereas RTP improves the performance at the cost of increased complexity in physical layer which is incurred by employing OC-enabled ONUs. Furthermore, while RT-P shifts its additional signalling (QIRs) to a separate channel, MT-P must include its REPORTs with data, hence creating additional overhead. On the downstream channel, since QIRs are merely live updates of the queue status that do not necessarily trigger GATE messages, RT-P transmits only one GATE per ONU in each polling cycle. In contrast, MT-P requires N GATE messages to be transmitted downstream for each polling cycle. Besides, MT-P does not remove the requirement imposed by conventional polling to wait for a REPORT before transmitting a corresponding GATE message (see Fig. 2 (a)). Rather, it alleviates its consequences by creating frequent REPORT and GATE instances. On the other hand, thanks to its real-time queue updating at the OLT, RT-P completely eliminates that constraint, hence simplifying downstream GATE scheduling. Due to the mentioned differences between RT-P and MT-P, RT-P yields improved frame delay and capacity performance as will be verified by simulation results in Section IV.

III. Delay Analysis Framework for Enhanced Real-Time Polling

We consider a modified version of RTP where a data transmission from an ONU consists of only a single packet. In other words, instead of sending QIR messages, each ONU is sending the actual size of each packet after its generation. We call this method enhanced RT-P (ERT-P). In ERT-P, a codeword of N bits should be sent to the OLT over the QIR channel to report the size of generated packet (out-of-band reporting). The size of codewords (N) depends on the maximum size of generated packet. The N bits (N pulses) are On/Off-modulated to represent the length of the packet. This architecture allows the OC-enabled ONU to use the QIR channel for reporting the exact frame size instead of queue increments. This setting is feasible and will add minor complexity to QIR messaging approach. We elaborate more on ERT-P in numerical results in Section IV.

E-RTP can be modeled as M/G/1 queuing system [14]. This is because packet arrivals at the ONUs can be modeled as Markovian arrivals (exponential inter-arrival times (M)). The service time of each packet can have a general distribution (G), consisting of reporting, scheduling, granting, and transmission.

Furthermore, the system includes one server which is a single channel carrying one packet at a time. The offered service consists of reporting, scheduling, granting, and transmission of packets. Our analysis assumes the stability condition, i.e., the arrival rate λ is always less than the service rate μ, and hence there will be no packet loss. The traffic intensity ρ is defined as $\frac{\lambda}{\mu}$. We define Δ as the delay component at the OLT, from the arrival of the REPORT to transmission of the GATE messages. Assuming that all ONUs are equally loaded, the average end-to-end delay is (approximately) given by

$$\Delta_T = 3T_p + \Delta, \tag{1}$$

where T_p is the one-way propagation delay from ONU to OLT and Δ is the time from REPORT arrival at the OLT until the GATE message is sent to the ONU. Note that the beginning of the transmission is immediately placed at the arrival of the GATE message to the ONU. In this case, Δ is the average waiting time in an M/G/1-queue and can be obtained from the following equation which is a variation of the Pollaczek-Khinchine formula [14]

$$\Delta = \rho \frac{(\mu_p + \varepsilon)\left(1 + \frac{\sigma_p^2}{(\mu_p+\varepsilon)^2}\right)}{2\,(1-\rho)}. \tag{2}$$

Here, ρ is the traffic intensity including the overhead ϵ, $\mu_p = \mathbf{E}(P)$ is the average packet size, and $\sigma_p^2 = \mathbf{E}\left((P - \mathbf{E}(P))^2\right)$ is its variance. In ERT-P, we can assume that the overhead ϵ per transmission is actually an overhead per packet and thus add ϵ to the packet length. Equation (2) is obtained under stability condition where each of the N ONUs generates packets at a rate $\lambda > 0$. Substituting Equation (2) in (1), we find the following entity for the average frame delay (in seconds):

$$\Delta_T = 3T_p + \left[\frac{(\mu_p+\varepsilon)^2 + \sigma_p^2}{2(\mu_p+\varepsilon)}\right]\frac{\rho}{1-\rho}. \tag{3}$$

IV. Numerical Results

We evaluate average frame delay of each polling method using OMNeT++ simulation environment for an EPON with OLT-ONU distance of 20km and 100km, i.e., one-way propagation delay of 100μs and 500μs, respectively. Other network parameters are listed in Table I.

TABLE I
Network Parameters

Upstream data rate	1 Gbps
Number of ONUs	16
ONU buffer size	10 Mbytes
Guard bandwidth between adjacent slots	1μs
Queue increment size	64 bytes
QIR period	5μs

We consider an EPON standard with 1 Gbps transmission rate and Ethernet packets uniformly distributed between 64 bytes and 1518 bytes. In this setting, we have $\mu_p \simeq 6.33 \times 10^{-6}\mu sec$ and $\sigma_p^2 \simeq 1.35 \times 10^{-10}\mu sec^2$. Assuming a guard time overhead $\epsilon = 1\mu sec$, the end-to-end delay in expression (3) can be re-written as follows (in μsec):

$$\Delta_T \cong 3\tau + 12.874 \times \frac{\rho}{1-\rho} \tag{4}$$

This equation presents average frame delay in ERT-P where each packet is reported individually. Note that 11 bits (pulses) are required to report the exact packet size up to 1518 bytes. This is comparable to average number of QIR pulses which are required for RT-P method. The average size of Ethernet packets is 791 bytes. Assuming one QIR message for every 64 bytes, 13 QIR pulses are required on average for reporting each packet, i.e., 65 μsec pulse train when $T_s = 5\mu$sec.

Similar to [10], the 95% confidence intervals of our simulation results were smaller than 5% of the mean value. Fig. 3 illustrates average frame delay for conventional polling (IPACT) and ERT-P methods in a normal range and LR-PON with 16 ONUs. We employ "gated service" for conventional polling; i.e., OLT grants the whole request of each ONU. We notice that the simulation results for ERT-P perfectly match the analytical results derived from equation 4 for both short range EPON and LR-PON. Also we notice a significant delay reduction for ERT-P compared to IPACT.

Next, we compare RT-P with MT-P and conventional polling in a PON with 16 ONUs. The number of threads in our MT-P implementation is three which yields the highest performance as stated in [9]. As mentioned in Section II, we consider JIT framework for MTP. Similarly, our RTP implementation is based on JIT scheduling framework in a round-robin fashion. It means that the OLT updates queue length of each ONU using QIR and REPORT messages while monitoring the order of REPORT messages using a REPORT list. This procedure continues until the OLT learns that the upstream channel will be free in one link delay from the current time. At this moment, the OLT schedules next ONU in the REPORT list and sends a grant based on its REPORTed request and QIR messages which have been received after the REPORT. This setting allows OLT to collect as much information as possible before scheduling the ONUs.

Fig. 5 illustrates the obtained average delay for three polling methods when "gated service" is applied. We notice that MTP does not outperform IPACT in this setting, as the multi-thread framework downgrades to single thread when gated service is in use [9]. The RT-P has significantly smaller delay for various loads compared to conventional polling. The corresponding results with "limited service" are presented in Fig. 6. Here, we recognize remarkable improvement in MT-P and RT-P for both low and high loads.

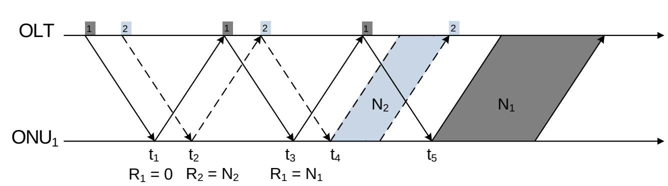
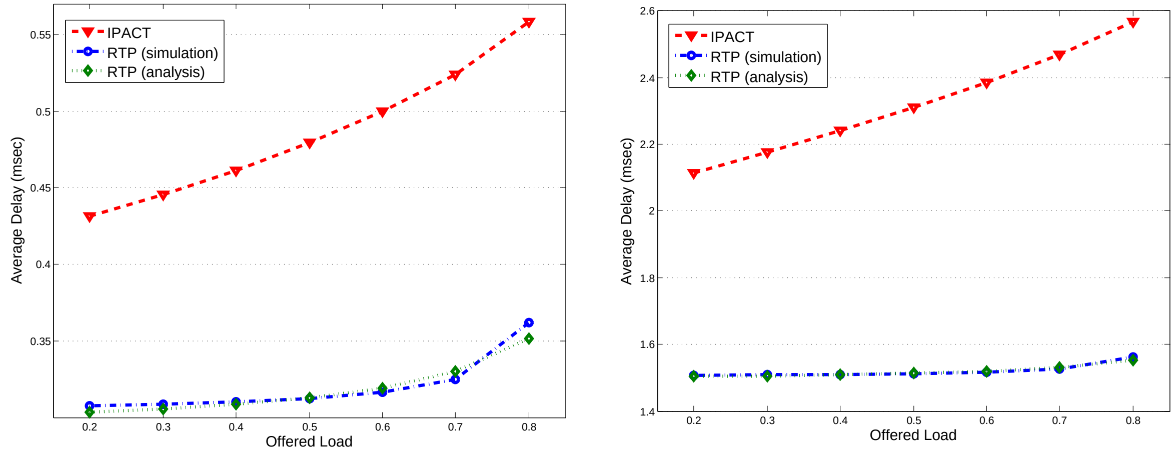


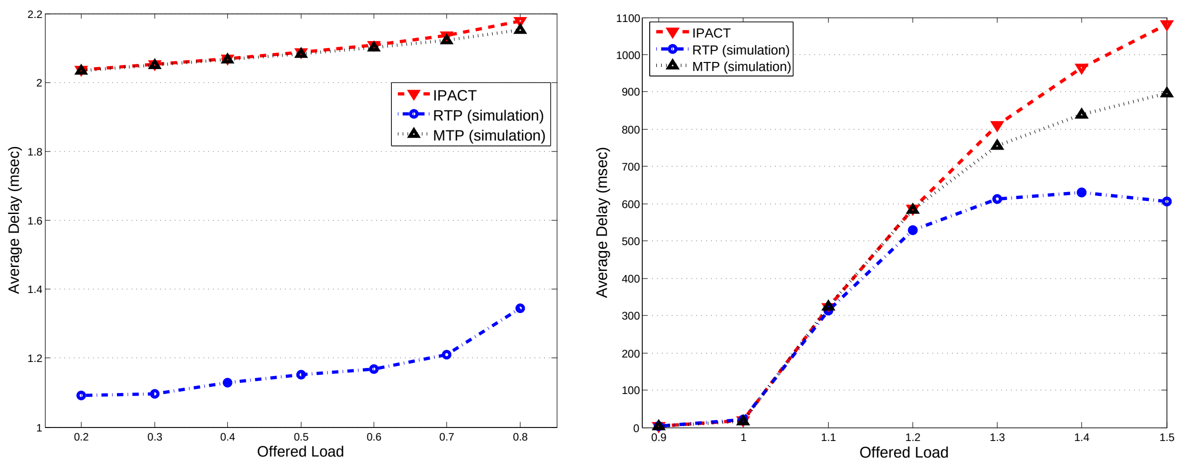
Fig. 4. Low load effect on delay performance in MTP

An interesting observation is that frame delay can be below 3τ for low loads in MT-P and RT-P. This is due to "post-reporting" condition which is illustrated in Fig. 4 for MT-P.



(a) One-way propagation delay is $100\mu s$ (20km) (b) One-way propagation delay is $500\mu s$ (100km)

Fig. 3. Delay comparison for conventional polling and ERT-P with gated service



(a) Low load (b) High load

Fig. 5. Delay performance of DBA methods with gated service when $\tau = 500\mu s$ (100km)

Assume that ONU_1 sends a request of $R_1 = 0$ for thread 1 and $R_2 = N_2$ for thread 2 at times t_1 and t_2, respectively. Using the just-in-time framework, the OLT receives the requests and sends gates with size 0 for thread 1 and N_2 for thread 2. Then the ONU sends the new request of $R_1 = N_1 > N_2$ at time t_3. However, N_2 bytes of this new request have already been reported in previous thread 2 and will be sent at time t_4. Therefore, when the first thread starts transmitting N_1 bytes at time t_5, it will transmit N_2 bytes from packets which have been generated in the interval of t_2 and t_5, thus resulting to delays less than 3τ. The same reasoning can be given for RT-P. This "post-reporting" circumstance is more pronounced in RT-P method due to higher frequency of sending incremental reports.

V. Conclusion

We studied the delay performance of three polling methods for next generation LR-PON with large propagation delay. Employing just-in-time framework for MT-P and RT-P schemes, we compared these polling techniques with conventional polling in terms of average upstream transmission delay in traditional and long-reach PONs. Also, we introduced an enhanced real-time polling (ERT-P) method and derived an analytical framework for its delay based on M/G/1 queuing system. The simulation results for ERT-P rely within 4% of the derived analytical expression for both ordinary and long-reach PONs. We conclude that both RT-P and MT-P achieve significantly lower frame delay compared to IPACT. Also, our results indicate that RT-P imposes the lowest frame delay in LR-PONs compared to MT-P and conventional polling methods.

Acknowledgement

The authors would like to thank Dr. Tomaz Berisa for his valuable insights on this work.

References

[1] 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, Feb. 2002.

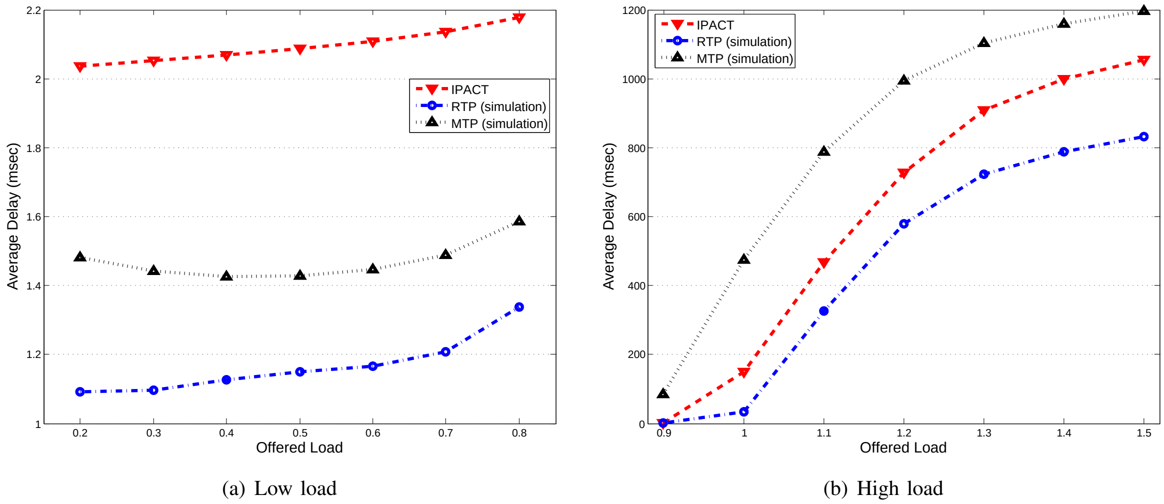


Fig. 6. Delay performance of DBA methods with limited service when $\tau = 500\mu s$ (100km)

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

[3] A. Dhaini, C. Assi, M. Maier, and A. Shami, "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 25, no. 1, pp. 277–286, Jan. 2007.

[4] B. Skubic, J. Chen, J. Ahmed, L. Wosinska, and B. Mukherjee, "A Comparison of Dynamic Bandwidth Allocation for EPON, GPON, and Next-Generation TDM PON," *IEEE Communications Magazine*, vol. 47, no. 3, pp. S40–S48, Mar. 2009.

[5] S. Bhatia and R. Bartos, "Closed-form expression for the collision probability in the IEEE Ethernet Passive Optical Network registration scheme," *OSA Journal of Optical Networking*, vol. 5, no. 1, pp. 1–14, Dec. 2005.

[6] B. Lannoo, L. Verslegers, D. Colle, M. Pickavet, M. Gagnaire, and P. Demeester, "Analytical model for the IPACT dynamic bandwidth allocation algorithm in EPONs," *OSA Journal of Optical Networking*, vol. 6, no. 6, pp. 677–688, May 2007.

[7] F. Aurzada, M. Scheutzow, M. Herzog, M. Maier, and M. Reisslein, "Delay Analysis of Ethernet Passive Optical Networks with Gated Service," *OSA Journal of Optical Networking*, vol. 7, no. 1, pp. 25–41, Jan. 2008.

[8] F. Aurzada, M. Scheutzow, M. Reisslein, N. Ghazisaidi, and M. Maier, "Capacity and Delay Analysis of Next-Generation Passive Optical Networks (NG-PONs)," *IEEE Transactions on Communications*, vol. 59, no. 5, pp. 1378–1388, May 2011.

[9] H. Song, B. W. Kim, and B. Mukherjee, "Multi-thread Polling: A Dynamic Bandwidth Distribution Scheme in Long-Reach PON," *IEEE Journal of Selected Areas in Communications*, vol. 27, no. 2, pp. 134–142, Feb. 2009.

[10] K. Fouli, T. Berisa, and M. Maier, "Optical Coding for Enhanced Real-Time Dynamic Bandwidth Allocation in Passive Optical Networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 27, no. 23, pp. 5376–5384, Dec. 2009.

[11] D. P. Shea and J. E. Mitchell, "Long-Reach Optical Access Technologies," *IEEE Network*, vol. 21, no. 5, pp. 5–11, Sept./Oct. 2007.

[12] M. P. McGarry, M. Reisslein, M. Maier, and A. Keha, "Bandwidth Management for WDM EPONs," *Journal of Optical Networking*, vol. 5, no. 9, pp. 637–654, Sept. 2006.

[13] T. Berisa, K. Fouli, M. Maier, and A. Bazant, "Adaptive Grant Extension in OC-Enhanced PONs," *IEEE Communications Letters*, vol. 15, no. 4, pp. 458–460, Apr. 2011.

[14] L. Kleinrock, *Queueing Systems: Volume I-Theory*, Wiley Interscience, New York, 1975.