Performance of the IEEE 802.3 EPON Registration Scheme Under High Load

Swapnil Bhatia and Radim Bartoš

Department of Computer Science, University of New Hampshire, Durham, NH 03824, USA

ABSTRACT

The proposed standard for the IEEE 802.3 Ethernet Passive Optical Network includes a random delayed transmission scheme for registration of new nodes. Although the scheme performs well on low loads, our simulation demonstrates the degraded and undesirable performance of the scheme at higher loads. We propose a simple modification to the current scheme that increases its range of operation and is compatible with the IEEE draft standard. We demonstrate the improvement in performance gained without any significant increase in registration delay.

Keywords: EPON, Registration, Discovery, Contention, Performance evaluation, Simulation

1. INTRODUCTION

This paper reports the results of our study of the performance of the point to multipoint Ethernet Passive Optical Network (EPON)[1] node registration process (also referred to as node discovery process) and proposes a scheme to address the issues that were discovered during the evaluation. EPON technology seems to be the next step in the evolution of networks delivering connectivity to the residential customers.[2] It is clear that the such networks are starting to be used not only for best-effort data but also to deliver services, such as voice and video that place strict constraints on the quality of service (QoS). There have been numerous studies of QoS of EPON networks published recently,[3, 4] concentrating in most cases on scheduling of data transmissions by the headend node. Relatively little attention has been paid to the process by which new nodes are discovered and registered with the headend. The performance of the registration process impacts the speed with which users gain access to the network. Furthermore, ONUs may get de-registered due to internal clock drift and enter the registration process even in the midst of a data transmission requiring QoS guarantees. Especially from this perspective, it is important to ensure that the discovery is handled as quickly as possible.

Earlier versions of the EPON standard considered an exponential backoff-based scheme to handle collisions during the registration process.[5] However, as a result of a study,[6] exponential backoff has been dropped from the standard and a preference has been given to a scheme that allows nodes to choose randomly when they transmit during an enlarged window of opportunity, termed *discovery slot* or *discovery window*. The main argument for this decision was the reduction in the average wait for the registration.

This paper presents a simulation study of registration process performance under sustained load for three categories of node distance distributions. The results show that for higher loads and cases where there is a higher percentage of nodes with similar distances from the OLT, the performance of the registration process degrades significantly.

To address the observed shortcomings of the current protocol, the paper proposes a modification of the way the OLT schedules the discovery slot. The length of the discovery slot is dynamically adjusted based

---

Further author information: (Send correspondence to Swapnil Bhatia). Swapnil Bhatia: E-mail: sbhatia@cs.unh.edu, Telephone +1 603 828 5513. Radim Bartoš: E-mail: rbartos@cs.unh.edu, Telephone: +1 603 862 3792.

This paper will be published in the conference on *Performance, Quality of Service, and Control of Next Generation Communication Networks II* at *Optics East 2004 (ITCom)* symposium and is made available as an electronic preprint with permission of SPIE.

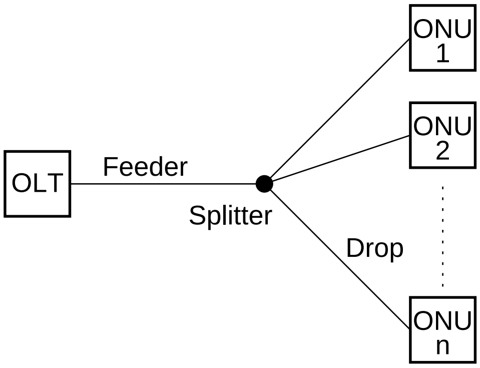


Figure 1. PON topology example.[1]

on whether collisions were observed by the OLT during a discovery cycle. The proposed scheme does not fundamentally deviate from the proposed standard. The results of a simulation study show that the scheme improves performance when compared to fixed discovery window size while limiting the increase in overhead associated with larger constant discovery slot sizes. The scheme does not suffer from the increase in wait time that has been observed in the exponential backoff-based schemes that were a part of the earlier version the standard.

2. THE IEEE 802.3 EPON REGISTRATION SCHEME

The IEEE 802.3ah is a proposed standard for Ethernet Passive Optical Network as a solution for the first mile access bottleneck. Although the draft proposes several architectures, this paper will focus on the point to multipoint tree topology. Figure 1 shows the topology of a point to multipoint EPON. In this architecture, a pair of uni-directional optical channels, no more than 20 km in length, serve as the medium of transmission. The *Optical Line Terminator* (OLT) serves as the edge forwarder for the traffic leaving and entering the EPON. The OLT is typically located at the service provider's premises. Subscriber devices, called *Optical Network Units* (ONUs), are connected as leaf nodes to the split and branched optical channels. The channel directed to the OLT is referred to as the *upstream* whereas the channel directed towards the ONUs is called the *downstream* channel. The downstream channel is broadcast whereas the signal transmitted on the upstream is received only by the OLT and not by any of the ONUs. The OLT arbitrates the timing of transmissions by the ONUs on the upstream. It receives reports about the traffic load at various ONUs through periodic REPORT messages and based upon this information grants transmission opportunities to the ONUs. The ONUs are informed about the precise time of their scheduled transmission opportunities through GATE messages sent by the OLT sufficiently ahead of time. Thus, normal data transmission is collision free. The scheduling scheme used by the OLT for allocation of upstream bandwidth is not standardized and is open to the implementor's designs.

However, for the OLT to be able to direct ONU transmissions, the ONUs and the OLT need to be time synchronized. Furthermore, even if the ONUs and OLT are synchronized, they cannot remain so for long periods of time due to clock drift. The OLT sends periodic timestamps to the ONUs in order to correct their clocks for drift error. Since the EPON can reach a maximum length of 20 km, the distance and hence the round trip times from the OLT to various ONUs can vary over a large range of values. Hence, the OLT must send individually adjusted timestamps to guarantee the correctness of the timestamp when it arrives at the ONU. This requires the OLT to know the round trip time (RTT) to every individual ONU. Without this information, the OLT cannot arbitrate the upstream channel in a collision free manner.

The OLT calculates the RTT to an ONU during, what is known as, the discovery process. When an ONU first joins the EPON, it is required to register itself with the OLT. The discovery process, illustrated in Figure 2 is a three-way handshake consisting of a REGISTRATION-REQUEST message from the ONU to the OLT which is replied to by the OLT in the form of a REGISTER message which in turn is acknowledged by the ONU through

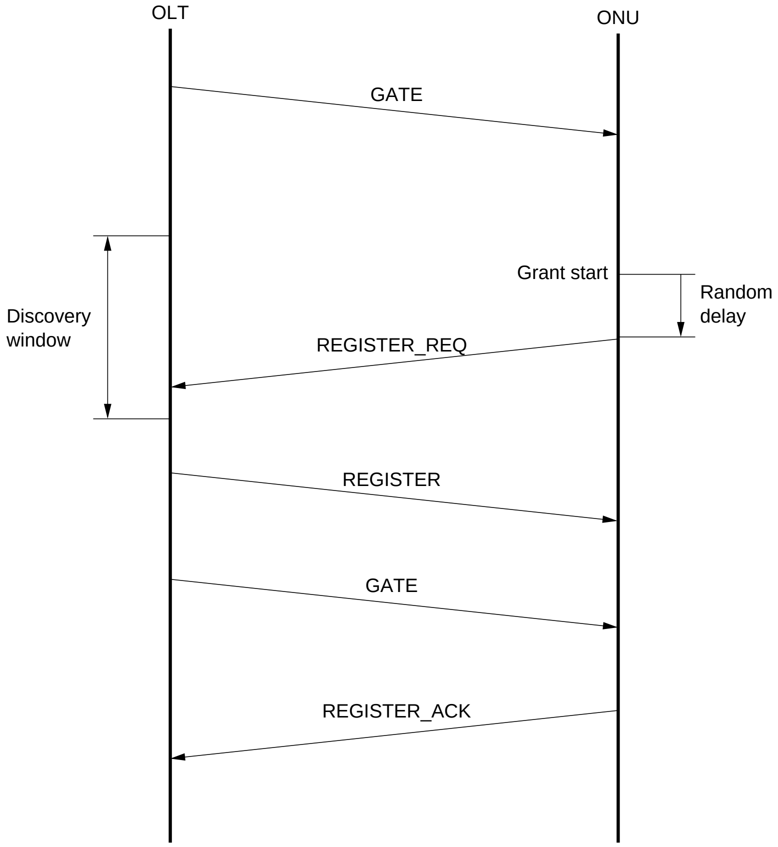


Figure 2. Discovery handshake message exchange.[1]

a REGISTER-ACK message. Through the REGISTER message, the OLT provides the ONU with identification and physical layer parameters. Next, the OLT provides the ONU with a unicast grant to allow it to transmit the REGISTER-ACK message. However, this not the case with the initial REGISTRATION-REQUEST message. For the initial message, since the OLT has no information about the distance from which and when new ONUs would transmit their requests, it reserves the upstream for a special contention-based broadcast grant called the DISCOVERY-WINDOW. New ONUs are informed about the starting time of this window through a special DISCOVERY-GATE message on the downstream channel. New ONUs wait for the window to begin and then as per the current IEEE draft, after a uniformly random delay into the window, transmit the REGISTRATION-REQUEST message. Since the discovery window is contention-based, REGISTRATION-REQUEST messages from multiple ONUs may overlap in time and collide. The random wait has been proposed as a simple scheme to minimize the probability of such a collision.

3. MOTIVATION

It is worthwhile to note the effect of distance of the ONU from the OLT on the discovery process. Firstly, regardless of the size of the discovery window advertised by the OLT to the ONUs, the OLT needs to additionally reserve transmission time equal to the maximum RTT (i.e., equivalent to 20 km), or about 200 μs on the upstream to allow a REGISTRATION-REQUEST message from the farthest ONU to reach the OLT in full before subsequently scheduled transmissions begin. In contrast, the typical length of messages exchanged during the discovery process is of the order of 1-2 μs. Thus, the guard time forms a constant but considerable fraction of the overhead associated with the discovery process.

Secondly, two new ONUs may decide to begin transmitting the REGISTRATION-REQUEST message at the same time but their messages may not overlap and collide if they are sufficiently separated in distance from each other. Thus, the distribution of distances of the ONUs from the OLT serves as a second randomizing term, in addition to the random wait, in the calculation of transmission time of the REGISTRATION-REQUEST. However, this holds only under the assumption that the distances of the ONUs from the OLT are uniformly randomly distributed. Previous studies,[6] have more or less, accepted this to be a truism. But this may not be the case, and in fact may be an exception rather than the rule. It is plausible that subscribers may be concentrated in clusters at random distances from the OLT over the 20 km span of the EPON. In this scenario, ONUs within a cluster may not be sufficiently separated in distance and hence their transmissions may not be distributed over a wide window in time. As a result, the proposed simplistic random wait scheme may prove to be inadequate at reducing collisions and increasing the number of successful registrations.

Thirdly, contrary to intuition, a larger discovery window size may actually hinder performance. Intuitively, two nodes that are sufficiently separated in distance cannot cause collisions and affect performance when using a sufficiently small discovery window. For a sufficiently large discovery window, the two distant nodes may now be forced to choose random transmission times within their windows of time that now overlap owing to their large size. Thus, any window size larger than the average distance between nodes or between clusters will worsen performance given that the probability of collision depends exponentially on the number of ONUs attempting to transmit in a given discovery window. The effect on performance may be less significant for uniform distribution of ONUs but will be heavy for the clustered case.

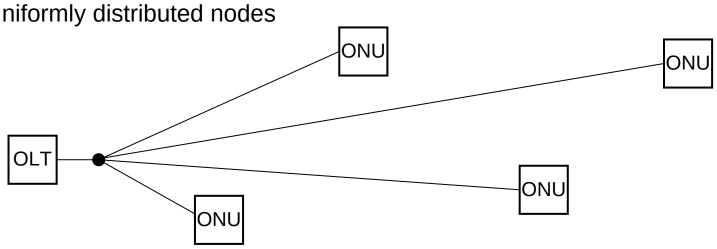
Another drawback of the random wait scheme is that it has a small region of optimal operation situated only at lower loads for any given discovery window size. We define load as the rate of arrival of new ONUs wishing to register with the OLT. At higher loads, the scheme is unable to manage the given resource (i.e., bandwidth inside the window) efficiently and as a consequence, fails to deliver the performance possible with the given resources. Since the size of the discovery window is fixed, the scheme plays a passive role in bandwidth management and cannot react to a changing load. On the other hand, owing to the simplicity of the scheme, the load remains and may become increasingly aggressive over time – retrying backlogged ONUs cause a cumulative buildup of load. Passive bandwidth management coupled with its tendency to create and inability to control cumulative load buildup make this scheme a poor choice except in cases where the load is predictable and discovery window overhead is not an issue. Current intended deployments for the EPON may appear to satisfy these criteria. However, historically, applications, traffic patterns and bandwidth have proven to be the three most unpredictable and elusive variables and have been primarily responsible for fueling the race to build next generation networks. Owing to their cost and capacity, since EPONs are intended to provide a long-term solution to the first mile access bottleneck, characterization of their scaling performance at higher loads deserves in-depth study. Moreover, better alternatives to the current scheme are well-known. We present a simple modification to the current scheme in this paper. In the next section we describe our simulation experiments to support the above discussion.

4. SIMULATION EXPERIMENTS

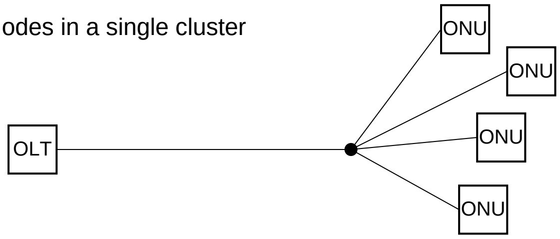
All the experiments in this study were conducted using a point to multipoint EPON MAC-layer simulator. The simulator was written in C and is capable of simulating the discovery and normal traffic transmission behavior of the point-to-multipoint EPON. Work is in progress to add QoS simulation capability to the simulator. Various parameters such as ONU distance distributions, message lengths, window size and frequency are easily configurable. The simulator produces a detailed trace along with grant snapshots and other desired measurements. In addition, it also produces a visualization trace compatible with the NS-2/NAM trace file format. Work is also underway to build a unique time-line tool for automatic illustration of simulation events in the familiar, textbook-style cause-effect arrow diagram similar to Figure 2.

As discussed in the previous section, the distribution of distances of the ONUs from the OLT plays an important role on the performance obtained from a given discovery window size. Therefore, to investigate the performance of the current registration scheme under high load we consider three different categories of distances in our study. In the first case, ONUs are uniformly distributed between a 100 m and 20 km from the OLT (Figure 3-a). In the second case, we study a single cluster of diameter 600 m situated 15 km from the OLT

(a) Uniformly distributed nodes



(b) Nodes in a single cluster



(c) Nodes in multiple clusters

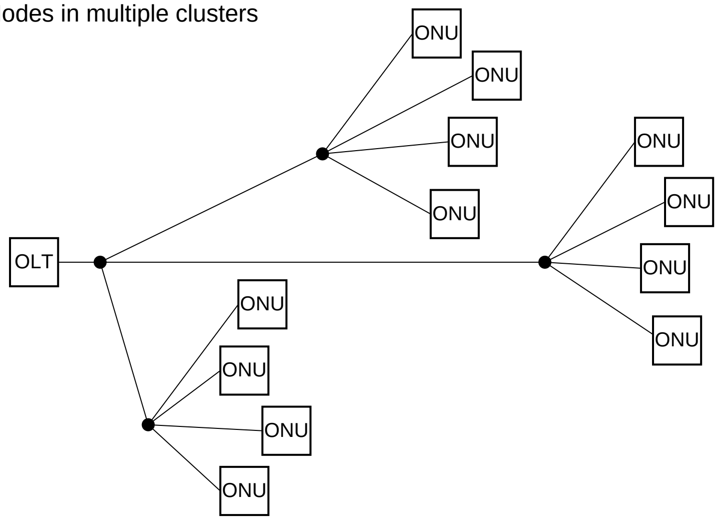


Figure 3. Configurations of ONU distances used in experiments.

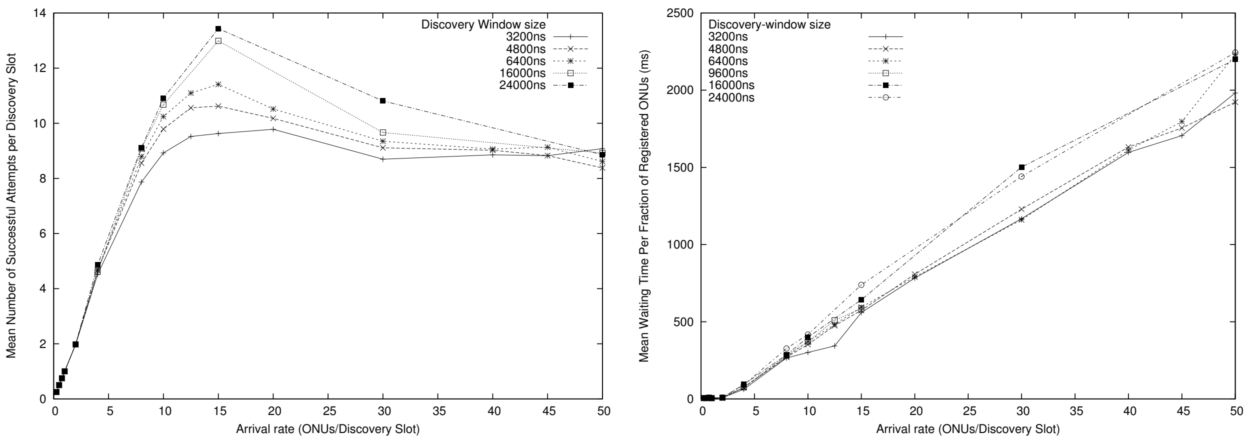


Figure 4. Uniformly distributed nodes: Mean number of successful registration attempts and mean waiting time.

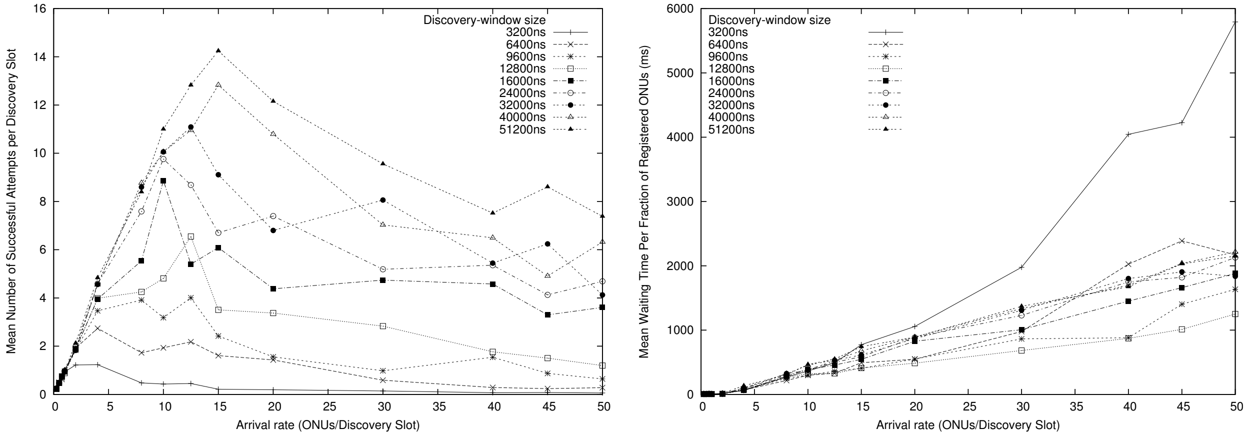


Figure 5. Nodes in multiple clusters: Mean number of successful registrations attempts and mean waiting time.

(i.e., ONUs are randomly placed between 15 km to 15.6 km away from the OLT, similar to Figure 3-b). In the third case, ONUs are distributed in between 1 and 10 non-overlapping clusters of an average diameter of 200 m situated at uniformly random distances from the OLT (Figure 3-c). For each configuration, we vary the load, defined as the rate of arrival of new ONUs, and the discovery window size and measure the mean number of ONUs successfully receiving the REGISTER message reply from the OLT. This will be the metric of performance comparison in our study. Each simulation was repeated for twenty random distance configurations for each load and window size value in each of the three categories.

Figure 4 shows the performance of the current random wait scheme on a configuration with ONUs distributed at uniformly random distances from the OLT. Each curve in Figure 4 represents a different discovery window size. For low loads (i.e., ≤ 10 arrivals/discovery slot), on an average, all new arrivals are serviced without any buildup of backlogged ONUs. Thus, for low loads the current random wait scheme delivers optimal performance. However, for medium to high loads (i.e., > 10 arrivals/discovery slot), the performance of the current scheme deteriorates. Since the scheme fixes the size of the discovery window regardless of the load, the maximum performance achievable is limited. Moreover, for loads at the higher end, due to cumulative buildup, performance suffers and begins to drop even lower than that possible with the fixed window size*.

Figure 5 shows the performance of the current scheme on a configuration with ONUs distributed in between

---

*Simulations for even higher loads and larger window sizes are currently in progress but are delayed due to their heavy computational demands. Please check `http://www.cs.unh.edu/cnrg/epon` for updated results.

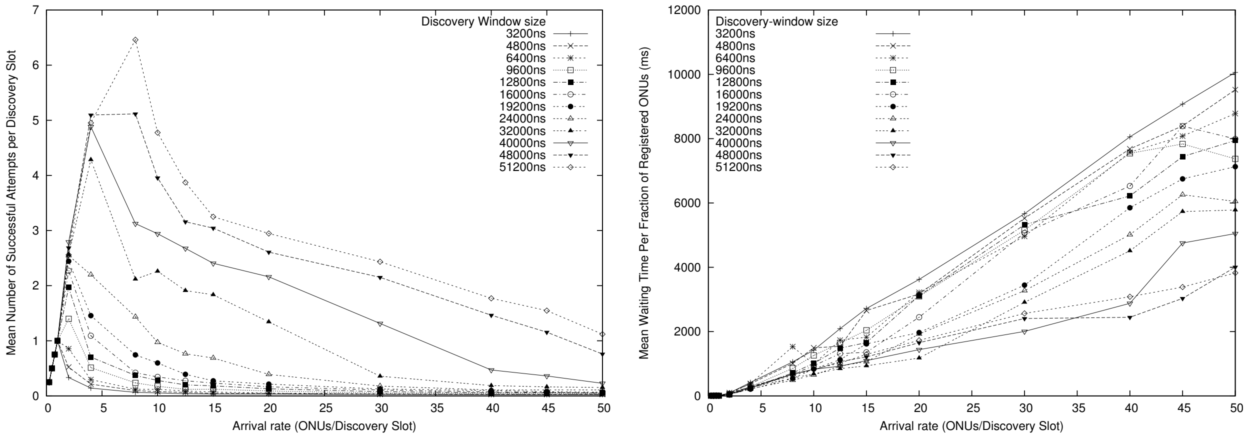


Figure 6. Nodes in a single cluster: Mean number of successful registrations attempts and mean waiting time.

1 and 10 clusters at uniformly random distances from the OLT. Within each cluster, ONUs are randomly placed at an average distance of 200 m from the center of the cluster. Figure 5 clearly illustrates the degradation in performance at higher loads (i.e., > 10 arrivals/discovery window). The scheme fails to service the cumulative backlog, and as a result also fails to completely utilize the available window size. Consequently, performance at higher loads is much lower than expected for the allocated discovery window bandwidth.

Figure 6 shows the performance of the random wait scheme on a configuration with ONUs arriving at a single cluster 15 km away from the OLT. New ONUs arrive at random distances from the center of the cluster within an average diameter of 600 m. As is clear from the downward trend in performance beyond very low loads (i.e., ≤ 5 arrivals/discovery window) the random wait scheme is unable to deliver expected performance for higher loads. The region of operation is particularly small and the drop in performance particularly quick for the single cluster scenario since, owing to a much smaller variance in ONU distances, the scheme can no longer benefit from the added dispersion in transmission times due to random distancing of new ONUs from the OLT. Again, this is a plausible scenario – the assumption that ONU distances from the OLT, when deployed in the first mile, would be uniformly randomly distributed is questionable.

5. PROPOSED MODIFICATION

We present a simple modification to the random wait scheme in order to improve its performance under high loads. Our proposal stems from the observation that the current scheme degrades in performance at high loads because it does not have any built-in mechanisms to react to changes in bandwidth or load. To remedy this, we propose that the OLT, upon observing collisions in a discovery window, increase the size of the following window. If a discovery window finishes collision-free, the OLT should decrease the size of the next window. The increment, decrement, maximum and starting sizes for this *Dynamic Window Sizing* scheme remain configurable.

For low loads, the modified scheme advertises a configurable minimal discovery window. For medium loads, the modified scheme dynamically adjusts the window size to match the offered load. For high loads, the new scheme increases the window size to minimize the number of collisions. Since the dynamic window sizing selects and matches the window size to the offered load, it can reduce the overhead of discovery windows for low to medium loads. For high loads, whereas the current random wait scheme degrades in performance, the modification would allow it to scale according to the load.

In order to test these hypotheses, we conducted another set of simulations with the dynamic window sizing modification added to the discovery module the OLT. The minimum and maximum window sizes for the dynamic window were set to 1.6 μs and 1.6 ms respectively. The increment and decrement for dynamic window sizing were set to 12.8 μs and 25.6 μs respectively.

Figures 7, 8, and 9 show the performance of the random wait scheme with dynamic window sizing for different distance configurations. In each case, dynamic window significantly improves (doubles) the region of operation of

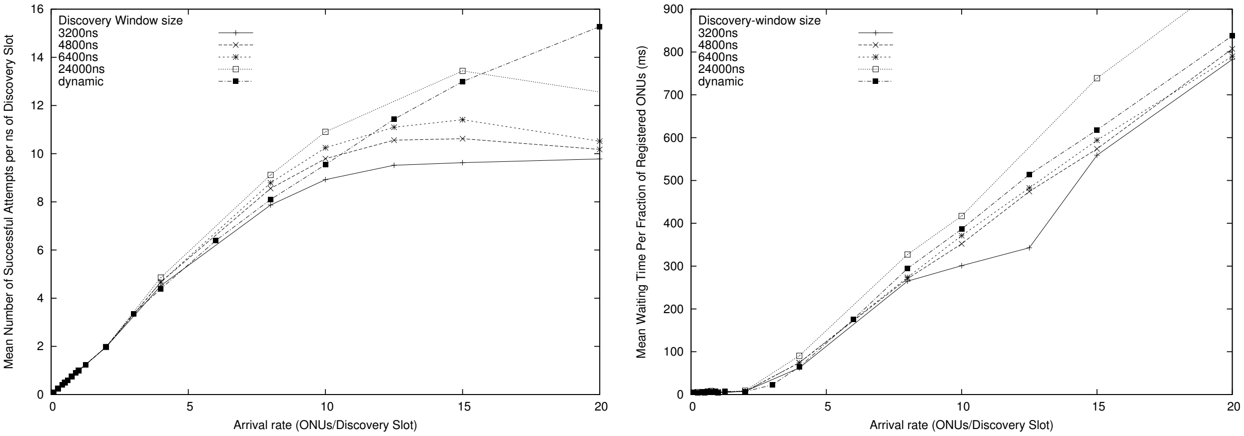


Figure 7. Uniformly distributed: Mean number of successful registrations attempts and mean waiting time with dynamic window sizing.

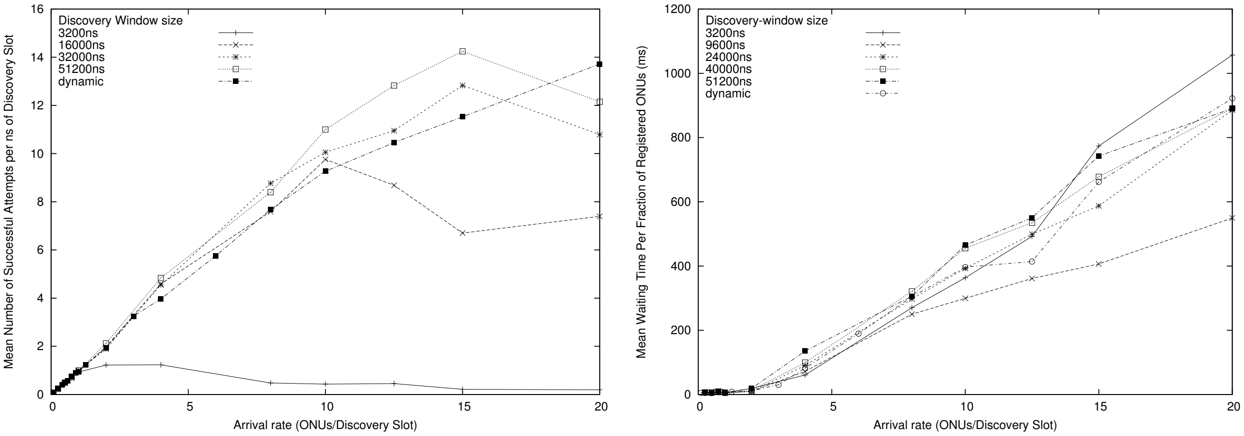


Figure 8. Clustered nodes: Mean number of successful registrations attempts and mean waiting time with dynamic window sizing.

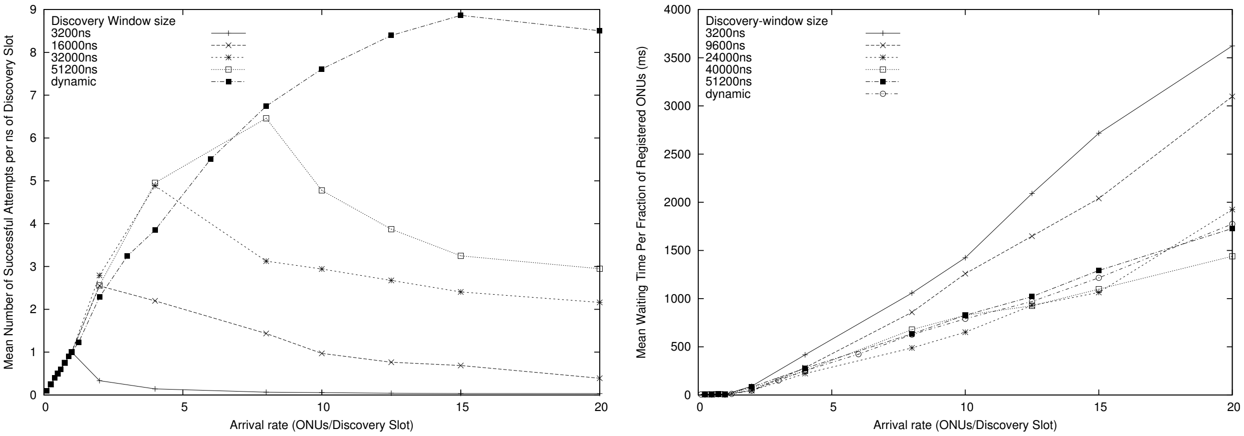


Figure 9. Nodes in a single cluster: Mean number of successful registrations attempts and mean waiting time with dynamic window sizing.

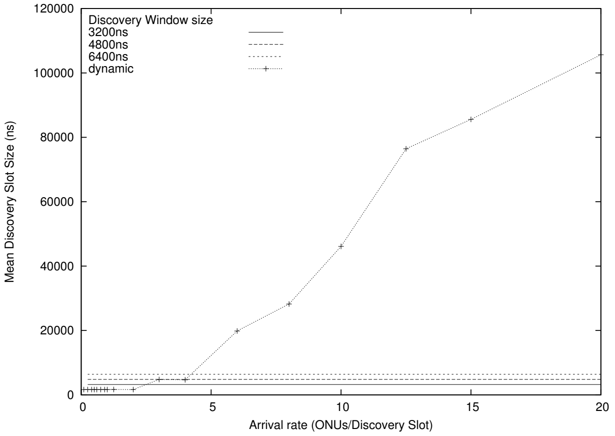


Figure 10. Mean slot size of dynamic window: Uniformly distributed configuration.

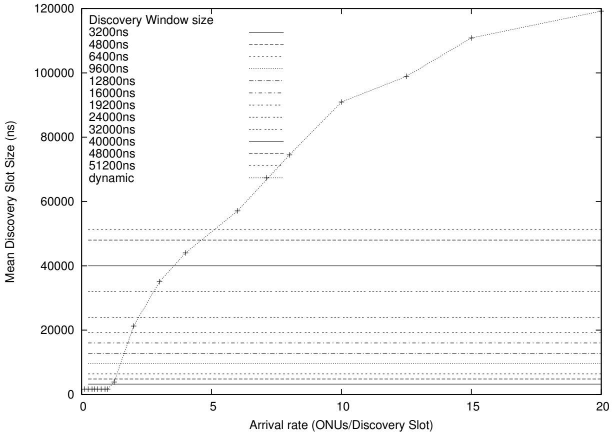


Figure 11. Mean slot size of the dynamic window: Uni-clustered configuration.

the random wait scheme. Moreover, the waiting time graphs show that apart from the gain in the mean number of successful registrations, the scheme is also able to maintain the low average waiting time of the original scheme. Thus, the new scheme offers the advantage of load management of a backoff scheme but without its steep rise in latency.

Figures 10, 11, and 12 show the average size of the discovery window used by the dynamic window scheme. This, represents the overhead of the proposed modification. As expected, the dynamic scheme adapts to high loads and uses large window sizes in an attempt to deliver acceptable performance. At low and medium loads, as illustrated in Figures 10, 11, and 12, the dynamic scheme selects the appropriate window size to deliver optimal performance. The performance delivered by the dynamic scheme can be improved even further by changing the increment and decrement values per discovery window. A higher increment will allow the scheme to adapt to high loads faster, thus delivering performance competitive with the original scheme. A higher decrement value on the other hand will allow the scheme to curb excessive use of larger window sizes thus achieving optimal reductions in overhead.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we investigated the performance of the IEEE 802.3 EPON discovery and registration scheme under high loads. We studied the performance of the scheme for three different distributions of nodes: uniformly random, uni-clustered and multi-clustered. Our simulations demonstrate that the performance of the random

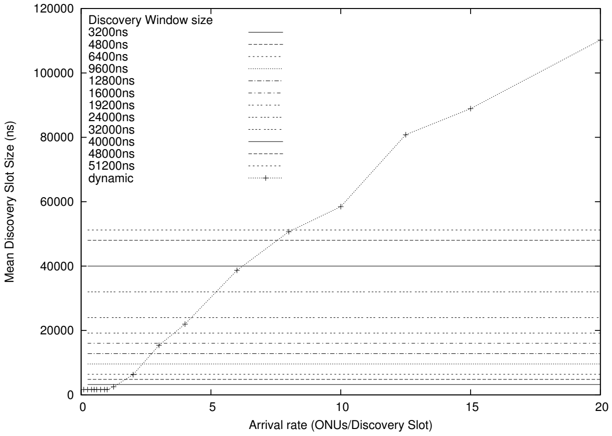


Figure 12. Mean slot size of the dynamic window: Multi-clustered configuration.

wait scheme degrades significantly for medium to high loads for all three distributions. The results show that the random wait scheme delivers acceptable performance only in a very small load region and does not scale well to higher loads. At high loads, the advantage of low waiting time is no longer sufficient to offset the degraded performance. To remedy this, we proposed a simple modification to the scheme involving dynamic sizing of the discovery window to match the load and demonstrated the improvement in performance due to the new scheme. Our simulations show that dynamic sizing of the window allows the random wait scheme to scale to higher loads. Moreover, unlike backoff schemes, it does not increase waiting time significantly.

The simulation results discussed in this paper are promising. Our proposed modification increases the region of operation of the random wait scheme. However, for high loads, the dynamic window increases to large values. The effect of various window size increments is still under study. To decrease overhead further, the scheme can be modified to increase the window size sublinearly. For example, the scheme can be tailored to tolerate a healthy number of collisions by maintaining current window size for a number of discovery windows despite collisions. This headend analogue of a backoff scheme is currently under study. A priority scheme is another alternative to manage high loads. In such a scheme, certain discovery windows or portions of every discovery window could be reserved for ONUs whose waiting time has exceeded a configurable threshold. Such a scheme would be capable of bounding the waiting time for registration – a useful property for QoS sensitive applications.

Another aspect of characterizing the behavior of the random wait scheme is the nature of the load offered. In this study, a sustained, constant load was used. However, the experiments need to be run against more realistic load models as well. For example, a bursty (self-similar) load would be very useful in demonstrating the reduction in overhead due to our proposed modification.Work on analytical verification of our proposal is currently underway.

ACKNOWLEDGMENTS

This project was supported in part by a grant from Cisco Systems' University Research Program. The authors would also like to thank the UNH InterOperability Laboratory (`http://www.iol.unh.edu`) and its Ethernet in the First Mile consortium led by Eric Lynskey for continued support.

REFERENCES

1. *IEEE Draft P802.3ah/D3.3: Draft Amendment to Carrier Sense Multiple Access with Collision Detection (CSMA/CD) access method and physical layer specifications (Amendment to IEEE Std 802.3-2002)*, Institute of Electrical and Electronics Engineers, New York, NY, April 19, 2004.
2. A. Gumaste and T. Anthony, *First Mile Access Networks and Enabling Technologies*, Cisco Press, Indianapolis, IN, 2004.

3. G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A dynamic protocol for an Ethernet PON (EPON)," *IEEE Communications* **40**, pp. 74–80, February 2002.
4. M. Ma, Y. Zhu, and T. H. Cheng, "A bandwidth guaranteed polling MAC protocol for Ethernet passive optical networks," in *Proceedings of IEEE Infocom 2003*, (San Francisco, CA), April 2003.
5. G. Kramer, B. Mukherjee, and A. Maislos, "Ethernet passive optical networks," in *Multiprotocol over DWDM: Building the Next Generation Optical Internet*, S. Dixit, ed., John Wiley & Sons, Inc., March 2003.
6. G. Kramer, "Presentation on Contention Resolution." (http://grouper.ieee.org/groups/802/3/efm/public/nov02/p2mp/index.html), November 2002.
7. G. Kramer and G. Pesavento, "Ethernet Passive Optical Network (EPON): Building a next-generation optical access network," *IEEE Communications* **40**, pp. 66–73, February 2002.