Closed-Form Expression for the Collision Probability in the IEEE EPON Registration Scheme

Swapnil Bhatia and Radim Bartoš
Department of Computer Science
University of New Hampshire
Durham, NH 03824, USA
Email: {sbhatia, rbartos}@cs.unh.edu

***Abstract*— We derive a closed-form expression for the message collision probability in the IEEE 802.3ah Ethernet Passive Optical Network (EPON) registration scheme. The expression obtained, although based on an approximation, shows a good match with simulation results. We use the results of our analysis to compute the size of the most efficient contention window and the most efficient number of nodes serviced by a given window size.**

I. Motivation

Protocols for emerging access network technologies such as DOCSIS [1], EPON [2] and some wireless technologies include a preliminary phase where the subscriber device must register with a headend or base station residing at a central office. Since this is the first communication between the headend and the subscriber device, no information about key parameters such as latency or timing is available to either party. Subscriber devices may be located at random distances unknown to the headend. As a result, most protocols rely on some collision avoidance scheme in order to reduce contention in the use of the communication channel. The recently adopted IEEE 802.3ah EPON standard prescribes the Random Delay scheme for this purpose. In this scheme, the headend broadcasts the size of a contention interval. The nodes, upon receiving this message, wait for a uniformly random interval and then transmit their registration message. In this paper, we derive a closed-form expression for the probability of message collision in the this scheme and validate our result through simulation.

To our knowledge, this is the first attempt at computing the probability of collision for the IEEE EPON registration scheme. While previous work in this area [3], [4] serves as an excellent general reference, its focus has primarily been on the stability and throughput of multiaccess schemes. Many of its assumptions (Poisson arrivals, backlogged nodes etc.) are either not relevant to or are out of scope of our current work. Our own past work [5] focuses on the high load performance characterization of the IEEE EPON registration scheme through simulations. A more recent analysis [6] is restricted only to the simpler case of identically distanced nodes. We propose a more generic model applicable to identically distanced as well as randomly distributed nodes and extensible to multiple node clusters. Inclusion of the random round trip delay together with the random contention window size and parameterization based on message size, contention window size and round trip time makes our work directly applicable to a practical analysis of the IEEE EPON registration protocol.

II. Derivation of Collision Probability

Let X, Y and Z be random variables. Let X, Y be independent and have a uniform distribution with $X \in \text{Uniform}[0, M]$, $Y \in \text{Uniform}[0, m]$, $M \geq m \geq 0$. To simplify the exposition, we first consider $m > 0$ and add $m = 0$ as a separate case later. Let f_X, f_Y denote their probability mass function of X and Y respectively. Thus, $f_X(x) = \frac{1}{M}$ and $f_Y(y) = \frac{1}{m}$. Let $Z = X + Y$. Since X and Y are independent, their joint density $f_{XY}(x, y) = \frac{1}{Mm}$. Let $F_Z(z) = P(Z \leq z)$ denote the cumulative distribution function (CDF) of Z. We compute the CDF of Z, by integrating $X + Y = Z$ w.r.t z [7]. Our main derivation involves the computation of many such integrals and we omit the details of these calculations due to space constraints. The limits used for each integral are specified in the accompanying figures and should aid the reader in computing the integrals, if desired. Thus, we have:

$$F_Z(z) = \begin{cases} 0 & \text{if } z < 0 \\ \frac{z^2}{2Mm} & \text{if } 0 \leq z \leq m \\ \frac{2z-m}{2M} & \text{if } m < z \leq M \\ \frac{-(m^2+z^2-2zM+M^2-2zm)}{2mM} & \text{if } M < z \leq M+m \\ 1 & \text{if } z > M+m. \end{cases} \quad (1)$$

To find the probability mass function $f_Z(z)$, we differentiate w.r.t. z to get:

$$f_Z(z) = \begin{cases} f_1(z) = \frac{z}{mM} & \text{if } 0 \leq z \leq m \\ f_2(z) = \frac{1}{M} & \text{if } m < z \leq M \\ f_3(z) = \frac{(m+M-z)}{mM} & \text{if } M < z \leq M+m \end{cases} \quad (2)$$

and 0 otherwise. Next, consider two i.i.d. random variables Z_1 and Z_2 with probability mass functions $f_{Z_1}(z_1)$ and $f_{Z_2}(z_2)$ as described by Eqn. (2). Due to the piecewise structure of $f_Z(z)$, the joint density of Z_1 and Z_2 will comprise of nine regions defined by the three cases each for Z_1 and Z_2 as shown in each of the Figs. 2—8. The shaded area in each figure shows the region where $|Z_1 - Z_2| \leq k$. Let P be the event $|Z_1 - Z_2| \leq k$. To find the probability of the event P we must consider several cases arising from the magnitudes of M, m and k relative to each other. In each case, we also

0-7803-9415-1/05/$20.00 (C) 2005 IEEE
This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2005 proceedings.

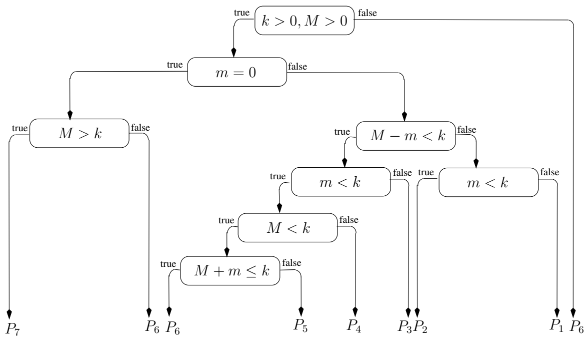


Fig. 1. $P(|Z_1 - Z_2| \leq k)$

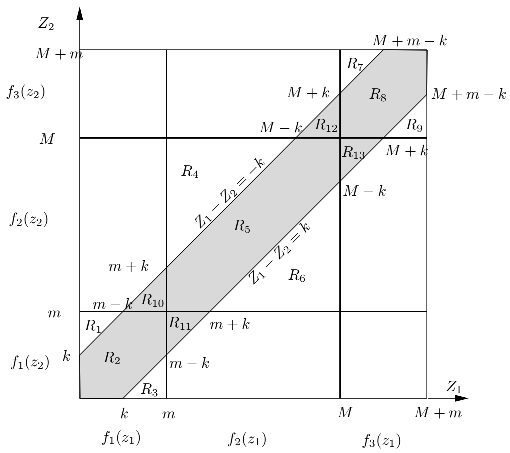


Fig. 2. P_1: $m \geq k, M > m$ and $M - m \geq k$

need to consider whether $m, M \leq k$ or $m, M > k$. Finally, $M + m < k$ or $m = 0$ are other special cases. Together, all the scenarios result in a function of the form expressed in Fig. 1. Figs. 2—8 illustrate each of the P_i in Fig. 1. Using these figures, we calculate the probability contained in the shaded region for each P_i. Integrating piecewise within the limits assigned to the shaded region in each figure, we obtain the following expressions for each of the P_i in Fig. 1:

$$P_1 = \frac{k(k^3 - 4m^3 - 4mk^2 + 12Mm^2)}{6m^2M^2} \quad (3)$$

$$P_2 = \frac{(12Mk - m^2 - 6k^2)}{6M^2} \quad (4)$$

$$\begin{aligned} P_3 \quad = \quad & \left[\frac{12Mkm^2 + 12mkM^2 - 12mk^2M}{12m^2M^2}\right. \\ & + \frac{6m^2M^2 + m^4 + 3k^4 - 4km^3}{12m^2M^2} \\ & + \frac{6k^2m^2 - 4mk^3 - 4kM^3}{12m^2M^2} \\ & + \frac{6k^2M^2 - 4Mk^3 + M^4}{12m^2M^2} \\ & \left. - \frac{4mM^3 + 4Mm^3}{12m^2M^2}\right] \end{aligned} \quad (5)$$

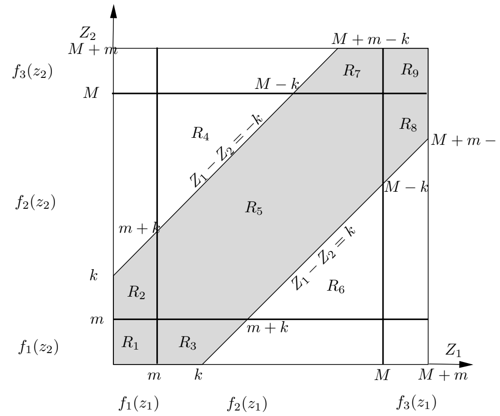


Fig. 3. P_2: $m < k, M > m$ and $M - m \geq k$

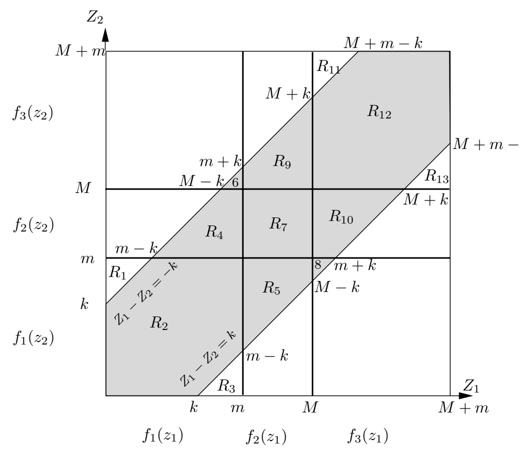


Fig. 4. P_3: $m \geq k, M \geq m$ and $M - m < k$

$$\begin{aligned} P_4 \quad = \quad & \left[\frac{6m^2M^2 - 4mM^3 + 12mM^2k}{12m^2M^2}\right. \\ & - \frac{12mMk^2 - 4mk^3 + m^4}{12m^2M^2} \\ & + \frac{12Mkm^2 - 4M^3k + 6M^2k^2}{12m^2M^2} \\ & - \frac{4Mk^3 - 4km^3 + 4Mm^3}{12m^2M^2} \\ & \left. + \frac{k^4 + M^4 - 6k^2m^2}{12m^2M^2}\right] \end{aligned} \quad (6)$$

$$\begin{aligned} P_5 \quad = \quad & \left[\frac{12km^2M + 12kmM^2 - 12k^2mM}{12m^2M^2}\right. \\ & + \frac{6m^2M^2 - k^4 - 4m^3M - 4mM^3}{12m^2M^2} \\ & + \frac{4mk^3 + 4Mk^3 - M^4 - m^4}{12m^2M^2} \\ & \left. + \frac{4km^3 + 4kM^3 - 6k^2m^2 - 6k^2M^2}{12m^2M^2}\right] \end{aligned} \quad (7)$$

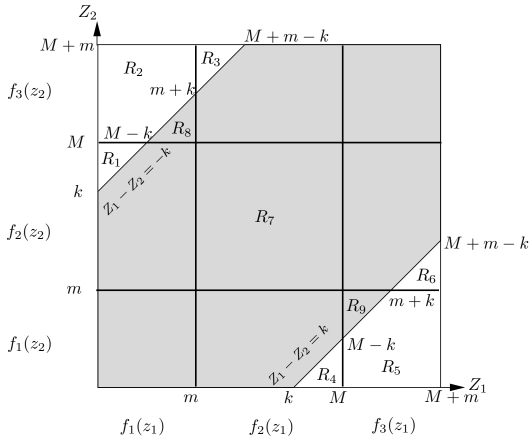


Fig. 5. P_4: $m < k, M > m, M \geq k$ and $M - m < k$

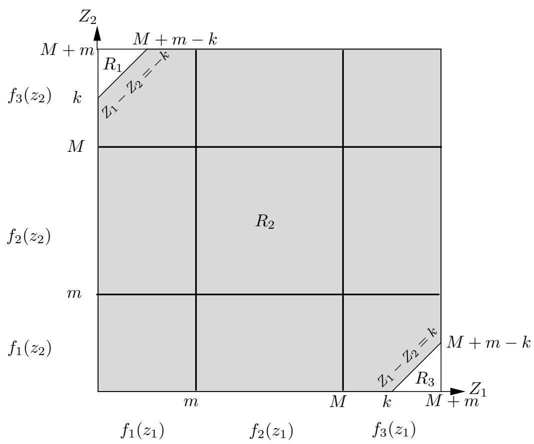


Fig. 6. P_5: $m < k, M \geq m, M < k$ and $M - m < k, M + m > k$

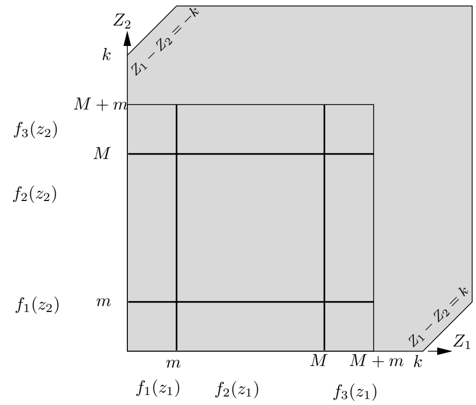


Fig. 7. P_6: $m < k, M \geq m, M < k$ and $M - m < k, M + m \leq k$

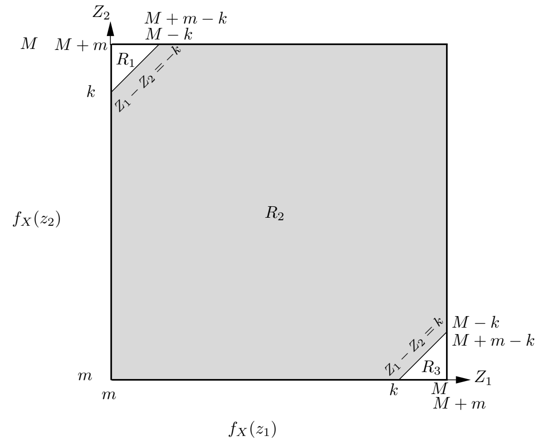


Fig. 8. P_7: $m = 0, M > k$

$$P_6 \quad = \quad 1 \tag{8}$$

If $m = 0$, then $Y \in$ Uniform$[0, 0]$ and hence $Z = X$. Therefore, $f_Z = f_X$.

$$P_7 = \frac{k(2M - k)}{M^2} \tag{9}$$

III. Probability of Message Collision in the IEEE EPON Registration Scheme

We can now use the derivation in the previous section as a backdrop to cast the contention in the IEEE EPON registration scheme. Due to technological constraints on the power and reach of a transmitted signal, the IEEE EPON standard [2] fixes the maximum distance from the headend at which a node may be located. We assume the maximum reach of our network to be such as to result in a maximum one-way propagation time of p and the maximum random wait time to be w (also fixed by the headend). As per the EPON protocol, the headend broadcasts a discovery message to signal the beginning of a special interval reserved for new-node registration. A new node, upon receiving it, replies with a registration request message transmitted after a random wait. If two such registration messages, say of length k each, arrive at the headend overlapping in time, then there is a collision. Thus, to detect a collision we must model the arrival time of a message at the headend. We observe that this arrival time is a sum of the one-way propagation time of the broadcast message from the headend to the node, a random wait at the node and another one-way propagation time of the registration request message from the node to the headend. Thus, the arrival time can vary between a minimum of 0 and a maximum of $2p + w$. We let $M = \max(2p, w)$ and $m = \min(2p, w)$ to ensure $M \geq m$. Thus, $X \in$ Uniform$[0, M]$ and $Y \in$ Uniform$[0, m]$ model the two-way propagation time and the random wait. Thus, $Z = X + Y$ models the arrival time of a message from a node at the headend. For two nodes $Z_1 = X_1 + Y_1$ and $Z_2 = X_2 + Y_2$, $|Z_1 - Z_2| \leq k$ represents the condition indicating a message collision at the headend. We already computed the probability of this event in the previous section and is given by the function shown in Fig. 1.

Fig. 9 (left) shows the probability of collision for two nodes participating in the IEEE EPON registration scheme with a message length $k = 2.528$ μs as specified in the IEEE EPON standard. While the value for the parameter p is also specified in the standard as 100 μs, the range of values for p in the figure

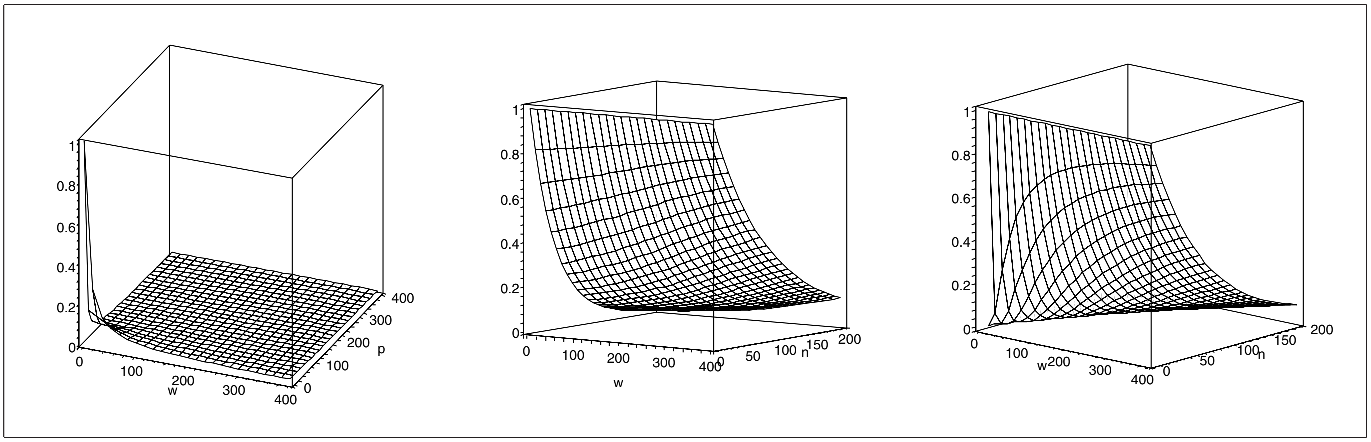


Fig. 9. *Left:* Probability of collision with 2 randomly distanced nodes. Probability of successful transmission with n randomly distanced *(center)* and n identically distanced *(right)* nodes.

allows us to use the same model to compute the probabilities for clustered nodes or nodes situated at an identical distance. The range of values for the wait period w is unspecified by the standard and is open to various implementation schemes.

We now extend our 2-node model to n nodes. (Note that n denotes the number of nodes attempting to register in a contention window and may not be the total number of devices in the EPON.) Let $P_s(k)$ and $P_c(k)$ denote the probability of successful and unsuccessful transmission (i.e. collision) respectively for a node in presence of $k-1$ other nodes. The probability of successful transmission in the 2-node case is thus $P_s(2) = 1 - P_c(2)$. A successful transmission by a node in the presence of $n-1$ other nodes implies that its transmission did not collide with any of the other $n-1$ nodes. If we assume independence of each pairwise collision event, we can write:

$$P_s(n) = P_s(2)^{n-1}. \tag{10}$$

Fig. 9 (center) shows the probability of successful transmissions for 1 to 200 nodes for a range of waiting times. The propagation time p is set to 100 μs. We can also formulate the scenario where all the nodes are at an identical distance by setting $p = 0$. Fig. 9 (right) shows the performance of the scheme for 1 to 200 nodes located at identical distances. Figs. 10 and 11 compare the results from simulation plotted with those from our closed form expression. Our model matches the simulation precisely except for a small range of window sizes in the uniformly random case when the number of devices is very large. This error results from our assumption about the independence of two or more collision events. Our simulations show that the error introduced is negligible and is present only for a small range of window sizes. Work on an alternative derivation free of this assumption is currently in progress.

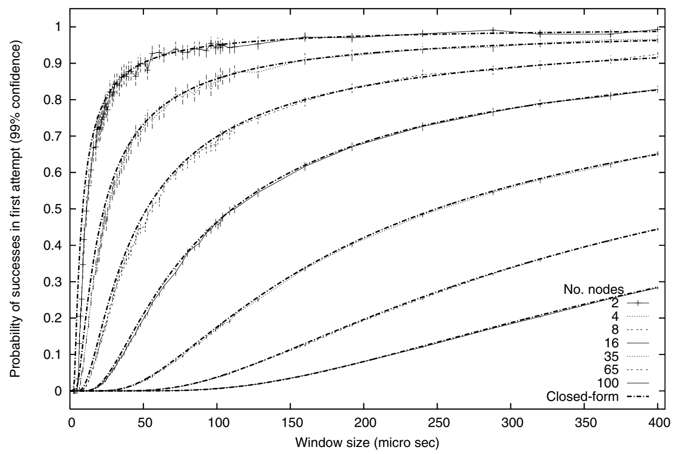


Fig. 10. Comparison of the probability of success for n identically distanced nodes obtained from simulation and closed-form expression. Thick and thin lines show the value from the closed-form expression and simulation respectively in each curve. Vertical lines show 99% confidence intervals.

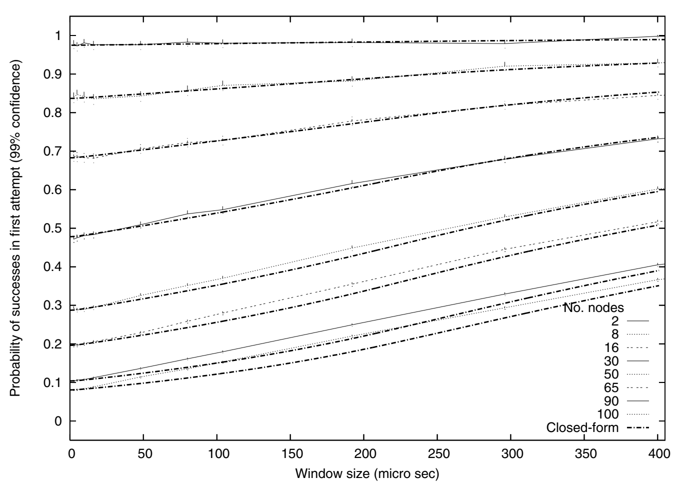
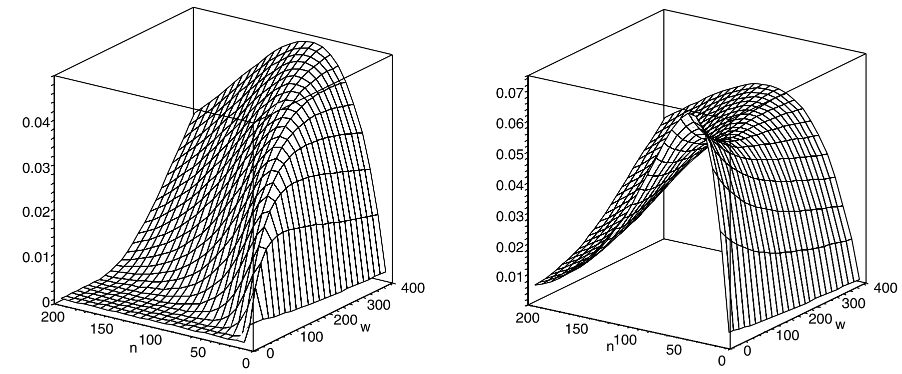


Fig. 11. Comparison of Probability of Success for n uniformly randomly distanced nodes obtained from simulation and closed-form expression. Thick and thin lines show the value from the closed-form expression and simulation respectively in each curve. Vertical lines show 99% confidence intervals.

IV. EFFICIENCY OF THE CONTENTION WINDOW

In the IEEE EPON registration scheme, the headend must reserve the communication link every time it wishes to allow new nodes to register. Thus, a valuable portion of the available bandwidth is used at every such discovery cycle. The headend

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE GLOBECOM 2005 proceedings.



# devices ($p = 0$)	Most efficient window size (μs)
2	35.82
4	64.63
8	105.20
10	122.39
16	168.43
32	273.77
64	459.65
100	655.74
200	1179.31

Fig. 12. Contention window efficiency for n identically distanced *(left)* and randomly distanced *(center)* nodes. *Right:* The most efficient contention window size for identically distanced nodes.

must reserve the channel for a duration of $2p_{\max} + w$ where $p_{\max} = 100\ \mu$s as specified by the IEEE standard. It is desirable to minimize this duration when the channel is exclusively used for discovering new devices—regular traffic cannot be transmitted. To take this criterion into account, we can define a measure for the efficiency of a particular contention window size as the ratio of the average number of successful registrations to the size of the duration of the reservation [6]. Thus, efficiency

$$\rho = \frac{n \cdot P_s(n)}{2p_{\max} + w}. \tag{11}$$

We use our n-node model to relate efficiency to window size and node number. Fig. 12 (left, center) shows the variation of efficiency with the window size and node number for the identically distanced and the uniformly randomly distanced cases respectively. For the identically distanced case, Fig. 12 (right) shows the most efficient window size for a given number of nodes, i.e., the smallest window size that maximizes the success probability. Due to the shape of the surface in Fig. 12 (center) equivalent maxima cannot be obtained for the uniformly random case. However, Fig. 13 shows the most efficient number of nodes that can be serviced by a contention window of a given size.

V. Summary and Future Work

We derived a closed form expression for the probability of message collision in the IEEE EPON registration scheme. We compared the probability computed by the expression with simulation results and obtained a reasonably precise match. We used our results to compute the most efficient contention window sizes for identically and randomly distributed nodes.

The current model for the n-node case relies on the assumption of independence of the collision event. Despite the assumption, the probabilities obtained match rather well with simulation results. However, any two collision events say $|Z_1 - Z_2| \leq k$ and $|Z_1 - Z_3| \leq k$ are not independent. Thus, for n devices, a conditional distribution of the collision event can provide an exact model.

Acknowledgment

This project was supported in part by a grant from the Cisco Systems University Research Program. We sincerely thank Glen Kramer for providing valuable feedback. We thank the reviewers for providing valuable feedback and helping us improve the quality of this paper.

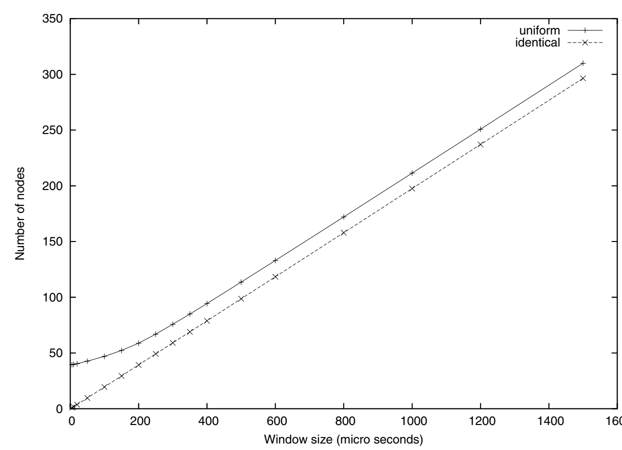


Fig. 13. Number of nodes served by a window size with maximum efficiency.

References

[1] *Data-Over-Cable Service Interface Specifications-Radio Frequency Interface Specification*, Cable Television Laboratories, Inc. Standard, Rev. SP-RFI v1.1-I06-001215, Version 1.1, 2000.
[2] *IEEE Draft P802.3ah (tm)*, IEEE Standard, 2004.
[3] S. Tasaka, *Performance Analysis of Multiple Access Protocols*. MIT Press, Computer Systems Series, 1986.
[4] D. Bertsekas and R. Gallager, *Data Networks*, 2nd ed. Prentice Hall, 1992, ch. 4.
[5] S. Bhatia and R. Bartoš, "Performance of the IEEE 802.3 EPON Registration Scheme Under High Load," in *Proc. of the SPIE Optics East Conference on Performance, Quality of Service and Control of Next-Generation Communication Networks II*, Philadelphia, USA, Oct. 2004.
[6] G. Kramer, *Ethernet Passive Optical Networks*. McGraw-Hill Professional, March, 2005.
[7] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th ed. McGraw Hill, 2002, ch. 5, 6, 7.