IPACT: A Dynamic Protocol for an Ethernet PON (EPON)

Glen Kramer and Biswanath Mukherjee, University of California, Davis
Gerry Pesavento, Alloptic, Inc. Livermore

ABSTRACT

We investigate design issues for access networks based on passive optical network technology. A PON based on polling, with data encapsulated in Ethernet frames, possesses many desirable qualities, such as dynamic bandwidth distribution, use of a single downstream and a single upstream wavelength, ability to provision a fractional wavelength capacity to each user, and ease of adding a new user. To support dynamic bandwidth distribution, we propose an interleaved polling algorithm called IPACT. We also suggest a scheme for in-band signaling that allows using a single wavelength for both downstream data and control message transmission. To obtain realistic simulation results, we generated synthetic traffic that exhibits the properties of self-similarity and long-range dependence. We then analyzed the network performance under varying offered loads.

INTRODUCTION

Passive optical network (PON) is a technology viewed by many as an attractive solution to the last mile problem [1, 2]. A PON is a point-to-multipoint optical network with no active elements in the signals' path from source to destination. The only interior elements used in PON are passive combiners, couplers, and splitters.

Advantages of using PON for a subscriber access network include large coverage area, reduced fiber deployment, multicast and broadcast capabilities, reduced cost of maintenance (due to devices being passive), and ease of upgrades to higher bit rate or additional wavelengths.

The work of standardizing Ethernet PON (EPON) as a solution for access networks is ongoing in the IEEE 802.3ah Ethernet in the First Mile (EFM) Task Force. The group's focus is on ensuring interoperability by specifying parameters and operations of physical and data link layers. Particular bandwidth allocation algorithms are considered to be vendor-specific and outside of scope of the task force. In this study we describe a new dynamic bandwidth allocation algorithm and analyze its performance in an EPON.

In [3] we discussed the advantages of using time-division multiple access (TDMA) in a PON, namely the scalability and ability to provide a fraction of a wavelength capacity to a user, a single wavelength for all upstream channels, a single receiver in the head-end, and so on. However, we also showed that a considerable amount of bandwidth was wasted due to time slots not being filled to capacity. To make the cost of a PON-based access network lower, it is very important to utilize bandwidth efficiently.

In this study we present a new protocol called Interleaved Polling with Adaptive Cycle Time (IPACT, pronounced *eye-pact*). The PON-based network under consideration uses a polling scheme to deliver data encapsulated in Ethernet packets from a collection of optical network units (ONUs) to a central optical line terminal (OLT) over the PON access network. The OLT, in turn, is connected to the rest of the Internet.

To avoid the accumulation of walk times (switchover times) associated with polling, we employ an *interleaved* scheme where multiple polling requests are overlapped in time. We then discuss an efficient way to use in-band control signaling to perform the polling.

We present simulation results to demonstrate system performance such as bounds on packet delay, queue occupancy, and packet-loss probability.

DESIGN OF AN ACCESS NETWORK BASED ON PON TECHNOLOGY

Figure 1 shows a typical tree-based PON topology suitable for the access network. Other PON topologies include a ring and a bus. A PON can also be deployed in a redundant configuration as a double ring or a double tree. All transmissions in a PON are performed between OLT and ONUs. Therefore, in the downstream direction (from OLT to ONUs), PON is a point-to-multipoint network, and in the upstream direction it is a multipoint-to-point network.

The OLT resides in the local exchange (central office), connecting the optical access network to an IP, ATM, or SONET backbone. The ONU is located at either the curb (FTTC solution), or the end-user location (FTTH, FTTB solutions), and provides broadband voice, data, and video services.

The Full Service Access Network (FSAN) standard defines an optical access network that uses asynchronous transfer mode (ATM) as its transport protocol. However, ATM equipment is significantly more expensive than Ethernet [4]. Thus, FSAN's future, relying on ATM transport, looks rather grim. EPONs, on the contrary, appear to be the preferred choice. Newly adopted quality of service (QoS) techniques have made Ethernet networks capable of supporting voice, data, and video. Ethernet is an inexpensive technology that is interoperable with a variety of legacy equipment. In this study we will focus on EPONs.

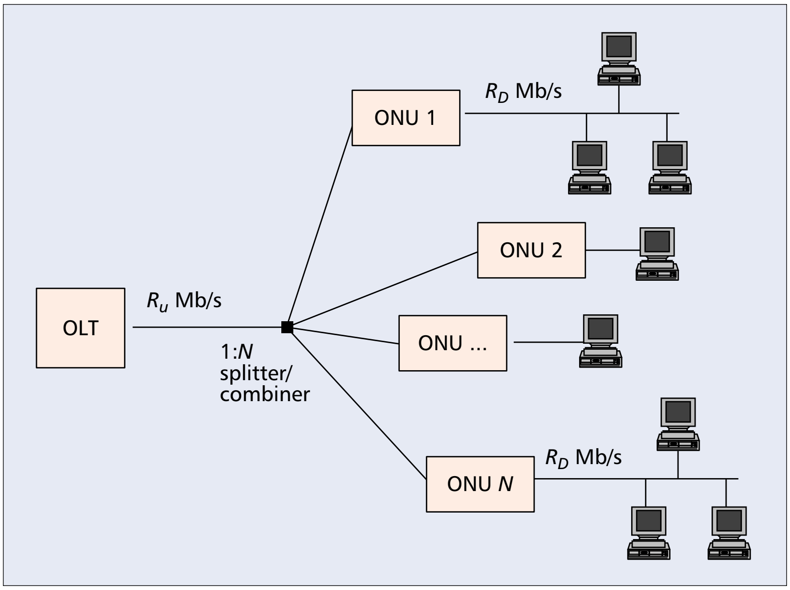


Figure 1. *A typical access network based on tree PON topology.*

CHANNEL MULTIPLEXING

Because Ethernet is broadcast by nature, in the downstream direction (from network to user) it fits perfectly with the EPON architecture: packets are broadcast by the OLT and extracted by their destination ONU based on the media access control (MAC) address.

In the upstream direction (from user to network), the ONUs should share the channel capacity and resources. We believe the TDMA approach is a preferred method of channel sharing in an access network since it allows using a single upstream wavelength and results in a very cost-effective solution. However, in [3] we also showed the limitation of TDMA approach: the lack of statistical multiplexing.

The burstiness of network traffic results in a situation where some time slots overflow even under very light load, resulting in packets being delayed for several time slot periods. It is also true that some time slots remain underutilized (not filled completely) even if the traffic load is very high. This leads to the PON bandwidth being underutilized.

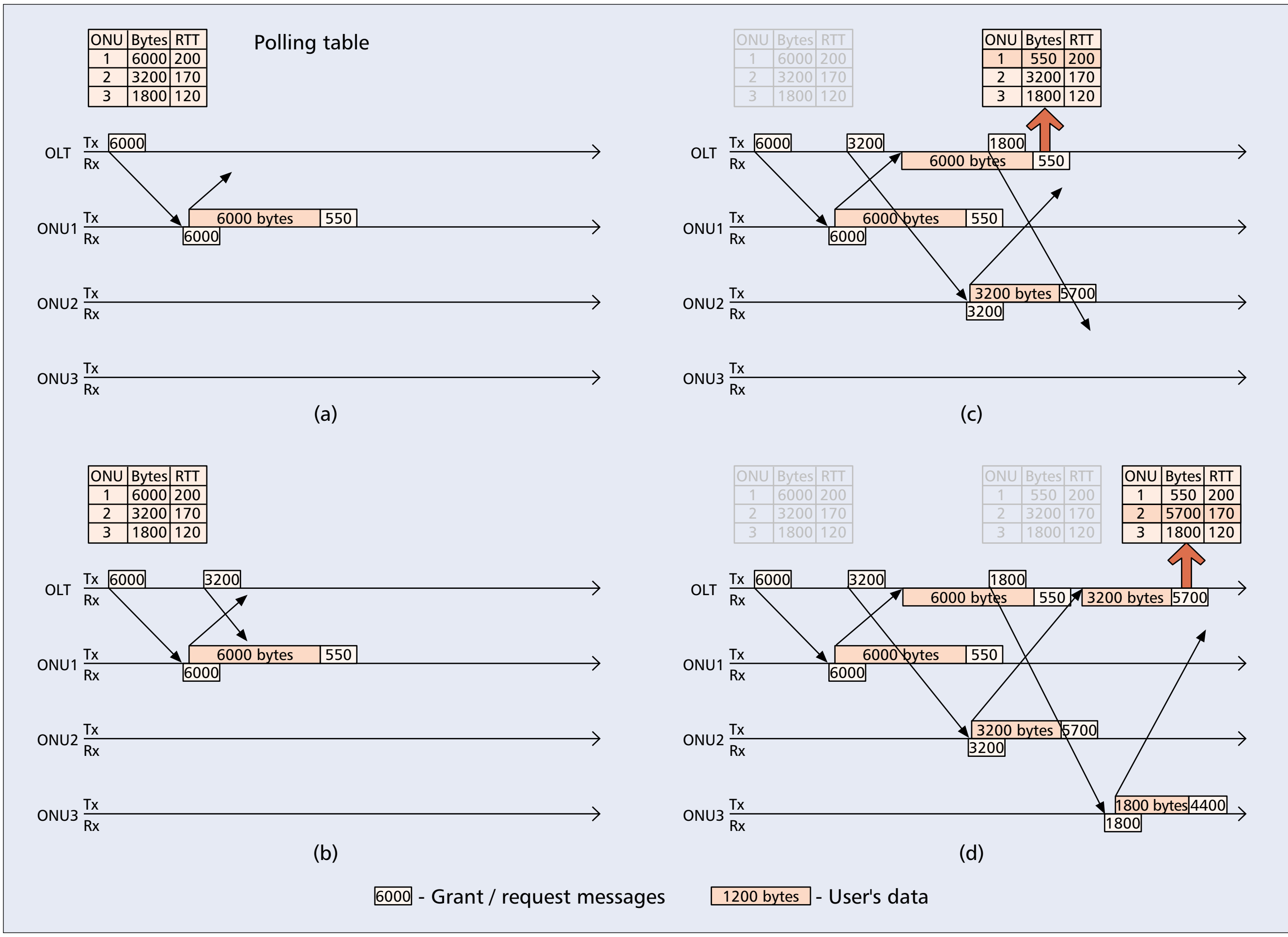
A dynamic scheme that reduces the time slot size when there is no data would allow the excess bandwidth to be used by other ONUs. The challenge of implementing such a scheme is in the fact that the OLT doesn't know exactly how many bytes of data each ONU has.

Below we present an OLT-based polling scheme, similar to hub polling [5]. Our algorithm uses an interleaved polling approach where the next ONU is polled before the transmission from the previous one has arrived. This scheme provides statistical multiplexing for ONUs and results in efficient upstream channel utilization.

INTERLEAVED POLLING WITH ADAPTIVE CYCLE TIME

In this section we give a high-level overview of the proposed algorithm. For simplicity of illustration, we will consider a system with three ONUs:

1. Assume that at some time t_0 the OLT knows exactly how many bytes are waiting in each ONU's buffer and the round-trip time (RTT) to each ONU. OLT keeps this data in a polling table, shown in Fig. 2a. At time t_0 the OLT sends a control message to ONU1, allowing it to send 6000 bytes (Fig. 2a). We call such a message a *Grant*. Since, in the downstream direction, the OLT broadcasts data to all ONUs, a Grant should contain the ID of the destination ONU, as well as the size of the granted window (in bytes).
2. Upon receiving the Grant from the OLT, ONU1 starts sending its data up to the size of the granted window (Fig. 2b), in our example up to 6000 bytes. At the same time the ONU keeps receiving new data packets from users. At the end of its transmission window, ONU1 will generate its own control message (*Request*). The Request sent by ONU1 tells the OLT how many bytes were in ONU1's buffer at the moment the Request was generated. In our case there were 550 bytes.
3. Even before the OLT receives a reply from ONU1, it knows when the last bit of ONU1's transmission will arrive. This is how OLT knows it:
   (a) The first bit will arrive exactly after the RTT time. The RTT in our calculation includes the actual RTT, Grant processing time, and a preamble for the OLT to perform bit and byte alignment on received data, that is, exactly the time interval between sending a Grant to an ONU and receiving data from the same ONU.
   (b) Since the OLT knows how many bytes (bits) it has authorized ONU1 to send, it knows when the last bit from ONU1 will arrive. Then, knowing RTT for ONU2, the OLT can schedule a Grant to ONU2 such that the first bit from ONU2 will arrive soon after the last bit from ONU1, with only a small guard interval in between (Fig. 2b). The guard intervals provide protection for fluctuations of RTT and control message processing time of various ONUs. Additionally, the OLT receiver needs some



■ **Figure 2.** *Steps of the Polling Algorithm.*

time to readjust its sensitivity due to the fact that every ONU is located at a different distance from the OLT (far-near problem).

4. After some time, the data from ONU1 arrives. At the end of the transmission from ONU1, there is a new Request that contains information on how many bytes remained in ONU1's buffer when the Request transmission began. The OLT will use this information to update its polling table (Fig. 2c). By keeping track of times when Grants are sent out and data is received, the OLT constantly updates the RTT entries for the corresponding ONUs.
5. Similarly to the above step, the OLT can calculate the time when the last bit from ONU2 will arrive. Hence, it will know when to send the Grant to ONU3 so that its data is tailed to the end of ONU2's data. After some more time, the data from ONU2 will arrive. The OLT will again update its table, this time the entry for ONU2 (Fig. 2d).

Note that if an ONU emptied its buffer completely, it will report 0 bytes back to the OLT. Correspondingly, in the next cycle, the ONU will be granted 0 bytes, that is, it will be allowed to send a new Request, but no data.

It should be clear from the above description that there is no need to synchronize the ONUs, nor is there a need to perform a ranging (making ONUs to appear equidistant from the OLT by delaying the response from ONUs by a specific amount of time) traditionally employed in TDMA schemes. Every ONU executes the same procedure driven by the Grant messages received from the OLT. The entire scheduling and bandwidth allocation algorithm is located in the OLT. Thus, it is easy to adaptively change the scheduling at run-time based on some network conditions; the ONUs don't need to negotiate or acknowledge new parameters, nor do they need to switch to new settings synchronously.

In the above algorithm, we started with the OLT already having its table populated. During system initialization, since round-trip times are unknown, OLT should poll each ONU one at a time. The reader is referred to [6] for a detailed description of cold start and ONU initialization procedures.

Control-Message Format — The Request and Grant messages should only contain two pieces of information: ONU's node identification (NID) and requested/granted window size (WS). Format of control messages may have high impact on overall system performance. For example, if the control messages are encapsulated in Ethernet frames, the two problems become apparent:

• Grant blocking behind long downstream frame increases guard band and consequently degrades upstream utilization.

- In case of asymmetric load (high downstream load and light upstream load) cycle time becomes very short, which results in more frequent Grants consuming more downstream bandwidth.

Fortunately, another solution is available. First, we note that an Ethernet frame with its 64-byte minimum size is overkill for a control message consisting only of 1-byte NID and 2-byte WS fields. The solution we propose is to embed the Grant messages inside the downstream data packets using escape sequences. To understand this approach, first recall that Ethernet uses 8-to-10 bit encoding as defined by IEEE standard 802.3. However, not all of 10-bit values are valid encoding of an 8-bit value. One or more of these "non-valid" codes can be chosen to represent an escape code. Thus, the control message (either Grant or Request) will be 4 bytes long and will consist of 1-byte escape code (ESC), 1-byte node ID value, and 2 bytes of window size. Such a control message can be inserted in the middle of an Ethernet frame or between the frames. The receiver will recognize the beginning of the embedded control sequence by reading the ESC code. It will then extract the 3 bytes that follow the ESC byte before passing the rest of the received data to a standard Ethernet MAC.

Upstream control messages (Requests) will use the same 4-byte format and they will be sent at the end of the transmission from a given ONU.

Scheduling a Control Message — Grants are always scheduled one cycle ahead. The objective of Grant scheduling is to achieve a situation when transmissions from all ONUs arrive at OLT in order and without overlaps (collisions). In fact, to allow the receiver in the OLT adjust to a new power level and synchronize on new bit/byte boundary, we require a minimum gap (guard time) between transmissions from deferent ONUs.

The Grants are scheduled with regard to the corresponding round-trip times and granted window sizes. As a result, the order of Grants may be different in every cycle. Scheduling the Grant to ONU $i+1$ ahead of the Grant to ONU i is not a problem as the order of Grants is determined in a cycle prior to when they should be transmitted.

Scheduling the Grants as described above may result in a Grant scheduling conflict. The conflict occurs when two Grants are scheduled less than a Grant transmission time apart. Adapting to the message format described above and assuming a transmission speed of 1 Gb/s, the Grant transmission time is equal to 32 ns. To resolve such a conflict, the Grant that is scheduled last should be delayed till the end of transmission of the previous Grant. Grant delay has no significant effect on system performance; the only consequence of it is the corresponding delay of transmission from the ONU (i.e., increase of the guard time before that transmission). Of course, after the conflicting Grant is delayed, it may collide again with another Grant that was scheduled before. In the extreme case, some Grant may collide with at most $N - 1$ other Grants, where N is the number of ONUs. For the case of $N = 16$, the maximum acquired delay is 0.48 μs. Obviously, this solution does not introduce any scalability issues in terms of the value of N because multiple collisions only negligibly increase the guard time.

Maximum Transmission Window — If the OLT authorizes each ONU to send its entire buffer contents in one transmission, ONUs with high data volume could monopolize the entire bandwidth. To avoid this situation, the OLT will limit the maximum transmission size: every ONU gets a Grant to send as many bytes as it has requested, but no more than some limit (maximum transmission window size). There could be various schemes for specifying the limit. It can be fixed, say, based on a Service Level Agreement (SLA) for each ONU, or dynamic — based on network conditions. Let us denote an ONU-specific maximum transmission window size $W_{MAX}^{[i]}$. The choice of specific values of $W_{MAX}^{[i]}$ determines the maximum polling cycle time T_{MAX} under heavy load conditions. Making T_{MAX} too large will result in increased delay for all the packets, including high-priority (real-time) packets. Making T_{MAX} too small will result in more bandwidth being wasted by guard times.

Also, the $W_{MAX}^{[i]}$ value determines the guaranteed bandwidth available to ONU-i, that is, the ONU is guaranteed to be able to send $W_{MAX}^{[i]}$ bytes in at most T_{MAX} time. Of course, an ONU's bandwidth will be limited to its guaranteed bandwidth only if all other ONUs in the system also use all their available bandwidth. If at least one ONU has less data, it will be granted a shorter transmission window, thus making the cycle time shorter, and therefore the available bandwidth to all other ONUs will increase proportionally to their $W_{MAX}^{[i]}$. This is the mechanism behind dynamic bandwidth distribution: by adapting the cycle time to the instantaneous network load (i.e., queue occupancy), the bandwidth is automatically distributed to ONUs based on their loads.

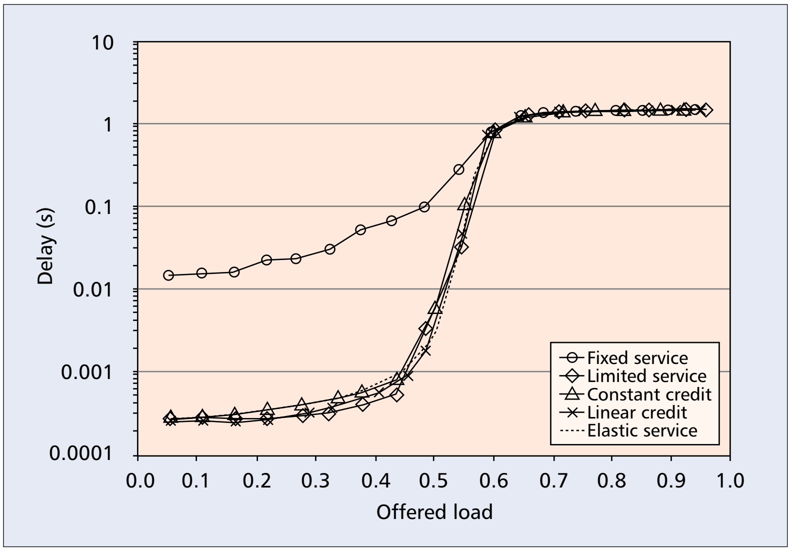
In our simulations we assume that all ONUs have the same maximum transmission window, that is, $W_{MAX}^{[i]} = W_{MAX}, \forall i$. We believe $T_{MAX} = 2$ ms and guard time of 5 μs are reasonable choices. That made $W_{MAX} = 15{,}000$ bytes. With that choice of parameters, every ONU will get a guaranteed bandwidth of 60 Mb/s. It can be shown that when only one ONU has data to send, the maximum bandwidth available to that ONU is 600 Mb/s [6].

The remaining question is how the OLT should determine the granted window size if the requested window size $W^{[i]} < W_{MAX}$. Below we describe a few approaches the OLT may employ in making its decision.

Fixed service ignores the requested window size and always grants the maximum window. As a result it has a constant cycle time T_{MAX}. Essentially this approach corresponds to the fixed TDMA PON system [3]. It is shown here only for comparison.

Limited service grants the requested number of bytes, but no more than W_{MAX}. It is the most conservative scheme and has the shortest cycle of all the schemes.

The Grants are scheduled with regard to the corresponding round-trip times and granted window sizes. As a result, the order of Grants may be different in every cycle.



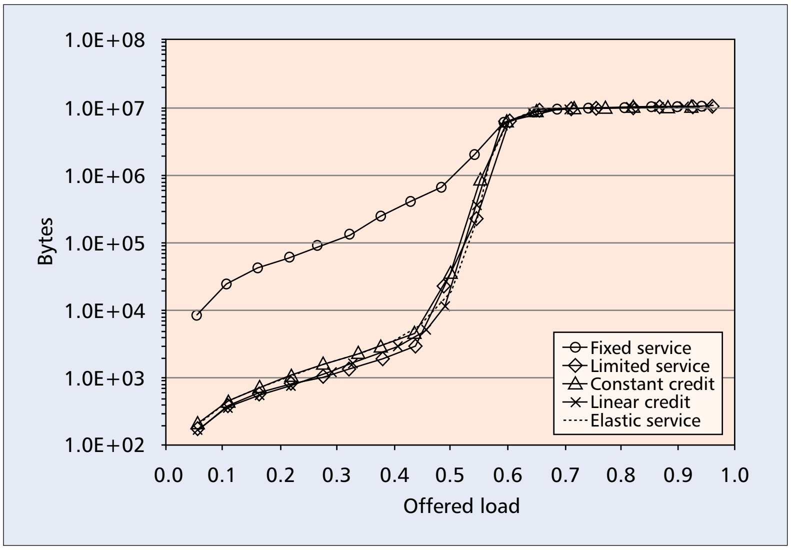
■ **Figure 3.** *Mean packet delay.*

The *Constant Credit* scheme adds a constant credit to the requested window size. The idea behind adding the credit is the following: assume x bytes arrived between the times an ONU sent a Request and received the Grant. If the granted window size equals requested window + x (i.e., it has a credit of size x), these x bytes will not have to wait for the next Grant to arrive; they will be transmitted with the current Grant, and the average packet delay will be shorter.

The *Linear Credit* scheme uses a similar approach as the Constant Credit scheme. However, the size of the credit is proportional to the requested window. The reasoning here is the following: network traffic possesses a certain degree of predictability [7] specifically; if we observe a long burst of data, this burst is likely to continue for some time into the future. Correspondingly, the arrival of more data during the last cycle may signal that we are observing a burst of packets.

Elastic service is an attempt to get rid of a fixed maximum window limit. The only limiting factor is the maximum cycle time T_{MAX}. The maximum window is granted in such a way that the accumulated size of last N Grants (including the one being granted) does not exceed $N \times W_{MAX}$ bytes (where N is the number of ONUs). Thus, if only one ONU has data to send, it may get a Grant of size up to $N \times W_{MAX}$.

It is worth mentioning that because each polling table entry should be updated before issuing a Grant to the corresponding ONU, the cycle time can never be less than maximum RTT to any of the ONUs in the PON. If an ONU is ready to be polled but the previous Request has not arrived yet, the algorithm will have to pause. To remedy such inefficiency, the algorithm may grant larger than requested transmission slots to ensure that sum of all granted slots and corresponding guard times is at least as large as the maximum RTT. We have not considered the effects of such optimization in this study.



■ **Figure 4.** *Average queue size.*

Model Description

In this study we consider a PON access network consisting of an OLT and N ONUs (Fig. 1). Every ONU is assigned a downstream propagation delay (from the OLT to the ONU) and an upstream propagation delay (from the ONU to the OLT). To keep the model general, we assume independent upstream and downstream propagation delays and select them randomly (uniformly) over the interval [50 µs, 100 µs]. These values correspond to distances between the OLT and ONUs ranging from 10 to 20 km.

The transmission speeds of the PON and user access link may not necessarily be the same. In our model we consider R_D Mb/s to be the data rate of the access link from a user to an ONU, and R_U Mb/s to be the rate of the upstream link from an ONU to the OLT (Fig. 1). We should mention here that if $R_U \geq N \times R_D$, the bandwidth utilization problem does not exist, since the system throughput is higher than the peak aggregated load from all ONUs. In this study we consider a system with $N = 16$ and R_D and R_U being 100 Mb/s and 1000 Mb/s, respectively. Every ONU has a finite memory buffer of size Q. In our simulations, Q was set to 10 Mbytes.

To obtain an accurate and realistic performance analysis, it is very important to simulate the system behavior with appropriate traffic injected into the system. Our simulation was performed using synthetic traffic traces that exhibit the properties of self-similarity and long-range dependence (LRD).

To generate self-similar traffic, we used the method described in [8], where the resulting traffic is an aggregation of multiple streams, each consisting of alternating Pareto-distributed ON/OFF periods. In our implementation every stream generates Ethernet packets that are transmitted in packet trains (bursts). The number of packets per burst (ON period) follows the Pareto distribution with a minimum of 1 (i.e., the smallest burst consist of only 1 packet) and shape parameter $\alpha = 1.4$. The choice of α was prompted by measurements on actual Ethernet traffic performed by Leland *et al.* [9]. They reported the measured Hurst parameter of approximately 0.8 for moderate network load.

The relationship between the Hurst parameter H and the shape parameter α is $H = (3 - \alpha)/s$ [8]. We have verified that the Hurst parameter of the resulting data stream indeed equals 0.8.

Simulation Results

In Fig. 3 we present the mean packet delay for different Grant scheduling services as a function of an ONU's offered load. In this simulation all ONUs had identical load.

As can be seen in the figure, all granting services except fixed service have almost coinciding plots. We will discuss fixed service results below. As for the rest of them, no other method gives a detectable improvement in packet delay. The explanation of this lies in the fact that all these methods are trying to send more data by way of increasing the granted window size. While that may clear the queue in fewer polling cycles, the polling cycle itself will increase. Overall, all these services have negligible effect on packet delay.

The fixed service plot is interesting as an illustration of the traffic long-range dependence. Even at the very light load of 5 percent, the average packet delay is already very high (~15 ms). This is because most packets arrive in very large packet trains. In fact, the packet trains were so large that the 10-Mbyte buffers overflowed and about 0.14 percent of packets were dropped. Why do we observe this anomalous behavior only with fixed service? The reason is that all other services have much shorter cycle times; there is just not enough time in a cycle to receive more bytes than W_{MAX}; thus, the queue never builds up. In fixed service, on the other hand, the cycle is large (and fixed) from the very beginning, and several bursts that arrive close to each other can easily overflow the buffer.

We want to note here that the reduced cycle time that adapts exactly to the amount of data available in the ONUs is the main advantage of the proposed algorithm.

Figure 4 presents the simulation results for the average queue size and is similar to the mean delay plot. Again, fixed service has a larger queue, which comes as no surprise. That leads to a conclusion that neither of the discussed service disciplines is better than limited service. As such, for the remainder of this study we will focus our attention on the limited service discipline.

Performance of Limited Service

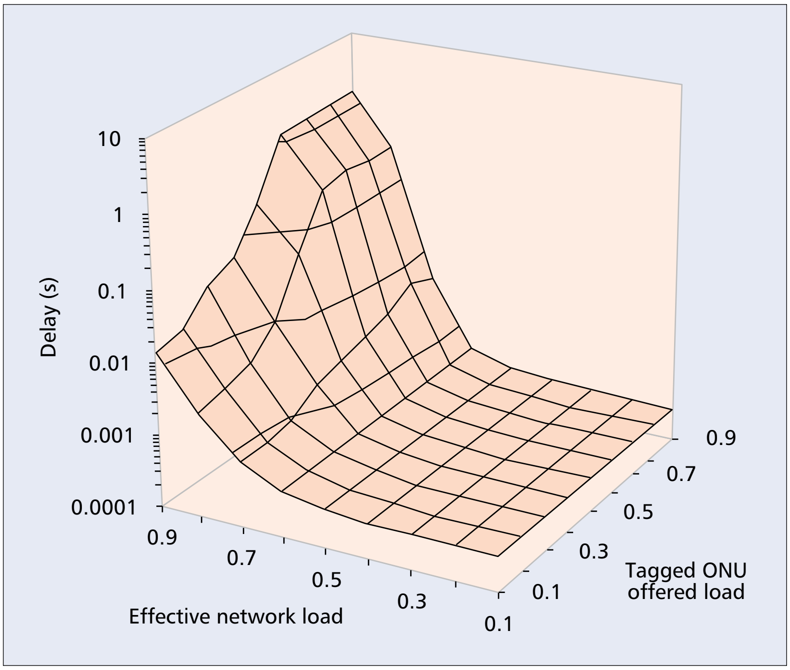
In this section we analyze the performance of a tagged ONU i as a function of its offered load and the effective load of the entire network. In Fig. 5, we present the average packet delay.

When the effective network load is low, all packets in a tagged source experience very little delay, no matter what the ONU's offered load is. This is a manifestation of dynamic bandwidth allocation — when the network load is low, the tagged source gets more bandwidth.

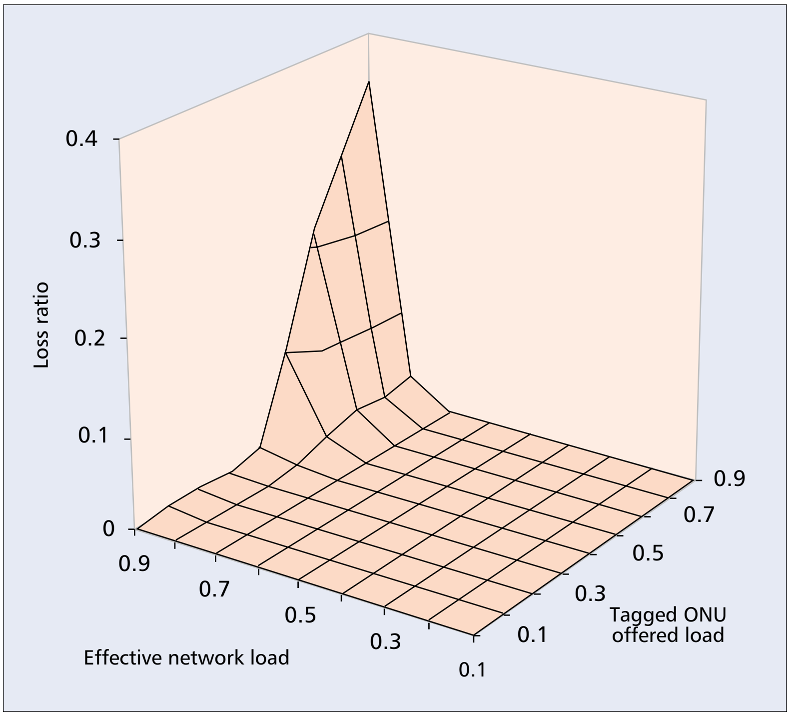
The opposite situation — low offered load at the ONU and high effective network load — results in higher delay. The only reason for this is the burstiness (i.e., long-range dependence) of the traffic. This is the same phenomenon observed with fixed service; the only difference is that in fixed service the cycle time is large and fixed independent of the effective load. In the current case, the cycle time is large because of increased effective network load. This cycle time is large enough to receive more than W_{MAX} bytes of data during a burst. Hence, the queuing delay for some packets will increase beyond one cycle time.

Figure 6 shows the probability of a packet loss in a tagged ONU i as a function of its offered load and the effective load of the entire network.

Once again we observe that packet loss is



■ **Figure 5.** *Average packet delay as a function of effective network load and ONU offered load.*



■ **Figure 6.** *Packet loss probability as a function of effective network load and ONU offered load.*

We believe that a PON based on polling and with data encapsulated in Ethernet frames possesses very desirable qualities, such as use of a single downstream and a single upstream wavelength, ability to provision a fractional wavelength capacity to each user, and so on.

zero or negligible if the effective network load is less than 80 percent. When the network load is above 80 percent and the tagged ONU offered load is above 50 percent (50 Mb/s), we observe considerable packet loss due to buffer overflow, even though the guaranteed bandwidth available to tagged ONU is 60 Mb/s. This is again the consequence of traffic burstiness.

CONCLUSION

In this study we discuss and evaluate design issues that must be dealt with in a PON access network. Specifically, to drive the cost of an access network down, it is very important to have an efficient scalable solution. We believe that a PON based on polling and with data encapsulated in Ethernet frames possesses very desirable qualities, such as use of a single downstream and a single upstream wavelength, and the ability to provision a fractional wavelength capacity to each user.

We present a simple algorithm for dynamic bandwidth allocation based on an interleaved polling scheme with an adaptive cycle time. We suggest a novel approach for an in-band signaling that allows use of a single wavelength for both downstream data and Grant transmission. Also, we showed this approach to be scalable with the number of ONUs in the system.

Since each ONU uses the window size required at the moment, the polling cycle time *adapts* to the instantaneous queue loads, leading to an adaptive cycle time. This is the basic idea behind the fair unused bandwidth redistribution: reduced cycle time leads to an increase in the amount of best-effort bandwidth available to busy ONUs. This increase is proportional to their bandwidth needs.

REFERENCES

[1] G. Pesavento and M. Kelsey, "PONs for the Broadband Local loop," *Lightwave*, vol. 16, no. 10, Sept. 1999, pp. 68–74.

[2] B. Lung, "PON Architecture 'Futureproofs' FTTH," *Lightwave*, vol. 16, no. 10, Sept. 1999, pp. 104–7.

[3] G. Kramer, B. Mukherjee, and G.Pesavento, "Ethernet PON (ePON): Design and Analysis of an Optical Access Network," *Phot. Net. Commun.*, vol. 3, no. 3, July 2001, pp. 307–19.

[4] S. Clavenna, "Metro Optical Ethernet," *Lightreading* (www.lightreading.com), Nov. 2000.

[5] J. L. Hammond and P. J. P. O'Reilly, *Performance Analysis of Local Computer Networks*, Addison Wesley, 1987.

[6] G. Kramer and B. Mukherjee, "Interleaved Polling with Adaptive Cycle Time (IPACT): Protocol Design and Performance Analysis," Tech. rep. CSE-2001-4, Dept. of Comp. Sci., UC Davis, Aug. 2001.

[7] K. Park and W. Willinger, "Self-Similar Network Traffic: An Overview," K. Park and W. Willinger, Eds., *Self-Similar Network Traffic and Performance Evaluation*, Wiley Interscience, 2000.

[8] W. Willinger *et al.*, "Self-Similarity through High-Variability: Statistical Analysis of Ethernet LAN Traffic at the Source Level," *Proc. ACM SIGCOMM '95*, pp. 100–13.

[9] W. Leland *et al.*, "On the Self-Similar Nature of Ethernet Traffic (Extended Version)," *IEEE/ACM Trans. Net.*, vol. 2, no. 1, Feb. 1994, pp. 1–15.

BIOGRAPHIES

GLEN KRAMER (glen.kramer@alloptic.com) is an R&D engineer in the Advanced Technology Laboratory of Alloptic, Inc. He received a B.S. degree in computer engineering from Kishinev Polytechnic Institute, Moldova, in 1992, and an M.S. degree from the University of California, Davis in 2000. He is a member of the Networks Research Laboratory at UC Davis. His research interests are in the area of PONs, broadband access network architectures and protocols, QoS in packet networks, traffic characterization, and modeling.

BISWANATH MUKHERJEE (mukerjee@cs.ucdavis.edu) received his B.Tech. (Hons.) degree from the Indian Institute of Technology, Kharagpur, in 1980, and his Ph.D. from University of Washington, Seattle, in June 1987. At Washington he held a GTE Teaching Fellowship and a General Electric Foundation Fellowship. In July 1987 he joined the University of California, Davis, where he has been professor of computer science since July 1995, and chairman of computer science since September 1997. He is co-winner of paper awards presented at the 1991 and 1994 National Computer Security Conferences. He serves on the editorial boards of *IEEE/ACM Transactions on Networking*, *IEEE Network*, *ACM/Baltzer Wireless Information Networks*, *Journal of High-Speed Networks*, *Photonic Network Communications*, and *Optical Network Magazine*. He also served as editor-at-large for optical networking and communications for the IEEE Communications Society. He served as Technical Program Chair of the IEEE INFOCOM '96 conference. He is author of the textbook *Optical Communication Networks* (McGraw-Hill, 1997), which received the Association of American Publishers, Inc.'s 1997 Honorable Mention in Computer Science. His research interests include lightwave networks, network security, and wireless networks.

GERRY PESAVENTO (gerry.pesavento@alloptic.com) is the founder and vice president of R&D at Alloptic, Inc. He is also the current chair of the EPON track within the IEEE P802.3ah Ethernet in the First Mile standards effort. His research interests include fiber optic components and optical Ethernet networks. He received a B.S.E.E. from Rensselaer Polytechnic Institute (RPI) and an M.B.A. from the University of California at Berkeley.