How efficient is EPON?

Glen Kramer, Teknovus, Inc.

INTRODUCTION

EPON efficiency depends on many parameters, such as packet size distribution, configuration of the scheduler, and the speed of the laser driver and clock recovery circuits, etc. Making unrealistic assumptions about any of these parameters can result in efficiency numbers being far off from the true value.

It is, therefore, clear that to answer the question of EPON efficiency, one has to come up with an unambiguous set of EPON operational parameters and traffic characteristics. In this article, we attempt to identify all the parameters affecting the efficiency and will justify the chosen values for these parameters.

WHAT IS EPON EFFICIENCY?

By network efficiency we usually mean the *throughput efficiency*, also called *utilization*. Throughput is a measure of how much user data (application-level data) the network can carry through in a unit of time. Throughput efficiency is a ratio of maximum throughput to the network bit rate.

Perhaps, the easiest way to calculate the efficiency is to find the overhead components associated with encapsulation and scheduling.

ENCAPSULATION OVERHEAD

The Ethernet encapsulation (framing) overhead is a result of adding 8-byte frame preamble, 14-byte Ethernet header, and 4-byte FCS field to MAC Service data units (m_sdu) comprised of user's data. Additionally, at least 12-byte minimum inter-frame gap (IFG) should be left between two adjacent frames[1]. Thus, the absolute overhead per one frame is constant and equal to 38 bytes[2] (see Figure 1). **This encapsulation overhead is not specific to EPON, but a property common to all Ethernet networks.**

[1] IFG is specified as 96-ns time interval, which is equal to 12 byte-transmission times in 1Gbps (1000BASE-X) Ethernet.

[2] Short payloads are padded to a minimum length of 46-bytes. This also contributes to the Ethernet encapsulation overhead and is counted in our calculations.

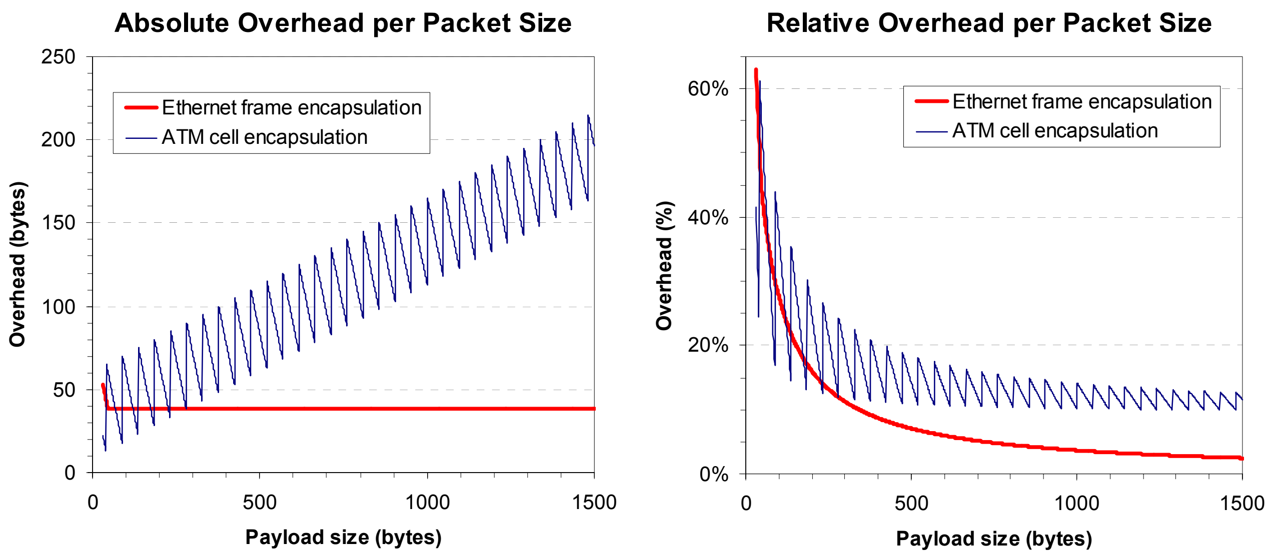

Figure 1: Comparison of Ethernet framing overhead and ATM cell tax.

In ATM networks, the user's data units, such as IP datagrams, should be broken in multiple cells. The ATM encapsulation overhead (also known as *cell tax*) comprised of multiple cell headers, 8-byte ATM Adaptation Layer 5 (AAL5) trailer, and variable-size padding. The AAL5 trailer is needed for proper IP-datagram reassembly, and the padding is used to fill any remaining portion of the last cell. As is seen in Figure 1, the ATM encapsulation overhead depends on the payload size and is considerably higher than the Ethernet overhead.

The average value of the encapsulation overhead depends on the distribution of packet (m_sdu) sizes. The distributions of packet sizes were reported in the literature. These distributions have a tri-modal shape and are similar for backbone networks [1] and access networks [2].

The value of average overhead can be obtained using the following formula:

$$\text{avg\_overhead} = \frac{\sum_s E(s) \times f(s) - \sum_s s \times f(s)}{\sum_s E(s) \times f(s)} \qquad (1)$$

where s = size of the payload (m_sdu), $f(\cdot)$ = probability distribution function, and E(s) is the size of an encapsulated payload s. For the Ethernet encapsulation, the E(s) function is

$$\begin{aligned} E_{\text{Ethernet}}(s) &= \max\{s, \text{min\_payload}\} + \text{header} + \text{FCS} + \text{preamble} + \text{IFG} \\ &= \max\{s, 46\} + 38 \end{aligned}$$

Using IP packet-size distribution obtained in a head end of a cable network [2], we get the Ethernet encapsulation overhead to be

$$\text{avg\_overhead} = 1 - \frac{\sum_s s \times f(s)}{\sum_s (\max\{s, 46\} + 38) \times f(s)} = 1 - \frac{509.02}{549.79} = 7.42\%$$

For the ATM cell encapsulation, the function E(s) looks like the following:

$$E_{ATM}(s) = \left\lceil \frac{s + AAL5}{\text{cell\_payload}} \right\rceil \times \text{cell\_size} = \left\lceil \frac{s+8}{48} \right\rceil \times 53$$

Using the same distribution of packet sizes [2] makes the average ATM encapsulation overhead equal to

$$\text{avg\_overhead} = 1 - \frac{\sum_s s \times f(s)}{\sum_s \left\lceil \frac{s+8}{48} \right\rceil \times 53 \times f(s)} = 1 - \frac{509.02}{586.57} = 13.22\%$$

The above calculations show the advantages of using variable-sized Ethernet frames to carry variable-sized IP packets. **The Ethernet frame encapsulation overhead of 7.42% is significantly lower than the ATM cell encapsulation overhead of 13.22%.**

SCHEDULING OVERHEAD

The scheduling overhead in EPON consists of control message overhead, guard band overhead, discovery overhead, and frame delineation overhead. Some of the parameters affecting the overhead, such as cycle time or frequency of the discovery attempts are outside the scope of IEEE802.3ah. Therefore, in some cases we will present multiple values of overhead for the different choices of configuration parameters.

Control message overhead represents bandwidth lost due to use of in-band control messages such as GATEs and REPORTs. The amount of overhead depends on the number of ONUs and cycle time, i.e., an interval of time in which each ONU should receive a GATE message and send a REPORT message. We make an assumption here that the scheduling algorithm requires only one GATE message and one REPORT message to be exchanged between each ONU and the OLT in one cycle time. The ITU-T recommendation G.114 "One way transmission time" specifies the delay for voice traffic in access network at 1.5 ms. To achieve the average delay of 1.5 ms for frames carrying voice data, the cycle time should be about 1 ms. If the maximum delay is to not exceed the 1.5 ms limit, the cycle time should be fixed at

750 μs. We, therefore, present the control message overhead values for both of these cycle times.

The control message overhead is calculated as:

$$\text{control\_overhead} = \frac{\text{message\_size} \times N_{ONU}}{\text{cycle\_time} \times \text{EPON\_rate}}$$

where message_size is the size of GATE or REPORT message (including preamble and IFG), NONU = number of ONUs (i.e., number of messages sent in one cycle time), and EPON_rate = 1Gbps.

Table 1 presents the results of control message overhead for different number of ONUs and various cycle times. This overhead is present in both upstream and downstream directions.

	Cycle = 1ms	Cycle = 750 μs
16 ONUs	1.08%	1.43%
32 ONUs	2.15%	2.87%

Table 1: Control message overhead.

Guard band overhead depends on PMD and PMA parameters such as Laser ON/OFF times, Automatic Gain Control (AGC) and Clock-and-Data Recovery (CDR) times. The draft IEEE 802.3ah D1.414 specifies four possible values (classes) for the AGC and CDR parameters: 96 ns, 192 ns, 288 ns, and 400 ns. The laser ON/OFF times are fixed at 512 ns. In addition, guard bands should include a 128-ns dead zone to allow for timing variability of the Multi-Point Control protocol[3]. As is shown in Figure 2, the Laser OFF time may partially overlap the laser ON time of the next ONU.

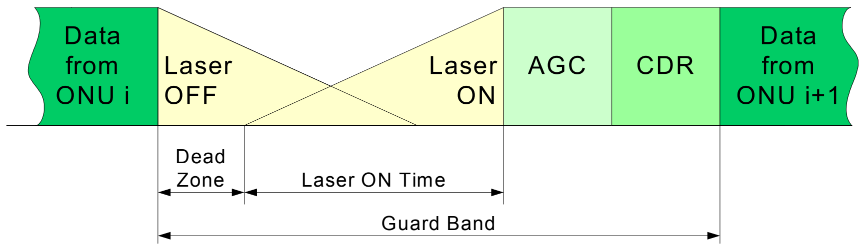

Figure 2: Structure of the guard band.

[3] The presented timing values are based on IEEE802.3ah draft D1.414.

The guard band overhead is calculated using the following formula:

$$\text{guard\_overhead} = \frac{(\text{laser\_ON} + \text{dead\_zone} + \text{AGC} + \text{CDR}) \times N_{ONU}}{\text{cycle\_time}}$$

Table 2 shows the guard band overhead for different numbers of ONUs and different cycle times. We present the overhead values for the maximum and minimum values of AGC and CDR lock times. This guard band overhead is present only in the upstream direction.

	Cycle = 1ms	**Cycle = 750 μs**
16 ONUs **AGC = 96 ns** **CDR = 96 ns**	**1.33%**	**1.77%**
16 ONUs **AGC = 400 ns** **CDR = 400 ns**	**2.30%**	**3.07%**
32 ONUs **AGC = 96 ns** **CDR = 96 ns**	**2.66%**	**3.55%**
32 ONUs **AGC = 400 ns** **CDR = 400 ns**	**4.61%**	**6.14%**

Table 2: Guard band overhead.

Discovery overhead represents the bandwidth lost due to allocation of a discovery window. The discovery window should be larger than the maximum round-trip time of 200 μs. In our calculations we assume the discovery window of size 300 μs. Frequency of the discovery attempts is not specified in the IEEE 802.3ah draft. Intelligent algorithms may detect a situation when all ONUs are operational and cease all discovery attempts. We, however, will assume a simpler algorithm that performs periodic discovery regardless of the number of registered ONUs. The discovery period can be very large, for example 1 second or more.

With a 1-second discovery period, the discovery overhead is equal 300 μs / 1 second = 0.03%.

Frame delineation overhead is associated with the fact that variable-sized frames may not be able to completely occupy the fixed-sized cycle. Grants to ONUs are based on their reported queue lengths. However, multiple grants with their associated guard bands may not fill the fixed cycle time exactly. The expected size of the unused remainder can be calculated using the following formula:

$$avg\_remainder = \frac{\sum_{r=1}^{S^{MAX}-1} r \times \left[1 - F_{Ethernet}(r)\right]}{\sum_{s=S^{MIN}}^{S^{MAX}} s \times f_{Ethernet}(s)} \qquad (2)$$

where S^{MIN} and S^{MAX} are the minimum and the maximum Ethernet frame sizes, $f_{Ethernet}(\cdot)$ = probability distribution function for Ethernet frames, $F_{Ethernet}(\cdot)$ = cumulative distribution function for Ethernet frames.

Using the packet size distribution from [2], we get the average remainder approximately equal to 595 bytes. That means than we should expect, in average, 595 bytes wasted due to variable-sized frames not packing the fixed cycle completely. Table 3 shows the average frame delineation overhead for different cycle times.

Cycle = 1ms	Cycle = 750 μs
0.48%	0.63%

Table 3: Frame delineation overhead.

Table 4 summarizes the values of various upstream overhead components and calculates the combined upstream overhead.

	Min. overhead	Max. overhead
Control message overhead	1.08%	2.87%
Guard band overhead	1.33%	6.14%
Discovery overhead	0.03%	0.03%
Frame delineation overhead	0.48%	0.63%
Total upstream scheduling overhead	2.92%	9.67%

Table 4: Upstream overhead summary.

In the downstream direction, only the control message overhead is present. Table 5 shows the total scheduling overhead in the downstream direction.

	Min. overhead	Max. overhead
Control message overhead	**1.08%**	**2.87%**
Total downstream scheduling overhead	**1.08%**	**2.87%**

Table 5: Downstream overhead summary.

Thus, in the upstream direction the total scheduling overhead can be anywhere between 2.92% and 9.67%. In other words, EPON efficiency is 90.33% to 97.08% compared to 1GbE point-to-point link. In the downstream direction, EPON efficiency reaches 97.13% to 98.92% of the efficiency of a point-to-point 1GbE link.

It is possible that a particular scheduling algorithm or implementation will have lower efficiency, however, that would only be a result of particular design decisions and not an intrinsic overhead of EPON architecture.

ABSOLUTE EPON EFFICIENCY

Combined EPON efficiency is just a product of encapsulation efficiency and scheduling efficiencies. The combined efficiency multiplied by the EPON rate of 1Gbps gives us the *net* EPON throughput, i.e., the application-level throughput. Table 6 summarizes the efficiency and net throughput for upstream and downstream directions.

	Downstream	Upstream
Minimum absolute efficiency	**89.92%**	**83.63%**
Maximum absolute efficiency	**91.58%**	**89.88%**
Minimum efficiency (relative to 1GbE PtP link)	**97.13%**	**90.33%**
Maximum efficiency (relative to 1GbE PtP link)	**98.92%**	**97.08%**
Minimum net throughput	**899.2 Mbps**	**836.3 Mbps**
Maximum net throughput	**915.8 Mbps**	**898.8 Mbps**

Table 4: EPON efficiency and througput.

The maximum values represent an EPON system with 16 ONUs, 96-ns AGC and CDR times, and the cycle time of 1 ms. The minimum values are calculated for an EPON with 32 ONUs, 400-ns AGC and CDR times, and the cycle time equal to 750 μs.

In all our calculations we considered the overhead and efficiency at the GMII interface. Some may argue that Ethernet 8B/10B line coding contributes additional 20% of overhead if each transition on the line is considered a bit. Of course, this would result in a lower percental value of efficiency, however, coupled with the 1.25Gbps line rate would result in the same net throughput as shown in Table 6.

REFERENCES

[1] K. Claffy, G. Miller, and K. Thompson, *"The nature of the beast: Recent traffic measurements from an internet backbone,"* in Proceedings INET '98, (Geneva, Switzerland), July 1998.

[2] D. Sala and A. Gummalla, "*PON Functional Requirements: Services and Performance,*" presented at IEEE 802.3ah meeting in Portland, OR, July 2001. Available at http://grouper.ieee.org/groups/802/3/efm/public/jul01/presentations/sala_1_0701.pdf.

[3] ITU-T Recommendation G.114, *One-Way Transmission Time*, in Series G: Transmission Systems and Media, Digital Systems and Networks, Telecommunication Standartization Sector of ITU, May 2000.