

# Performance Evaluation of Ethernet Frame Burst Mode in EPON Downstream Link

---

Wen-Kang Jia and Yaw-Chung Chen

**We apply IEEE 802.3 frame burst mode (FBM) to the Ethernet passive optical network (EPON) downstream link and compare its performance with non-frame burst mode for various traffic patterns. Although in light traffic loads ( $\rho < 0.5$ ) the efficiency of the FBM mechanism is not significant, it does feature high throughput, small jitter, low queue occupancy, and short queuing delay in optical line terminals under various traffic loads with various numbers of optical network units (ONUs). The FBM performance always approaches that of full-duplex mode, especially under heavy traffic loads ( $\rho > 0.5$ ). Moreover, an increase in number of ONUs will decrease the burst performance. Our work shows that FBM scheme is very useful for EPON transmission and has low design complexity.**

**Keywords:** EPON, FBM, burst mode, full-duplex mode, traffic aggregate.

## I. Introduction

Metro Ethernet Network (MEN) is an access technology based on the Ethernet standard. It will be used as a metropolitan access network to connect subscribers to a public domain such as the Internet, or private sites such as corporate Intranet systems. Large businesses can also use MEN to connect branch offices to their Intranets. MEN is also called a fiber-to-the-home (FTTH) solution. It provides point-to-point (P2P) and point-to-multipoint (P2MP) broadband access to triple-play services for the home or business [1]-[4].

To fulfill the bandwidth demand of immense traffic volume on the subscriber side, the IEEE Standards Board approved IEEE 802.3ah, or Ethernet in the first mile (EFM), which is also known as Ethernet passive optical network (EPON) (alternatively GE-PON). It is a collection of protocols specified in IEEE 802.3, defining the Ethernet as the access networks, namely, first/last mile [4]-[6].

Most studies on the traffic engineering (TE) of EPON focus on the upstream link because downstream links are insufficient and scarce. Many proposed solutions, such as MPCP, DBA, IPACT, and DGP attempt to optimize the bandwidth utilization of EPON upstream links. However, the traffic behavior in the access network is asymmetric. Most users need more downstream bandwidth than upstream bandwidth [7], [8]; therefore, we have ADSL technology. For these reasons, we focus on 802.3 frame burst mode (FBM) of EPON downstream links in this paper. Nevertheless, our discussion might be more valuable for facilitating network assessment and for precisely planning the administration cost of optical lines, users, bandwidth, equipment, and so on.

---

Manuscript received Aug. 02, 2007; revised Feb. 20, 2008

This work was partially supported by the National Science Council of Rep. of China under Grant No. 96-2752-E-009-006-PAE.

Wen-Kang Jia (phone: + 886 920 227 405, email: wkchia@cs.nctu.edu.tw) and Yaw-Chung Chen (email: ycchen@cs.nctu.edu.tw) are with Department of Computer Science, National Chiao Tung University, Hsinchu, Taiwan, Rep. of China.

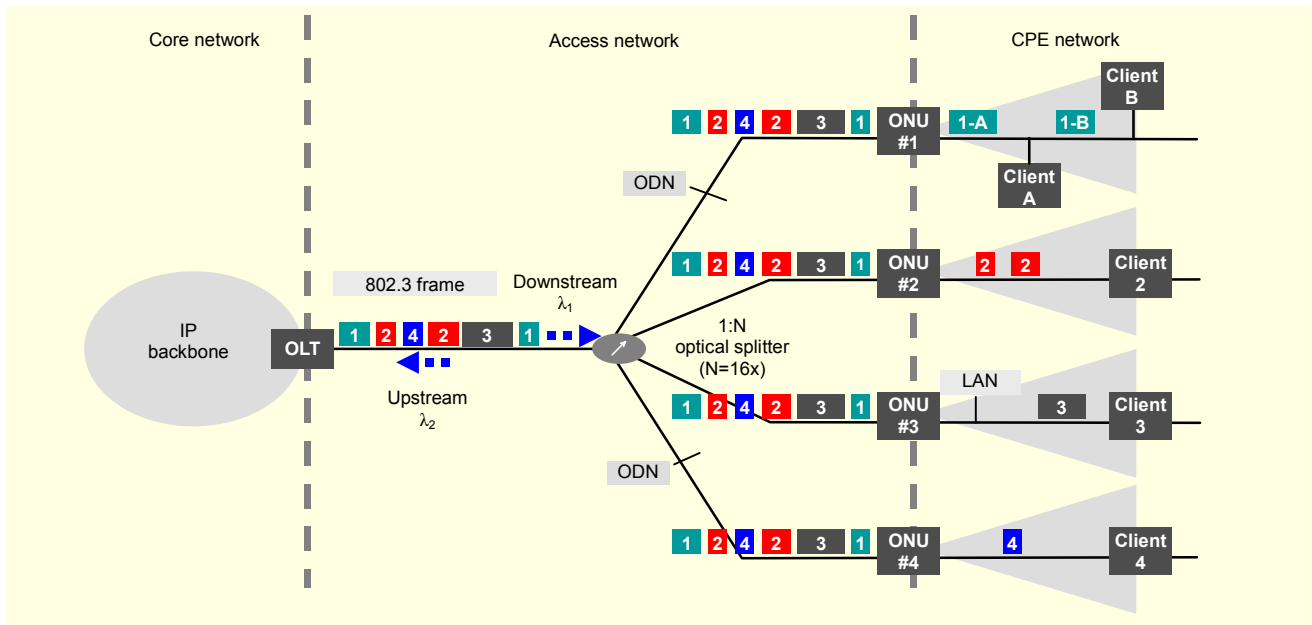


Fig. 1. Network architecture of downstream links in EPON with one OLT and 4+ ONUs.

## II. Ethernet Passive Optical Network

### 1. EPON Architecture

The Medium Access Control (MAC) protocol of EPON uses IEEE 802.3 framing and line coding, and packets transmitted from the Ethernet are broadcast by the optical line terminal (OLT) and selectively extracted by their corresponding destination optical network unit (ONU) through the optical distributed network (ODN). Figure 1 shows an EPON architecture with one OLT and 4+ ONUs. A frame broadcast from the OLT is sent to subscribers through the trunk fiber and 1:N passive optical splitter, which delivers  $N$  copies of the data frame over  $N$  ONUs. At the OLT, the LLID tag is added to the preamble of each frame and extracted at the ONU in its reconciliation sub-layer. Thus, each ONU receives all frames transmitted from the OLT and extracts only its own frames with matching LLID. Frame extraction is based only on the LLID because the ONU's MAC is in promiscuous mode and accepts all frames. Because of the downstream broadcasting nature of EPON, an encryption mechanism is often considered a requirement [1]-[3], [6].

While moving in the upstream direction, packets from an ONU reach only the OLT but not other ONUs, as in a point-to-point transmission. However, since multiple ONUs try to access one OLT simultaneously, there is unavoidable collision if no arbitration mechanism is used; hence, all ONUs are synchronized to a common clock source, and each ONU is allocated a timeslot. It is necessary for an ONU to buffer frames received from a subscriber until its timeslot is available, as in a time-division multiplexing (TDM) scheme. In short,

there is no difference between EPON and a shared-medium Ethernet LAN in an upstream link, but we do not address this here [1]-[4], [9].

The physical layer of EPON is different from that of traditional Ethernet. All ONUs share one fiber in both upstream and downstream links, and all clients share one downstream wavelength ( $\lambda_1$ ) and one upstream wavelength ( $\lambda_2$ ) [1]-[3], [6].

### 2. Frame Burst Mode

A burst mode (BM) is usually implemented by allowing a network node to seize control of the network medium and not permitting other nodes to interrupt. A node transmits data continuously without waiting for another node to terminate. In Ethernet, FBM is implemented by automatically fetching the next queued packet with the same destination prior to its turn. This is essentially the same technique used by various types of I/O devices including random access memory (RAM), hard disk drive (HDD) interfaces; accelerated graphics port (AGP) processors, and so on. One characteristic that all burst modes have in common is that they are temporary and unsustainable. They allow faster data transfer rates than normal, but only for a limited period of time and only under certain special conditions [10], [11].

As defined in IEEE 802.3 serial standard, there are two basic operation modes on the Ethernet, one is carrier-sense multiple-access/collision detection (CSMA/CD), also known as half-duplex (HDX) mode and the other is full-duplex (FDX) mode with CSMA/CD disabled. The most obvious difference between these two modes is that HDX uses a shared medium

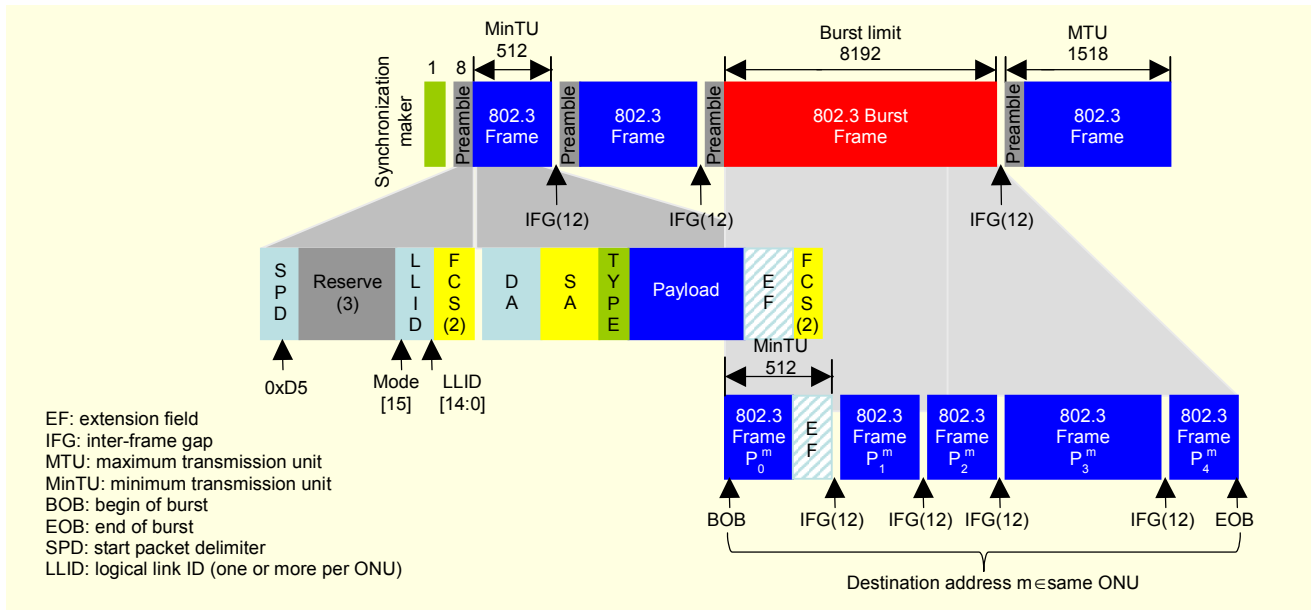


Fig. 2. EPON packet encapsulation.

and FDX uses point-to-point links attributed to the usage of switches. Ethernet MAC can operate in either of these two modes [10], [12]. While EPON is operating in HDX mode, a carrier extension (CE) mechanism is performed, and an extension field (EF) may be appended to the end of the Gigabit Ethernet frame if the original frame length is less than 512 bytes. This is to ensure that the frame is long enough for collision detection. In Gigabit Ethernet, although the meaningful payload may be only 46 bytes, the minimum length of the transmitted frame must be 512 bytes [6], [10]-[14]. An EF field is appended as needed so that the carrier is extended to the required minimum slot time [10], [12]. The EF field is not used in FDX mode because the CSMA/CD [5] protocol is not used in this mode. It is for point-to-point transmission [10]-[12].

Carrier extension is a simple solution, but it may waste bandwidth when the data packet is very small. Up to 448 padding bytes may be sent and result in very low throughput. In fact, for a large number of small packets, the throughput of Gigabit Ethernet is only marginally better than Fast Ethernet. In order to solve the problem of such inefficiency, an optional FBM that allows a node to transmit consecutive frames without relinquishing control of the medium has been specified. FBM is only available under HDX operation, which is an extension of CE. Figure 2 illustrates an EPON packet/frame encapsulation structure with an EF appended. After successful transmission of a frame, a sending node may use burst mode to continuously transmit additional frames until it reaches a burst limit of 8,192 bytes. An inter-frame gap (IFG) period is then inserted between each frame in the burst.

The first frame of a burst is transmitted as normal, and it

includes an EF as required. Subsequent frames in the burst do not require an EF, and they are transmitted back-to-back with the minimum IFG until a burst timer expires. If a collision occurs, only the first frame in the burst will be affected and retransmitted [5], [10]-[12].

EPON employs FBM for data packets in both downstream and upstream instead of using continuous data packets as in frame continuous mode (FCM). FBM improves EPON's throughput, and saves ONU power by allowing an ONU to power off the laser driver during idle periods [5]-[6], [13]. In the non-frame burst mode (NBM) downstream link, an OLT employs the original FIFO discipline. An Ethernet frame is encapsulated in an IEEE 802.3 frame immediately after its arrival and is put into a global FIFO queue waiting for transmission.

In the FBM downstream link, packets coming from core networks are aggregated in an OLT and are dispatched to ONUs. If an OLT supports FBM, it reschedules the packets in a queue, puts together those packets with the same LLID, and transmits them continuously. Packets with different LLIDs will be postponed until their leading burst phase is terminated. Sometimes, many destination addresses (DAs) may be under a single LLID, which means that multiple clients share one ONU. This still allows the OLT to aggregate all DA traffic to one burst phase [5], [6], [14].

As a result of FBM, each ONU receives a beginning of burst (BOB) packet transmitted from the OLT and extracts only its own frames with matching LLID in the preamble. Thus, it is unnecessary to add the preamble for each frame again before the end of burst (EOB), as shown in Fig. 2. Here, the preamble

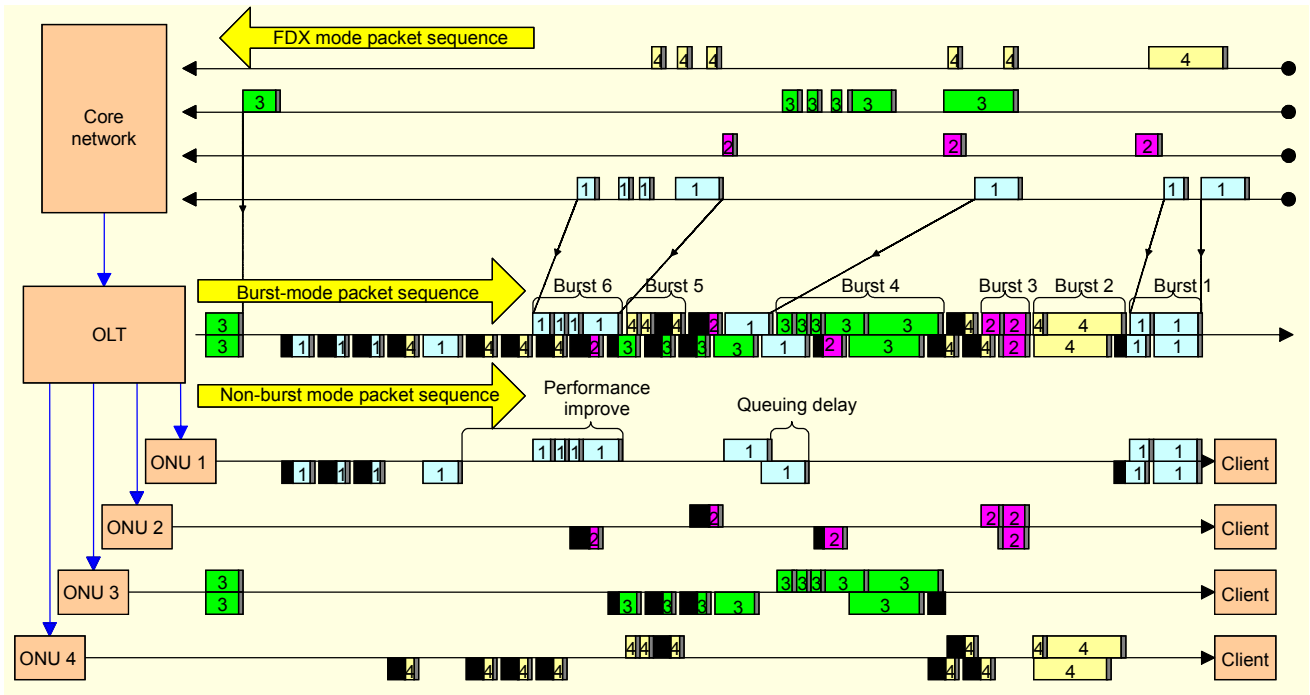


Fig. 3. ON-OFF model example of FBM in EPON downstream link.

occupies 64 bit slots, which are 64 nanoseconds in 1 Gbps EPON [1], [2], [6].

Figure 3 illustrates the ON-OFF model [15] of the burst mode performed in EPON downstream link. Frame bursting substantially reduces the bandwidth usage and increases the throughput by eliminating the EF and preamble. Nevertheless, the improvement may lead to a tradeoff between packet transmission latency and jitter. Also, more buffer space is needed in an OLT. Since short packets frequently appear in layer 3 protocols, burst mode could be used to improve the performance of EPON quite substantially [5], [6], [11], [14].

### III. Modeling of FBM in EPON Downstream Link

Our FBM traffic model is based on the following assumptions:

- Outbound port is an HDX EPON optical line
- Neither collision nor bit-error will occur in outbound port
- Inbound port is an FDX Gigabit Ethernet link
- Infinite queue size
- Single input multiple outputs (SIMO)
- Poisson distribution in frame inter-arrival time
- Uniform distribution in frame length
- FIFO queuing discipline
- M/M/1 queuing model
- Switching mechanism of OLT is store and forward

For an incoming frame sequence  $F_n, n=0,1,\dots,k$ , the service

time of a frame is assumed to be identical to its length divided by the transmission rate. Since the minimum transfer unit ( $MinTU$ ) is 4,096 bit times (512 bytes/slot) in EPON, a packet whose length is smaller than  $MinTU$  must be appended an extension to make it up to  $MinTU$ . We set the  $L(F_n)$  as the layer 2 frame length of  $F_n$ , and  $L(P_n)$  as the layer 3 packet length of  $P_n$ . When a layer 3 packet  $P_n$  (46 to 1,500 bytes) is encapsulated by a layer 2 frame (64 to 1,518 bytes), we have the service time of the layer 2 frame  $F_n$  as

$$T(F_n) = \max(T(P_n) + MACHeader, MinTU), \quad (1)$$

where  $MACHeader$  is equal to 144 bit times (18 bytes). It is necessary to keep a minimum idle period  $T_{IFG}$  between two consecutive frame transmissions. This so-called inter-frame gap (IFG) provides a short recovery time between frames to allow interface devices to prepare for reception of the next frame. The minimum IFG is 96 bit times, which equals 96 ns for 1 Gbps EPON [10]-[12]. In the EPON devices, the OLT/ONU also needs processing time  $T_{FD}$  to forward a frame to its corresponding output port. Assume the transmission time of  $F_n$  is  $T(F_n)$  when consecutive  $k$  frames form a burst arrival. We can calculate the total service time  $X$  in the output queue as

$$X = \begin{cases} \sum_{n=0}^k (T(F_n) + T_{PRE} + T_{IFG}) + T_{FD}, & \text{if } T_{IFG} > T_{FD}, \\ \sum_{n=0}^k (T(F_n) + T_{PRE} + T_{FD}), & \text{otherwise.} \end{cases} \quad (2)$$

Here,  $T_{PRE}$  is the Ethernet preamble time. In general,  $T_{FD}$  is smaller than  $T_{IFG}$  in FE networks (960 ns). More specific studies indicate that the IPv4 routing delay lasts approximately 250 ns to 350 ns; and the average IPv6 routing delay lasts from around 460 ns to 960 ns [16], in which case,  $T_{FD}$  could be ignored. Since the line speed is upgraded nowadays (96 ns in GbE and 9.6 ns in 10 GbE), we are not sure whether the  $T_{FD}$  is still bigger than  $T_{IFG}$  in existing layer 3 network devices, which may have poor forwarding performance [15], [16].

Where the parameter packet interval  $V(P)$  is regarding destinations  $m$  and throughput  $\rho$ , suppose that all destinations have equal loads. We could obtain the relationship between the ON-period  $T(P)$  and OFF-period  $V(P)$  in a long enough sequence of packets 0 to  $k$  as

$$\sum_{n=0, \forall m}^k V(P_n^m) = \left( m + \left( \frac{m - m\rho}{\rho} \right) - 1 \right) \times \sum_{n=0, \forall m}^k T(P_n^m). \quad (3)$$

As an example, for a single destination ( $m=1$ ) and a full offered load ( $\rho=1$ ), we obtain  $V(P)$  as 0 ( $m + ((m - m\rho)/\rho) - 1 = 0$ ) based on (3). However, the necessary condition for a frame to burst in sequential packets when the leading frame is being sent out is that the subsequent frame should already be in the queue. If this process continues, the totally transmitted frame length reaches the *BurstLimitation* (65,536 bit times = 8,192 bytes). To ensure that the carrier sense mechanism will work normally, the leading packet  $P_0$  must be no smaller than *MinTU*.

In a real system we should consider the time interval between two consecutive frame arrivals. This interval may affect the probability of a burst, so we can obtain the following:

$$\begin{aligned} \forall P_k^m, \exists m : & \left[ V(P_{k-1}^m) + T(P_{k-1}^m) + T_{PRE} + T_{IFG} + T(F_{k-1}^m) \right] \\ \geq & \left[ V(P_{k-1}^m) + T(P_{k-1}^m) + V(P_k^m) + T(P_k^m) + (2 \times T_{PRE}) + T_{IFG} \right] (4) \\ \Rightarrow & T(F_{k-1}^m) \geq \left[ V(P_k^m) + T(P_k^m) + T_{PRE} \right], \end{aligned}$$

where  $P_k^m$  means the  $k$ -th packet in a sequence of packets with the  $m$ -th ONU as their destination, and the frame  $F_{k-1}$  is the BOB frame  $F_{BOB}$ . Now, we reset  $F_{k-1}$  to  $F_0$  so that the burst is continued if the subsequent frames fulfill the same condition, and the sending time stamp of frame  $T_{sendout}(F_k)$  is

$$T_{sendout}(F_k) = \begin{cases} T(F_{k-1}^m) + T(P_k^m), & \text{if (4) is true,} \\ V(P_k^m) + T(P_k^m) + T_{PRE} + T(F_k^m), & \text{otherwise.} \end{cases} \quad (5)$$

Hence, the leading frame  $F_{BOB}$  of a burst incorporates a frame bursting sequence until those frames in the queue up to an aggregated length of 8,192 bytes that have the same destination are sent out. If (4) is no longer held, it means that either the queue is empty or there is no available packet with the same destination in the queue. In that case, the burst is terminated

and packet  $P_{k-1}$  is the EOB packet  $P_{EOB}$ . Packet  $P_k$  is then appended to an EF if its packet length is smaller than *MinTU*, and it may become a new  $P_{BOB}$ :

$$\begin{aligned} \forall P_k^m, \exists m : & \sum_{n=0}^{k-1} \left( V(P_n^m) + T(P_n^m) + (2 \times T_{PRE}) + T_{sendout}(F_{n-1}) + T_{IFG} \right) \\ & + \left( T(F_0^m) - T(P_0^m) \right) - T_{IFG} \\ \geq & \sum_{n=0}^k \left( V(P_n^m) + T(P_n^m) + T_{PRE} + T_{IFG} \right) - T_{IFG}. \end{aligned} \quad (6)$$

Sometimes, the burst phase is stopped when it reaches the burst limitation:

$$\sum_{n=0}^k \left( T(P_n^m) + T_{IFG} \right) + \left( T(F_0^m) - T(P_0^m) \right) \leq \text{BurstLimitation}. \quad (7)$$

We assume that packet length is uniformly distributed. Packets are randomly sent to  $m$  destinations with equal probability. We can obtain the probability of BOB as

$$\begin{aligned} P(b) = & \left( \sum_{i=MinTU+1}^{MTU} \frac{i}{(MTU - MinFrame)^2} \right. \\ & \left. + \sum_{MinFrame}^{MinTU} \frac{MinTU}{(MTU - MinFrame)^2} \right) \times \frac{1}{m^2}. \end{aligned} \quad (8)$$

Theoretically, the burst size (number of frames per burst) may reach a range  $\theta$  as

$$2 \leq \theta \leq INT \left( \frac{\text{BurstLimitation} - MinTU - T_{IFG}}{MinFrame + T_{IFG}} \right) + 2, \quad (9)$$

where we may obtain  $\theta$  as large as 102 frames per burst (FPB), which means that 102 frames are aggregated in one burst process under the extreme condition. Also, we can obtain the largest possible *BurstLength* of 9,735 by (9) when  $F_k = MTU$  (1,518 bytes), where  $F_k$  is the last frame in the burst. Based on IEEE 802.3 rules, the last frame that encounters the *BurstLimitation* is allowed to transmit over the *BurstLimitation* until the frame runs out. This means that *BurstLength* permission is bigger than the *BurstLimitation*. In some extreme situations, the *BurstLength* is just equal to the *BurstLimitation*. We obtain the *BurstLength* as follows, where the packet  $P_{k+1}$  is the last frame ( $P_{EOB}$ ) in this burst:

$$\text{BurstLength} = \sum_{n=0}^{k+1} \left( T(P_n^m) + T_{IFG} \right) + \left( T(F_0^m) - T(P_0^m) \right). \quad (10)$$

Hence, the burst is stopped if all frames with the same destination and  $F_{BOB}$  in the queue are sent out, the queue

becomes empty, or the burst length reaches the *BurstLimitation* [10], [12].

During a frame bursting period of  $P_{0-k}$ , by reducing the EF and preamble from whole slot times, we can calculate the percentage of utilization reduction in an FBM and NBM as

$$\frac{\sum_{n=1}^k (T(P_n^m) + T_{IFG}) + T(F_0^m) + T_{PRE}}{\sum_{n=0}^k (T(F_n^m) + T_{IFG} + T_{PRE}) - T_{IFG}}. \quad (11)$$

Assume that the OLT is in store-and-forward mode after an idle period to ensure that the buffer is empty, for any packet  $F_0^m$  whose subsequent packet fulfills the following condition:

$$\forall F_0^m, \exists m : (2 \times T(F_0^m) + T_{PRE} + T_{IFG}) \geq \sum_{n=0, \forall m}^{k=MaxQsize} (V(P_n^m) + T(P_n^m) + T_{PRE} + T_{IFG}) - T_{IFG}. \quad (12)$$

We can obtain  $k$ , and before  $F_n^m$  is sent out, the concurrent queue size (bytes) will be

$$L(F_n) = L_q(F_n) + L_s(F_n) = \sum_{n=0, \forall m}^k L(P_n^m) + L(F_n^m). \quad (13)$$

After an idle period, ensure that the buffer is empty, from the first packet  $F_0$  to any current packet  $F_k$  that fulfills the following condition:

$$\forall F_k^m, \exists m : \sum_{n=0, \forall m}^{k-1} (V(P_n^m) + T(P_n^m) + (2 \times T_{PRE}) + T_{sendout}(F_{n-1}) + T_{IFG}) + (T(F_0^m) - T(P_0^m)) - T_{IFG} < \sum_{n=0, \forall m}^k (V(P_n^m) + T(P_n^m) + T_{PRE} + T_{IFG}) - T_{IFG}. \quad (14)$$

We can obtain the current queuing delay time  $W(P_k)$  in nanoseconds as

$$W(F_k) = \left[ \sum_{n=0, \forall m}^{k-1} (V(P_n^m) + T(P_n^m) + (2 \times T_{PRE}) + T_{sendout}(F_{n-1}) + T_{IFG}) + (T(F_0^m) - T(P_0^m)) - T_{IFG} \right] - \left[ \sum_{n=0, \forall m}^k (V(P_n^m) + T(P_n^m) + T_{PRE} + T_{IFG}) - T_{IFG} \right]. \quad (15)$$

The delay time  $W(P_k)$  is 0 (no queuing delay) when (14) no longer holds.

## IV. Simulations

In this section, we present a scheduling algorithm which is designed to improve the performance over the original FIFO discipline in terms of overall throughput, queue occupancy, queuing delay, and delay variation. We evaluate the performance of the FBM in EPON through simulation. We assume that there are four factors which influence the results.

**Input rate ( $\lambda$ ):** This is also called the traffic load or traffic intensity. In general, a high arrival rate causes a short inter-frame interval (excluding IFG) or even no interval, which means that packets have a greater opportunity to form a burst. On the other hand, extremely long inter-frame intervals may lead to an emptied queue in which frame bursting is unlikely to

**Algorithm 1.** C code for the burst-mode-scheduling algorithm.

```

Set_Interrupt(PKT_SEND_COMPLETE());
Set_Interrupt(PKT_RECV_COMPLETE());
while(!EXIT)
{
if (Check_Input_Queue()) /*return concurrent queue size*/
{Pkt=Pop_Input_Queue(); /*Read top one*/
If (Pkt.Length<512) {Pkt.Length=Add_CE(Pkt);}
Pkt_Send(Pkt.OutputPort,Pkt,Pkt.Length);
Burst_Length=Pkt.Length; Burst_size=1;
Burst_Destination=Pkt.OutputPort; }
}
Interrupt PKT_SEND_COMPLETE()
{
while (queue_size=Check_Input_Queue())
{
for (i=0;i++;i<queue_size)
{
Pkt=Pop_Input_Queue();
If(Pkt.OutputPort==Burst_Destination)
{ /*Same Destination Gen. burst*/
if((Pkt.Length+ Burst_Length)<=8192)
{ /* Init Burst */
Pkt_Send(Pkt.OutputPort,Pkt,Pkt.Length);
Burst_Length= Pkt.Length+Burst_Length;
Burst_size++; }
else /*Burst finish by Reach Burst Limitation*/
{ Record_Burst_Size&Length();
Burst_size=0; }
}
else /*Store temporary*/
{Pop_Buffer(); temp_buffer++;}
}
} /*Burst finish by Queue Empty*/
while(tempbuffer--)
{Pkt=Push_Buffer(); Push_Input_Queue(Pkt);}
Record_Burst_Size&Length();Burst_size=0
/*Clear Burst count*/
return;
}
Interrupt PKT_RECV_COMPLETE()
{ /*Add Packet to Queue Tailer*/
Add_Input_Queue(PKT_RECV());
}
}

```

happen [8], [17].

**Packet length distribution:** Theoretically, if the frame length distribution is centered at a mean value as in standard normal distribution, it is unlikely to lead to frame bursting. Similarly, if packet length distribution throughout the traffic is dispersed, the long frames and the short frames are more likely to be aggregated together for frame bursting. In other words, one long frame will lead many short frames to form a frame burst [14], [15], [17].

**Number of inputs:** If an OLT has more than one uplink interface, traffic from different inputs may arrive simultaneously. This aggregated traffic fills up the queue temporarily. When a queue has many packets waiting, the probability of frame bursting increases.

**Number of outputs (ONUs):** Traffic aggregated in the queue is sent out to destination ONUs, and FBM performance may be impacted because many active ONUs may cause packets sent to the same destination to be interleaved with other packets that have different destinations.

We set up a simple environment as shown in Fig. 1 to evaluate the single-input multiple-output (SIMO) scenario. Algorithm 1 could be used in the example shown in Fig. 1.

### 1. Performance of FBM in Multiple ONUs

For FBM, we are most interested in the relationship between the traffic load and performance improvement. In a light load, the packets may be dispersed by long inter-arrivals, and this makes frame bursting unlikely to happen.

Here we setup a simulation based on a uniformly distributed traffic pattern, using a single input with 1 Gbps FDX line speed. The input rate ranges from 5% to 100% of line capacity, and output is set with  $m=1, 2, 4, 16, 32,$  and  $64$  ONUs; thus, we can observe the relationship between the traffic load and the probability of frame bursting.

Figure 4 shows the burst mode performance improvement on multiple destinations with  $m=1, 2, 4, 16, 32,$  and  $64$  ONUs. The dotted line shows the NBM performance for reference purpose. The other lines represent the results of different numbers of ONUs. Comparing FBM with the curve of NBM, we find that as the input load ( $\rho$ ) increases, burst performance also increases until it gradually approaches FDX efficiency with  $\rho=100\%$ . With a full input load ( $\rho=1$ ), the output load increases from  $\rho=0.92947$  ( $m=1$ ) to  $\rho=0.93464$  ( $m=64$ ), which means that FBM bandwidth usage is reduced by 7.053% to 6.536% under burst mode. In other words, since NBM experiences 108.425% bandwidth usage, FBM offers 100.7778% to 101.3389% output bandwidth occupancy for FDX mode. When the input rate is reduced to  $\rho=0.15$ , the output utilization after a burst is between  $\rho=0.162117$  and

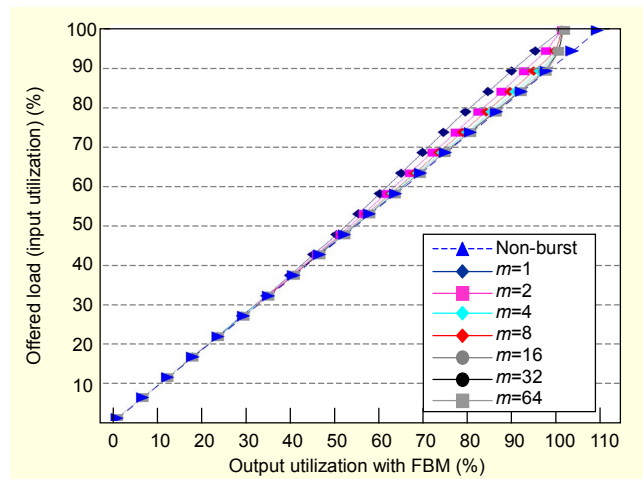


Fig. 4. Performance improvement of FBM with various numbers of ONUs (uniform traffic pattern input).

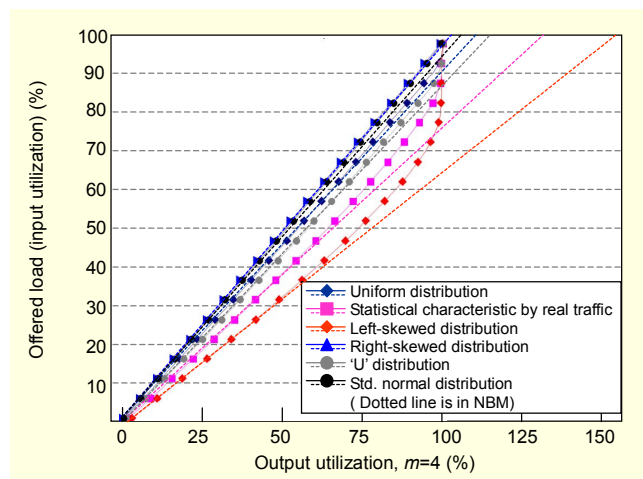


Fig. 5. Performance improvement of FBM and bandwidth usage of NBM with various traffic patterns.

0.1626347; therefore, the burst performance is just 0.996804 to 0.999983 for only NBM. The traffic load is sufficient to form a sequence of bursts. More ONUs cause packets with the same destination to be interleaved with other packets that have different destinations. As a result, burst performance declines significantly. It falls close to NBM when  $m=64$  ONUs. The FBM performance always approaches FDX mode with any number of ONUs.

### 2. Performance of FBM with Various Traffic Patterns

Based on this modeling, we setup a simulation based on the following 6 input traffic patterns: uniform distribution, standard normal distribution, U distribution, left-skewed distribution, right-skewed distribution, and the statistical characteristic of real traffic patterns. We use a single input with an FDX line speed of 1 Gbps. The input rate is fixed at 90%, and the output

is set to  $m=4$  ONUs. Then, we observe the relationship between traffic patterns and the probability of FBM [17].

Figure 5 shows the outbound traffic utilization in FBM. All of these results are based on different traffic patterns with the input rate varying from 5% to 100% through a single gigabit input. Figure 6 shows the bandwidth usage of NBM for reference. All curves go up in heavy utilization compared with FDX mode. Apparently, the left-skewed packet length distribution has both the poorest NBM performance and the best FBM performance. The same effect appears in other packet-length distributions. However, the burst performance always approaches FDX mode with any traffic pattern.

We also calculated the maximum input rate limitation through an FDX link under various traffic patterns, in which packet drop starts when the input load exceeds the limitation. Therefore, when packets from an HDX link enter an FDX link, the limitation will be inverted. For instance, the limitation for uniform packet length distribution is 1.80425, which means the offered input load should not exceed 0.922297 to avoid packet loss.

We observe that the burst performance compared with NBM has *Left-Skewed*>*Real traffic*>'U'>*Uniform*>*Normal*>*Right-Skewed* distribution. The performance improvement with normal and right-skewed distribution is insignificant. Real burst performance compared with an FDX input link has *RightSkewed*>*Normal*>*LeftSkewed*>*Uni-form*>'U'>*Real traffic* distribution.

### 3. Queue Size and Queuing Delay of FBM with Various Numbers of ONUs

Figure 6 shows the queue occupancy with various offered traffic loads and different numbers of ONUs. We found that the average queue occupancy per step in FBM is shorter than that in NBMs. This is because the packets in a burst phase are sent out ahead of schedule. Some other packets of the current burst phase will be postponed. The total ahead time is slightly less than the total lag time because both the average queue occupancy and average queuing delay are slightly shorter than NBMs.

When the number of ONUs increases, apparently the OLT needs more queue space to buffer those postponed packets. However, the queue occupancy of FBM still does not exceed the queue occupancy of NBM. Apparently FBM does not increase the cost of OLT.

Figure 7 shows the average queuing delay of the offered input traffic load and various numbers of ONUs. We found a similar trend in the queue occupancy. When the number of ONUs is 64 we get an average queuing delay of 655 ns, which is still better than that in NBM (696 ns).

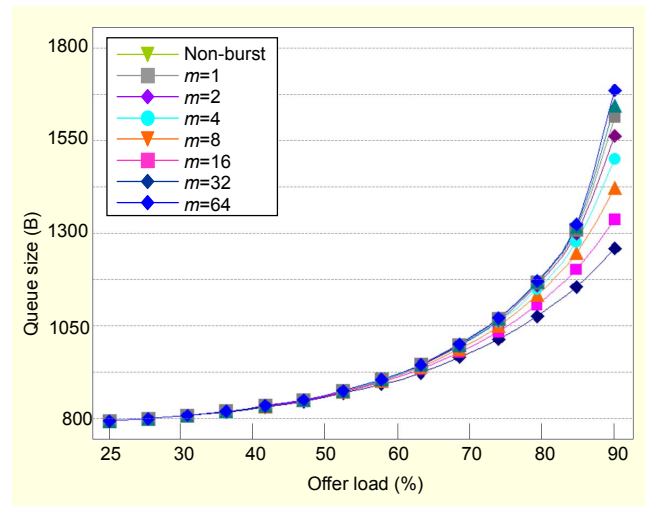


Fig. 6. Average queue occupancy of FBM in OLT.

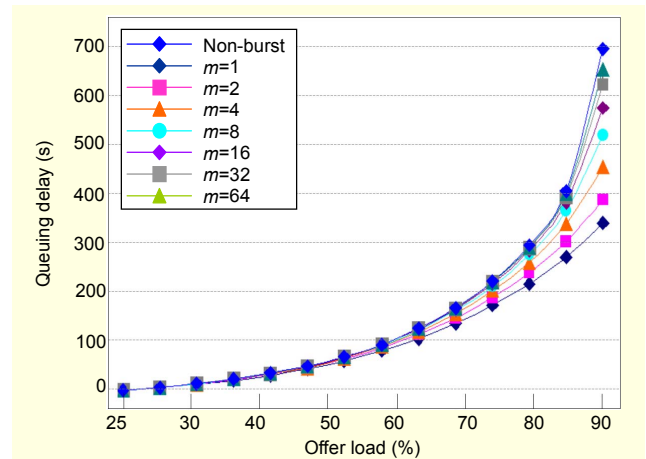


Fig. 7. Average queuing delay of FBM with various numbers of ONUs.

### 4. Queue Space and Queuing Delay of FBM in Various Traffic Patterns

Figure 8 shows the queue occupancy with offered input traffic load and different traffic patterns. The right-skewed input traffic pattern occupies the highest buffer space on average because most frames are long and cause long forwarding delay, queuing delay, and longer bursts. In the input traffic with left-skewed packet length distribution, the average packet length is too short; therefore, the buffer occupancy is less than the average. However, left-skewed packet length distribution achieves the best FBM performance. In this distribution, the buffer occupancy is raised quickly in a heavy load. In fact, congestion may occur at the same time.

Moreover, the queuing delay curves are totally different from the curves of queuing delay shown in Fig. 9. The queuing delay in FBM has a *Real traffic*>'U'>*Left-skewed*>*Uniform*>



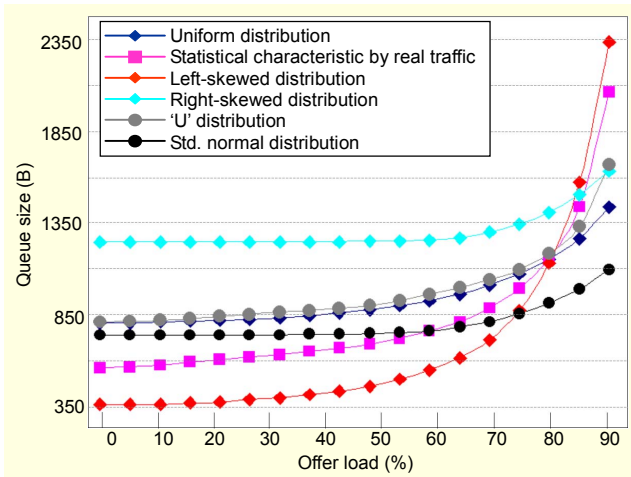


Fig. 8. Average queue occupancy of FBM in OLT.

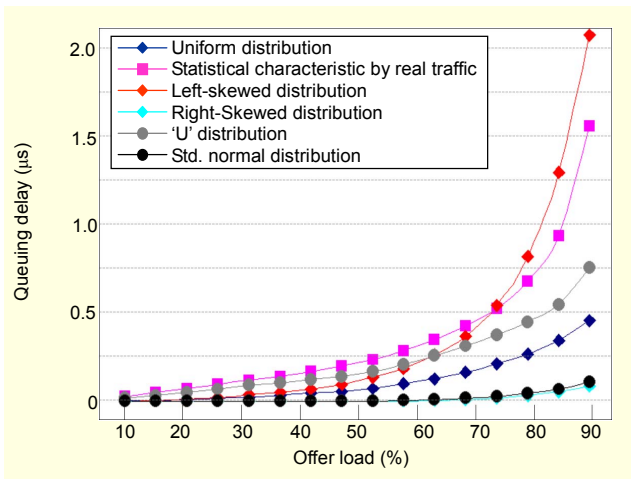


Fig. 9. Average queuing delay of FBM with various traffic patterns.

Normal > Right-skewed distribution on average. Compared with Fig. 8, the queue occupancy has a Right-skewed > 'U' > Uniform > Normal > Real traffic > Left-skewed distribution on average. Curves for all traffic patterns rise quickly under a heavy traffic load in FBM. This is not a complete figure because the available bandwidth is insufficient when there is a heavy load. Congestion occurs at the same time, and most traffic is dropped.

### 5. Analysis of Delay Variation

We setup a simulation with an offered load of  $\rho=0.9$  and 4 ONUs to observe the jitter difference between packets in FBM and NBM during a certain period of time. As shown in Fig. 10, under a particular point, FBM has more jitter than NBM and experiences higher jitter diversity. Nevertheless, there is less jitter in FBM than in NBM on average.

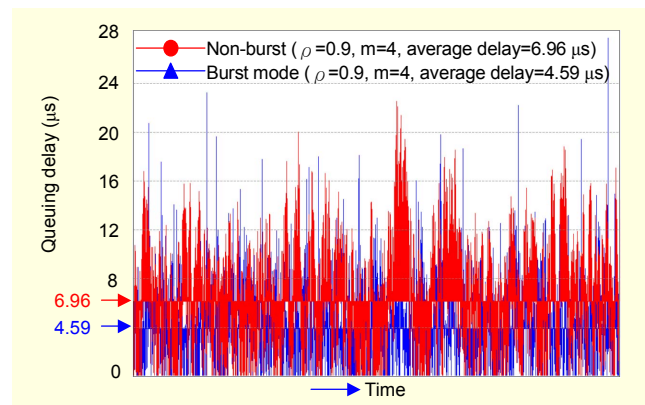


Fig. 10. Delay variation (jitter) of FBM vs. NBM.

We compare those simulation results with the theoretical analysis of FBM. According to the simulation results, under a heavy traffic load ( $\rho > 0.5$ ), the FBM performance always approaches full-duplex (FDX) mode and even higher with various numbers of ONUs and under various traffic patterns. Under a light traffic load ( $\rho < 0.5$ ), the FBM scheme has throughput similar to that of NBM. An increase in the number of ONUs worsens the burst performance. All these simulation results are consistent with the theoretical results presented in section III.

### V. Conclusion

In this paper, we investigated the current EPON technology and demonstrated the IEEE 802.3 FBM scheme in the EPON downstream link. A detailed process for setting up the demo configuration of FBM was provided, and simulations and results were presented. The design and performance analysis of the FBM scheduling scheme were presented, demonstrating the efficiency of FBM between the OLT and ONUs in the EPON downstream link. The FBM mechanism was proposed to resolve the inefficiency problem of EPON, which also impacts the fairness access principle of CSMA/CD. We observed the transitional effect from FDX to HDX with FBM/NBM networks. The analysis was based on the variance of four factors: traffic load, packet distribution, the number of input channels, and the number of output destinations. These four factors were applied to the measurement of throughput, queue occupancy, queuing delay, and jitter. The final results of the analysis were presented in terms of probability and distribution pattern of frame bursting.

The efficiency of FBM transmission is strongly related to packet length distribution. As data is encapsulated across multiple layers, different data lengths in different layers usually lead to complexity problems. For the application of bulk volume data transfer, longer frames would be more efficient.

Large frames could also be properly fragmented as the cross-layer requirement for shorter frames. For the transfer of shorter packets, the physical layer constraint is the major cause of inefficiency. For example, in VoIP packet transfer, the payload size generated by an 8k to 16k codec is typically around 40 to 80 bytes. Even less economic codecs, such as G711, which generates a relatively larger payload size of typically 200 to 300 bytes, are still considered small. Such small packets should be padded to 512 bytes in order to fulfill the MTU of EPON.

Nowadays, while using Ethernet Technology in a WAN or MAN, the mode of choice is always FDX. Indeed, 10GE (IEEE 802.3ae) will not even support MAC in HDX. Although GbE (IEEE 802.3z) supports HDX, very few implementations use HDX MAC. HDX is a historical relic of Ethernet that uses the CSMA/CD mechanism, which achieves only 30 to 40% utilization. On the other hand, it is easy for connections under FDX mode to reach 90% or even higher utilization. With FBM, HDX-mode EPON could improve its utilization to approach FDX performance without requiring a complex mechanism or high cost.

EPON has the potential to be a low-cost standard-based interoperable broadband access network. It represents one of many types of broadband access deployment, and may become an important architecture in an Ethernet-based access network. The EPON standard is in progress. Most of the technical considerations for EPON have already been defined, and the work now is focused on optimizing the protocol and ensuring technical feasibility, efficiency, and interoperability. While this work is ongoing, some of the key concepts have been addressed in this paper. Many upstream control protocols use polling schemes based on grants and requests, and this creates an overhead in limited downstream links. The impact of upstream protocols on downstream burst performance will be explored in future works. Moreover, the 802.3 FBM in the EPON downstream link is not fair and has not yet been optimized. To develop an efficient scheduling scheme during the burst period in accordance with these requirements is also required.

The development of an FBM scheme not only improves the performance of EPON, but can also be applied to Gigabit-capable PON (GPON), ATM PON (APON), GbE and 10GbE, as well as many wireless technologies, such as WiFi (802.11e), WiMAX (802.16), Bluetooth (802.15.1), WUSB, and Zigbee (802.15.4). With a design concept similar to that of FBM, jumbo frame will be a major technique for improving the network performance of next generation networks.

## References

[1] G. Kramer and G. Pesavento, "Ethernet Passive Optical Network

- (EPON): Building a Next Generation Optical Access Network," *IEEE Communications Magazine*, vol. 40, no. 2, Feb. 2002, pp. 66-72.
- [2] G. Kramer, B. Mukherjee, and G. Pesavento, "Ethernet PON (ePON): Design and Analysis of an Optical Access Network," *Photonic Network Comm.*, vol. 3, no. 3, July 2001, pp. 307-319.
- [3] G. Pesavento and M. Kelsey, "PONs for the Broadband Local-Loop," *Journal of Lightwave Tech.*, vol. 16, no. 10, Sept. 1999, pp. 68-75.
- [4] S. Halabi, *Metro Ethernet*, Cisco Press, 2003.
- [5] IEEE Draft P802.3ah/D1.0TM, *Media Access Control Parameters, Physical Layers and Management Parameters for Subscriber Access Networks*, Aug. 2002.
- [6] IEEE Std. 802.3ah-2004, Institute of Electrical and Electronics Engineers, New York, 2004.
- [7] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: A Dynamic Protocol for an Ethernet PON (EPON)," *IEEE Communications Magazine*, Feb. 2002, pp. 74-80.
- [8] V. Bommel, "MPCP and TDM Services," IEEE 802.3ah EFM, St. Louis, MO, Mar. 2002.
- [9] C.H. Foh and M. Zukerman, "A Novel and Simple MAC Protocol for High-Speed Passive Optical LANs," *Proceedings of Networking 2002*, Pisa, Italy, 2002.
- [10] C. Partridge, *Gigabit Networking*, Addison Wesley, Oct. 1993
- [11] M. Molle, M. Kalkunte, and J. Kadambi, "Frame Bursting: A Technique for Scaling CSMA/CD to Gigabit Speeds," *IEEE Network*, July/Aug. 1997, pp. 6-15.
- [12] R. Breyer and S. Riley, *Switched, Fast, and Gigabit Ethernet/3e*, New Riders Publishing, 1999.
- [13] C.A. Eldering, "Theoretical Determination of Sensitivity Penalty for Burst Mode Fiber Optic Receivers," *Journal of Lightwave Technology*, vol. 11, no. 12, 1993, pp. 2145-2149.
- [14] C. Su, L.K. Chen and K.W. Cheung, "Theory of Burst-Mode Receiver and Its Application in Optical Multi-access Networks," *Journal of Light-Technology*, vol. 15, no. 4, 1997, pp. 590-606.
- [15] E. Fuchs and P.E. Jackson, "Estimates of Distributions of Random Variables for Certain Computer Communication Traffic Models," *Comm. of ACM*, vol. 13, no. 12, Dec. 1970, pp. 752-767.
- [16] A. Belenkiy and N. Uzun, "Deterministic IP Table Look-Up at Wire Speed," *INET'99: The Internet Global Summit*, San Jose, CA, Jun. 1999.
- [17] K. Khalil, K. Luc, and D. Wilson, "LAN Traffic Analysis and Workload Characterization," *Proc. Local Computer Networks*, Minneapolis, MN, Sep. 1990, pp. 112-122.



**Wen-Kang Jia** received his BS degree in industrial management from the Chung-Hua University, and the EMBA degree in information management from Ming-Chuan University. He is also a PhD student with the Department of Computer Science, National Chiao-Tung University, Hsinchu, Taiwan. He

has fifteen years of networking experience in a variety of fields including network integration, IP protocol design, VoIP technology, NAT traversal, teletraffic engineering, and telecom industry analysis. He was previously a senior professional manager of the IT, Manufacture, Telecom, and Network Industry.



**Yaw-Chung Chen** received his PhD degree in computer science from Northwestern University, Evanston, Illinois in 1987. From 1987 to 1990, he worked at AT&T Bell Laboratories as a member of technical staff. Since 1990, he has been with the Department of Computer Science, National Chiao-Tung

University, Hsinchu, Taiwan, where he is currently a professor. His research interests include multimedia communications, P2P systems, mobility in wireless networks, and Internet QoS. He is a member of ACM and a senior member of IEEE.