Traffic Behavior Analysis of Frame Bursting for SISO IEEE 802.3z Networks

Wen-Kang Jia, Yaw-Chung Chen, Cheng-Yuan Ho
Department of Computer Science, National Chiao Tung University
1001 Tahsueh Road, Hsinchu, Taiwan, ROC
wkchia@cs.nctu.edu.tw , ycchen@cs.nctu.edu.tw, cyho@csie.nctu.edu.tw

ABSTRACT

This paper presents traffic analysis of Frame Bursting (FB) in Single-Input Single-Output (SISO) IEEE 802.3z (Gigabit Ethernet; GbE) Networks. This analysis characterizes the Frame bursting traffic behavior and performance measurement, shows its implications for future research issues. GbE is one of the world's largest Enterprise LANs solutions in terms of the number of concurrent users. FB Mechanism improves the performance of GbE CSMA/CD networks when transmitting short frames by reducing the Extension Field (EF) overhead, and tries to accord the controversy over the cross-layer (Layer2/Layer3) scheme. But FB get some negative effect like Jumbo Frame, long burst will affect the equal opportunity of other transmission. We simulate about 1 Tera-packets by random generation, let the packet pass through form non-FB domain to FB domain, observe that FB behavior like the burst per 100 packets (BPHP), packet per burst(PPB) and it's distribution, the burst length in different Input Rate, Input Channel, and Packet Length Distribution Input, The analysis of the burst performance reveals that the geometrical proportion between burst size and burst length, and the inverse proportion between the percentage of Burst and the output utilization after bursting. We find that there is a linear relationship between Input Rate and the burst behavior. As for the burst behavior is highly dynamic complex in variable Number of Inputs, and Packet Length Distribution types.

Keywords: Gigabit Ethernet, Frame Burst, Traffic Behavior

1. INTRODUCTION

As the new applications in web access, e-commerce, entertainment and high-quality streaming media over the Internet are ever increasing, massive demand for network bandwidth becomes a must and administrators of LANs have to seek and adopt new technologies. To fulfill the bandwidth demand of immense traffic volume, the IEEE Standards Board approved IEEE 802.3z Gigabit Ethernet in July 1998. It allows data transmission rate up to 1,000 Mb/s, 10 times higher than Fast Ethernet over LANs. [3]

The Gigabit Ethernet LAN solutions can be applied wherever Fast Ethernet works. It is a simple, cost-effective investment that can easily and quickly relieve bottlenecks of network connections. Since Ethernet links to network equipment have historically been significantly less expensive than TDM-based solutions, enterprise customers find an incentive to deploy Ethernet connecting to their ISPs. As a consequence, major carriers are exploring methods to provide Ethernet-based services in addition to traditional TDM-based private lines. [7][11]

Ethernet devices must allow a minimum idle period between two consecutive frame transmissions; this is so-called Inter-Frame gap (IFG) or Inter-Packet gap (IPG). It provides a short recovery time between frames to allow interface devices to prepare for reception of the next frame. The minimum IPG is 96 bit times, which is 9.6 microseconds for 10 Mb/s Ethernet, 960 nanoseconds for 100 Mb/s Fast Ethernet, and 96 nanoseconds for 1 Gb/s Gigabit Ethernet (GbE).[11]

In both Ethernet and Fast Ethernet, Layer3 payload (packet) is a sequence of n bytes ($46 \leqq n \leqq 1,500$) of any value. The minimum and maximum Layer2 frame sizes (Layer3 packet + overhead) are 64 bytes and 1,544 bytes respectively. The largest physical frame size is called Maximum Transmission Unit (MTU) [3][7][11], and frames larger than MTU are divided into smaller ones, called fragments, before being sent. Different networks may have different MTU. While the

Minimum Transmission Unit (MinTU) is often conforming to Layer 2 and Layer 3 protocol in Ethernet and Fast Ethernet, but this rule has been broken by GbE due to the introduction of Extension Field (EF) in the protocol.

Two operation modes, half-duplex and full-duplex, are specified in Gigabit Ethernet. While operating in half-duplex mode, an EF field may be appended to the end of the 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. Although the meaningful payload may be only 46 bytes, the minimum length of the transmitted frame must be 512 bytes. The EF field is not used in full-duplex mode, which is for point to point transmission. [7][9][11]

Carrier Extension is a simple solution, but it may waste bandwidth. Up to 448 padding bytes may be sent for small packets. This results in very low throughput. In fact, for a large number of small packets, the throughput is only marginally better than Fast Ethernet. In order to solve the problem of inefficiency, an optional burst mode that allows a station to transmit consecutive frames without relinquishing control of the medium was specified. Burst mode is only available on Gigabit and 10 GB/s Ethernet under half-duplex mode, as the CSMA/CD protocol is not used in full-duplex mode. In non-burst mode, an EF field is appended as needed so the carrier is extended to the required minimum slot time. [7][9]
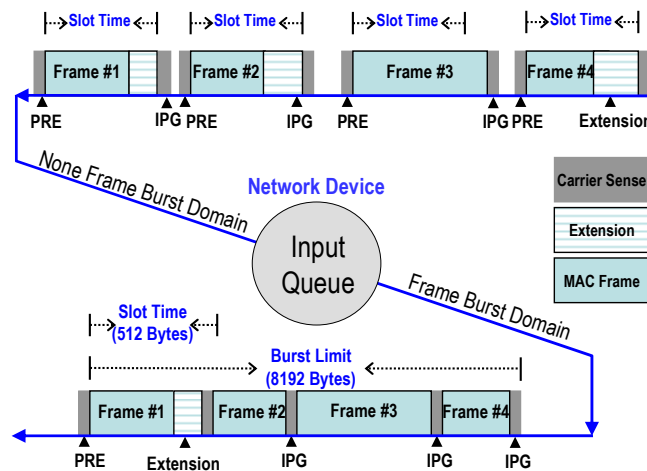

Figure (1) Burst mode Mechanism

Figure 1 illustrates a frame with an extension field appended. After successfully transmitting a frame, a station sending in burst mode may continue to 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. But instead of allowing the medium to go idle between frames, the transmitting station fills the IFG with extension symbols. These symbols are used for maintaining an active carrier, and are distinguished from normal data symbols. 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. If a collision occurs, only the first frame in the burst will be affected and require retransmission.

Frame Bursting is an extension of Carrier Extension. In other words, it is "Carrier Extension plus a burst of frames". When a station has a number of frames to transmit, the first frame is padded to the slot time using carrier extension if necessary. Subsequent frames are transmitted back to back with the minimum IFG until a burst timer expires. Frame Bursting substantially increases the throughput. Figure 1 shows how Frame Bursting works. Since short packets frequently appear in traditional Layer3 protocols, Burst mode could be used to improve the performance of GbE/10GbE when transmitting short frames. [9]

2. THE FRAME BURST MODEL

For frame bursting, we are most interested in the relationship between the traffic intensity and probability of frame bursting, as well as the performance improvement with frame bursting.

Our frame bursting traffic model is based on the following assumptions:

1. A half-duplex CSMA/CD switched network,
2. Neither collision and nor congestion will occur,
3. Infinite output buffer size,
4. Poisson distribution in frame inter-arrival time,
5. Normal distribution in frame size, and
6. FIFO-Queue/ Single Server Model.

For an incoming frame sequence F_n, $n= \{0, 1…k\}$, the service time of a frame is assumed identical to its length divided by the transmission rate. Since the Minimum Transfer Unit $MinTU$ is 4096 slot times (512 bytes) in GbE network, a packet with length smaller than $MinTU$ must be appended with Extension to make it up to $MinTU$, Let the $|F_n|$ is the Layer2 frame size of F_n, and the $|P_n|$ is the Layer3 packet size of P_n, so when a Layer3 packet P_n is sent to Layer2, we will have the Layer2 frame F_n as

$$|F_n| = \max(|P_n|, MinTU) \qquad (1)$$

It is necessary to keep a silence period T_{IPG} between frames, called Inter-Frame Gap, in the CSMA/CD network. The switch also need processing time T_{FD} to forward a frame to its corresponding output port. Assume the transmission slot time of F_n is $T(F_n)$, For consecutive k frame arrivals which may form a burst, we could calculate the total service time X in the output queue as follows:

$$X = \sum_{n=0}^{k}\left(T(F_n) + T_{IPG}\right) + T_{FD} \qquad (2)$$

In General, the T_{FD} is usually small than $T_{IPG,}$ within FE networks (960ns), More specific studies indicate that the IPv4 lookup delay lasts approximately 250~350ns; and the average IPv6 routing delay lasts around 460~960ns [1][6]. Where T_{FD} could be Ignores in the case, But the line speed is up nowadays (96ns in GbE and 9.6ns in 10GbE), we are not sure it's exist a L3 network device which have poor forwarding performance that the T_{FD} is big than T_{IPG}, where the X is

$$X = \sum_{n=0}^{k}\left(T(F_n) + \max(T_{IPG}, T_{FD})\right) \qquad (3)$$

However, the condition for a frame bursting to occur in sequential packets is that when the heading frame is sent out, the following frame is already in the queue. If the same frame transmission pattern repeats this until the total transmitted frame length reaches the $BurstLimit$ (65,536 slot times = 8,192 bytes), and in order to make sure the Carrier Sense is working well, the leading frame P_0 must be no smaller than $MinTU$. Then we could have the following:

$$\left[\sum_{n=1}^{k-1}\left(|P_n| + T_{IPG}\right) + |F_0| + T_F\right] \geq \sum_{n=1}^{k}\left(|F_n| + T_{IPG}\right) \qquad (4)$$

and

$$\left[\sum_{n=1}^{k}\left(|P_n|+T_{IPG}\right)+|F_0|+T_{FD}\right]\leq BurstLimited \qquad (5)$$

Where k is the burst size, and we obtain the *BurstLength* below, where the P_k is the last frame in the burst.

$$BurstLength = \sum_{n=1}^{k-1}\left(|P_n|+T_{IPG}\right)+|F_0|+|P_k| \qquad (6)$$

Theoretically, the burst size k may reach a range as shown in the equation below:

$$2 \leq k \leq \left[INT\left(\frac{ButstLimited - MinTU - |T_{IPG}||}{MinFrame + |T_{IPG}|}\right)+2\right] \qquad (7)$$

Where we get k are 102 frames per burst (FPB), which means there will be 102 frames aggregated in one burst process under the extreme condition. And we can get the largest possible *BurstLength* 9,735 by Eq.(6) when F_k=MTU (1,544 bytes).

Considerations in a real system should include time intervals between two consecutive frame arrivals, this interval will reduce the probability of burst. From Eq.(4), we obtain

$$\left[\sum_{n=1}^{k}\left(|P_n|+T_{IPG}\right)+|F_0|+T_{FD}\right]\geq\left[\sum_{n=1}^{k+1}\left(|P_n|+T_{IPG}+T_{INTERVAL}\right)\right] \qquad (8)$$

Where the $T_{INTERVAL}$ is considering that come from m input and in throughput ρ, example in the single input (m=1) and the fully load throughput (ρ=1), we obtain the $T_{INTERVAL}$ is zero by below equation:

$$T_{INTERVAL} = m + \left(\frac{(m-mp)}{\rho}\right) - 1 \qquad (9)$$

Hence, the leading frame F_0 is incorporating a frame bursting sequence until all frames in the queue are sent out and the queue becomes empty, or the burst length reaches the *BurstLimit*, and it must stop the burst mode. We could obtain the *BurstLength* as

$$BurstLength = \left(\sum_{n=0}^{k+1}|P_n|+T_{IPG}\right)-T_{IPG} \qquad (10)$$

Based on 802.3z, the last frame which encounters the *BurstLimit* are allowed to transmit over the *BurstLimit* until the frame is run out. This means the *BurstLength* permission is bigger than *BurstLimit*, but in some extreme situation the *BurstLength* is just equal to *BurstLimit* when the *BurstLength* reach the *BurstLimit*. In such case, we have

$$BurstLimited =$$
$$BurstLength = \left[\sum_{n=1}^{k}\left(|P_n|+T_{IPG}\right)+|F_0|\right] \qquad (11)$$

During a frame bursting period $P_{0\sim k}$, to reduce the extension from whole slot times, we obtain:

$$\sum_{n=1}^{k}\left(\left|F_n\right| - \left|P_n\right|\right) \qquad (12)$$

And we can calculate the percentage of utilization reduction in a frame bursting process:

$$\left.\sum_{n=1}^{k}\left(\left|F_n\right| - \left|P_n\right|\right) \middle/ \left[\sum_{n=1}^{k}\left(\left|P_n\right| + T_{IPG}\right) + \left|F_0\right| + T_{FD}\right]\right. \qquad (13)$$

3. SIMULATION

3.1 Simulation Environment

We evaluate the performance of the frame bursting in GbE network through simulation. We assume that three factors that influence the results:

(1) Frame arrival rate (λ): so-called traffic intensity, in general sense, high arrival rate causes smaller or even no inter-frame interval (excluding Inter-Frame Gap) in the traffic, that means packets have higher opportunity to be mixed together. On the other hand, extremely long inter-frame intervals may lead to an emptied queue in which frame bursting is unlikely to happen.

(2) Frame Length Distribution: Theoretically, if the frame length distribution is centered at a mean value or same value, like standard normal distribution, which difficult to make frame bursting, Similarly, if packet length distributed in the whole traffic, that mean the longer frames and the shorter frames will be working together to make frame bursting. In other words, one longer frame will lead many shorter frames to make a frame burst.

(3) Number of Inputs: Traffic from different inputs may be heading to a specific output queue and these traffic will be aggregated. When an output queue has many packets in waiting, the probability of frame bursting will be increased.

Algorithm (1) Calculating Burst Size and Length

```
Set_Interrupt(PKT_SEND_COMPLETE());
Set_Interrupt(PKT_RECV_COMPLETE());
While(!EXIT)
{    if (CHECK_Queue()) //Store and Forward
     { P=POP_Queue();
     PKT_SEND(P);
     Burst_Length=sizeof(P);
     Burst_size=1; }      }//Init Burst
Interrupt PKT_SEND_COMPLETE()
{    if (!CHECK_Queue())  //Burst finish by Queue
Empty
     { Record_Burst_Size&Length();
     Burst_size=0;       //Clear Burst count;
     return; }
   P=POP_Queue();
   if((Burst_Length+sizeof(P))<=BurstLimited))
   // Init Burst
     { PKT_SEND(P);
     Burst_Length=Burst_Length+sizeof(P);
     Burst_size++; }
   else //Burst finish by Reach Burst Limited
     { Record_Burst_Size&Length();
```

```
        Burst_size=0; }        }//Clear Burst count;
Interrupt PKT_RECV_COMPLETE()
{PUSH_Queue(PKT_RECV());}
```

First, we set a simple environment like that in Figure 1 to observe the result for the "Single Input Single Output" scenario, and Algorithm (1) is a simple illustration to example of doing equations.

From this, we setup a simulation by varying input traffic pattern with six types below, and use a single input channel, in which the traffic intensity is 100%.

(a) Uniform Distribution

(b) Standard Normal Distribution

(c) 'U' Distribution

(d) Left-Skewed Distribution

(e) Right-Skewed Distribution

(f) Statistical Characteristic by Real Traffic Pattern

The second Step, we choose the Real Traffic Pattern, and setup input rate varying from 10%~90% through a single input channel, then we observe the relation between traffic intensity and probability of frame bursting.

3.2 Burst Performance

Figure 2 shows the number of bursts per hundred packets (BPHP) in several of input rate of Real Statistical Characteristic Packet Length Distribution. As the input traffic (λ) increases, the BPHP also increases, and peaks to 24.11 times when $\lambda=80\%$, but drops to 13.32 times when $\lambda=100\%$ gradually, because the traffic intensity is plenty enough to congregate a sequence of bursts.
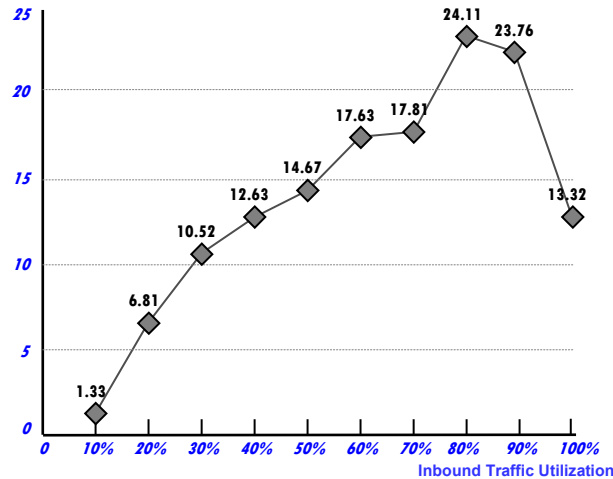

Figure (2) The Simulation Results of Burst times per 100 Packets (BPHP) in several of traffic intensity

Figure 3 shows curves of percentage for frame burst in the whole traffic flow (by frames).and a curve of percentage for burst length in the whole traffic length (by bytes). We may observe that both two above present the similar behavior.
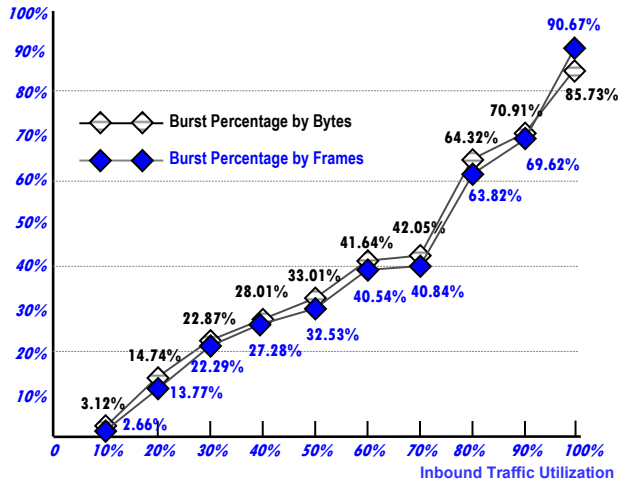
Figure (3) The Simulation Results of Percentage of Burst in whole Traffic (by frames and by bytes)

In Figure 4, it shows the outbound traffic utilization after the burst, all of these results are based on real traffic pattern, and with input rate varying from 10%~90% through a single input channel. We can observe that the full input rate makes the best performance with 79.85% outbound utilization, that means frame bursts reduce 20.15% transmission time, in other words, the frame burst improves 25.23% transmission performance. When the input rate is down to 10%, the output utilization after burst is 99.48%, We may say the burst performance is just 0.52% only.
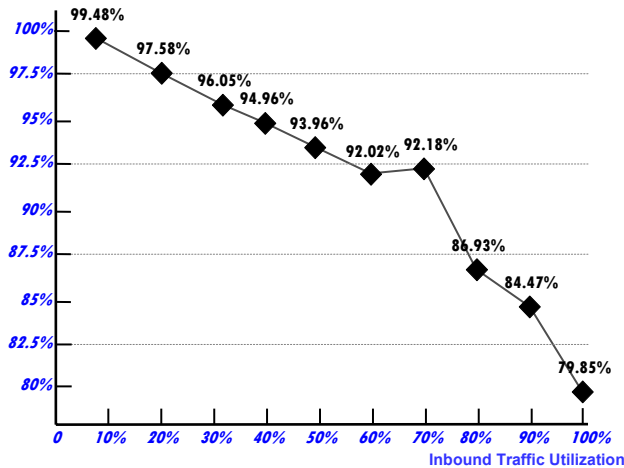
Figure (4) Outbound Traffic Utilization after bursting (Input non-burst Traffic to Output burst Traffic)

As the Figure 5 shows, the darker area is the input traffic; the lighter area is the length distribution of the first leading packet; which explains the behavior of burst leading packet distribution of the first leading packet's length when output traffic generates the Frame Burst. As we notice, the distribution which burst leading packet is generated, the longer packets has a higher opportunity than the shorter ones, which makes a slower slope of the burst leading packet Distribution curve starting from 434 Bytes, and climbs up to the right of the longer-length packets, at 989 Bytes to the peak, and falls to 1,359 Bytes with a much deeper slope caused by the limit of the input packet length. This simulation result has fully explained the behavior which longer packets may effectively trigger following packets to a burst behavior. And that mean the longer frames and the shorter frames will be working together to make frame bursting. In other words, one longer frame will lead many shorter frames to make a frame burst.
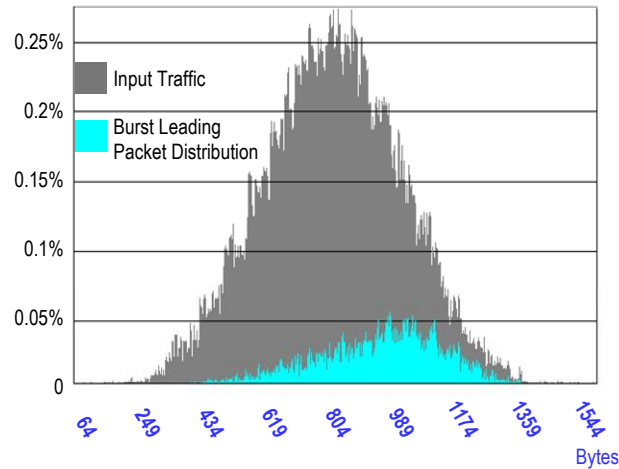
Figure (5) the Simulation Results of Burst Leading Packet Length Distribution (e.g. Standard Normal Distribution)

3.3 Various Behavior of Burst Length (PPS; Packets per Burst) Distribution

Figure 6 shows the Burst Length (PPS; Packets per Burst) distribution of the offered input traffic patterns and difference Traffic Intensity; Figure 6a is the result from a uniform distribution traffic input. It is apparent that Poisson distribution as PPB increases. Similarly, Figure 6f illustrates the same effect of homogenized packet length distribution.

Figure 6c shows a "U" distribution traffic input. It is apparently Poisson distribution as PPB increases too. The frame length disperses to either very short or very long because the longer frames have a greater chance to be aggregated with shorter frames to generate burst. On balance, both uniform and non-uniform frame length distribution have large chance in generating a burst.
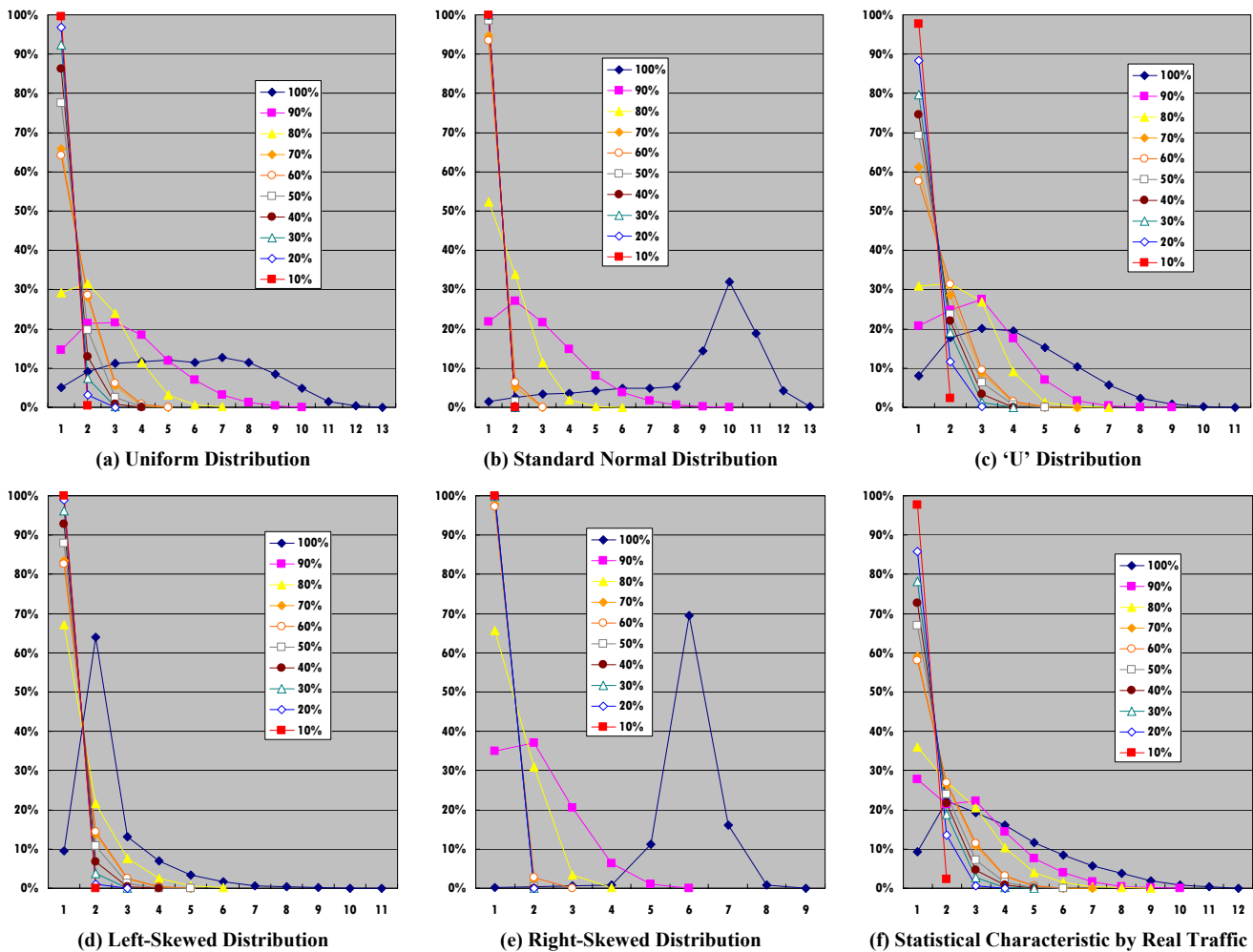
Figure 6b shows an input with standard normal distribution. It is apparent that most of the burst sizes are between 9~10 PPS. We assumed that such situation is most likely related to mean-length packets that produce a burst which lasts until the burst limit is reached.

Figure 6d shows a left-skewed distribution traffic input. That means most frames are short because no enough traffic intensity to make a long burst. So we obtain the burst sizes which are around 2 PPS and seldom extend up to 3 PPS or higher. And the higher number of frames per burst in 10 which has an average burst limit of just 5,691 bytes. This reveals that most burst is terminated due to queue empty before it reaches the burst limit.

Figure 6e shows a right-skewed distribution traffic input. It means that the most frames are large because they have high intensity to generate longer burst, as shown in Figure 2b. We obtain burst sizes which are mostly 6 PPS, because the average packet length is too long so that the burst reach the burst limit prematurely.

3.4 Various Behavior of Burst Size (BPS; Bytes per Burst) Distribution

Figure 7 shows the Burst Size (BPS; Bytes per Burst) distribution of the offered input traffic patterns and difference Traffic Intensity; we should find that the Average Burst Size per step is shorter in which Burst Length is longer. That's because the pervious burst power will be postponed to trade-in more Burst Length. In other words, we should denote that one Short Burst maybe consisted of a minority of the longer frame, which premature stop due to reach the burst limit. And one Longer Burst maybe consisted of most of the shorter frame, which have more chance to accumulating a huge mass of bytes to generate a huge Burst Length but not restricted in burst limit.

(a) Uniform Distribution

(b) Standard Normal Distribution

(c) 'U' Distribution

(d) Left-Skewed Distribution

(e) Right-Skewed Distribution

(f) Statistical Characteristic by Real Traffic

Axis Y: Percentage of Burst (Packet)

Axis X: Frame per Burst, '1' is no burst, '2' is 2 Frames in once Burst…etc.

Figure(6) The Simulation Results of Burst Length(PPB) Behavior on various types of Packet Length Distribution import

4. SUMMARY AND CONCLUSIONS

Transmission efficiency is highly related to frame length. [4][12][13] As data was encapsulated across multiple layers, different data sizes in different layers will cause the problem more complicated. For the application of bulk volume data transfer, larger frame size will be more efficient. Large frame packets could also be properly fragmented as the cross layer requirement demands for smaller frame size. For the application of small size packet transfer, physical layer constraint will be the major cause of inefficiency. For example, in VoIP packet transfer, payload size generated by G.723/G.729 codec is typically around 40~80 bytes, even for less economic codec such as G.711, it generates relatively larger payload size, typically 200~300 bytes only, are still considered as small. Such small packets should be padded to 512 bytes in order to fulfill the Minimum Transmission Unit of Gigabit Ethernet.

Frame bursting mechanism was proposed to solve above issues. As the frame bursting mechanism was applied to resolve the inefficiency problem of GbE, it will also impact the equal access principle of CSMA/CD. This article explores the microscopic phenomena of FBM by observing the transitional effect from non frame bursting network to framing bursting network, the analysis will be based on three variances; different traffic intensity(input rate), different packet distribution and different number of input channels. Two metric factors were applied as measurement of performance;

decreasing of utilization and packet length per burst. The final result of analysis is the formation of probability and distribution pattern of frame bursting.

The development of Frame bursting not only improves the performance of GbE and 10GbE, it also applicable in many wireless topics such as WiFi(802.11e), WiMAX, Bluebooth,WUSB and Zigbee. With similar design concept of frame bursting, jumbo frame will be a major technology for improving the network performance of next generation network.

## ACKNOWLEDGMENT

REFERENCES

1. A. Belenkiy and N. Uzun, "Deterministic IP Table Look-Up at Wire Speed", INET'99: The Internet Global Summit, San Jose, CA, Jun. 1999.
2. A. Feldmann, A.C.Gilbert, P. Huang and W. Willinger, "Dynamics of IP Traffic: A Study of the Role of Variability and the Impact of Control", Proceedings of the ACM/SIGCOMM'99, Cambridge, MA.
3. Breyer, Robert and Sean Riley, "Switched, Fast, and Gigabit Ethernet", New Riders Publishing, 1999.
4. C. Partridge, "The End of Simple Traffic Models", (Editor's Note), IEEE Network, Vol. 7, No. 5, September 1993, page3-3.
5. 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, December 1970, pages 752-767
6. D. Pao, C. Liu, A. Wu, L. Yeung, and K. Chan, "Efficient. Hardware Architecture for Fast IP Address Lookup", IEEE. Informatics and Communications Conference (INFOCOM), Vol. 2, June 2002, 555-561.
7. "Gigabit Ethernet: Accelerating the Standard for Speed", Gigabit Ethernet Alliance white paper, 1998.
8. Klivansky, S. K., Mukherjee,A. and Song, C. "On Long Range Dependence in NSFNET Traffic" Technical Report GIT-CC-94/61, Geogia Institute of Technology, Altlanta, GA 30332, USA, December 1994.
9. Molle, M., Kalkunte, M., and Kadambi, J., "Frame Bursting: A Technique for Scaling CSMA/CD to Gigabit speeds". IEEE Network. July/August 1997. pg. 6-15.
10. P. Abry and D. Veitch, "Wavelet Analysis of Long-range-dependent Traffic", IEEE Trans. on Information Theory, Vol. 44 NO. 1, Jan. 1998, page2-15.
11. Partridge C., "Gigabit Networking", Addison Wesley, October 1993
12. Rene L. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation", IEEE Transaction on Information Theory, Vol. 37, No. 1, January 1991, Pages 114-131.
13. Rene L. Cruz, "A Calculus for Network Delay, Part II: Network Analysis", IEEE Transaction on Information Theory, Vol. 37, No. 1, January 1991, Pages 132-141.
14. T. Tuan and K. Park, "Multiple Time Scale Redundancy Control for QoS-sensitive Transport of Real-time Traffic", Proc. IEEE INFOCOM'00, 2000.
15. W. E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, "On the Self-Similar Nature of Ethernet Traffic", IEEE/ACM Transactions on Networking, Vol.2, No.1, February 1994, pages 1-15
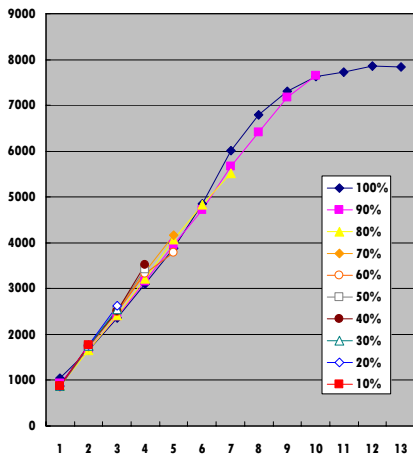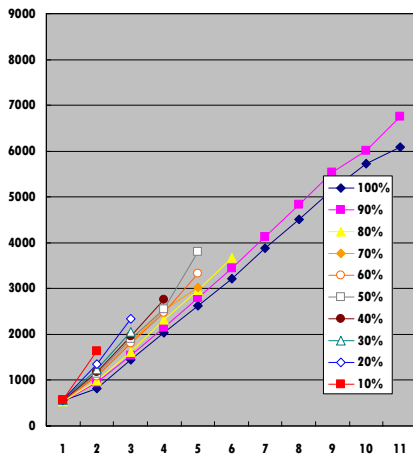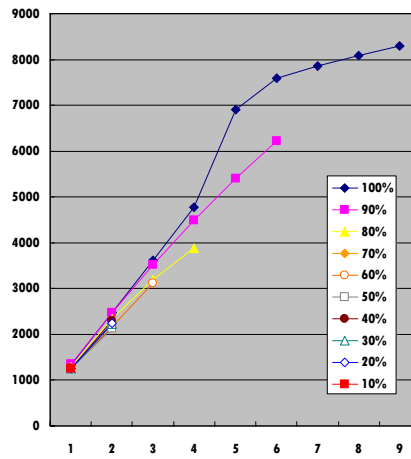
(a) Uniform Distribution
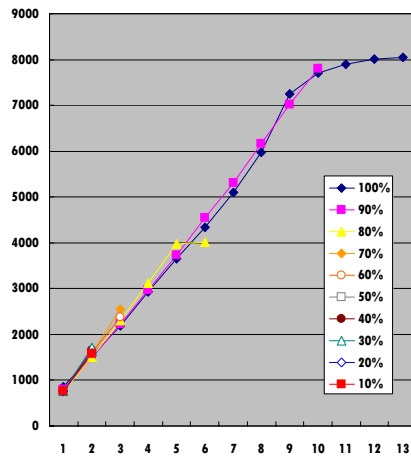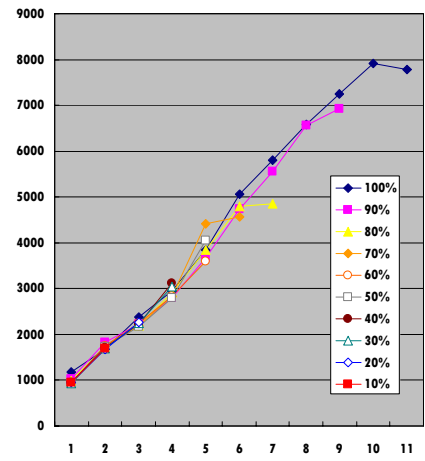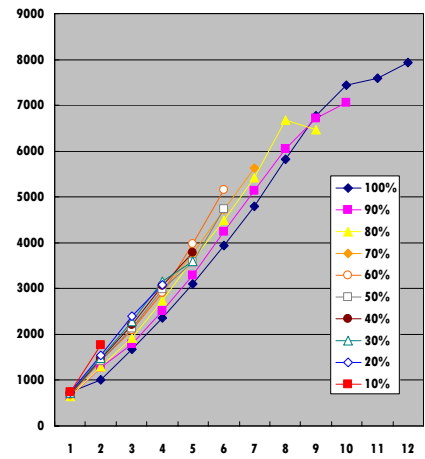
(b) Standard Normal Distribution

(c) 'U' Distribution

(d) Left-Skewed Distribution

(e) Right-Skewed Distribution

(f) Statistical Characteristic by Real Traffic

Axis Y: Average Burst Length (Bytes)

Axis X: Frame per Burst, '1' is no burst, '2' is 2 Frames in once Burst…etc.

Figure(7) The Simulation Results of Burst Size(BPB) Behavior on various types of Packet Length Distribution import