PERFORMANCE OF BUFFERED TAGLE SHARMA SWITCH UNDER RANDOM TRAFFIC

[1]Luisito I. Tabada and [2]Pierre U. Tagle

Abstract

The study proposed an enhancement to the Tagle-Sharma network, a high performance self routing fault tolerant switch which employed an enhanced scheme of the banyan network. The enhancement incorporated a shared buffer in each switching element (SE) of the network. Suitable management of incoming packets in a shared buffer switch is necessary to provide high buffer utilization and least amount of buffer space. Two buffer management configurations were proposed. The First-In-First-Out (FIFO) configuration used internal control and followed the FIFO policy in handling the buffer. The look-ahead configuration also followed all the features of the FIFO scheme but used the buffer only when needed to improve performance in terms of transmission delay. Simulation results showed that both configurations gave high throughput. Introduction of the deadline kept the delay of the network low. The look-ahead scheme offered better transmission delay performance than the FIFO approach. The Tagle-Sharma switch needed only a buffer size of 4 to achieve higher performance.

Keywords: *tagle-sharma, traffic, performance, network*

1.0 Introduction

Tagle-Sharma network illustrated in Figure 1a is a high performance self routing fault tolerant switch that employs an enhanced scheme of the banyan network. It consists of two banyan networks with links provided at every stage to allow packet transfer to and from each banyan plane, thereby offering multiple paths between each input-output pair and giving it a high degree of fault tolerance and overcoming the
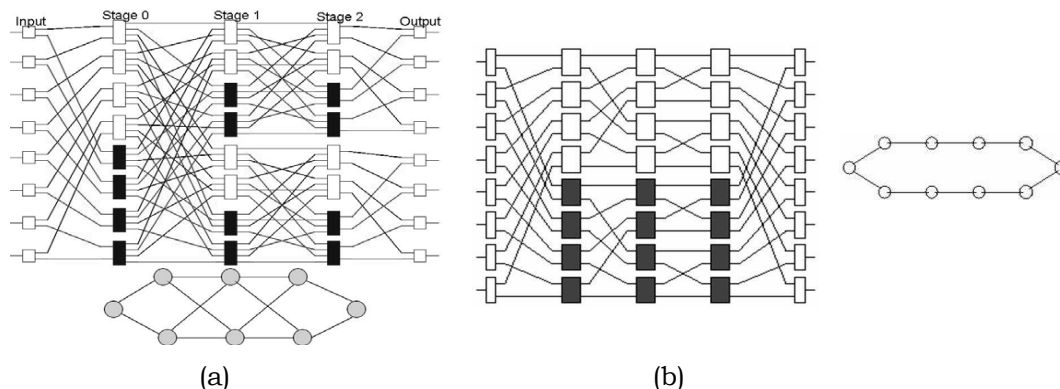

(a)   (b)

Figure 1. Photograph of an (a) 8x8 Tagle-Sharma switch and its channel graph; and (b) Parallel Banyan switch (lower bound Tagle-Sharma switch)

[1]*Caraga State University*
[2]*Ateneo de Manila University*

single path limitation of banyan networks. Even with the significant reduction in packet loss rate, the baseline (bufferless) Tagle-Sharma switch is still susceptible to packet loss especially at peak load condition because the switch can receive only one packet at a time inside each switching element. One of the best solutions is buffering. However, it has been noticed based on previous related studies conducted that buffering cannot totally eliminate packet loss, but it can reduce it to acceptable levels. Blocking is a problem with which every switch design must deal. Blocking can happen internally when packets contend for the same resource. Buffers provide physical storage to hold packets that cannot yet be sent to their desired output ports.

The aim of the study was to enhance the Tagle-Sharma switch by internal buffering using shared buffer approach. Shared buffer provides high buffer utilization and needs less amount of buffer space. Several efforts have been reported to study the performance of Multistage Interconnection Networks (MINs) with internal buffering. Studies also agreed that shared buffer has the best performance among all other internal buffering strategies.

The First-In-First-Out (FIFO) and look-ahead buffer handling techniques for internal buffering using shared buffer strategy were proposed. In this study, the Tagle-Sharma network is modeled to evaluate the performance of the shared buffer SEs of the network. The FIFO scheme employs backpressure flow control mechanism and follows the FIFO policy. The look-ahead configuration also applies the FIFO scheme, but utilizes the buffer only when needed. The performance of the upper bound (Figure 1a) and lower bound (Figure 1b) networks of Tagle-Sharma were also evaluated.

2.0 Materials and Methods

The goal of the research was to enhance the Tagle-Sharma switch by incorporating a shared buffer in each switching element to improve its performance in terms of increased throughput, and keeping transmission delay within acceptable levels. Look-ahead and FIFO buffer management schemes were proposed to handle the buffer efficiently.

The Tagle-Sharma switch consisted two planes each with n stages (n = $\log_2 N$), where N is the network size. Figure 2 illustrates
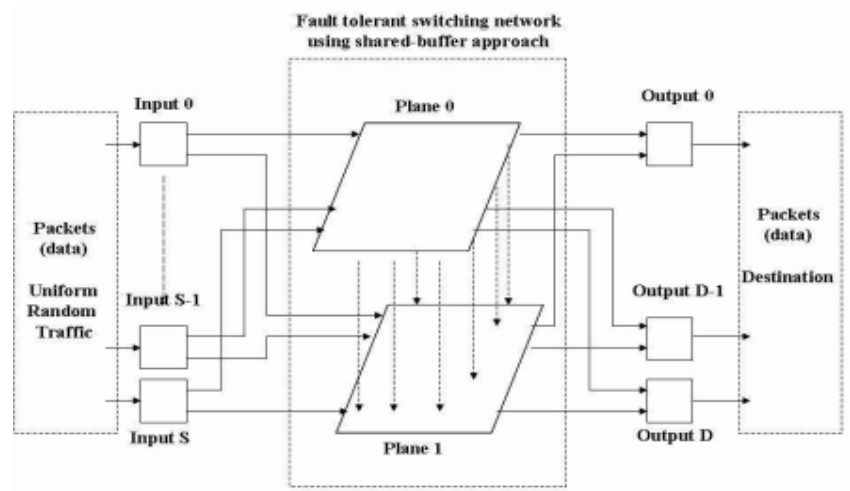

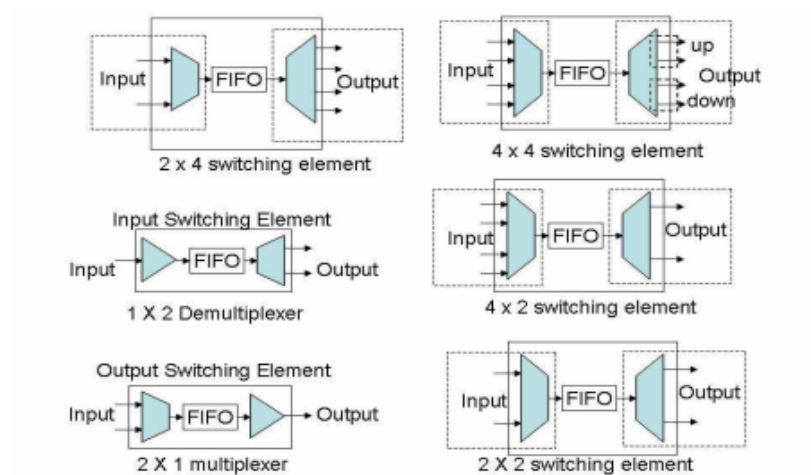Figure 2. Conceptual Framework

Figure 3. Shared buffer switching elements (SE)

the conceptual framework of the study.

The first stage was an array of 2x4 Switching Element (SE), succeeding stages were arrays of 4x4 SEs, and the last stage, an array 4x2 SE. A 1x2 demultiplexer was installed in each input port of the network that functions as source and a 2x1 multiplexer was placed in each output port that acted as sink. Each SE had a buffer of size m to store packets. All SEs in each stage were provided with a simulated counter to monitor the incoming and outgoing of packets.

Basically, routing of the Tagle-Sharma switch still followed the algorithm of the bufferless design as presented in Figure 4.

There was a slight change with the incorporation of buffer between the input and output terminals. When the packet enters the input terminal, the packet is routed to the buffer before it is sent to the desired output terminal of the switching element (SE) depending on its destination address.

2.1 Buffer Handling Techniques: FIFO Configuration

The FIFO configuration was simple and straightforward. Packets entering each SE store them immediately in the buffer following the FIFO policy. Back pressure

```
Input:    Packet
Output: Packet's next destination terminal/ node

Begin
    RT = destination = d_{n-1} d_{n-2}, … d_0
    for i = 1 to n-1
        case d_i  // examine bit d_i at switch [p, s, sn]
                0: choose upper output terminal of SE
                1: choose lower output terminal of SE
        end case
        if switch [p, s+1, next sn] is Faulty then
                send packet to [!p, s+1, next sn]
        else
                send packet to [p, s+1, next sn]
end
```

Figure 4. Routing algorithm of the Tagle-Sharma switch

the system by checking if the buffer of the next SE is available (not full); if it is, the packet at the head of the buffer is sent, otherwise, it will stay for a certain period of time. A deadline time of the packet at the head of the buffer was set in an attempt to reduce delay in transmission. As soon as the deadline expires, the packet has to be sent whether or not the buffer of the next SE is available.

2.2 Buffer Handling Techniques: Look-ahead Configuration

The Tagle-Sharma buffered switch with look-ahead capability only used the buffer when needed. Look-ahead configuration also employed the same FIFO policy and back pressure flow control mechanism. However, before the packets were stored in the buffer, the system checked the following parameters: (1) buffer is empty, (2) the output terminal of that SE is available, and (3) the buffer of the next SE is available or not full. If these conditions are met, the packet will bypass the buffer, otherwise, it will follow the FIFO configuration. Checking that the buffer in the SE is empty will avoid out-of-sequence packets; while checking for availability of output terminal in the SE will prevent blocking of packet; and checking the availability of buffer in the next SE avoids overflow in its buffer.

2.3 Simulation

The simulators were designed to model the properties of Tagle-Sharma switch. Simulation was performed under uniform random traffic. The simulator was packet based. Packets were generated at the sources, and were independent of other packets generated at previous cycles and other terminals. Each packet was provided with a tag that determined its destination

and was used to route it. The tag was modeled using binary elements of $n = \log_2 N$. Each place value of the binary number represented a network stage: the first stage corresponded to the most significant bit until the least significant bit which corresponded to the last stage. A new set of packet will be generated at the source and sent to the first stage if there is available space in the buffer; otherwise, the packet is dropped. Any packet not accepted is dropped. If two or more packets are destined to the same output or resource, one is chosen at random. A backpressure mechanism was applied to ensure that no packet is lost inside the SE. This means that a packet can leave the SE if the next stage SE is available. The sink (SE2x1) can always accept a packet. A simulated clock was placed to measure instantaneous throughput and transmission delay. Throughput was measured by counting the successful packets, that is, packets that arrive at the sink or its destination. Transmission delay or latency was measured by counting the number of hops, meaning, counting the number of nodes the packet was routed before it arrived to its destination. Sizes of buffer were varied from 4 packets up to 1024 packets capacity. The study also dealt with network sizes (NxN) which range from 8x8 up to 256x256, to look into the scalability aspect of the buffered approach. Computer simulations were conducted under uniform random traffic to see how the parameters like the buffer size and the network size affect the performance in terms of throughput and transmission delay. Packet losses were also monitored at every stage of the network to investigate the effect of bypassing the buffer when next destination terminal is available.

Simulations were done in three trials. The maximum throughput of the trials for each case was selected. Network sizes were

3.0 Results and Discussion

The Tagle-Sharma switch offered less than 1% packet loss at increasing load and buffer sizes. This was shown in figure 5(a) and 5(b). The lower bound switch gives higher packet loss than the upper bound network for both small and large buffers. Based on the same figures, it can be deduced that packet loss increases as the applied load increases.
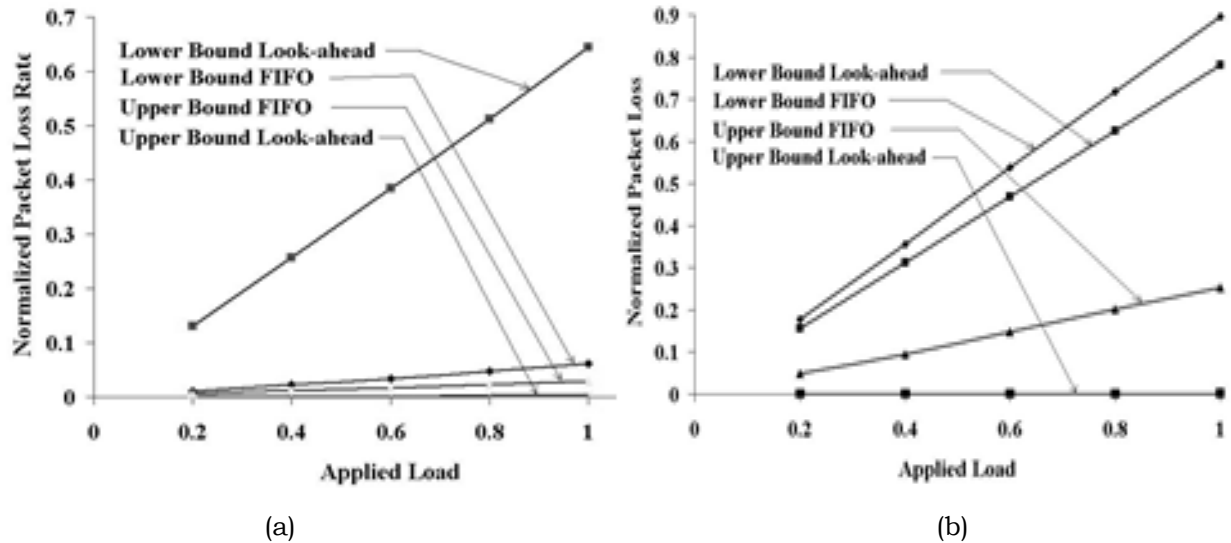

(a)

(b)

Figure 5. Applied load vs. packet loss at (a) m=8 and (b) m=2048

Likewise, the shared buffer approach of the Tagle-Sharma network offered 99% level throughput even for the lower bound FIFO model at deadline time less than 8 ms as shown in Figure 6. The upper bound is not affected by the deadline but the lower bound reaches its maximum throughput at time between 6 to 8 ms for both buffer handling techniques. This is because aside from the buffer, upper bound is fault tolerant at every stage of the network which enhances its throughput performance.
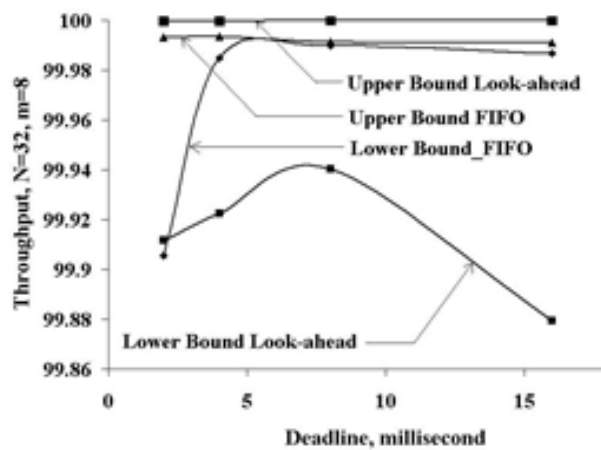

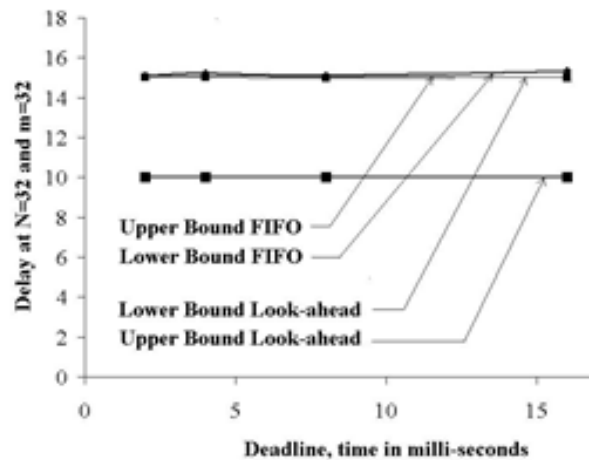Figure 6. Throughput VS. Deadline Time at N=32, m=8

Figure 7. Delay VS. Deadline Time at N=32, m=32

It was observed that the look-ahead scheme provided better delay performance than the FIFO configuration. As shown in figure 7, there was no effect on the transmission delay of the packets for the look-ahead scheme but a slight increase was seen on the FIFO configuration beyond 10 ms. This is because the look-ahead scheme used the buffer only when needed (can bypass the buffer), thus decreasing the transmission delay. It can be inferred that the optimum deadline time to achieve high performance in the Tagle-Sharma switch was 8 ms.

The throughput of the upper bound switch employing the look-ahead scheme is not affected by the increase of buffer size as shown in figure 8. However, a decrease in the thr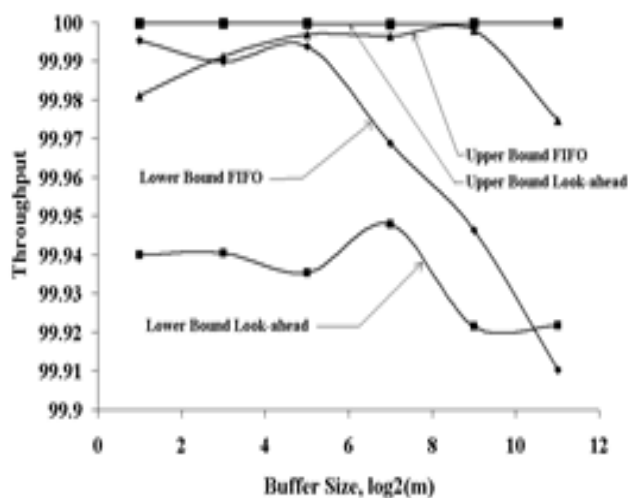oughput is observed in the lower bound network using the FIFO configuration for a certain level of buffer size. Its throughput is not affected for small buffer sizes, but starts to decrease when the buffer becomes larger (buffer size is greater than 32). The upper bound using FIFO configuration started to decrease its throughput when the buffer size is already large (greater than 128). The transmission delay of the network using the look-ahead configuration is not affected by the increase in buffer size. The FIFO scheme exhibits a decline of the delay performance when the buffer size is larger. The upper bound showed a decline of the delay performance when the buffer is large. Based on the results as shown in the figures, increasing the buffer size exhibits a diminishing benefit.
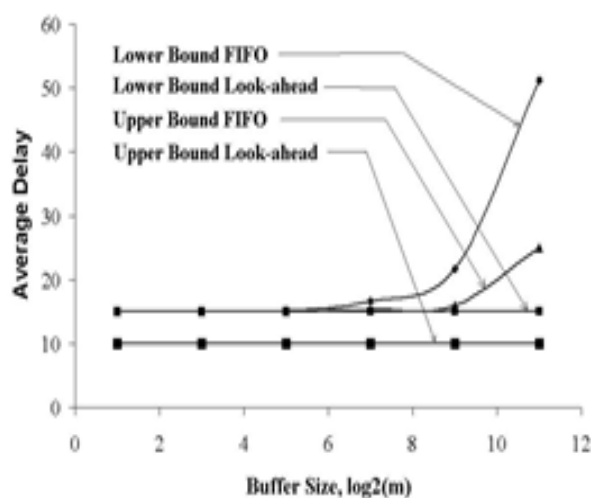

Figure 8. Throughput VS. Buffer size at N=32

Figure 9. Delay vs. Buffer Size at N=8

The drawback of look-ahead configuration is that when the packet bypasses the buffer of the current SE, back pressure mechanism is not applied and the packet reaches its destination faster, thus could cause an overflow of buffer in the next stage of the network. This is shown in figure 10, which illustrates the packet loss in every stage of the network. In the upper bound look-ahead configuration, packet loss is high at the first stage then dramatically drops in the second stage and increases again in the third stage and drops again in the fourth stage and so on. Since the packet has two nodes to go, there is higher possibility that packets will bypass the buffer. The effect is that packets arrive in its destination faster than packet that is routed in the buffer. As a result, buffer in the next stage would easily get full and overflow would likely to occur.

The lower bound look-ahead starts with a lower packet loss but depicted a dramatic increase in the packet loss, more than the packet losses of the FIFO scheme, then starts to decrease its packet loss in the succeeding stages. This is because the lower bound is fault tolerant only at the first stage of the network. When packet has only one node to go, there is a high probability that it can detect collision thus buffer is always utilized and its working in FIFO configuration, which demonstrates a decreasing packet loss in its succeeding stages.
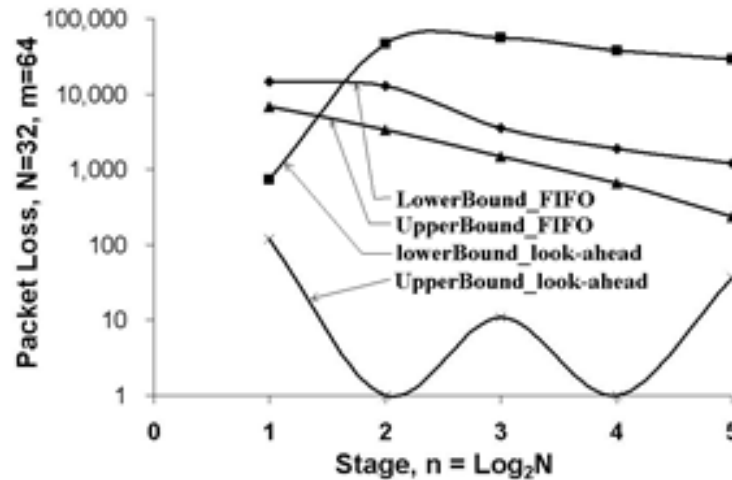

Figure 10. Packet loss at every stage of the Tagle-Sharma network for N =32 and buffer size = 64.

4.0 Conclusion

Two buffer handling techniques for shared buffer internal buffering strategy in the Tagle-Sharma switch were proposed. The FIFO scheme employed back pressure packet flow control mechanism and followed FIFO policy. The look-ahead scheme applied the features of the FIFO scheme but used the buffer only when needed. One drawback of the look-ahead strategy was that when the packet bypasses the buffer of the current SE, back pressure mechanism is not applied and the packet reaches its destination faster, thus could cause an overflow of buffer in the next stage of the network. Both the upper and lower bounds of the Tagle-Sharma network were evaluated. The switch is fault tolerant at every stage of the network, hence the deadline times did not have significant effect on the Tagle-Sharma network in terms of throughput and transmission delay. Simulations under uniform traffic showed that the Tagle-Sharma network using both schemes achieved high throughput and minimal transmission delay. Furthermore, higher throughput can be achieved and practical shared buffer switches can be built by using small buffer sizes for a fault tolerant switch such as the Tagle-Sharma switching network.

Transmission delay increases as network size increases. Networks using the FIFO scheme demonstrated higher transmission delay than the look-ahead scheme especially for small network sizes (n < 5). Increasing the buffer size of the networks using the look-ahead configuration does not give significant benefit to the network. The Tagle-Sharma network needed only a buffer size of 4 to achieve higher performance.

References:

Blake, J. and Trivedi, K.S. (1989). Reliability analysis of interconnected networks using hierarchical composition. *IEEE transactions on reliability, 38,* (*1*).

Boppana, R. and Raghavedra, C. (1999) *Designing efficient benes and banyan based input-buffered atm switches.* ICC, Vancouver, B.C., Canada.

Hluchyj, M., and Karol, M. (1988). Queuing in high-performance packet switching. *IEEE J. Select Areas Communication (6), 1587–1597.*

Kumar, V., Grama, A. Gupta, A. and Karypis, G. (1994). *Introduction to parallel computing: Design and analysis of algorithm.* The Benjamin Cummings Publishing Company, Inc., 24-30.

Lagman, C. and Tagle, P. (2002). Aggressive *transmission in the interconnected parallel banyan network.* IEEE Computer Society – ISPAN. 02. 2002.

Mun, Y. and Youn, H.Y. (1994). Performance analysis of finite buffering multistage interconnection networks. *IEEE Transactions on Computers 43(2), 153-162.*

Patterson, D. and Hanness, J. (1996). *Computer architecture: A quantitative approach.* Morgan Kauffmann Publishers, Inc., San Francisco California, 2nd Edition. 521-522.

Sharmaon, S. and Viniotis, Y. (1995). *Optimal buffer management policies for shared-buffer atm switches.* .IEEE Transactions, Networking 7, 4 (Aug. 1999), 575–587.

Zhou, I. and Atiquzzaman, M.(2002). A performance comparison of four buffering schemes for multistage interconnection networks. *International Journal of Parallel and Distributed Systems and Networks, 5 (1).*