Comparative Study of blocking mechanisms for Packet Switched Omega Networks

D. C. VASILIADIS
University of Peloponnese
Department of Computer
Science and Technology
GR-221 00 Tripolis
GREECE

G. E. RIZOS
University of Peloponnese
Department of Computer
Science and Technology
GR-221 00 Tripolis
GREECE

S. V. MARGARITI
University of Ioannina
Computer
Science Department
GR-451 10 Ioannina
GREECE

L. E. TSIANTIS
Hellenic Open University
Faculty of
Science and Technology
GR-262 22 Patra
GREECE

Abstract: Omega Networks are a famous subclass of blocking Multistage Interconnection Networks (MINs). They have been recently identified as an efficient interconnection network for a switching fabric of communication structures such as gigabit ethernet switch, terabit router, and ATM switch. Interconnection network performance is also a key factor when constructing multiprocessor systems. In this paper we are interested in studying the influence of the blocking mechanisms on the main performance parameters of a typical 8x8 Omega Network with finite buffer size queues. We investigate the packet loss probability , the throughput and the latency of an Omega Network using both the Back-pressure and the Block-and-lost blocking Models respectively. This study can be used in future in order to analyse the performance of an actual MIN , where lost packets are resubmitted in the MIN.

Key–Words: Multistage Interconnection Networks, Omega Networks, Banyan Switches, Performance Analysis.

1 Introduction

Multistage Interconnection Networks (MINs) with the Banyan property as Omega Networks are proposed to connect a large number of processors to establish a multiprocessor system. They are also used as interconnection networks in Gigabit Ethernet and ATM switches. Non-Banyan MINs, are in general , more expensive than Banyan networks and more complex to control. There is a great interest about Switching Systems and especially for self-routing systems called Banyan Switches. Thus, MINs have received considerable interest in the development of networks. The main parameter is their low cost, taking into account the performance they offer. The performance of MINs has a direct effect in the overal performance of communication between networks. It can be investigated either by time-consuming simulations or approximated by mathematical models.

During the last several decades, a number of studies and approaches have been publihsed. Hsiao [3] and Theimer [10] studied MINs with uniform load traffic on inputs. Merchant [7] and Zhou [16] used Markov chains in order to approximate the behavior of them under different buffering schemes. Hot spot traffic performance was also examined by Jurczyk [4] and Turner [11] deals with multicast in Clos networks as a subclass of MINs. Yang [15] approximated their behavior using mathematical methods. Group communication in circuit switched MINs was investigated by applying Markov chains as a modeling technique. Moreover, Merchant calculated the throughput of finite and infinite buffered MINs under uniform and non uniform traffic. Atiquzzaman [1, 2] focused only on non uniform arriving traffic schemes. Kleinrock [6] discusses approaches that examine the case of Poisson traffic on inputs of a MIN. Rehrmann [9] , makes an analysis of communication throughput of single-buffered multistage interconnection networks consisting of $2X2$ switches with maximum arrivals of packets , using relaxed blocking model. Furthermore, there are studies that deal with self-similar traffic on inputs.

In this paper , our analysis focused on the performance of Omega Networks. They were defined first by Lawrie [5] and they are a subset of the Delta Networks family proposed by Patel [8] , which is a bit-controlled interconnection networks family. The Omega differs from the Delta Network in the pattern of interconnections between the stages. Omega Networks use the ”perfect shuffle” algorithm for destination routing. In our study , we evaluate their performance under different size internal queuing configurations using the Back-pressure and the Block-and-lost blocking Models respectively. Based on the blocking mechanism , when a packet meets the next buffer po-

sition occupied then it cannot be routed and is thus either blocked (Back-pressure Model) or lost (Block-and-lost Model). At first we present and analyze a typical 8x8 Omega Network. Then, we explain the performance criteria and parameters of this. Finally we present the results of our simulation experiments and provide the concluding remarks.

2 Analysis of an Omega Network

A MIN can be defined as a network used to interconnect a group of N inputs to a group of M outputs using several stages of small size Switching Elements (SEs) followed (or leaded) by link states. It is usually defined by , among others , its topology , routing algorithm , switching strategy and flow control mechanism. They are typical multistage self-routing switching fabrics. That means, every SE that accepts a packet in one of its input ports can decide in which of its output ports to forward this packet depending only on the destination address of it. Thus , every SE of stage k can decide in which output port to send it based on the k^{th} bit of the destination address. If this bit is 0, then the packet is forwarded to the upper output port and if the bit is 1 packet is forwarded to the lower output port. They are also characterized by the property that there is exactly one path from any input to any output. An Omega MIN of dimension NXN is constructed by $n = log_c N$ stages of cxc SEs where c is the degree of them. A SE consists of c input ports and c output ports and there are exactly $\frac{N}{c}$ SEs at each stage. So, the total number of SEs of a MIN is $\frac{N}{c} log_c N$ and there are $N log_c N$ interconnections among all stages, as opposed to the crossbar network which requires $O(N^2)$ SEs and links. In this paper we study a typical 8X8 Omega MIN that consists of $2X2$ SEs. A configuration of a typical 8X8 Omega MIN with finite buffer size queues is shown below. It is assumed to operate under the following conditions :

- The network clock cycle consists of two phases. In the first face flow control information passes through the network from the last stage to the first stage. In the second phase packets flow from one stage to the next in accordance with the flow control information.

- A typical 8X8 Omega Network consists of 2X2 SEs (2-input ports and 2 output ports). It has the ability to accept packets in every input port and send packets randomly to one of two output ports. It also operates with switching packets and static routing. The messages are transferred as simple cells and the route is based on

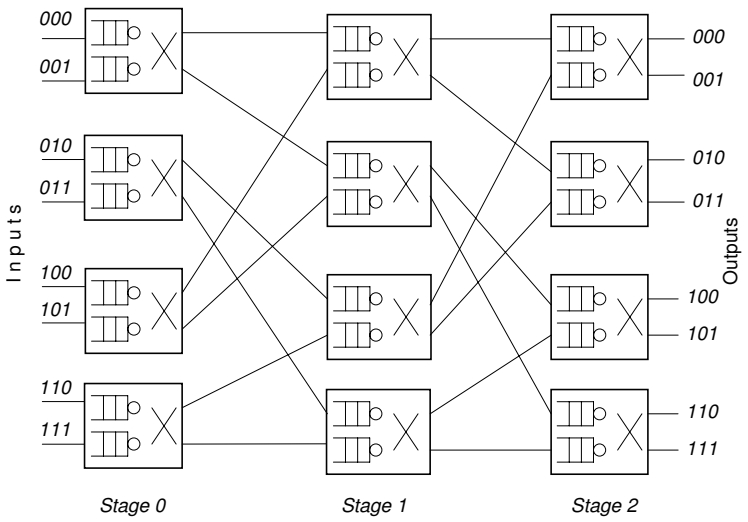

Figure 1: An $8X8$ Omega Network

k^{th} bit of the destination address for each stage ($k = 0, 1, 2$) of the MIN.

- The arrival process of each input of the network is a simple Bernoulli process , i.e. , the probability that a packet arrives within a clock cycle is constant and the arrivals are independent of each other.

- Each SE has the ability to send only one packet to the next stage queues in a time cycle. The packets are uniformly distributed across all the destinations and each queue uses a FIFO policy for all output ports.

- When two packets in the k^{th} stage contend for the same buffer in the $k + 1^{th}$ stage and there is not adequate free space for both of them to be stored, there is a conflict. In this case, one of them will be accepted at random and the other will be rejected by means of upstream control signals. When a MIN uses the mechanism of the Back-pressure blocking the rejected packet stays in a buffer of k^{th} stage and would have to try to make request again in the next time slot. On the other hand , when a MIN uses the mechanism of Block-and-lost blocking the rejected packet is discarded.

- Finally, all packets in input ports contain both the data to be transferred and the routing tag. In order to achieve synchronously operating SEs , the MIN is internally clocked. As soon as packets reach a destination port they are removed from

the MIN. Thus, packets cannot be blocked at the last stage.

3 Performance evaluation methodology and parameters

In order to evaluate the performance of an NXN MIN with $n = \log_c N$ intermediate stages of cxc SEs , we use the following metrics. Let a relatively large time divided into υ discrete time intervals $(1\tau, 2\tau, \ldots, \upsilon\tau)$.

- *packet loss probability (p_ℓ)* is the probability of lost packets at any stage of the MIN. Implementations based on the Back-pressure Model have lost packets only on inputs. On the other hand , when a MIN uses the mechanism of Block-and-lost blocking Model there are also lost packets at all intermediate stages except last. Formally , p_ℓ can be expressed by

$$p_\ell = p_\ell(i) + \sum_{k=0}^{n-2} p_\ell(k) \tag{1}$$

  where $p_\ell(i)$ denotes the *packet loss probability* on inputs and $p_\ell(k)$ denotes the *packet loss probability* on a queue of the k^{th} intermediate stage of a n-stage MIN (there is no blocking on last stage $n-1$).

- *Average throughput* (T_{avg}) is the average number of packets accepted by destinations per network cycle. This metric is also referred as *bandwidth*. Formally , T_{avg} can be defined as

$$T_{avg} = \lim_{\upsilon\to\infty} \frac{\sum_{i=1}^{\upsilon} n(i)}{\upsilon} \tag{2}$$

  where $n(i)$ denotes the number of packets that reach their destinations during the i^{th} time interval τ.

- *Normalized throughput (T)* is the ratio of the *average trhoughput* T_{avg} to network size N

$$T = \frac{T_{avg}}{N} \tag{3}$$

- *Average latency* (L_{avg}) is the average number of network cycles a packet needs to arrive to its destination. Formally , L_{avg} can be expressed by

$$L_{avg} = \lim_{\upsilon\to\infty} \frac{\sum_{i=1}^{\upsilon_a} t(i)}{\upsilon_a} \tag{4}$$

  where υ_a denotes the total number of packets accepted by destinations into υ time intervals and $t(i)$ represents the total number of network cycles needs the i^{th} packet to arrive to its destination. We consider $t(i) = tw(i) + tr(i)$, where $tw(i)$ denotes the total number of network cycles for i^{th} packet waiting at any stage for the availability of an empty buffer at the next stage of the network. The second term $tr(i)$ denotes the total number of network cycles needs the i^{th} packet to be transmited to its destination without any queuing delay, that is just $n\tau$, where τ is the network cycle of a n-stage network.

- *Normalized latency (L)* is the ratio of the L_{avg} to the *minimum latency*, which is simply the total number of network cycles needs a packet to be transmited to its destination , that is just $n\tau$. Formally , L can be defined as

$$L = \frac{L_{avg}}{n\tau} \tag{5}$$

The following parameters affect the *loss probability* , the *throughput* and the *latency* of a MIN.

- *probability of arrivals (λ)* is the steady-state fixed probability of arriving packets at each queue on inputs. In our simulation λ is assumed to be $\lambda = 0.1, 0.2, \ldots, 0.9, 0.99$.

- *buffer size (β)* is the maximum number of packets that a buffer of a SE can hold. In our case β is assumed to be $\beta = 0, 2, 8$. At first we analyzed the behavior of an unbuffered SE. Then we chose to investigate SEs with *buffer size* $\beta = 2$, that is sufficient for medium traffic loads [14]. Finally, we considered setups with *buffer size* $\beta = 8$, since network *throughput* is better exploited , while *latency* can be tolerated.

4 Mathematical Model for the loss probability on inputs

The *packet loss probability* $p_\ell(i)$ on inputs of an unbuffered Omega MIN using the mechanism of Block-and-lost blocking Model is given by binomial distribution. Let a be the random variable denoting the count of arrivals of packets at the end of a network cycle on a queue of a cxc SE at the first stage of the MIN.

$$x_{c,a} = \begin{cases} \binom{c}{a} \left(\frac{\lambda}{c}\right)^a \left(1 - \frac{\lambda}{c}\right)^{c-a} & \text{for } 0 <= a <= c \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

When two packets arrive on the same unbuffered SE of a MIN at the end of a network cycle then one gains the buffer and the other is discarded. Consequently , $p_\ell(i)$ can be defined as

$$p_\ell(i) = x_{2,2} = \lambda^2/4 \tag{7}$$

where λ is the steady-state fixed *probability of arrivals* of packets on inputs. In this case , validity of the mathematical model for the *packet loss probability* $p_\ell(i)$ was tested by comparing the results of the above equation (7) with those of simulation experiments which were found to be in close aggreement (within 1%).

5 Simulation and performance results

The performance of MINs is usually determined by modeling, using simulation [13] or mathematical methods [12].

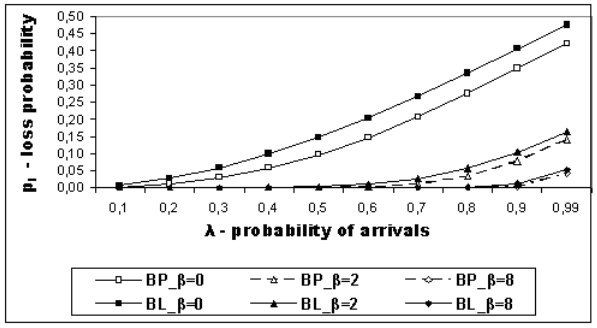

Figure 2: Loss probability

In our study we implemented two simulators for an 8X8 Omega Network using internal FIFO queuing configurations with non-shared buffers : one with the mechanism of Back-pressure blocking and the other with the mechanism of Block-and-lost blocking Model.

We performed extensive simulations to validate our results. Especially , in the case of unbuffered MINs using the mechanism of Block-and-lost blocking , the results of *packet loss probabilities* on inputs were also validated by equation (7). The number of simulation runs was adjusted to ensure a stady-state operating condition for the MIN. In this paper, we present the results of simulation experiments in accordance with the parameters which were defined in section 3.

Figure 2 presents the *packet loss probability* as defined by equation (1). We notice here that the *packet loss probability* of a MIN using the mechanism of the Back-pressure blocking Model is lower than using the mechanism of the Block-and-lost Model.

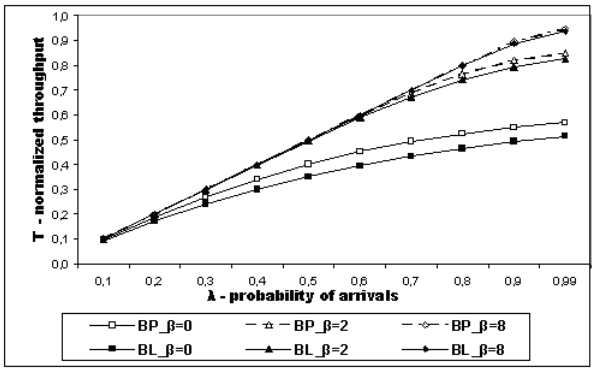

Figure 3: Normalized throughput

Figure 3 illustrates the variation of *normalized throughput* as defined by equation (3). We notice here that *normalized throughput* of a MIN using the mechanism of the Back-pressure blocking Model is higher than using the mechanism of the Block-and-lost Model. We also notice that *throughput* is not satisfactory for $(\beta = 0)$, for both configuration setups.

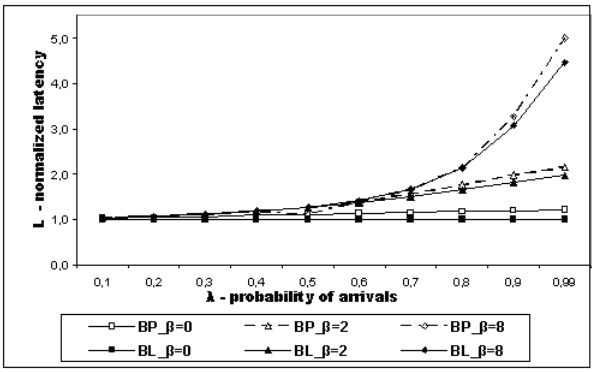

Figure 4: Normalized latency

Figure 4 presents the *normalized latency* as defined by equation (5). We notice here that *normalized latency* of a MIN using the mechanism of the Block-and-lost blocking Model is lower than using the Back-pressure Model. We also notice that *normalized latency* is very high for $(\beta = 8)$, for both configuration setups.

6 Conclusion

In this paper , we used simulations for Omega Networks in order to study the effect of blocking mechanisms on their performance. The simulations yielded performance measures such as *packet loss probabilities* , *throughput* and *latency* for Omega Networks.

We noticed that , all values of performance parameters were increasing as the values of *arrivals of packets* on inputs were increasing. In our study , we analyzed two major groups of performance parameters : parameters to be maximized as *throughput* and parameters to be minimized as *packet loss probabilities* and *latency* of MINs. In the case of *packet loss probabilities* and *throughput* the mechanism of Back-pressure blocking Model seems to be better than the mechanism of Block-and-lost blocking Model. On the other hand , the mechanism of the Block-and-lost blocking Model produces lower (better) values for *latency*. We also noticed that the use of buffers in front of SEs decreases the values of *packet loss probabilities* and increases the *throughput* of MINs. On the other hand, large-sized buffer configurations produce higher values for *latency*, especially under very heavy traffic.

Our contribution in analysis of a typical Omega Network using different blocking mechanisms can be used in analysis of several types of networks in order to study the performance of transport packets from network to network via multistage switches. This study can also be used in future in order to analyse the performance of an actual MIN , where lost packets are resubmitted in the MIN.

References:

[1] M. Atiquzzaman and M.S. Akhatar, Effect of Non-Uniform Traffic on the Performance of Multistage Interconnection Networks, *9th International Conference on System Engineering*, Las Vegas, pp. 31-35, July 1993.

[2] M. Atiquzzaman and M.S. Akhatar, Efficient of Non-Uniform Traffic on Performance of Unbuffered Multistage Interconnection Networks, *IEE Proceedings* Part-E, 1994.

[3] S.H. Hsiao and R. Y. Chen, Performance Analysis of Single-Buffered Multistage Interconnection Networks, *3rd IEEE Symposium on Parallel and Distributed Processing*, pp. 864-867, December 1-5, 1991.

[4] M. Jurczyk. Performance Comparison of Wormhole-Routing Priority Switch Architectures. *In Proceedings of the International Conference on Parallel and Distributed Processing Techniques and Applications* 2001 (PDPTA'01); Las Vegas, pp. 1834..1840, 2001.

[5] D. A. Lawrie. Access and alignment of data in an array processor. *IEEE Transactions in Computers*, C-24(12): pp. 1145..155, Dec. 1975.

[6] T. Lin, L. Kleinrock, Performance Analysis of Finite-Buffered Multistage Interconnection Networks with a General Traffic Pattern, Joint International Conference on Measurement and Modeling of Computer Systems, *Proceedings of the 1991 ACM SIGMETRICS conference on Measurement and modeling of computer systems*, San Diego, California, United States, pp. 68..78, 1991.

[7] A. Merchant, A Markov chain approximation for analysis of Banyan networks, in Proc. *ACM Sigmetrics Conference* On Measurement and Modelling of Computer systems, 1991.

[8] J.H. Patel. Processor-memory interconnections for mutliprocessors. *In Proceedings of 6th Annual Symposium on Computer Architecture* New York, pp. 168..177, 1979.

[9] R. Rehrmann, B. Monien, R. Luling, R. Diemann, On the communication throughput of buffered multistage interconnection networks, *in ACM SPAA* 96 pp. 152..161, 1996.

[10] T.H. Theimer, E.P. Rathgeb, and M.N. Huber, Performance Analysis of Buffered Banyan Networks, *IEEE Transactions on Communications*, vol. 39, no. 2, pp. 269..277, February 1991.

[11] J. Turner, R. Melen. Multirate Clos Networks. *IEEE Communications Magazine*, 41, no. 10: pp. 38..44, 2003

[12] D. Tutsch, G. Hommel. Generating Systems of Equations for Performance Evaluation of Buffered Multistage Interconnection Networks. *Journal of Parallel and Distributed Computing*, 62, no. 2: pp. 228..240, 2002.

[13] D. Tutsch, M. Brenner. A Multistage Interconnection Network Simulator. In 17th European Simulation Multiconference: *Foundations for Successful Modelling & Simulation (ESM'03);* Nottingham, SCS, pp. 211..216, 2003.

[14] D.C. Vasiliadis, G.E. Rizos Simulation for Multistage Interconnection Networks using relaxed blocking model. *Proceedings of the ICCMSE 2006 conference*, Greece, vol. 7, pp. 570..575, 2006.

[15] Y. Yang, J. Wang. A Class of Multistage Conference Switching Networks for Group Communication. *IEEE Transactions on Parallel and Distributed Systems*, 15, no. 3: pp. 228..243, 2004.

[16] B. Zhou, M. Atiquzzaman. A Performance Comparison of Four Buffering Schemes for Multistage Interconnection Networks. *International Journal of Parallel and Distributed Systems and Networks*, 5, no. 1: pp. 17..25, 2002.