Performance Analysis of the Bidirectional Ring-Based Multiprocessor*

Hitoshi Oi and N. Ranganathan
Department of Computer Science and Engineering,
University of South Florida
Tampa, FL 33620
Email: {oi, ranganat}@csee.usf.edu

Abstract

In this paper, we propose and evaluate a bidirectional slotted ring architecture for the distributed shared memory multiprocessor model. The performance of the bidirectional ring is compared against the unidirectional ring model using analytical models. The effects of workload parameters such as miss rate, communication locality and system size on processor utilization and effective computation power are analyzed. Our study shows that bidirectional slotted rings provide about 15 to 35% better processor utilization than unidirectional rings with the same link width.

Keywords: Interconnection networks, shared memory multiprocessors, performance evaluation, slotted ring.

1 Introduction

Distributed shared memory (DSM) multiprocessors provide a view of a single globally shared memory with memory units that are physically distributed among the processing elements (PE). The interconnection networks that have been used for DSM multiprocessors include the 2-D Mesh (Stanford FLASH [1]) and the fat-tree (TMC CM-5 [2]). Ring networks have the advantages of (1) fixed node degree (modular expandability), (2) simple network interface structure (fast operation speed) and (3) low wiring complexity (fast transmission speed).

There are several ring-based multiprocessors that have been proposed and/or actually been built. Holiday and Stumm explored the design space of hierarchical ring multiprocessors with parametric simulations [3]. Barroso and Dubois evaluated the performance of the small to medium scale slotted ring multiprocessors with a hybrid approach (combining trace-driven simulations and analytical models) [4]. KSR-1 of Kendall Square Research was a cache-only memory architecture (COMA) multiprocessor with hierarchical rings [5].

Most research in the past including the above are based on unidirectional rings. In a unidirectional ring, the messages have to traverse all the way through the ring even if the destination is within the local neighborhood. The length of traversal can be significantly reduced by using bidirectional rings. In this paper, we analyze the performance of bidirectional slotted rings against unidirectional rings keeping the width of the links constant. The performance improvement is quantitatively studied using analytical models.

2 Architecture Model

The architecture of the ring-based multiprocessor and its node are shown in Figure 1. Each PE consists of a processor, a cache, a memory unit and a network interface. The number of data lines (m in the figure) that connect neighboring PE's are assumed to be the same in the unidirectional and bidirectional rings. Thus, a packet on the unidirectional ring is equivalent to two packets on the bidirectional ring. Header messages (read request and invalidation) and data messages are one and eight packets long on the unidirectional ring respectively (doubled on the bidirectional ring). When a message consists of multiple packets, each packet is transmitted independently.

It is assumed that the architecture is based on the CC-NUMA model with directory based cache coherency protocol. The local memory within a PE is a part of the globally shared memory and is associated with directory entries corresponding to the part of global address space assigned to the PE. If a cache miss corresponds to a non-local memory address, it

*The full length version of this paper is available at `http://figaro.csee.usf.edu/~oi/RESEARCH`
This research is supported in part by a National Science Foundation Grant No. CDA-9522265.

will be first sent to the home node of the accessed data, and the home node responds with a data message unless the requested data is dirty. If the requested data is dirty, the read request is forwarded to the PE that owns a modified copy of the requested data in the cache, and the owner PE responds with a message containing the data. In the bidirectional ring, the network interface selects the shorter path to the destination of the message.

We use the following parameters to define the configuration of the system: The number of processors, N, is assumed to represent small to medium configurations. In this study we assume $N = 32$ as in the case of KSR-1 and also the cases for $N = 16$ and 64. The latency corresponding to the transfer of a packet to the adjacent node is represented by t_r and is assumed to be 2 processor clock cycles. Each memory access t_m corresponds to 30 processor clock cycles. The time for cache protocol handling, t_c, is assumed to be equivalent to 2 processor clock cycles. A cache hit corresponds to one processor cycle. The processor is assumed to block on a cache miss.

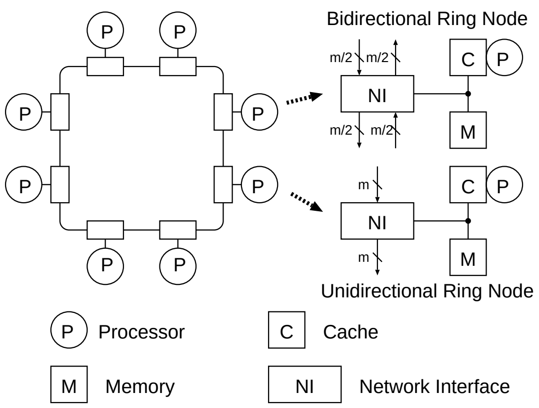

Figure 1: Ring-Based Multiprocessor Model

3 Analytical Model

In this section, we derive analytical models of both unidirectional and bidirectional slotted ring architectures for performance evaluations. We use the following parameters to represent the characteristic of the workload. G: group size, MR: miss rate, P_L: probability of local access and P_G probability of group access.

Since we assume the processor blocks on a miss, the cycle of computation and stall-on-miss is given by

$$1/MF = 1/MR + L \tag{1}$$

where MF is the miss frequency (miss per clock cycle) and L is the miss penalty. The processor performs either computation or data access from the cache for the period of $1/MR$ on the average, and then it stalls for the period L due to a cache miss.

Next, we break down L in to four components: The *transmission delay* T is the time for a PE to transmit a message (read request, data, or invalidation). To calculate T, we use approximated $M/G/1$ model by Bhuyan et al in [6] with slight modification to suit to our problem. The *propagation delay* P is the time for a message to reach the destination PE and is determined by number of hops. P is obtained from our model of communication locality and will be described below. The *memory access latency* M is the memory access time t_m defined in the previous section plus the average waiting time in the memory access queue. M is also calculated using a simple queue model. The *protocol handling latency* C is equal to t_c defined in the previous section.

The classification of non-local accesses is given in Figure 2. A node in the figure represents a class of access mode while the label on an edge represents the probability of the class. We assume a miss is either read or invalidation with probabilities $1 - P_W$ and P_W respectively. A write miss (a write access to a block not existing in the cache) is considered to generate both read and invalidation. The classification of the locality of data access is as follows: (i) *local*: accesses to local memory within the PE (ii) *group*: accesses to PE's within the distance G (within G hops) and (iii) *global*: all other accesses. It is assumed that the accesses within the same class are uniformly distributed. We used the term effective communication distance (ECD) to represent the average message traversal length, which is the sum of the leaf nodes in Figure 2 weighted by the probabilities on the edges from the root. By definition, the average of the transmission delay P appearing in all the access mode is equal to ECD.

From the equation (1), we get processor utilization $u = 1 - MFL$ and miss rate $MR = MF/(1 - MFL)$. We use the ratio of processor utilizations of the bidirectional and the unidirectional ring (PUR) for performance comparison.

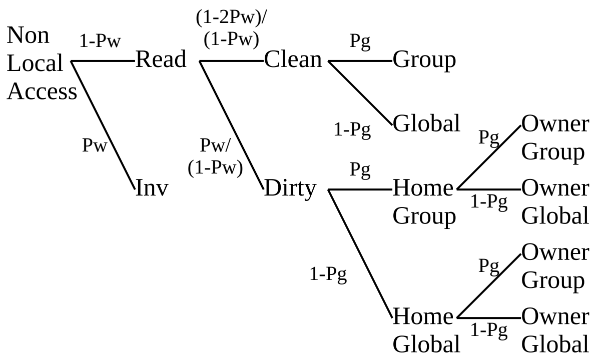

Figure 2: Classification of Non-local Access

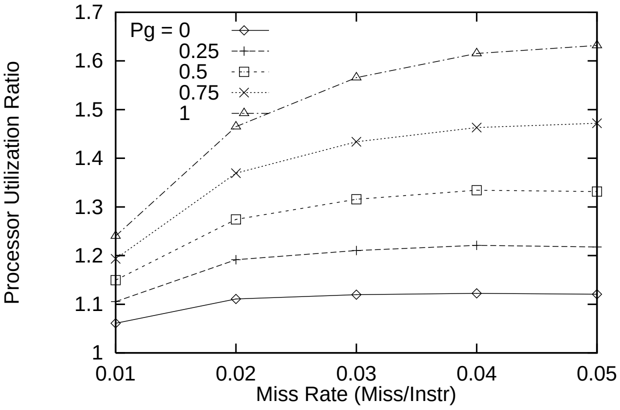

Figure 3: Effect of Group Access Probability

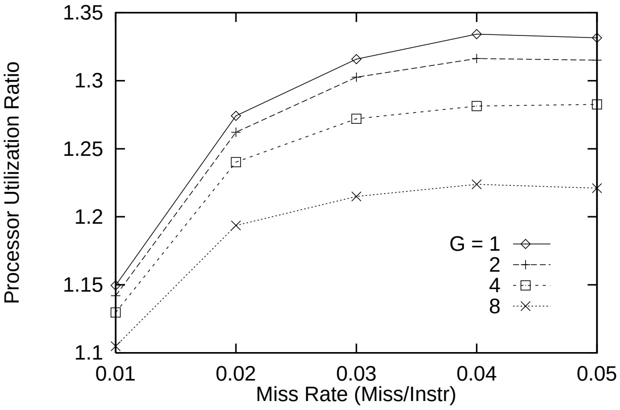

Figure 4: Effect of Group Size

4 Performance Comparison

In this section, we compare the performance of unidirectional and bidirectional ring-based multiprocessors using the analytical model derived in the previous section. We use the number of processors $N = 32$, and the write probability $P_W = 0.3$ as in previous studies including [3] unless otherwise stated.

4.1 Effect of Group Access Probability and Group Size

The ECD of the bidirectional ring is affected by many factors, such as the partitionability of the problem and the data placement policy. In this section, we investigate the effect of two parameters that represent communication locality, namely, the group access probability (P_G) and the group size (G).

Figures 3 and 4 show the effect of group access probability and group size on PUR. It is assumed that $P_L = 0.5$. In Figure 3, the PUR for different values of P_G (0, 0.25, 0.5, 0.75 and 1) for $G = 1$ are plotted. In Figure 4, the plots for $P_G = 0.5$ for different group sizes (G = 1, 2, 4 and 8) are given. It is observed that PUR looks to be more sensitive to P_G than G. Thus, it seems important to increase the amount of group-shared data than to decrease the number of PE's within the group in order to improve performance. Also note that even without preference to the group access ($P_G = 0$), the bidirectional ring performs by 6% to 12% better than the unidirectional ring (Figure 3).

4.2 Effect of System Size

Next, we consider the scalability of the ring-based multiprocessor. From the related study including [4], we use the following assumptions: (1) miss rate MR is a linearly increasing function of $\lg N$, (2) P_W is constant over the increase of N, and (3) P_G remains the same while G increases linearly with N. Although the choice of parameters are not based on actual programs, the parameters reflect the characteristics of typical applications in parallel processing to some degree.

We use four sets of parameters that have combinations of high and low values for P_W (0.3 and 0.1) and P_G (0.9 and 0.1) as shown in Table 1. The miss rate MR is assumed to be 0.03, 0.04 and 0.05 for N = 16, 32, and 64. The miss rate values are assumed to be relatively high in order to observe the effects of the ring interconnection network on the overall performance. Also, G = 1, 2 and 4 are used for N = 16, 32, and 64 respectively, for all four cases.

The effective computation power Nu (product of

Case	MR (%)	P_W	P_L	G	P_G
i	3/4/5	0.3	0.3	1/2/4	0.1
ii	3/4/5	0.3	0.3	1/2/4	0.9
iii	3/4/5	0.1	0.3	1/2/4	0.1
iv	3/4/5	0.1	0.3	1/2/4	0.9

Table 1: Workload Parameters for System Size

the number of processing element and the processor utilization) is plotted against N in Figure 5. It is not surprising that on the unidirectional Nu decreases against increasing N since this phenomena can be seen commonly in applications that are not well partitionable (e.g. MP3D). On the bidirectional ring, Nu is strongly affected by P_G. When the write probability is relatively high (case i and ii), Nu ranges from 3.8 to 5.7 for $N = 64$. The effect of P_G is yet stronger for lower P_W; Nu ranges from 4.1 to 6.9. In Figure 6, the PUR is plotted for cases i to iv. The bidirectional ring still achieves some level of speed up under the workload parameters that are not favorable to multiprocessors. Thus, the advantage of the bidirectional ring is more significant for larger N.

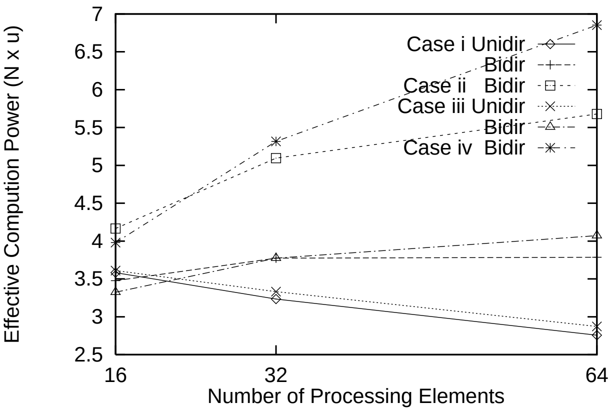

Figure 5: Effective Computation Power

5 Conclusions

In this paper, we presented a performance comparison of the unidirectional and the bidirectional slotted ring multiprocessors using an analytical model. The result of our study indicated that the bidirectional ring performs significantly better when the application programs exhibit high locality of access. Further investigations could include the extraction of workload parameters from applications and the analysis of dynamic system behavior using execution-driven simulations.

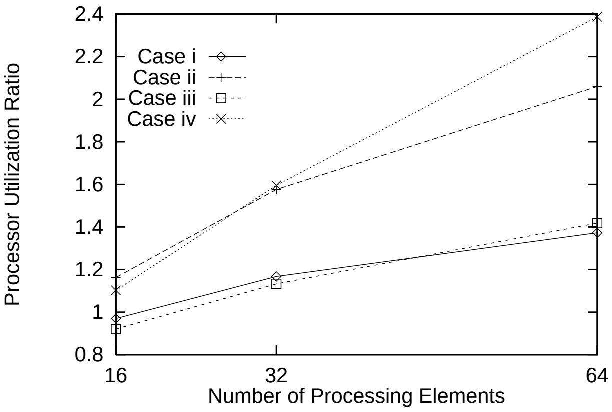

Figure 6: Effect of System Size on PUR

References

[1] Jeffrey Kuskin et al,, "The Stanford FLASH Multiprocessor", In *Proceedings of the 21st International Symposium on Computer Architecture*, 302–313, April 1994.

[2] C. Leiserson et al., "Network architecture of the Connection Machine CM-5", In *Proceedings of 4th Annual ACM Symposium on Parallel Algorithms and Architectures*, 272–285, June–July 1992.

[3] M. Holiday and M. Stumm, "Performance of Hierarchical Ring-Based Shared Memory Multiprocessors", *Transactions on Computers*, IEEE, Vol. 43, No. 1, 52–67, January 1994.

[4] L. A. Barroso and M. Dubois, "Performance Evaluation of the Slotted Ring Multiprocessor", *Transactions on Computers*, IEEE, Vol. 44, No. 7, 878–890, July 1995.

[5] Kendall Square Research Corporation, *Technical Summary*, 1992.

[6] L. Bhuyan, D. Ghosal and Q. Yang, "Approximate Analysis of Single and Multiple Ring Networks", *Transactions on Computers*, IEEE, Vol. 38, No. 7, 1027–1040, July 1989.