

R. 66
Sig 1203 Gon

1400008586
M-REPORT/362

**AN APPROXIMATE ANALYSIS OF
SYNCHRONOUS MULTIPLE BUS ARCHITECTURE
WITH DISTRIBUTED MEMORY**

Luis GONZALEZ
Emilio SANVICENTE

RT85/03

AN APPROXIMATE ANALYSIS OF A SYNCHRONOUS MULTIPLE
BUS ARCHITECTURE WITH DISTRIBUTED MEMORY.

Luis González (*)

Emilio Sanvicente (**)

(*) Facultad de Informática. Barcelona. Spain:
On leave from the Universidad Central de Venezuela.

(**) E.T.S.I. Telecomunicación. Barcelona. Spain.

AN APPROXIMATE ANALYSIS OF A SYNCHRONOUS MULTIPLE BUS
ARCHITECTURE WITH DISTRIBUTED MEMORY.

Abstract. This paper presents an approximate analytic model for evaluating the performance of a loosely coupled multiprocessor architecture whose memory, organized in modules, is shared by all the processors. Each memory module (M_i) is associated with a particular processor (P_i), and it may be accessed by this processor (local access) or by any other processor ($P_j, j \neq i$) through b multiple shared busses (external access). The performance indexes used in this paper are the memory bandwidth B and the delay to access a memory module. The system is evaluated for different values of p' ($p' \leq 1$), the probability of memory service requirement in each memory cycle; therefore, we allow internal processing. Also, we consider two kinds of system operation, depending upon the number of cycles needed to access the requested memory when the request is for an external module. These results are obtained rather easily by solving a set of algebraic equations that approximately describe the system in steady state, and that yield performance values extremely close to the simulation results.

Index terms: memory bandwidth, multiprocessors, interconnection networks, performance evaluation

I. INTRODUCTION

There is a growing interest in the field of multiprocessor architecture design due to the great number of computer applications (image processing, weather prediction, etc), that need a high instruction execution rate. Nowadays it is economically feasible to construct a concurrent system by interconnecting several memory modules, each of which can be accessed by all the processors in the system. In fact there already exist several such implementations [2,3], and others are being designed or proposed [4,5].

The structure of the network connecting the processors and shared memory is an important question to consider in the design of such systems. Many parameters have a bearing on this choice. Some of them are: reliability, cost, modularity, degradation, geometry and expandability. Several kinds of interconnection networks (IN) have been proposed for these systems. The crossbar [2] provides the largest potential performance because there is no degradation in the network. Unfortunately, it has a high cost which is prohibitive for a large number of processors. Consequently, other INs become attractive in terms of hardware cost, provided they do not significantly degrade the system performance [7,9].

Many different IN architectures are possible depending on the location of the shared memory and on the structure of the processing units. For these architectures, several approaches have been recently presented to evaluate its complexity and performance. Sanvicente,et.al.[1] presented exact and approximate models to evaluate the performance of a single bus multiprocessor system with distributed common memory; Marsan,et.al.[5] developed analytic models for the study of several single bus multiprocessor architectures; Marsan,et.al.[6] used Markovian models for analyzing the performance of multiprocessor system based on the availability of two global busses; Marsan,et.al.[7] developed an asynchronous model to analyze the performance of the multiple bus system; Valero,et.al.[10] used a synchronous model to study the multiple bus system; Bhuyan [12] used a probabilistic model to study two loosely coupled architectures based on a single shared bus; Hoener,et.al.[14] presented probabilistic analysis of bus contention in a single bus multiprocessor system; Willis [15] considered a simplified model of multiple bus systems, assuming no queueing for busy resources.

In this paper we study the loosely coupled architecture shown in fig.1, in which each processor (P_i) is associated with a particular memory module (M_i), that we will denote as local. Other memory modules ($M_j, j \neq i$) which we call external (with respect to the given

processor) are also available. A processor can access its own local memory, through a switch (Si) with arbitration and switching functions, and also the external modules through a multiple bus IN (MBIN). In this system we can choose two kind of priority policies to access a common memory, namely: it can be either granted to the local processor or to one of the requests coming from the external processors. Of the two, in this paper we choose the latter, since it seems reasonable to free common resources as soon as possible. Also, this fact has been validated by our simulations.

There are many different performance measures that can be used to evaluate such a system. Some measures that have been used are: the amount of processing work done per time unit, the average number of memory modules accessed per time unit and the average waiting time of a processor. All of these measures are related, and we use the effective memory bandwidth (B) and the mean access time as our performance indexes. In general, the B of such a multiprocessor depends on the load balance, the memory service time, the conflicts in the network, and the delay due to arbitration, propagation, address mapping, etc.

We evaluate the performance of this loosely coupled architecture considering several cases. To begin with, we assume that each processor issues a new request as soon as the previous one is serviced ($p'=1$). We, then, analyze the case where in each memory cycle, any

processor with no pending request asks for a new memory service with probability p' ($p' < 1$), and remains active with probability $q' = 1 - p'$. This is equivalent to saying that there is internal processing. When a processor (P_i) demands access to a memory module, it does so with probability q to its local memory (M_i) and $(1-q)/(n-1)$ to any other memory module ($M_j, j \neq i$). We also consider two modes of operation depending upon the number of cycles needed to access an external memory. Firstly we suppose that the time required to access any memory module does not depend upon whether the module is external or not (with respect to the processor). Later we assume that, when the referenced module is external, it takes 'k' extra memory cycles to have selected external memory request ready [3]. These extra cycles are needed to model the propagation time through the bus, the mapping between virtual and physical address, arbitration time, etc.

Section II describes the assumptions about the system operation, and the performance measures chosen are stated. In section III, the model is presented and all the cases above mentioned are studied. In section IV the values obtained using the approximate analytic model are compared to those given by simulation. We also discuss how to choose the number of busses for a given performance. Finally in section V conclusions are presented.

II. ASSUMPTIONS AND PERFORMANCE MEASURES

The following assumptions are made regarding the operation of the system.

1) The system consists of n statistically identical and independent processor-memory-switch modules connected by a MBIN with b busses.

2) The operation is synchronized. The system is observed at multiples of the memory cycle. The propagation delays and arbitration times associated with the MBIN are not included explicitly, but may be thought of as forming part of the memory cycle.

3) In each cycle, the busses are assigned randomly to the memory modules that have at least one outstanding external request. For a module that receives a bus, a processor is selected at random from those with outstanding external request for that module.

4) When a memory cycle begins and there are two requests for the same memory module, one local and the other external, an arbiter of fixed priority implemented in the switch S_i , grants the module to the external request.

5) At the beginning of each memory cycle, all the active processors and those that have just completed a memory access, generate a new demand to the shared modules with probability $p' \leq 1$. Any processor with a

pending request is blocked while awaiting service.

6) When non active, each processor (P_i) requests its local memory with static probability q , and any other memory ($M_{j,i \neq j}$) module with probability $(1-q)/(n-1)$, i.e: the requests are uniformly distributed among the external memory modules.

7) The memory access time is constant regardless of the type of operation (read,write). When a processor requires external access, it takes k ($k \geq 0$) extra memory cycles to have the selected external memory request ready. During this $k+1$ cycles, the bus is fully occupied but the external memory is busy for the last cycle only.

Many measures may be chosen to evaluate the system performance. Here we use the B and the mean access time. Related to these, we also have a few others, namely:

a) The dynamic access probability of each processor to its local memory, ℓ . This probability depends on the behavior of all processors in the system, whereas the static probability q depends only on the characteristics of the job executed, and is not affected by the other processors [13]. A measure related to ℓ is the average number of memory modules that are active in each cycle due to local request, B^ℓ . In fact, we have $B^\ell = n * \ell$.

b) The dynamic access probability of any processor to each external module, P_e . A measure related to this parameter is the average number of memory modules that are active due to external request B_e . As before: $B_e = n \cdot P_e$.

c) The system total bandwidth is $B = B_l + B_e$, the sum of the local and external bandwidths.

d) The average number of cycles needed for a processor to access its local memory, T_l , including the cycle in which the demand is granted. T_e is a similar parameter for a external request. Other related measures are the probabilities of a processor waiting to access its local memory or any external module (w_l and w_e , respectively).

III. PERFORMANCE EVALUATION

For the assumptions stated in II, an exact analysis using Markov chains could, in principle, be carried out. However, as the number of processors and busses increases, such an approach soon becomes tedious and computationally prohibitive [5]. Therefore, we do not even consider this approach here.

Instead, in this section we present an approximate method to evaluate the performance of the multiprocessor architecture described earlier.

Let us begin with the computation of 'e' for the case of no internal processing ($p=1$) and $k=0$.

If $p(i)$ represents the probability that i busses are busy, we have the following equalities:

$$e = \frac{\sum_{i=1}^b p(i) \binom{n-1}{i-1}}{\sum_{i=1}^b p(i) \binom{n}{i}} = \frac{1}{n} \sum_{i=1}^b i p(i) \quad (1)$$

where the sum,

$$\sum_{i=1}^b i p(i)$$

is the effective bandwidth of a multiple-bus network with no local memory, evaluated for n processors, n memory modules, b busses and an unknown dynamic probability, α , of a processor referencing its local memory (successfully or not). According to [10] this bandwidth can be obtained as

$$n \left[1 - \left(1 - \frac{1-\alpha}{n} \right)^n \right] - L(n,b,\alpha)$$

where the first term corresponds to the Strecker's formula, and the second one represents the losses with respect to the crossbar, calculated removing the queues as indicated in [10]. For completeness, we briefly sketch this procedure here.

We have:

$$L(n, b, \alpha) = \sum_{i=b+1}^n (i-n) p_{\alpha}(n, i)$$

where $p_{\alpha}(n, i)$ is the probability that the requests from the n processors (some of which remain in internal processing with probability α) go to exactly 'i' different memory modules.

$p_{\alpha}(n, i)$ is, therefore, given by:

$$p_{\alpha}(n, i) = \sum_{k=i}^n \binom{n}{k} (1-\alpha)^k \alpha^{n-k} p_0(k, i)$$

where,

$$p_0(k, i) = \binom{n}{i} \frac{k!}{n! \dots n!} \left(\frac{1}{n}\right)^k$$

with the summation above carried over all n_1, \dots, n_i such that $n_1 + \dots + n_i = k$.

Let λ be the dynamic access probability of a given processor to its local memory. In order to access its own module, that processor must be referencing it and find no interference from the other $n-1$. If $\Pi(i)$ represents the probability that the other $n-1$ processors generate i different external requests, we have

$$l = \alpha \sum_{i=0}^{n-1} \pi(i) \left(1 - \frac{1}{n-1} \min\{i, b\}\right)$$

which asymptotically (n large, but the numerical results show that $n > 2$ is enough) can be written as,

$$\begin{aligned} l &= \alpha \sum_{i=0}^{n-1} \pi(i) \left(1 - \frac{1}{n-1} \min\{i, b\}\right) = \\ &= \alpha \left(1 - \frac{1}{n-1} \sum_{i=0}^{n-1} \pi(i) \min\{i, b\}\right) \end{aligned} \quad (2)$$

The sum in the above expression is, again, the effective bandwidth of a MBIN, but now the number of processors is only $n-1$.

Therefore, we can write:

$$e = \frac{1}{n} B(n, b, \alpha) \quad (3)$$

$$l = \alpha \left(1 - \frac{1}{n-1} B(n-1, b, \alpha)\right) \quad (4)$$

where $B(n, b, \alpha)$ represents the bandwidth of a MBIN of the indicated parameters .

It may be instructive to particularize the above two equations for the case $b=1$. Assuming that the event "Pi references Mi" is independent of the event "Pj references Mj" ($j \neq i$), we have

$$B(n,1,\alpha) = 1 - \alpha^n$$

which gives:

$$e = \frac{1}{n} (1 - \alpha^n)$$

$$l = \alpha \left[1 - \frac{1}{n-1} (1 - \alpha^{n-1}) \right]$$

as it is stated in [1] under a slightly different form.

Another relation can be found between l and e using the equilibrium condition $l = (e + l)q$ (see figure 2). From that equation we obtain

$$\frac{l}{q} = \frac{e}{p} \quad (5)$$

Which establishes the equality between the ratios of dynamic and static probabilities.

Equation (5) together with (3) and (4) determine α , e and l .

The effective bandwidth is related to e and l as follows,

$$B = n^* \lambda + n^* e$$

where $n^* \lambda$ and $n^* e$ are the average number of memory modules that are busy due to local and external requests, respectively.

The mean time to access local memory, including the cycle in which access is granted, can now be computed as

$$T \lambda = \frac{\alpha}{\lambda}$$

Similarly, the mean time to access a external module is

$$T e = \frac{\beta}{e}$$

Where β represents the probability that a processor references an external memory module.

To allow for internal processing, let us now denote by 'a' the probability that a processor is in internal procesing. Equating probability ratios as in (5), we must have

$$\frac{a}{q'} = \frac{e}{p'p} = \frac{\lambda}{p'q} \quad (6)$$

Equations (3),(4) are still valid substituting $a+\alpha$ for α .

Solving these algebraic equations for α and 'a', the performance parameters for the model can be found easily.

Finally, the case $k > 0$ can be treated analogously if we introduce the additional simplifying assumption that all the requests through the bus are in the same cycle (synchronized). Obviously this is not true, but the errors incurred are practically negligible, as the simulation results show. This gives,

$$e = \frac{1}{k+1} \frac{1}{n} B(n, b, \alpha + a)$$

and

$$l = \alpha \left(1 - \frac{1}{k+1} \frac{1}{n-1} B(n-1, b, \alpha + a) \right)$$

which, together with (8), yield α , a , l and e .

The other performance indexes follow from the relations:

$$a + \alpha + \beta = 1$$

$$\alpha = l + w l$$

$$\beta = e + k * e + w e$$

The results of these computations will be commented and compared to the simulation values in the next section.

IV. NUMERICAL RESULTS.

In this section we comment upon the results obtained using the above model and compare them to the simulation values. We use the relation defined by Lavenverg[16] to

establish the number of simulation repetitions needed in terms of the required confidence level.

Before we discuss the results, we would like to point out the accuracy of the approximate model developed here for this architecture. Fig.3 presents the bandwidth obtained for $p'=0.5$, $q=0.5$ and different values of busses and k , as a function of the number of processors. A numerical comparison with the simulation results is given in table I. Observe the extremely good agreement between the two. Therefore, our model seem to be a valid approximation when rapid answers about the system behavior are desired.

Figures 4a and 4b show the number of busses vs EBW and busses vs access time, for given parameters n, p', q, k . The results obtained show that for the same values of n, p', q , when k is increased there are enormous performance losses. As can be seen for the same values of the parameters $n=4, p'=0.7, q=0.7, b=1$, when k is increased from 0 to 2, there is an average delay increment in the access time of 59%, and an average reduction in the B of the 180%. Observe also, that a saturation point is reached for each curve. For instance, if $n=4$ and $k=0$ the curve saturates when the number of busses is 2. However in this case, if the number of busses is reduced to only 1, the degradation in bandwidth is 11% and the time is 20% longer.

In Table II we have collected the recommended number of busses for different sets of parameters and accepted level of performance. For example, if the number of processors is 16, the hit to local memory is 0.7, the factor due to delays is $k=2$ (external service + overhead) and access times that are longer than 1.2 times the minimum are not desired, 6 busses are enough. The B obtained with them is 90% of saturation. If for the same parameters we want the system to work at the bandwidth saturation point, the number of busses has to be incremented to 8. These results can, therefore, provide valuable help in the design process.

V. CONCLUSIONS.

We have presented an analytic model to evaluate the performance of a loosely coupled multiprocessor system. Although for the hypotheses stated the model is approximate, the performance measures obtained using it are extremely close to the simulation values in all the computed cases.

In this Architecture the choice of the number of busses for specific parameters, such as concurrency level, speed of communication between tasks and locality of programs is a very important design decision. The results show that care must be exercised in this choice. At the bandwidth saturation point, for instance, the access time could be unacceptable and the number of busses may have to be increased to meet specific delay

bounds. Tables and curves are provided to help in the design process.

Our model seems to be rather realistic for the performance evaluation of a great number of present multiprocessor systems and it seems to be a valid approximation when rapid answers about the system behavior are desired.

REFERENCES.

[1] SANVICENTE, VALERO, LANG and ALEGRE : "Exact and approximate models for Multiprocessor systems with Single Bus and distributed Memory". MIMI'82, Paris, pp. 15-18

[2] WULF AND BELL : "Cmmp A modular Multi-mini-processor". AFIPS. Conf. Proc. Vol.41, 1972, pp. 765-777.

[3] Swan R.J. et al., "Cm*-A Modular Multiprocessor". AFIPS. Conf. Proc. Vol.46, 1977, pp. 637-644.

[4] PHILIPSON L., NILSSON, and BREIDEGARD : "A communication structure for a Multiprocessor Computer with distributed global memory", ACM 10th Inter. Symp. Computer Architecture, 1983, pp.334-339.

[5] MARSAN, CARRA : "Bus and Memory Interference in Double Bus Multiprocessor Systems", North-Holland Publishing Company MM 13 (1984).

[6] AJMONE, BALBO and CONTE : "Comparative Performance Analysis of Single Bus Multiprocessor Architectures", IEEE Trans. on Computers, vol. C-31, pp.1179-1191, December 1982.

[7] AJMONE and GREGORETTI : "Markov models for Multiple bus Multiprocessor Systems", IEEE Trans. on Computers, vol. C-31, pp.239-248, March 1982

[8] SIEGEL, McMILLEN, MUELLER : "A survey of interconnection methods for reconfigurable parallel processing systems", AFIPS Conf. Proc., NCC 1979, June 1979, pp.529-542.

[9] LANG, VALERO and ALEGRE : "Bandwidth of Crossbar and Multiple-bus connections for Multiprocessors", IEEE Trans. on Computers, vol. C-31, pp.1227-1234, December 1982.

[10] VALERO et al., "A Performance Evaluation of the Multiple Bus Network for Multiprocessor Systems", 1983 ACM Sigmetrics Conf. on Measurement and Modeling of Comp. Systems., pp.200-206, Aug. 1983.

[11] ALFARO, GONZALEZ : "Performance Evaluation of the Shuffle-Exchange Networks", MIMI'83, Italy (Bary)., pp.103-106.

[12] BHUYAN : "On the performance of loosely coupled Multiprocessor", ACM 11th Inter. Symp. Computer Architecture, 1984, pp.256-262

[13] HOOGENDORN,C. " A general model for memory interference in Multiprocessor", IEEE Trans. on Computers, vol. C-26, pp.998-1005, October 1977.

[14] HOENER and ROELDER : "Efficiency of a multi-microprocessor system with time shared busses", EuroMicro,pp.35-42, 1977.

[15]WILLIS P.j, "Derivation and Comparison of multiprocessor contention measures", IEE J. Comput. Digital Techniques, Aug. 1978.

[16] LAVENGERG,S et al., "Sequential stopping rules for the regenerative method of simulation", IBM J.Res.Develop.,vol.21,pp.545-558,1977.

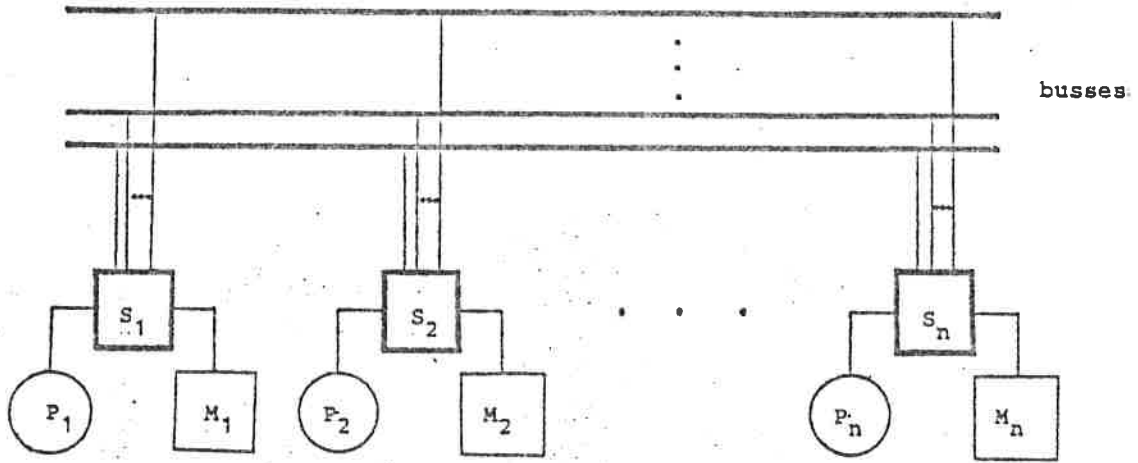


Fig. 11 : A Multiple Bus IN with distributed memory.

S_i Switch
 M_i Memory module i
 P_i Processor i

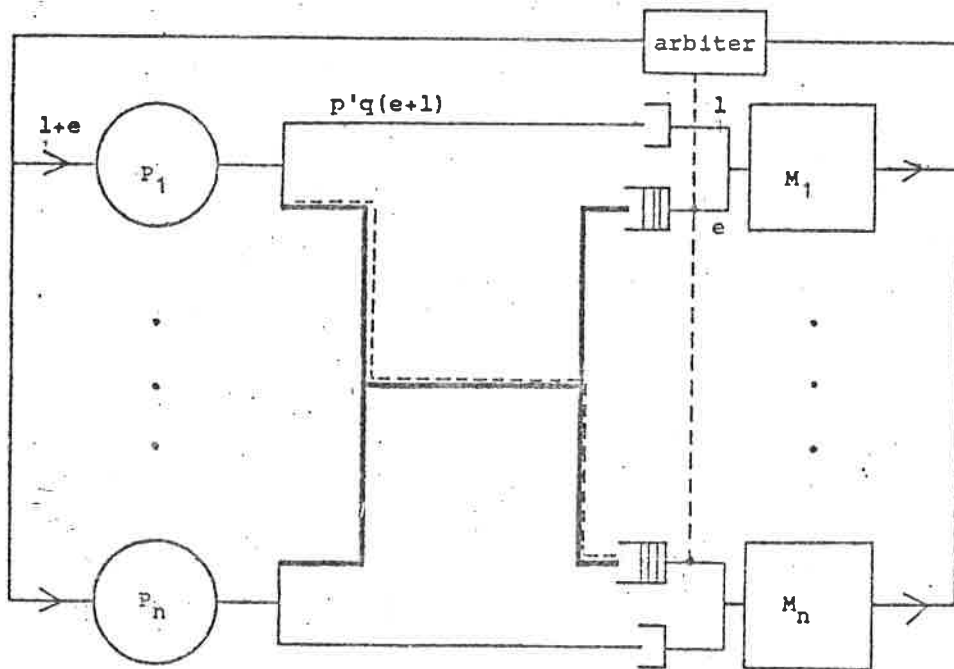


Fig. 2 : A network of queues for the system in fig. 1 .

1.0 1.2 1.5

N	FP	Q	K	sat.	BUSES			% (B vs .95 Bsaturate)		B						
					1	2	3	1.00	1.05	0.95	0.95	0.73	2.65	2.65	2.65	1.94
4	.90	.50	0	2	2	2	1	1.00	1.00	1.00	0.91	0.91	2.65	2.65	2.65	1.94
			1	3	3	2	1.00	1.00	0.91	0.91	1.97	1.97	1.78	1.78		
			2	3	3	2	1.00	1.00	0.86	0.86	1.47	1.47	1.27	1.27		
			0	2	2	2	1.00	1.00	0.89	0.89	2.95	2.95	2.63	2.63		
			1	2	2	2	1.00	1.00	1.00	0.69	2.34	2.34	2.34	1.62		
			2	3	3	2	1.00	1.00	0.94	0.94	1.96	1.96	1.84	1.84		
			0	1	1	1	1.00	1.00	1.00	1.00	3.30	3.30	3.30	3.30		
			1	1	1	2	1.00	1.05	1.00	1.00	2.95	3.08	2.95	2.95		
			2	2	2	2	1.00	1.00	0.91	0.91	2.83	2.83	2.57	2.57		
			8	.90	.50	0	3	4	2	1.00	1.05	1.00	1.00	0.78	5.07	5.31
1	5	5				3	1.00	1.00	0.95	0.79	3.77	3.77	3.58	2.96		
2	5	6				4	1.00	1.02	0.92	0.72	2.75	2.80	2.52	1.99		
0	3	3				2	1.00	1.00	0.94	0.94	5.84	5.84	5.48	5.48		
1	4	4				2	1.00	1.00	0.94	0.71	4.67	4.67	4.38	3.29		
2	4	5				3	1.00	1.03	1.00	0.87	3.67	3.79	3.67	3.20		
0	1	1				1	1.00	1.00	1.00	1.00	6.40	6.40	6.40	6.40		
1	2	2				2	1.00	1.00	1.00	0.80	6.02	6.02	6.02	4.80		
2	3	3				2	1.00	1.00	0.95	0.95	5.62	5.62	5.32	5.32		
16	.90	.50				0	6	7	4	1.00	1.04	0.93	0.77	10.35	10.72	9.61
			1	8	9	6	1.00	1.03	0.93	0.82	7.30	7.51	6.79	5.97		
			2	9	9	6	1.00	1.00	0.94	0.73	5.48	5.48	5.18	3.98		
			0	4	5	3	1.00	1.03	1.00	0.87	11.29	11.67	11.29	9.79		
			1	6	7	4	1.00	1.04	0.91	0.74	8.95	9.28	8.11	6.63		
			2	7	8	5	1.00	1.04	0.90	0.77	7.22	7.53	6.51	5.55		
			0	2	2	1	1.00	1.00	1.00	0.75	13.07	13.07	13.07	9.83		
			1	3	3	2	1.00	1.00	1.00	0.83	11.81	11.81	11.81	9.83		
			2	4	4	3	1.00	1.00	0.89	0.89	10.90	10.90	9.66	9.66		

Table I : Effective Bandwidth and access time of the model and simulation .

* * * $\rho=0.70$ $\rho=0.70$ * * *									
$n \times l$ b	$\tau=0.5$		$\tau=0.7$		$\tau=0.7$				
	*	*	*	*	*	*			
N=4	S	1.27 1.37	2.76 2.76	1.60 1.53	N=4	S	2.29 2.30	1.57 1.58	1.10 1.11
	S1	0.28 0.28	1.93 1.93	1.24 1.23		S1	1.60 1.41	1.12 1.11	0.77 0.79
	S2	0.29 0.28	0.31 0.32	0.26 0.26		S2	0.89 0.89	0.47 0.47	0.23 0.23
	S3	1.47 1.47	1.23 1.23	1.10 1.11		S3	1.24 1.21	1.14 1.17	1.09 1.12
N=8	S	1.29 1.27	3.33 3.32	4.31 4.28	N=8	S	3.30 3.33	1.66 1.69	1.11 1.13
	S1	0.28 0.28	2.23 2.22	4.22 4.23		S1	2.71 1.24	1.25 1.13	0.77 0.80
	S2	1.00 1.00	1.00 1.00	0.69 0.70		S2	0.29 0.29	0.50 0.50	0.23 0.23
	S3	1.17 1.17	1.17 1.17	1.10 1.10		S3	1.15 1.16	1.07 1.09	1.04 1.05
N=16	S	1.29 1.29	3.33 3.33	9.93 9.95	N=16	S	3.33 3.37	12.99 11.98	10.15 10.11
	S1	0.28 0.28	2.23 2.23	3.29 3.25		S1	1.72 1.73	1.44 1.45	1.21 1.21
	S2	1.00 1.00	1.00 1.00	1.00 1.00		S2	0.74 0.75	0.62 0.61	0.53 0.53
	S3	1.07 1.07	1.07 1.07	1.07 1.07		S3	1.24 1.23	1.17 1.19	1.14 1.15
N=32	S	14.34 14.74	13.50 13.51	6.40 6.41	N=32	S	2.47 2.49	2.07 2.07	1.74 1.74
	S1	0.28 0.28	2.23 2.23	3.29 3.25		S1	1.72 1.73	1.44 1.45	1.21 1.21
	S2	1.00 1.00	1.00 1.00	1.00 1.00		S2	0.74 0.75	0.62 0.61	0.53 0.53
	S3	1.07 1.07	1.07 1.07	1.07 1.07		S3	1.24 1.23	1.17 1.19	1.14 1.15
N=64	S	3.64 3.75	4.22 4.26	7.31 7.31	N=64	S	4.82 4.31	4.06 4.06	3.34 3.38
	S1	2.81 2.89	4.42 4.46	6.58 6.58		S1	3.44 3.44	2.84 2.84	2.33 2.37
	S2	2.33 2.28	1.90 1.91	0.73 0.73		S2	1.48 1.47	1.22 1.22	1.00 1.01
	S3	1.41 1.37	1.32 1.31	1.10 1.10		S3	1.24 1.23	1.17 1.19	1.14 1.15
N=128	S	1.22 1.21	1.13 1.12	1.04 1.04	N=128	S	1.10 1.14	2.40 2.38	1.89 1.90
	S1	9.79 9.84	12.44 12.52	14.60 14.62		S	9.67 9.57	6.37 6.60	4.44 4.47
	S1	4.89 4.92	3.71 4.79	13.14 13.16		S1	6.76 6.77	4.60 4.62	3.10 3.13
	S2	4.90 4.92	2.73 3.74	1.46 1.45		S2	7.21 2.90	1.38 1.93	1.33 1.33
N=256	S	1.50 1.48	1.22 1.21	1.10 1.10	N=256	S	1.24 1.23	1.14 1.15	1.09 1.10
	S1	1.77 1.77	1.21 1.20	1.05 1.05		S1	1.21 1.23	4.03 3.96	4.94 7.29
	S2	1.50 1.48	1.22 1.21	1.10 1.10		S2	1.24 1.23	1.14 1.15	1.09 1.10
	S3	1.77 1.77	1.21 1.20	1.05 1.05		S3	1.21 1.23	4.03 3.96	4.94 7.29

* Simulation + Model $n_p = b$

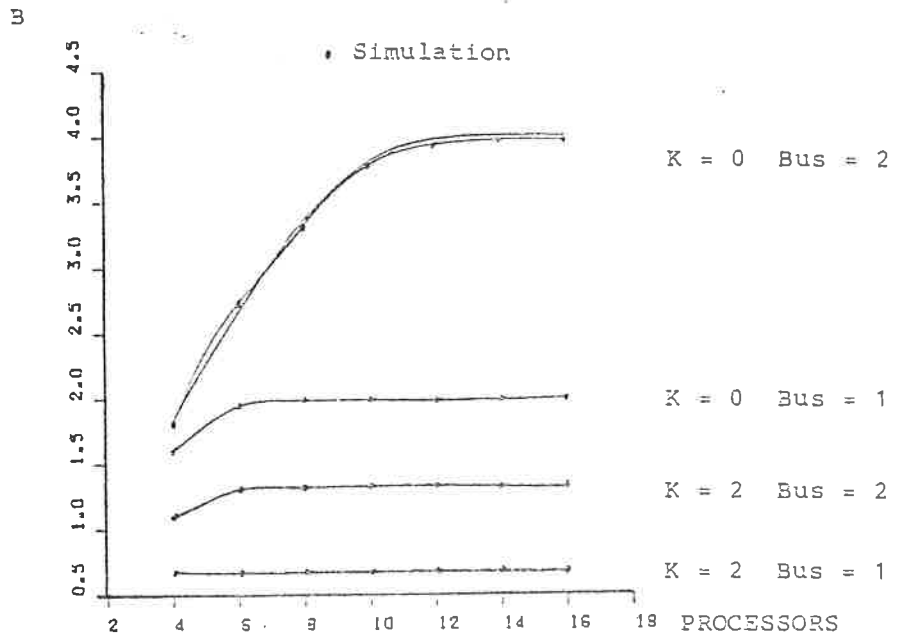


Fig. 3 : Comparison between model and simulation results.

Table II : Recommended number of busses for a given average access time .

M	Fp	Q	K	sat.	δ / δ_{min}			BUSES				% (B vs .95 Bsaturate)				B	
					1.0	1.2	1.5										
					4	.90	.50	0	2	2	2	1	1.00	1.00	1.00		0.73
					1	3	3	2	2	1.00	1.00	0.91	0.91	1.97	1.97	1.78	1.78
					2	3	3	2	2	1.00	1.00	0.86	0.86	1.47	1.47	1.27	1.27
		.70			0	2	2	1	1	1.00	1.00	0.89	0.89	2.95	2.95	2.63	2.63
					1	2	2	2	1	1.00	1.00	1.00	0.69	2.34	2.34	2.34	1.62
					2	3	3	2	2	1.00	1.00	0.94	0.94	1.96	1.96	1.84	1.84
		.90			0	1	1	1	1	1.00	1.00	1.00	1.00	3.30	3.30	3.30	3.30
					1	1	2	1	1	1.00	1.05	1.00	1.00	2.95	3.08	2.95	2.95
					2	2	2	1	1	1.00	1.00	0.91	0.91	2.83	2.83	2.57	2.57
8	.90	.50			0	3	4	3	2	1.00	1.05	1.00	0.78	5.07	5.31	5.07	3.95
					1	5	5	4	3	1.00	1.00	0.95	0.79	3.77	3.77	3.58	2.96
					2	5	6	4	3	1.00	1.02	0.92	0.72	2.75	2.80	2.52	1.99
		.70			0	3	3	2	2	1.00	1.00	0.94	0.94	5.84	5.84	5.48	5.48
					1	4	4	3	2	1.00	1.00	0.94	0.71	4.67	4.67	4.38	3.29
					2	4	5	4	3	1.00	1.03	1.00	0.87	3.67	3.79	3.67	3.20
		.90			0	1	1	1	1	1.00	1.00	1.00	1.00	6.40	6.40	6.40	6.40
					1	2	2	2	1	1.00	1.00	1.00	0.80	6.02	6.02	6.02	4.80
					2	3	3	2	2	1.00	1.00	0.95	0.95	5.62	5.62	5.32	5.32
16	.90	.50			0	6	7	5	4	1.00	1.04	0.93	0.77	10.35	10.72	9.61	8.00
					1	8	9	7	6	1.00	1.03	0.93	0.82	7.30	7.51	6.79	5.97
					2	9	9	8	6	1.00	1.00	0.94	0.73	5.48	5.48	5.18	3.98
		.70			0	4	5	4	3	1.00	1.03	1.00	0.87	11.29	11.67	11.29	9.79
					1	6	7	5	4	1.00	1.04	0.91	0.74	8.95	9.28	8.11	6.63
					2	7	8	6	5	1.00	1.04	0.90	0.77	7.22	7.53	6.51	5.55
		.90			0	2	2	2	1	1.00	1.00	1.00	0.75	13.07	13.07	13.07	9.83
					1	3	3	3	2	1.00	1.00	1.00	0.83	11.81	11.81	11.81	9.63
					2	4	4	3	3	1.00	1.00	0.89	0.89	10.90	10.90	9.66	9.66