Performability Measures Of Multiple Path Multistage Interconnection Networks

Dr. Sudarson Jena[1], G.Sri Sowmya[2] and P.Radhika[2]
Department of Comp. Science Engineering & IT,
GITAM University,
Hyderabad , India
sudarsonjena@gitam.edu, sowmya@gitam.edu, radhikapulicherla@gitam.edu

Abstract—**In this paper, attempts have been made to develop different combinatorial models for evaluation performability of various multiple path multistage interconnection networks (MINs). For the purpose here, two metrics of performability i.e., reliability and bandwidth availability (BA) in order to define the fault tolerance behavior of various multiple path MIN networks have been introduced. The results are compared and discussed.**

Keywords-Reliability, bandwidth availability, interconnection network, multiple path MIN, Performability.

I. INTRODUCTION

Multiprocessor interconnection networks have been applied in real time environment and they are becoming increasingly popular for large commercial applications as well. The prominent interconnection networks among these are Crossbar, Multiple shared bus, Omega Multi stage interconnection network, Extended omega network, Vertical stacked Multistage Interconnection network and Clos Network [2][5]. With advancement in research in the field of multiprocessing, an increased focus needs to be placed on the performability of its interconnection networks. Performability is a measure which addresses both reliability and performance. Therefore, it should be incorporated into a system right from the designing state itself. Evaluation of performability of multiprocessor interconnection network has been attempted by researchers in the past [1][3][5][6]. As the demand for these systems increase in society, it becomes imperative to build efficient and reliable multiprocessors. Efficiency is usually attributed to the performance while reliability is concerned with probability of success. To achieve these goals, quick and accurate evaluation of both the performance and reliability of parallel architectures become essential [4][7]. A reliability model represents the behavior of system in response to fault [4]. On the other hand, the performance model determines the levels of performance when the system is operational [6].

Meyer [3] developed a conceptual network of performability of a system which is defined as the accomplishment level of the problem over a specified time period *t*. Bhuyan [1] has proposed an analytical technique to evaluate performance of shared memory architectures. However their model does not include for performability evaluation of multiprocessor interconnection networks (MIN). Das and et.al.[2] have estimated the dependability modeling of multi processor networks. Their method does not provide a general solution to all multiprocessor interconnection network systems. In this paper we examined the performability aspects of Multipath Multistage interconnection networks (MINs) using combinatorial methods specifically. The study includes multipath MIN like Clos, Vertical stacked, Extra stage Multistage interconnection Networks. The analysis assumes that fault exists among processors, memory modules and interconnection networks. Our work was motivated by Das and Bhuyan [1][2]. They introduced two terms reliability and fault tolerance MIN system (Fig 1). The reliability of a system at time *t* is defined as the probability that the system is operational during [0,t]. Bandwidth availability (BA) is the expected value of available bandwidth in the system at time *t* .

Section 2 provides architectural description of multipath MIN system. Proposed strategy for multipath MIN is presented in section 3. In section 4, performability model of multipath MIN has been proposed. Numerical results and comparisons are presented in section 5 of this paper. Concluding remarks are presented in section 6.

II. ARCHITECTURE DESCRIPTION

In order to predict the performability of a system, one must consider its topological features in detail. This sub section discusses in brief the topological aspects of some candidate multiple path multistage interconnection systems. The networks considered are omega multistage interconnection network (MIN), Extra stage omega MIN, Vertical stacked MIN, and Clos network.

A. Multistage Interconnection Network (MIN)

Multistage interconnection networks (MINs) connect input devices to output devices through a number of switching stages, where each switch is a crossbar network. A multiprocessor system that includes N processors and M memory modules are interconnected via. NxM MINs. An M×N MIN consists of $log_a N$ stages of a×a switching elements (SE_s). Each stage of the network has N/a switches. The number of stages and the connection patterns between stages determine the routing capability of the networks. The BBN butterfly, IBM RP3, Cedar are the examples of multistage interconnection networks. Fig 1 shows an M×N multistage interconnection networks, where M=N=8.

The disadvantage of the MINs is their blocking property. A blocking network is unable to service all its required processor-memory connections and consequently blocks some of them. Also, the existence of one path between every processor memory path has a –ve effect on the fault tolerance of the network. Network details of different Mulitipath MIN systems are described in the following sections.

B. Extra stage MIN network

In extra stage MIN network, an additional stage of switching elements is added. In the case of an Omega interconnection network one of 2 ×2 Switching elements (SEs) results in two paths between any pair of processor and a memory module. If we add another stage, the paths become four (Fig.2). We can easily extend our reliability model to include the extra stage network because we can calculate the reliability of each path independently.
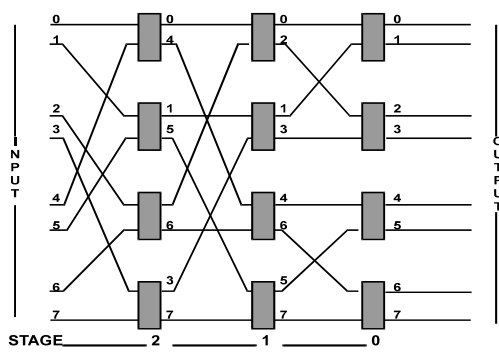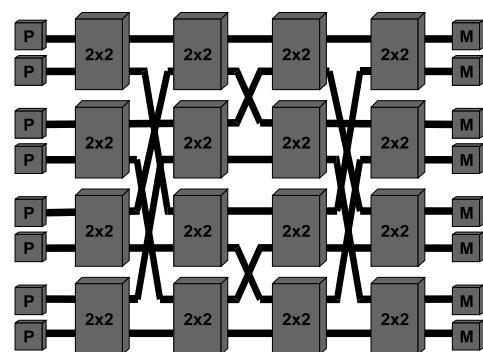

Figure 1. A MIN System

Figure 2. An Extra stage MIN system

C. Vertical stacked MIN

A different approach in improving the performance and fault tolerance behavior of a system that uses MIN is to employ two (or more) MINs in parallel (Fig. 3).The networks are the identical copies of a unique-path MINs,

such as Delta Network . In this way, each processor is able to use two or more independent MINs in order to send a request to a memory module.

D. Clos Interconnection Network

The Clos network is a special case of multiple-path MIN. It can be defined by three integer parameters m, n, r and can be denoted by *N (m, n, r)*. An N–input Clos network where $N = n * r$ incorporates three stages . The first and the third stages of the network each consists of r n×m SE's,which may be implemented as crossbars, or the other small Clos networks. The second stage consists of *m* rxr SEs, which are similarly implemented . The network has exactly one link between every two SE's situated in the consecutive stages.
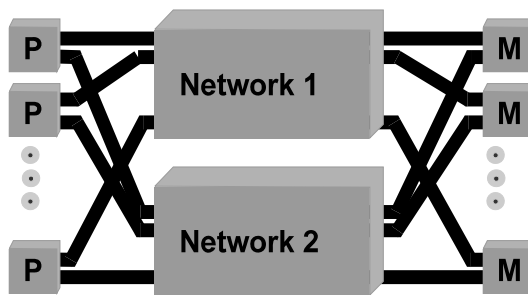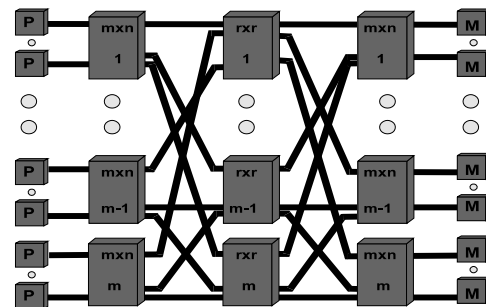

Figure 3. A two vertical stacked MIN system

Figure 4. A Clos Network system

Between every processor-memory pair, there are several disjoint paths, the number of which is determined by m (number of outputs in each SE of the first stage, or number of SE in the second stage). In Fig 4, a multiprocessor system with N processors and N memory modules that uses a N (m,n,r) Clos network is presented .

III.   PROPOSED STRATEGY FOR MULTIPLE PATH MIN

The strategy for using the analytical technique to design the Performability model to the desired system consists of the following steps:

1. Choose the desired metrics of performance and reliability requirements.
2. Construct performance model for a system using the analytical technique. This model represents the probabilistic nature of user demands (workload) and the variations in the internal states of the system.
3. Solve the performance model of the system for the steady state analysis to obtain the desired matrices.
4. Construct an analytical reliability model of the system. This model represents the changes of the structure of the system due to occurrence of faults and repairs.
5. Solve the reliability model.
6. Analyze both the performance model and the reliability model under similar assumptions.
7. Combine the features of reliability and performance model.

The proposed method finds the Performability of multistage interconnection network. The following assumptions are made throughout this paper for Performability analysis.

Reliability assumptions
The reliability analysis is based on the following assumptions:
   a) The reliability of each submodule is determined independently.
   b) The elements of each submodule are identical and have the same failure rate and
   c) The failures are exponentially distributed.

Performance Assumptions

Various analytical models developed for evaluating the performance of interconnection networks are based on the following assumptions:

a) All processors are synchronized.

b) The processor requests are independent and uniformly distributed.

c) The cycle time is constant and the same for all memory modules.

d) A processor issues a new request in each cycle according to the predominated request rate, after receiving a memory service.

e) The propagation delay and arbitration time are not considered separately, but assumed to be included in each memory cycle.

f) Requests, which are not accepted, are rejected. This assumption is unrealistic, but it leads to simpler analysis and it does not result in a substantial difference in the actual results.

g) At the beginning of every memory cycle, each processor generates a new request with probability y ($0 \leq y \leq 1$).

IV. PERFORMABILITY EVALUATION OF MULTIPLE PATH MIN SYSTEMS

In this section, the proposed model for studying the performability of Multiple path Multistage Interconnection Networks is presented. Performability addresses both the analysis of reliability and performance. First we define these measures and then discuss techniques for computing the reliability of multiple path MINs. Similarly, many performance parameters are applicable for interconnection networks. Memory bandwidth (BW) is the most common performance parameters used in our analysis.

A. Reliability model of Extra stage MIN

The reliability of extra stage MIN network can be found from the Reliability expression of MIN system with some modifications. In case of multistage system, we develop model for a task that requires at least I processors, J memory modules and the IJ paths that connect them. A state (i,j,k) means that there are I processors and J memory modules operational via. s paths. Let $P_{ijs}(t)$ be the probability that the system is in the state (i,j,s), then

$$P_{ijs}(t) = \binom{N}{i} C_p^{N-i} (R_p(t))^i (1-R_p(t))^{N-i} \binom{N*M}{s} C_{path}^{M*N-s} (R_{path}(t))^s (1-R_{path}(t))^{NM-s}$$

$$\binom{M}{j} C_m^{M-j} (R_m(t))^j (1-R_m(t))^{M-j} \tag{1}$$

where C_p, C_m, C_{path} are the coverage factors, $s = i*j$ and MN is the total number of paths that connects the processors to the memory modules. We assume that the SE is a×b crossbar network. Its reliability, when Q cross points are required to be operational is given by the relation

$$R_{SE}(t) = \sum_{K=Q}^{b} C_{cp}^{b-k} \binom{b}{k} (R_{cp}(t))^k (1-R_{cp}(t))^{b-k} \tag{2}$$

where R_{cp} is the reliability of a cross point of a SE and C_{cp} is the coverage factor. The reliability of each communication path is calculated independently by multiplying the reliability of the links (R_l) and SEs that it employs and is given by the following relation

$$R_{path}(t) = (R_l(t))^{n-l} * (R_{SE}(t))^n \qquad (3)$$

Therefore reliability of MIN system is

$$R_{MIN}(t) = \sum_{i=I}^{N} \sum_{j=J}^{M} \sum_{s=I*J}^{N*M} P_{ijs}(t). \qquad (4)$$

The Reliability of Extra stage MIN systems can be found from the above expressions with the following modifications

i) The number of cross points Q in relation (2) must be set equal to 1, for an SE that is situate an extra stage, because there are two independent paths from this SE to a memory module,

$$R'_{SE}(t) = \sum_{K=l}^{b} C_{cp}^{b-k} \binom{b}{k} (R_{cp}(t))^k (1 - R_{cp}(t))^{b-k} \qquad (5)$$

Thus the relation (3) that calculates the path reliability becomes

$$R'_{path}(t) = (R_l(t))^{n-l} * (R_{SE}(t))^{n-l} * (R'_{SE}(t)). \qquad (6)$$

ii) In relation (2), the number of paths becomes 2NM and 4NM for one extra and two extra stage MINs, respectively. Thus the probability $P_{ijs\,extra\,stage}(t)$ that the system is in state (i,j,s) is

$$P_{ijs}(t) = \binom{N}{i} C_p^{N-i} (R_p(t))^i (1 - R_p(t))^{N-i} \binom{2N*M}{s} C_{path}^{M*2N-s} (R'_{path}(t))^s (1 - R'_{path}(t))^{2NM-s}$$

$$\binom{M}{j} C_m^{M-j} (R_m(t))^j (1 - R_m(t))^{M-j} \qquad (7)$$

The reliability of the extra stage system $R_{MINextra}$ is given by

$$R_{extrastage}(t) = \sum_{i=I}^{N} \sum_{j=J}^{M} \sum_{s=I*J}^{2N*M} P_{ijsextrastage}(t). \qquad (8)$$

The reliability of the extra stage system is higher than the reliability of a MIN system because the system is redundant to R-1 path failures (due to fault SEs or faulty links), where R is the number of paths from a processor to a memory module.

B. *Performance model of Extra Stage MIN*

The performance of the network can be easily evaluated by applying the performance analysis of MIN presented by Patel [ ]. Patel used a recursive equation, which calculates the output rate of the final stage, given the input rate of the first stage. The expected number of requests that pass to b outputs is obtained by setting N=a and M=b.

If r_i is the probability that there is a request at the output of switch at stage i, then

$$r_i = 1 - (1 - \frac{r_{i-1}}{b})^a \quad , \quad \text{for } 1 \le i \le n , \qquad (9)$$

The bandwidth of an Extra stage MIN network can be evaluated using the above equation with the following modifications. (i) In the stages where two paths begin, we split the input rate in half, because there are two paths

leading to the same destination. (ii) Consequently, in the stage, where the two paths are joined together, we double the input rate. Thus,

$$r_1 = 1 - (1 - \frac{r_0}{b})^a \quad r_2 = 1 - (1 - \frac{r_1/2}{b})^a \quad \text{and} \quad r_3 = 1 - (1 - \frac{2*r_2}{b})^a$$

where b=2 (omega network) and the bandwidth of the system is

$$BW_{ij} = b^3 * r_3$$

The Bandwidth Availability (BA) of an extra stage omega network is

$$BA_{extrastage}(t) = \sum_{i=I}^{N} \sum_{j=J}^{N} \sum_{s=I*J}^{2N*M} P_{ijs}(t) * BW_{ij}. \tag{10}$$

C. Reliability model of Vertical Stacked MIN

The proposed method can be easily modified to include such networks, as follows:
- (a) We must include two reliability terms that correspond to the reliability of each IN.
- (b) In each IN, we require at least half of the paths to be operational (we assume that we have a uniform distribution of traffic between the two INs.).
- (c) In our calculations, we must also include the reliability of the multiplexer (that splits the traffic from the processor to the two networks) and the reliability of the demultiplexer (that joins the traffic from the two networks into the memory modules).

Thus, the reliability of a multiprocessor system that employs two vertical stacked MINs can given by:

$$R_{MIN}(t) = \sum_{i=I}^{N} \sum_{s=\frac{I*J}{2}}^{M*N} \sum_{y=\frac{I*J}{2}}^{N*M} \sum_{j=J}^{M} P_{isyj}(t) = \sum_{i=I}^{N} C_p^{N-i} \binom{N}{i} (R_p(t))^i (1-R_p(t))^{N-i}$$

$$R_{mp}(t) \sum_{s=\frac{I*J}{2}}^{NM} C_{path}^{MN-s} \binom{N*M}{s} (R_{path}(t))^s (1-R_{path}(t))^{NM-s} \sum_{y=\frac{I*J}{2}}^{NM} C_{path}^{MN-s} \binom{N*M}{y} (R_{path}(t))^y (1-R_{path}(t))^{NM-y} R_{dmp}(t)$$

$$\sum_{j=J}^{M} C_m^{M-j} \binom{M}{j} (R_m(t))^j (1-R_m(t))^{M-j} \tag{11}$$

The reliability of a duplex component, like the two independent MIN , has been reported in Ref.[1]. It is equal to 1-(1-R)[2], where R is the reliability of each component. If we employ this approach, we obtain similar results. However, our analysis has the advantage that it is more general and can easily be extended to include vertical stacked multiple path MIN.

D. Performance model of Vertical Stacked MIN

Here we assume that each interconnection network accepts half the request rate of the processors. By applying the output rate of the final stage of each network, r_n is calculated

$$r_i = 1 - (1 - \frac{r_{i-1}}{b})^a \quad , \quad r_0 = \Psi/2 \quad i = 0 \ldots \ldots n \tag{12}$$

Because each memory module is connected to both the networks, the probability that a memory module is requested by a processor is equal to $1-(1-r_n)^2$. Also,

$$BW = b^n (1 - (1 - r_n)^2)$$
(13)

By combining the bandwidth of the system with its reliability analysis, we can calculate its Bandwidth availability (BA) as:

$$BA_{MIN}(t) = \sum_{i=I}^{N} \sum_{s=\frac{I*J}{2}}^{M*N} \sum_{y=\frac{I*J}{2}}^{N*M} \sum_{j=J}^{M} P_{isyj}(t) * BW$$
(14)

E. Reliability model of Clos Network

For the reliability analysis of the Clos network assumptions similar to that of multiple path MINs are followed with the following exceptions. (i) The number of crosspoint Q in relation (2) must be set equal to 1 for the calculation of the reliability of the SE situated in the first stage, from which there are m independent paths leading to each memory module. (ii)The total number of paths that connects the processor to the memory modules is N×N×m. Thus the reliability of a Clos system is given by the relation

$$R_{clos}(t) = \sum_{i=I}^{N} \sum_{j=J}^{M} \sum_{s=I*J}^{D} P_{ijs}(t) = \sum_{i=I}^{N} C_p^{N-i} \binom{N}{i} (R_p(t))^i (1 - R_p(t))^{N-i}$$

$$\sum_{s=IJ}^{NM} C_{path}^{D-s} \binom{D}{s} (R_{path}(t))^s (1 - R_{path}(t))^{D-s} \sum_{j=J}^{M} C_m^{N-j} \binom{N}{j} (R_m(t))^j (1 - R_m(t))^{N-j}$$
(15)

where $P_{ijs}(t)$ is the probability that the Clos system is in (i,j,s) state (there are i processors, j memory modules and s paths operational), and D = N×N×f is the total number of paths that connect the processors to the memory modules .

F. Performance model of Clos Network

The performance analysis here is based on the following discussions. The rate of requests on any one of the m outputs of each SE situated in the first stage is

$$y_1 = 1 - (1 - \frac{y}{m})^n$$
(16)

The bandwidth analysis must take into account the variation of the traffic rates in each network stage, caused by the existence of m paths between every processor memory pair. More specifically, the rate of requests at every input of the second level is y_1 / m, and thus,

$$y_2 = 1 - (1 - \frac{y_1 / m}{r})^r$$
(17)

Consequently, in the stage, where the m paths are joined together, we multiply the input rate with m, and the output of the final stage becomes

$$y_3 = 1 - (1 - \frac{m * y_2}{n})^m$$
(18)

Thus, the bandwidth of the system is given by

$$BW_{ij} = N * y_3 \tag{19}$$

The BA of the Clos interconnection network is

$$BA_{clos}(t) = \sum_{i=I}^{N} \sum_{j=J}^{M} \sum_{s=I*J}^{D} P_{ijs}(t) * BW_{ij}$$

RESULTS AND DISCUSSIONS

This section considered the performability evaluation of some important Multipath multiprocessor interconnection networks. Fig.5 compares the reliability of various multiple path MIN systems. Specifically we evaluate the reliability of an 8x8 extra stage MIN system, 8x8 vertical stacked MIN system and 8(4,4,8) Clos system. It is noticed that Clos system exhibit considerable better reliability than the extra stage and Vertical stacked MIN. This can be easily explained by the existence of four alternative paths between every input-output pair in the Clos network. The addition of extra stages in each interconnection networks of the vertical stacked system can further improve the reliability of the system. For the given set of failure rates, all the multiple path MIN systems produce the same reliability at the beginning but after a time period ,the system with higher no of alternative paths prove more reliable. In fig.6, it is noticed that the number of processing elements and memory modules are increases then the reliability of Clos system increases. In Fig.7, we compare the bandwidth availability of various multipath MIN systems. The Clos network and extended MIN system produces the higher BA for t=0. But after certain time period, the Clos system proves to be the most attractive. Fig. 8 shows the bandwidth availability of Clos system with different operational requirements.
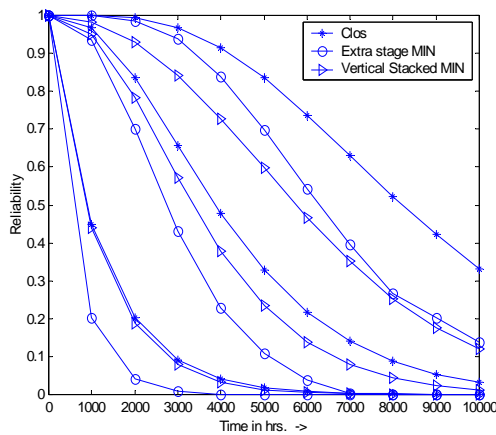

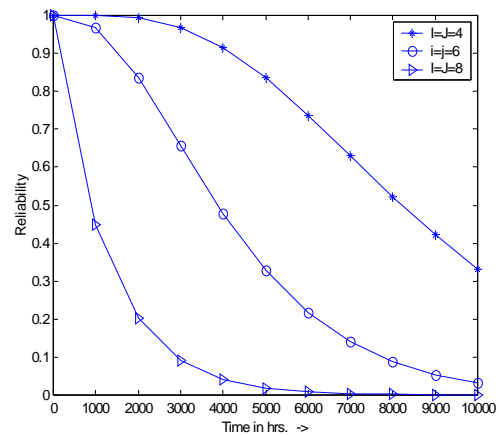
Figure5. Reliability comparison of MIN systems

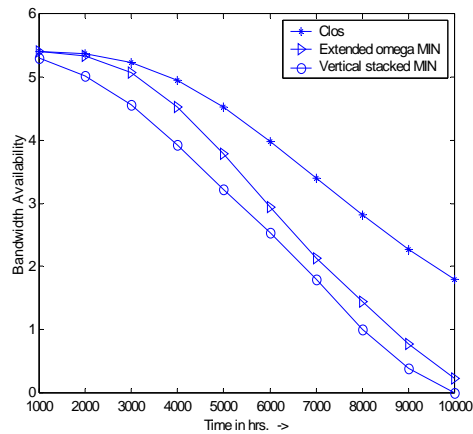Figure 6. Reliability Comparison of Clos systems

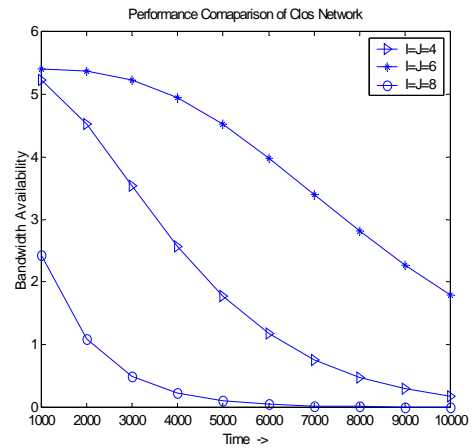Figure 7. Performance Comparison of Multiple path MINs

Figure 8. Performance Comparison of Clos Networks

CONCLUSIONS

In this paper, the performability model, evaluation and comparison of various types of multiple path MIN system have been carried out. General expressions that permit the calculation of reliability of multiprocessor systems, which employ different multiple path MINs, have been developed. The reliability and performance of such systems are comparable for substantial period of time. In multiple path MIN systems, it is clear that the existence of extra paths, between each processor and each memory increases the performability of the system. The Clos system is found the best among the multiple path MIN systems.

REFERENCES

[1]   I.N.Bhuyan, Analysis of Interconnection network, *IEEETrans. Computers,*C-34(3),279-283,1985
[2]   C.R.Das et al., Dependability modelling for multiprocessors, *IEEE Trans. Computers*, (23)7-19,1990.
[3]   F.Meyer, On evaluating the Performability of Degradable Computing Systems, *IEEE Trans. Computers*, 29(8),720-731,1980.
[4]   J.T.Blake, K.S.Trivedi, Multistage Interconnection network Reliability, *IEEE Trans. Computers,* 38(11),1600-1604,1989.
[5]   S.K.Bhogavilli, H.Abu-Amara, Design and Analysis of High Performance Multistage Interconnection Networks, *IEEE Trans. Computers*, 46(1),110-117,1997.
[6]   J.H.Patel, Performance of Processor-Memory Interconnections for Multiprocessors, *IEEE Trans. Computers*, 30(10),771-780-1981.
[7]   J.Arlat, K.Kanoun and J.C.Larpie, Dependability modelling and evaluation of software fault-tolerant systems, *IEEE Trans. Computers,* 39(4),504-514,1990.
[8]   C.R.Triparthy, R.N.Mohapatra and R.B.Misra, Reliabilty analysis of Hypercube Multicomputers, *Microelectronics and Reliability: An InternationalJournal ,*37(6),885-891,1997.
[9]   H.J.Siegel, "Study of multistage SIMD interconnection networks," in *Proc.5thAnnu.Symp.Comput.,Arch.,Apr.*1978,pp. 223-229.
[10]   D.Y.Chang, D.J.Kuck and D.H.Lawrie," On the effective Bandwidth of parallel memories," *IEEE Trans .Comput.*, vol. C-26,pp.480-490,May1997.