RELIABILITY ANALYSIS OF MULTI PATH MULTISTAGE INTERCONNECTION NETWORKS

Dr. Sudarson Jena[1], G.Sri Sowmya[2], P.Radhika[2] and P.Venkat Reddy[2]

[1]Department of Computer Science Engineering and Information Technology, GITAM University, Hyderabad
sudarsonjena@gitam.edu
[2]Department of Computer Science Engineering and Information Technology, GITAM University, Hyderabad
sowmya@gitam.edu, radhikapulicherla@gitam.edu, patlollavenkat83@gmail.com

ABSTRACT

As systems have grown more complex, the consequences on their reliable behavior have become severe in terms of cost, life, size etc, and the interest in accessing system reliability and the need for improving the reliability of systems have become very important. This paper presents an analytical model for the reliability evaluation of multipath multistage interconnection networks (MINs). In order to define the fault tolerance, behavior of various multi path MIN networks has been introduced. The results are compared and discussed.

KEYWORDS

Reliability, interconnection network, multi path MIN, Fault tolerant system, Multiprocessor system.

1. INTRODUCTION

In the race towards achieving high computational power, shared memory, multi processors systems appear to be the most promising field of research and development. Considerable effort has been focused on designing the most efficient system in terms of performance and reliability. A very crucial factor in the operation of multi processor systems is the mechanism needed for the information transfer among processors and memory modules, namely the interconnection networks(IN), many types of IN have been proposed in the past few years[3,6,11,12]. Multistage Interconnection Networks (MINs) have been proposed for use in multiprocessor systems, because they provide a compromise between networks of high latency and low cost such as shared bus and networks of low latency and high cost such as Crossbar [1,2]. The MINs have simple routing control, modular construction and can be pipelined. A multi-path MIN is composed of SEs that has at least 3 inputs and 3 outputs. The multiple paths between each input-output pair of the network are provided by connecting switching elements within the same stage [7,18]. The vertical stacked MIN, Extra Stage MIN and Clos Network consists of three stages, which is widely used in telecommunications and interprocessor communications in parallel computers and has been the focus of much research [9,10,15,16,19]. These networks provide multiple paths between any processor–memory module pair, and thus exhibits superior performance and fault tolerance in

DOI : 10.5121/ijcsit.2012.4105

comparison with the conventional MINs [5,10]. Reliability of an interconnection system depends upon the reliability of its components. However, for large parallel processing systems, it is very difficult to evaluate the probability of many failures since they have never occurred before[17,18]. The reliability of a system can be defined as the probability that the system has operated successfully over specific interval of time under static conditions. As the size and complexity of a system increases, reliability becomes an important issue. Evaluation of reliability of multiprocessor interconnection network has been attempted by researchers in the past [6,11,13]. A reliability model represents the behavior of system in response to fault [4]. Multistage Interconnection Networks (MINs) are designed to provide fast and efficient communication at a reasonable cost. Hwang & Chang [12] proposed a combinatorial method using graph-theoretic techniques to evaluate the reliability of multiprocessor structures such as crossbar switches, timeshared buses and multiport memories. However, their model does not include MINs. Cherkassy & Malek [13] estimate the network reliability for a non redundant (1-path) Banyan network. Their model uses the graph-theoretic technique to obtain a model for MINs, which is used to predict the network reliability. Though this model accounts for switch fan out, the analysis depends on technology and design. Cherkassy et al.[8] analyze the switching-banyan network with added stages and considers an (n-1) fault tolerant network. In reliability analysis, classical techniques have been used that included modeling the reliability of these systems from knowledge of failure rates of its individual components. Most of the studies were based on the evaluation of terminal or network reliability. Terminal reliability is the probability that at least one fault free path exists between a given input-output pair, and network reliability is the probability that atleast one fault free path exists between every input-output pair. All these analyses cover systems that comprise one type of IN. This means that we cannot directly compare (using the same analysis) the reliability of systems with different INs. One other problem that the above methods cannot address is the reliability evaluation of multipath MIN system.

In this paper we examined the reliability aspects of Multipath Multistage interconnection networks (MINs) using combinatorial methods specifically. The study includes multipath MINs like Clos, Vertical stacked, Extra stage Multistage interconnection Networks. The analysis assumes that fault exists among processors, memory modules and interconnection networks.

The rest of the paper is organized as follows. Section 2 provides architectural description of multipath MIN systems. Proposed strategies and assumptions are presented in section 3. In section 4, Reliability 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.

2. MULTI PATH MIN SYSTEMS

In order to predict the reliability of a system, one must consider its topological features in detail. This section discusses in brief the topological aspects of some multi path multistage interconnection network systems and the properties of non redundant and redundant MINs. 1-path MINs are non redundant, 2-path and multipath MINs are redundant.

1-Path MIN

A MIN design which uses 2 input 2 output switches to provide a single unique pat, from an input I_j to an output O_k, is a non redundant ,I-path MIN. This type of network provides a

simple and inexpensive connection. However single failure disrupts the communication path. An N-input MIN contains n= logN stages. Figure 1 depicts such a 1-path MIN.

2-Path MIN

MINs that provide 2 unique paths from input side to the output side of MIN are 2-path MINs. The purpose of redundant paths is to provide an alternate data path from any input to any output, in the event of failure of any SE . Figure 2 shows a 2-path MIN which contains 1/2N(n-2)SEs with n=(logN+1) stages. The addition of extra stage(thus providing multiple copies of MIN, and increasing the size of SE by adding extra inputs and outputs) has been proposed as a technique for providing redundant paths.

Multi-Path MIN

A multi-path MIN is composed of SEs that have atleast 3 inputs and 3 outputs. The multiple paths between each input-output pair of the network are provided by connecting switching elements within the same stage.

The multipath networks considered are Extra stage omega MIN, Vertical stacked MIN, and Clos network. 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 (SEs). Each stage of the network has N/a switches. 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.

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.

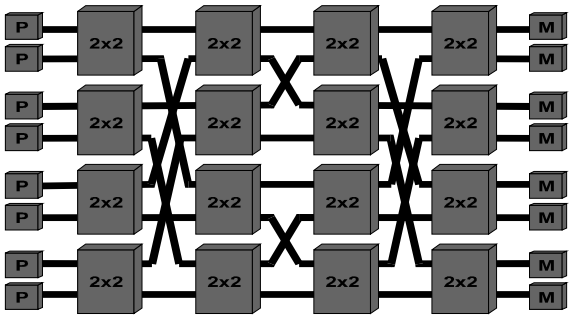

Figure 1. An Extra stage MIN system

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.

Clos Interconnection Network

The Clos network is a special case of multi-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.

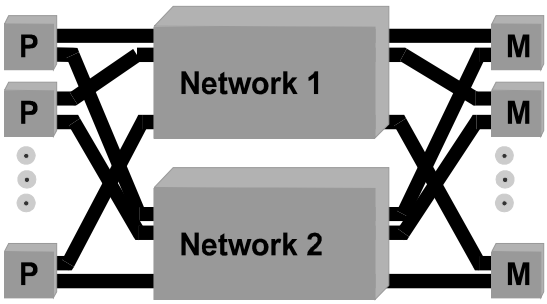

Figure 2. A two vertical stacked MIN system

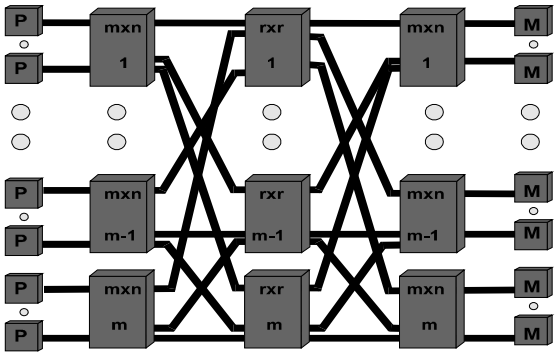

Figure 3. A Clos Network system

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 3, a multiprocessor system with N processors and N memory modules that uses a N (m,n,r) Clos network is presented .

3. PROPOSED STRATEGIES AND ASSUMPTIONS

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

1. Choose the desired metrics of reliability requirements.
2. 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.
3. Solve the reliability model.
4. Analyze the features of reliability model under similar assumptions.

The following assumptions are made throughout this paper for Reliability analysis.

Assumptions

a) The reliability of each sub module is determined independently.
b) The elements of each sub module are identical and have the same failure rate and
c) The failures are exponentially distributed.

4. RELIABILITY MODELING

In this section, the proposed model for studying the reliability of Multi path Multistage Interconnection Networks is presented. The computation of the reliability of multi path MINs are discussed by using analytical techniques.

4.1. Reliability model of Extra stage MIN

In Omega interconnection network, one extra stage of 2x2 SE's results in two paths between any pair of a processor and a memory module. If we add another stage, the paths become four. We can easily extend the reliability model to include the extra stage network because we calculate the reliability of each path independently. 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

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

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

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

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

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

Therefore reliability of MIN system is

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

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,

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

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

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

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

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

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

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

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

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.

4.2. Reliability model of Vertical Stacked MIN

The vertical stacked MINs are identical copies of a unique path MIN such as a Delta network [19]. In this way, each processor is each able to use two or more independent MINs in order to send a request to a memory module. 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 is given by:

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

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

$$\sum_{j=J}^{M} C_{m}^{M-j} \binom{M}{j} (R_{m}(t))^{j} (1-R_{m}(t))^{M-j} \qquad (9)$$

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 multi path MIN.

4.3. Reliability model of Clos Network

We examine a multiprocessor system that includes N processors and N memory modules interconnected via an N(f,n,r) Clos network. The switching elements(SE) in the network are implemented as crossbar networks. In order to study the reliability of the system, we divide it into three sub modules: processors, memory modules and communication paths.

Let λ_p, λ_m, λ_l and λ_{cp} be the failure rates of a processor, a memory module, a link and a crosspoint of SE respectively. Let $R_p(t)$, $R_m(t)$, $R_l(t)$ and $R_{cp}(t)$, represent the processor, the memory, the link and a crosspoint reliabilities respectively. For exponential distribution of failure time with rates λ_p, λ_m, λ_l and λ_{cp} are

$$R_p(t) = e^{-\lambda_p t},\ R_m(t) = e^{-\lambda_m t},\ \ R_l(t) = e^{-\lambda_l t} \text{ and } R_{cp}(t) = e^{-\lambda_{cp}\ t}$$

One request, in-order to reach its destination, has to travel through three SE and two links. We assume that the reliabilities of the links from the processors to the network and from the network to the memory modules are included in the reliability factors of the processors and memories, respectively. We assume that SE is in axb cross bar network. The reliability of an SE when Q crosspoints are required to be operational is given by

$$R_{SE}(t) = \sum_{K=Q}^{b} C_{cp}^{b-k} \binom{b}{k} (R_{cp}(t))^k (1-R_{cp}(t))^{b-k} \tag{10}$$

where C_{cp} is the coverage factor.
The reliability of each communication path is calculated independently by multiplying the reliability of links and the SE's that it employs and is given by the relation

$$R_{path}(t) = e^{-\lambda_{path}} (R_l(t))^2 * \prod_{i=1}^{3} R_{SE\,i}(t) \tag{11}$$

Where λ_{path} is the failure rate of the path, and $R_{SEi}(t)$ is the reliability of the SE situated in the *i* stage. The reliability analysis of the Clos network assumptions are similar to that of multi path MINs, followed with the following exceptions. (i) The number of crosspoint Q in relation (ii) 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. (iii)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

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

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

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 .

5. RESULTS AND DISCUSSIONS

This section considered the performability evaluation of some important Multipath multiprocessor interconnection networks. Fig.4 compares the reliability of various multi path MIN systems. Specifically we evaluate the reliability of an 8x8 extra stage MIN system, 8x8 vertical stacked MIN system and 8(4,4,2) 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 extra stage Omega system and the two vertical stacked Omega system appear to have similar behavior. The existence of two independent interconnection networks makes the system tolerant to the faults of links and SE's that makes communication paths unusable. 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 multi 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. This is caused by the fact that the failure rates of the links and cross points are assumed to be half the failure rate of the processor and the memory modules. Thus, it takes some times for a considerable no of costs to become faulty and to degrade the reliability of the system with the fewer alternative paths. Based on the above analysis, we perform a reliability comparison among systems that employ different Clos Networks. In Fig.5, we compare the reliability of three Clos systems for a given failure rate (The coverage parameters will be assumed to be equal to one.) and for various processor-memory requirements. The plots indicate that the reliability of systems with greater f is higher especially for I, J equal to 4. This was expected since f is the no of alternative paths that connects each processor memory module pair and by increasing f, we increase the redundancy of interconnection networks. It is noticed that the number of processing elements and memory modules increases, then the reliability of Clos system increases.

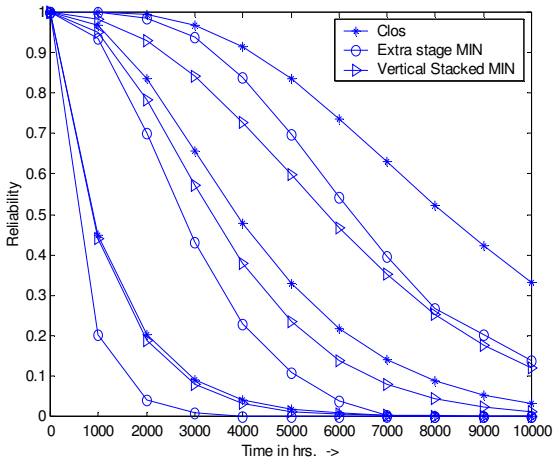

Figure 4. Reliability comparison of multipath MIN systems

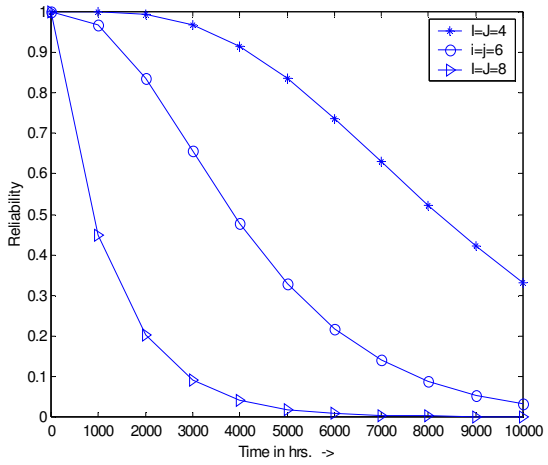

Figure 5. Reliability Comparison of Clos systems

6. CONCLUSIONS

In this paper, the reliability model, evaluation and comparison of various types of multi path MIN system have been carried out. General expressions that permit the calculation of reliability of multiprocessor systems, which employ different multi path MINs have been developed. With the help of these expressions, we are able to study some crucial reliability aspects of general clos networks. The proposed analysis can help us to select the system with the best reliability with the response to the coverage of the network. From the result, it is clear that the Clos networks appear to possess superior reliability and fault tolerant behavior in comparison with extra stage and vertical stacked MIN systems. The reliability of such systems are comparable for substantial period of time. In multi path MIN systems, it is clear that the existence of extra paths, between each processor and each memory increases the reliability of the system. A further increase in the fault tolerant of a network can be achieved with the incorporation of identical copies of multi path MIN networks.

REFERENCES

[1] I.N.Bhuyan, Analysis of Interconnection network, *IEEE Trans. Computers,*C-34(3),279-283,1985.

[2] Duato J, Yalmanchali S, Ni LM. "Interconnection networks an engineering approach", CA: *IEEE Computer Society*; 1997.

[3] J.T.Blake, K.S.Trivedi, Multistage Interconnection network Reliability, *IEEE Trans. Computers,* 38(11),1600-1604,1989.

[4] J.H.Patel, Performance of Processor-Memory Interconnections for Multiprocessors, *IEEE Trans. Computers*, 30(10), 771-780-1981.

[5] J.Arlat, K.Kanoun and J.C.Larpie, Dependability modeling and evaluation of software fault-tolerant systems, *IEEE Trans. Computers,* 39(4),504-514,1990.

[6] 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.

[7] H.J.Siegel, "Study of multistage SIMD interconnection networks," in *Proc.5th Annu.Symp.Comput.,Arch.,Apr. Proc.5thAnnu.Symp.Comput.,Arch.,Apr.*1978,pp. 223-229

[8] V. Cherkassy, E. Opper, M. Malek, "Reliability and fault diagnosis analysis of fault tolerant Multistage Interconnection Networks", *14th Int'l Symp. Fault-Tolerant Computing*, 1984 Jun, pp246- 251.

[9] V.E.Benes, on Re arrange able three- stage connecting networks, Bell Systems Tech. J.XLI(Sept 1962) 1481-1492.

[10] B.D.Douglass and A.Y.Oruc ,on self routing in Cl;os Connection Networks. IEEE Trans. Comm.. 41(Jan 1993) 121-124.

[11] J. Blake, K. Trivedi, "Multistage interconnection reliability", *IEEE Trans .Computers,(to appear)* .

[12] K. Hwang, T. Chang, "Combinatorial reliability analysis of multiprocessor computers", *IEEE Trans. Reliability vol R-31, 1982 Dec, pp469-473*

[13] V. Cherkassy, M. Malek, "Reliability and fail-softness analysis of Multistage Interconnection Networks", *IEEE Trans. Reliability*, vol R-34, 1985 Dec, pp 524-527.

[14] Y.Chang and L.N. Bhuyan, "Combinatorial analysis of subsube reliability in hypercubes," Tech. Report TR #94-059, Texas, A&M Univ., 1994

[15] Chuan-Bi Lin; Rojas-Cessa, R, "Module Matching Schemes for Input-Queued Clos-Network Packet Switches", IEEE, Volume 11, Issue 2, pp. 194~196, Feb.2006.

[16] Lee H.Y., Hwang F. K., Carpinelli D., "A new decomposition algorithm for rearrangeable Clos interconnection net-works". IEEE Trans on Commun., vol. 44, no. 11, pp. 1572-1578, 1996.

[17] Yuanyuan Yang, Jianchao Wang, " A Fault-Tolerant Rearrangeable Permutation Network", IEEE Trans. On Computers, vol. 53, no. 4, April 2004.

[18] Adams III B,et al. A survey and comparision of fault-tolerant multistage interconnection networks. IEEE Computer Magazine, 1987:14-27.

[19] Lea CT, Shyy DJ.Tradeoff of horizontal decomposition versus vertical stacking in rearrangeable Non blocking networks. IEEE Trans Commun 1931;39(6): 899-904.

Authors

1. **Sudarson Jena** received Ph.D degree in Computer Science from Sambalpur University, Odissa in 2008. He is currently working as Associate Professor, Department of Computer Science Engineering & IT, GITAM University, Hyderabad (AP), India. He is a life member of CSI and ISTE and member of IEEE. He has published/ presented more than 25 papers in reputed National, International Journals and Conferences. Currently he is guiding 4 Ph.D scholars under JNTU Hyderabad. His research interests include Parallel and Distributed Systems, Interconnection Networks, Soft Computing, Wireless and Sensor Networks etc.

2. **G. Sri Sowmya** is presently working as Assistant Professor in the Department of Computer Science Engineering & IT, GITAM University, Hyderabad. She received M.Tech degree in the year 2008 from Andhra University. Her areas of interests include Computer Networks, Parallel and Distributed Systems.

3. **P.Radhika** is presently working as Assistant Professor in the Department of Computer Science Engineering & IT, GITAM University, Hyderabad. She received M.Tech degree in the year 2007 from Andhra University. Her areas of interests include Parallel and Distributed Systems and Data Mining.

4. **P.Venkat Reddy** is presently working as Assistant Professor in the Department of Computer Science Engineering & IT, GITAM University, , Hyderabad. He received M.Tech degree in the year 2010 from Nagarjuna University. His areas of interests include Image Processing, Parallel and Distributed Systems.