

Design and Reliability Analysis of a Class of Irregular Fault-tolerant Multistage Interconnection Networks

Rinkle Rani Aggarwal

Department of Computer Science & Engineering,
Thapar University, Patiala –147004 (India)

ABSTRACT

Parallel processing can be used to design high performance computing systems. In a parallel computer inter-connecting processors and linking them efficiently to the memory modules is not an easy task. Therefore, there is a requirement of an interconnection network that provides the desired connectivity and performance at minimum cost. Multistage interconnection networks (MINs) provide cost-effective, high-bandwidth communication between processors and/or memory modules in comparison to bus and crossbar interconnection networks. In this paper a new irregular MIN named IFTN (Improved Four Tree Network) has been proposed. The performance of IFTN has been measured in terms of reliability and cost. It has been proved that the proposed network IFTN provides much better fault-tolerance and reliability at lesser cost In Comparison To Four Trees.

Keywords: Multistage Interconnection Network, Reliability, Four Tree Network, Fault-tolerance, MTTF.

1. INTRODUCTION

Today is the era of parallel processing and building of multiprocessor system with hundred processors is feasible. Advances in LSI and VLSI technology are encouraging greater use of multiple-processor systems with processing elements to provide computational parallelism and memory modules to store the data required by the processing elements. Interconnection Networks (INs) play a major role in the performance of modern parallel computers. Many aspects of INs, such as implementation complexity, routing algorithms, performance evaluation, fault-tolerance, and reliability have been the subjects of research over the years. There are many factors that may affect the choice of appropriate interconnection network for the underlying parallel computing environment. Though crossbar is the ideal IN for shared memory multiprocessor, where N inputs can simultaneously get connected to N outputs, but the hardware cost grows astronomically. Multistage Interconnection Networks (MINs) are recognized as cost-effective means to provide programmable data paths between functional modules in multiprocessor systems [1]. These networks are usually implemented with simple modular switches, employing two-input two-output switching elements. Most of the MINs proposed in the literature have been constructed with 2×2 crossbar switches as basic elements, and have $n = \log_2 N$ switching stages with each stage consisting of $N/2$ elements, which makes the cost of this network as $O(N \log N)$, as compared to $O(N^2)$ for a crossbar [2]. The pattern of interconnection may be uniform or non-uniform, which classifies the MINs to be regular or irregular respectively. In the case of irregular networks, the path length varies from any

input to any output, in contrast with regular networks, where it is the same. Fault-tolerance in an interconnection network is very important for its continuous operation over a relatively long period of time. Many networks have been designed and proposed to increase the fault-tolerance in the literature [3], [4,5,6,7]. Permutation capability and other issues related to routing have also been extensively researched [8,9,10]. Various routing schemes have also been studied in-depth [11,12,13]. However, little attention has been paid to the computation of reliability of these networks. Reliability is measured in terms of Mean Time to Failure (MTTF), which is evaluated using simple series-parallel probabilistic combinations. This analysis is based upon the lower and upper bounds of the network reliability. This paper has been organized into five sections whose details are as follows.

Section 1 introduces the subject under study. Section 2 describes the structure and design of network. Section 3 focuses on the routing scheme. Section 4 concentrates on the reliability analysis of IFTN network. Section 5 describes the cost effectiveness. Finally, the conclusion has been presented.

2. STRUCTURE AND DESIGN OF IFTN

The IFTN (Improved Four Tree Network) network is derived from FT network [14]. IFTN has one stage less in comparison to FT. An IFTN of size 16×16 ($N=16$) is shown in figure. 1. It contains total $(2^{n+2} - 8)$ switches, where $n = \log_2 N$. There are 2^{n-1} switches of size 2×2 and rest of size 3×3 . This network is constructed with the help of two identical groups, which are arranged one above the other. The two groups are formed based on the most significant bit (MSB) of the source destination terminals. Every 3×3 SE (switching element) in a stage forms a loop with the corresponding numbered 3×3 SE of other sub-network in the same stage. Every source and destination is connected to both the subgroups by means of multiplexers and demultiplexers

An IFTN network being an irregular network supports multiple paths of different path lengths. It also inherits the property of regular network since the numbers of switches are same in all the stages except the first and last stage of the network. The main advantage of this network is that if both the switches in a loop are simultaneously faulty even then some sources are connected to the destination. Therefore IFTN is more fault-tolerant. Here the maximum path length is 4 which is one less compared to FT. Thus there is a benefit in terms of cost and path length.

The sources S and destination D be represented in binary code as:

$$S = S_{n-1} \dots S_1 S_0$$

$$D = D_{n-1} \dots D_1 D_0$$

The sources and destinations are connected to the multiplexers and demultiplexers are connected as follows:

- i If $(S_{n-2} \dots S_1 S_0)$ bits are same for two sources, then these two sources are linked through the same pair of multiplexers.
- ii If $(D_{n-2} \dots D_1 D_0)$ bits are same for two destinations, then these two destinations are linked through the same pair of demultiplexers.

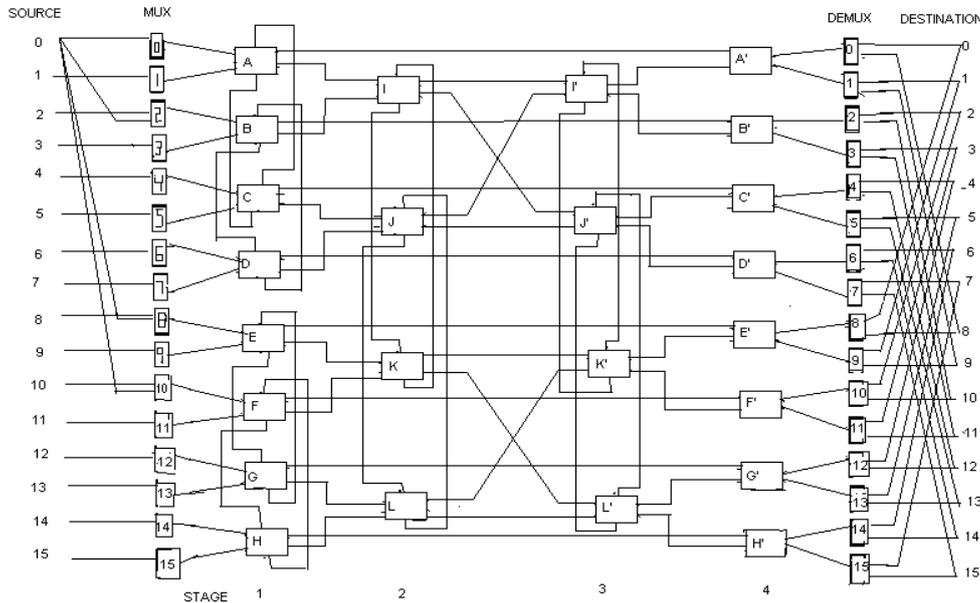


Fig 1: An IFTN of size 16 x 16

3. ROUTING SCHEME

3.1 Redundancy graph

A redundancy graph depicts all the available paths between a source and a destination in a MIN. It offers a convenient way

to study the properties of a multi-path MIN, such as the number of faults tolerated and the type of rerouting possible. It consists of two distinguished nodes source S and destination D. The rest of the nodes correspond to the switches that lie along the paths between S and D as shown in figure 2.

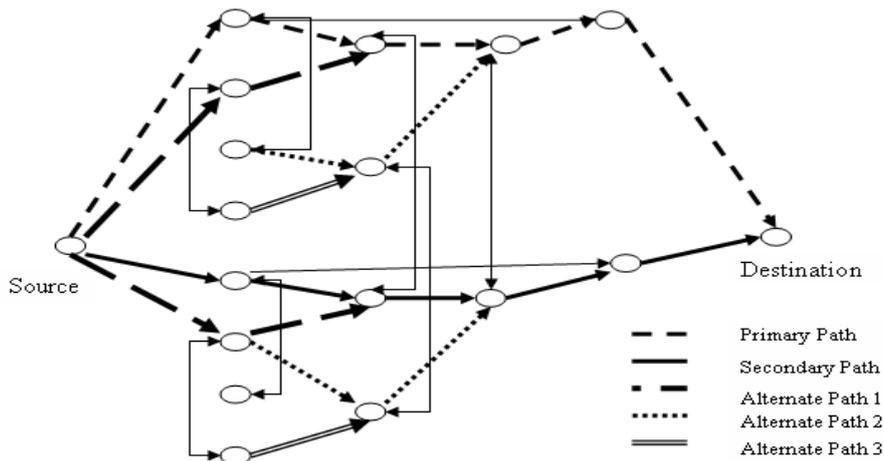


Fig 2: Redundancy Graph for IFTN Network

A request from source S to a given destination D is routed through the IFTN as:

- i The source S selects one of the sub network G^i based on the most significant bit of the destination D ($i=d_0$).
- ii Each source attempts an entry into the IFTN via its primary path. If the primary path is faulty (i.e. either multiplexer or primary switch or both are faulty), then the request is routed to the secondary path.
- iii To route a request through a network there are favourable path and less favourable paths. The path

length algorithm is used to determine whether a request can be routed through the most favourable path or not. If most favourable path is not available or is busy then there is need of alternate path. If the alternate path is also not available, because it is busy or faulty, then drop the request.

- iv A fault in the demultiplexer at the output of a switch in a stage $(2m - 1)$ is regarded as a fault in that switch. From the demultiplexer, the request is routed to the upper or lower destinations according to the least significant bit of the tag (i.e. d_{n-1}).

3.2 Path Length Algorithm

The algorithm for allocation of path length gives the information about different possible paths between a source-destination pair. The possible path lengths between a particular source-destination pair varies from 2 to 2_{m-1} for a $2^n \times 2^n$ network, depending upon the addresses of the source and destination terminals. For a given source-destination pair, there exist multiple paths of different lengths in IFTN network.

The path length algorithm is:

```

If
    [(Sn-2 ⊕ Dn-2) + (Sn-3 ⊕ Dn-3) + ..... +
    (S1 ⊕ D1)] Is zero
    (⊕ Represents an exclusive-OR and
    + represents a logical OR operator)
Then
    Minimum path length is 2 and all other paths of
    different lengths are possible i.e. paths of
    length 2, 4, 6.... (2m-2), (2m-1).
Else
    If
        [(Sn-2 ⊕ Dn-2) + (Sn-3 ⊕ Dn-3) + ..... +
        (S2 ⊕ D2)] Is zero
Then
    All paths of length equal to or greater than 4 are
    possible
Else
    If
        [(Sn-2 ⊕ Dn-2) + (Sn-3 ⊕ Dn-3) + ..... +
        (Sj ⊕ Dj)] Is zero
        {where 1 ≤ j ≤ (n-2)}
    Then
        All paths of length equal to or greater
        than 2j are possible.
Else
    
```

Path of length 2_{m-1} (i.e. longest path) is possible only.

3.3 Routing Tag Algorithm

The following algorithm is used to generate control tag for IFTN network which is required to establish a path between any source-destination terminal pair for a given path length (if it exists).

```

If
    2 ≤ x ≤ (2m-1)
Then
    Routing tag = 0.D1.Dn-1
Else
    If
        x = (2m-1)
    Then
        Routing tag = 1.D2.D1.D0.Dn-1
    Else
        No tag is possible.
    
```

Example : Let Source=0000 and Destination=0100 then routing tag =1.1.0.0.0. Following are the possible paths between source 0000 and destination 0100 as shown in figure 3.

Primary Path:

```

0 → MUX(0) → A → I → J' → C' → DEMUX(4) → 4
0 → MUX(2) → B → I → J' → C' → DEMUX(4) → 4
    
```

Secondary Path:

```

0 → MUX(8) → E → K → L' → G' → DEMUX(12) → 4
0 → MUX(10) → F → K → L' → G' → DEMUX(12) → 4
    
```

Alternate Path:

```

0 → MUX(0) → A → C → J → J' → C' → DEMUX(4) → 4
0 → MUX(2) → B → D → J → J' → C' → DEMUX(4) → 4
0 → MUX(8) → E → G → L → L' → G' → DEMUX(12) → 4
0 → MUX(10) → F → H → L → L' → G' → DEMUX(12) → 4
    
```

From figure 3 it is clear that there exist eight paths between a given source and destination pair in IFTN network. Whereas in case of FT there exists only 2 paths. Therefore the proposed network IFTN can entertain more number of requests even under faults in comparison to FT. Thus IFTN is more fault-tolerant than FT.

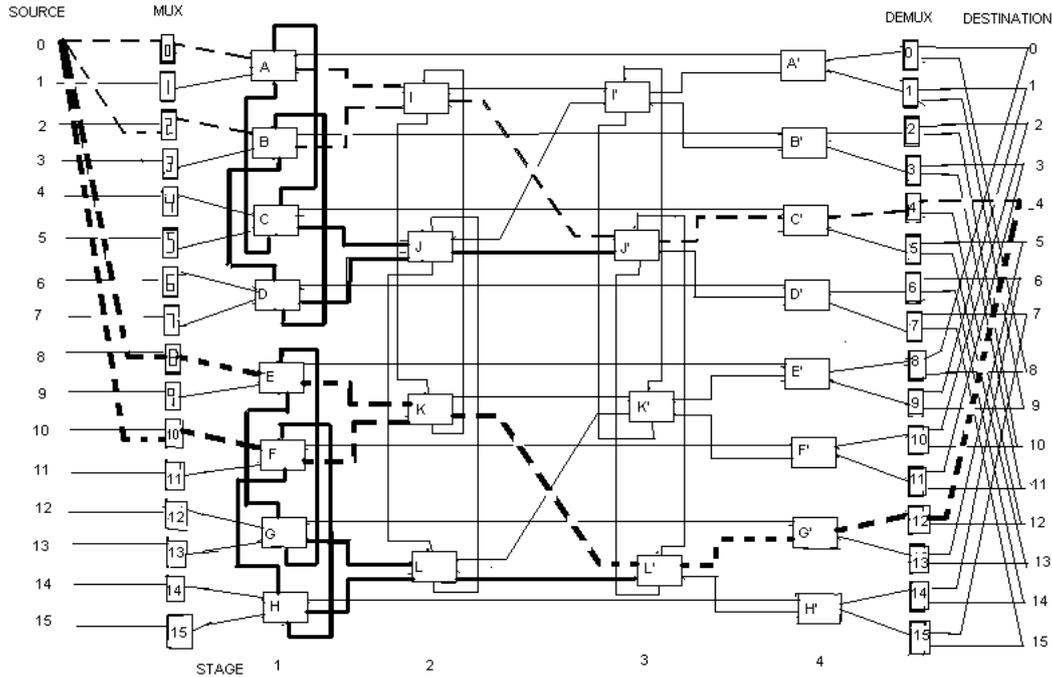


Fig 3 : Routing in IFTN

4. RELIABILITY ANALYSIS

Reliability of IFTN network is analyzed in terms of Mean time to Failure (MTTF) using simple series-parallel probabilistic models. The assumptions used in the analysis on the failure rates of the components are given below:

- i Failure rate of 2x2 crossbar switches is λ (Where, $\lambda = 10^{-6}$ per hour) [15].
- ii The failure rate of a component can be derived from its gate count. The number of gates in a 2x2 crossbar switch is approximately equal to that in a 2x1 MUX or a 1x2 DEMUX. Thus $\lambda_m = \lambda_d = m\lambda/2$ for a mx1 MUX, where λ_m and λ_d are failure rates of MUX and DEMUX respectively.
- iii The adaptive routing scheme considers a 2x2 switch in the last stage and its associated DEMUX as a series system, so based on gate count, a failure rate of $\lambda_{2d} = 2\lambda$ can be assigned to this group of elements.
- iv Let λ_3 is the failure rate for 3x3 switches, then $\lambda_3 = 2.25\lambda$ and $\lambda_{3m} = 4.25\lambda$.

4.1 Optimistic (Upper Bound) Analysis of IFTN network

In IFTN each source is connected to two multiplexers and each SE in the first stage has a conjugate pair. The switches in a conjugate pair are connected to the same switches in the next stage So, to estimate the upper bound, it is assumed that the IFTN is operational as long as one of the two multiplexers attached to a source is operational and both components in a conjugate pair are not faulty, then the network permit as many

as one-half of the system components to fail and the IFTN can still be operational. The reliability block diagram for the upper bound is shown in figure 4.

Reliability Equations are

$$R_{\text{optimistic}}(t) = f1 * f2 * f3 \quad \text{where,}$$

$$f1 = \left[1 - (1 - e^{-\lambda_m t})^2 \right]^{(N/2)}$$

$$f2 = \left[1 - (1 - e^{-\lambda_{3t}})^2 \right] \left\{ \binom{N/4}{1} + \binom{N/4}{2} + \dots + 1 \right\}^{\Delta}$$

$$f3 = \left[1 - (1 - e^{-\lambda_{2d} t})^2 \right]^{(N/4)}$$

* is the term added for path-length>2
 ° is the term added for the largest path-length
 △ is the term added for path-length>2

$$MTTF = \int_0^{\infty} R_{\text{Optimistic}}(t) dt$$

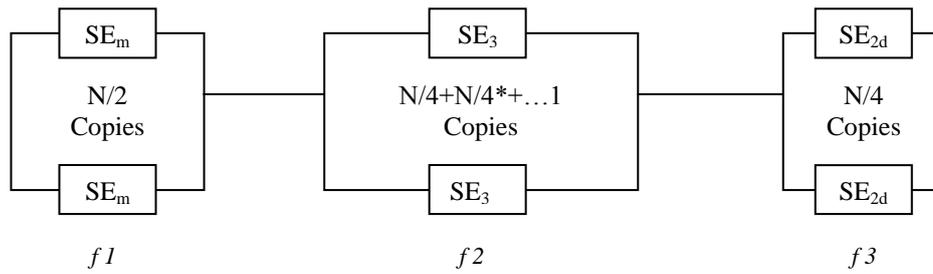


Fig 4: Reliability Block Diagram for the Upper Bound of IFTN

4.2.2 Pessimistic (Lower Bound) Analysis of IFTN network

At the input side of the IFTN, the routing scheme does not consider the multiplexers to be an integral part of a 3 x 3 switch. For example, as long as at least one of the two multiplexers attached to a particular switch is operational, the switch can still be used for routing. Hence, if we group two multiplexers with each switch in the input side and consider them a series system (SE_{3m}), their aggregate failure rate will be λ_{3m} = 4.25λ. Finally these aggregated components and the switches in the intermediate stages can be arranged in pairs of conjugate loops. To obtain the pessimistic (lower) bound on the reliability of IFTN, we assume that the network is failed whenever more than one conjugate loop has a faulty element or more than one conjugate switch in the last stage fails. The reliability block diagram is shown in figure 5. Reliability Equations are

$$R_{\text{pessimistic}}(t) = f1 * f2 * f3$$

$$f1 = \left[1 - (1 - e^{-\lambda_{3m}t})^2 \right]^{N/4}$$

$$f2 = \left[1 - (1 - e^{-\lambda_3t})^2 \right]^{\left\{ \left(\frac{N}{4} \right)^* + \dots + 1 \right\}^A}$$

$$f3 = \left[1 - (1 - e^{-\lambda_{2d}t})^2 \right]^{N/4}$$

$$MTTF = \int_0^{\infty} R_{\text{Pessimistic}}(t) dt$$

The results of the MTTF Reliability equations have been shown in Table 1.

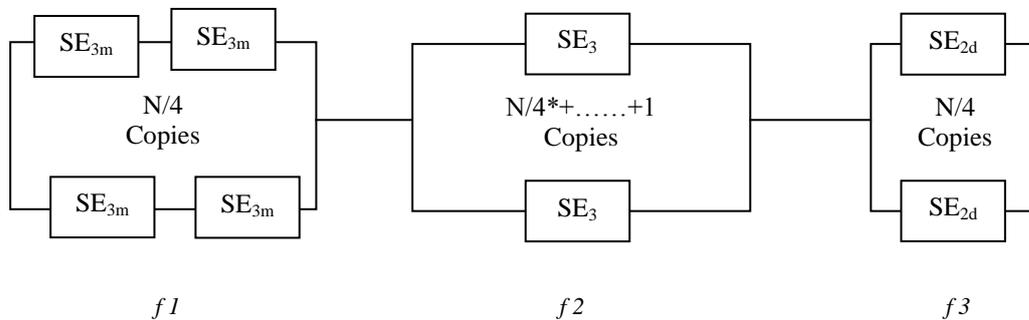


Fig 5: Lower Bound for IFTN

Table 1. MTTF of FT and IFTN

Network Size (LogN)	FT		IFTN	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound
4	5.15455	5.15455	5.76648	5.78874
5	4.97848	4.97848	5.56741	5.59737
6	4.80956	4.80956	5.37547	5.41119
7	4.64585	4.64585	5.18877	5.22874
8	4.48591	4.48591	5.00587	5.04894
9	4.32869	4.32869	4.82573	4.87104
10	4.17339	4.17339	4.64755	4.69449

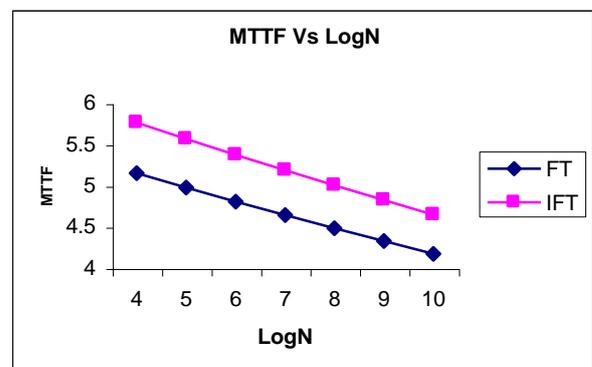


Fig 6. MTTF (Lower Bound) comparison of FT and IFTN

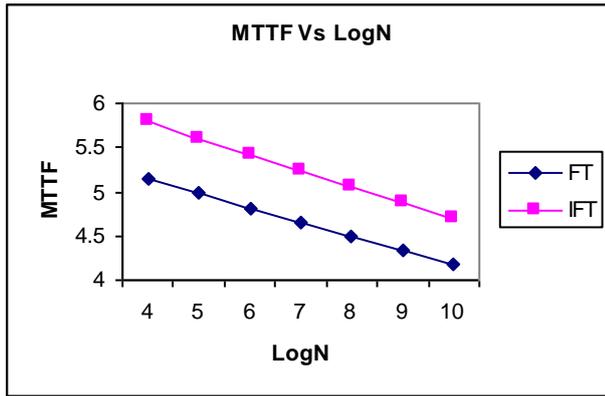


Fig 7. MTTF (Upper Bound) comparison of FT and IFTN

Figures 6 and 7 depict that the proposed network IFTN is more reliable than the existing FT network for both the upper and lower bounds of reliability.

5. COST ANALYSIS

In order to estimate cost of a network it has been assumed that the cost of a switch is proportional to the number of cross-points within that switch [16]. Therefore, a 2x2 switch has 4 units of hardware cost whereas cost of a 3x3 switch is 9 units. The cost functions for the proposed and existing MIN are given in the Table 2.

Table 2 . Cost Functions for Networks

Network	Cost Function
FT	$(9.75 \cdot 2^{n+1} - 54)$ (Bansal et al., 1992)
IFTN	$(9.75 \cdot 2^{n+1} - 9N)$

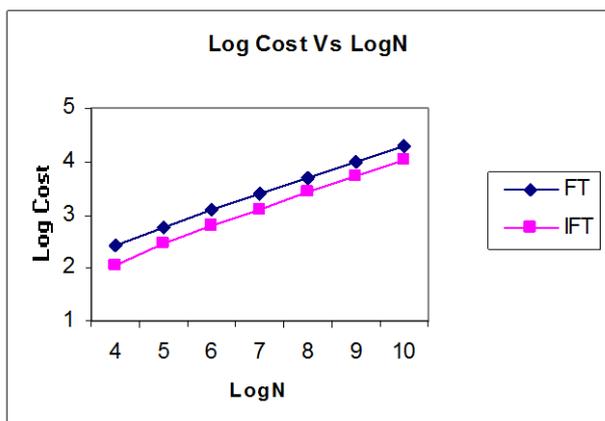


Fig 8: Cost Comparison of FT and IFTN

From figure 8 it is clear that IFTN is more cost effective than FT.

6. CONCLUSION

An Improved Four Tree Network (IFTN) is designed using existing Four Tree (FT) network. It has comparatively has one lesser stage and thus reduced cost than FT. IFTN is a dynamically re-routable network that provides multiple paths of varying lengths between a source-destination pair. It has been found that in IFTN, there are eight mutually exclusive distinct possible paths between any source-destination pair, whereas FT has only two such paths. Thus the new network IFTN can entertain larger number of requests even under faults. The upper and lower bound reliability analysis shows that IFTN is more reliable than FT. Thus the new network IFTN provides better fault-tolerance and reliability than the existing FT with lesser cost.

7. REFERENCES

- [1] Bhuyan, L.N., Qing, Y., Agrawal, D.P., 1989. Performance of Multiprocessor Interconnection Networks. *Computer*, 22(2), 25-37.
- [2] Gupta, A., Bansal P.K., 2008. Evaluation of Fault-tolerant Multistage Interconnection Networks. *CSI-Communication*, 1-16.
- [3] Sengupta, J., Bansal, P.K., 2001. High speed Dynamic Fault-Tolerance. IEEE Region 10 International Conference on Electrical and Electronic Technology, 669-675.
- [4] Bansal, P. K., Singh, K., Joshi, R.C., Siroha, G. P., 1991. Fault-tolerant Augmented Baseline Multistage Interconnection Network. IEEE International Conference TENCON-91, 200-204.
- [5] Sharma S., Kalhon K.S., Bansal P.K., Singh K., 2008a. Improved Irregular Augmented Shuffle Exchange Multistage Interconnection Network. *International Journal of Computer Science and Security*, 2(3), 28-33.
- [6] Nitin, 2006. On Analytic Bounds of Regular and Irregular Fault-tolerant Multistage Interconnection Networks. International Conference PDPTA-06, 221-226.
- [7] Hou, G., Yang, Y., 2000. Super Recursive Baselines: A Family of New Interconnection Networks with High Performance/Cost Ratios. IEEE International Symposium on Parallel Architectures, Algorithms and Networks ISPAN- 2000, 260-265.
- [8] Sengupta, J., Bansal, P.K., Gupta, A., 2000. Permutation and Reliability Measures of Regular and Irregular MINs. IEEE International Conference TENCON-2000, 531-536.
- [9] Sharma, S., Kahlon, K.S., Bansal, P.K., 2008b. On a class of Multistage Interconnection Networks in Parallel Processing. *International Journal of Computer Science and Network Security*, 8(5), 287-291.
- [10] Subramanyam, A., Prasad, E.V. & Nadamuni, R. (2006) Permutation Capability and Connectivity of Enhanced Multistage Interconnection Network (E-MIN). IEEE International Conference on Advanced Computing and Communications, 8-11.

- [11] Seok, S., You, B., Youm, S., Kim, K., Kang C., 2010. A Heuristic Multi-path routing scheme for online traffic in MPLS Networks. *International Journal of Computer Systems Science and Engineering*, 25(1).
- [12] Tang, Y., Zhang, Y., Chen, H., 2009. A parallel shortest path algorithm based on graph-partitioning and iterative correcting. *International Journal of Computer Systems Science and Engineering*, 24(5).
- [13] Zhen, C., Zengji L. Zhiliang, Q., Peng, C., Xiaoming, T., 2005. Balance Routing Traffic in Generalized Shuffle-Exchange Network. *Journal of Electronics*, 22(4), 345-350.
- [14] Bansal, P. K., Singh, K., Joshi, R.C., 1992. Routing and path length algorithm for a cost-effective Four-Tree Multistage Interconnection Network. *International Journal of Electronics*, 73(1), 107-115.
- [15] Bansal, P. K., Joshi, R.C., Singh, K., 1994. On a Fault-tolerant Multistage Interconnection Network. *International Journal of Electronics and Electrical Engineering*, 20(4), 335-345.
- [16] Aggarwal, H., Bansal, P.K., 2002. Routing and Path Length Algorithm for Cost-effective Modified Four Tree Network. IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering, 293-297.