Permutation & Reliability Measures of Regular and Irregular Multistage Interconnection Networks

A Dissertation Submitted
in partial fulfillment of the
requirements for the degree of
MASTER OF ENGINEERING
IN
COMPUTER SCIENCE

by

AJAY GUPTA

Supervised by

Dr. P.K. Bansal

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
THAPAR INSTITUTE OF ENGINEERING & TECHNOLOGY
(Deemed University)
PATIALA – 147 001

September, 1999

CERTIFICATE

I hereby certify that the work presented in the thesis entitled "Permutation & Reliability Measures of Regular and Irregular Multistage Interconnection Networks" in fulfillment of the requirement for the award of the Degree of Master of Engineering in Computer Science is being submitted in the Department of Computer Science & Engineering, TIET, Patiala.

It is further certified that the work carried out for the thesis is an authentic record done under the supervision of Dr. P.K. Bansal. The matter presented in this dissertation has not been submitted by me for the award of any other degree of this or any other university.

(AJAY GUPTA)

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

PKBansal
22/9/99
(Dr. P.K. Bansal)
Professor
Department of
Computer Sc.& Engg.
TIET, Patiala

RKSharma
22.9.99
Head
Department of
Computer Sc. & Engg.
TIET, Patiala

22.9.99.
Dean
Academic Affairs
TIET, Patiala

ACKNOWLEDGEMENTS

I would like to thank Thapar Institute of Engineering and Technology, Patiala for providing me an opportunity to carry out the thesis on **"Permutation & Reliability Measures of Regular & Irregular Multistage Interconnection Networks"** as a part of ME degree program.

I consider it a great privilege to express my deep sense of gratitude to Dr. P.K. BANSAL, for his invaluable guidance, constant inspiration and critical review throughout the course of this work. He has always been more than generous in sparing his very valuable time.

The cooperation and help extended by the Head, Department of Computer Science & Engineering, is gratefully acknowledged.

I would fail in my duty if I don't avail this opportunity to thank my parents for their encouragement and moral support.

I am also thankful to all my fellow friends Anurag Joshi, Vikas Kalra, Himanshu Aggarwal, Rajesh Bhatia, who always showered their good wishes on me. I am also grateful to Mrs. Jyotsna Sengupta for her valuable suggestions time to time.

Finally, my sincere appreciation's goes to my wife Anjali and my little son Ankit, for their cooperation and untiring patience in many odd hours during the period.

----- Ajay Gupta

ABSTRACT

The basic way to achieve high-performance, high-reliable computing is to use multiple processors linked by one or more shared buses. Such multiple bus systems offer the advantages of low hardware cost, high communication bandwidth, and graceful degradation in the presence of faults. The performance of multiprocessor systems rests primarily on the design of the Interconnection Network (IN). Multistage Interconnection Networks (MINs) provides a compromise between the time shared and the crossbar networks.

In this thesis, Irregular and Regular MINs are analyzed in terms of permutation passable, reliability and cost. Also a new MIN named Smart Four Tree (SFT) has been proposed in this work modifying the existing Four Tree (FT) MIN.

The permutation passable determines the data routing capability, which dominates performance for large sizes. This parameter is analyzed in terms of identity and incremental permutation layouts and results shows that Irregular MINs have better permutations as compared to Regular MINs. The results of the Reliability evaluated in terms of Meantime to Failure (MTTF) shows that FT, FDOT have better reliability than ASEN, ABN MINs. It is also found that MTTF-Cost ratio is better for all Irregular MINs as compared to Regular MINs, thus showing the supremacy of Irregular over Regular networks.

The results of the analysis of SFT shows that it is better in terms of permutation passable than the existing Regular and Irregular MINs. Reliability of SFT is almost comparable to that of FT. Thus the proposed SFT provides an improvement over the performance of existing Irregular MINs.

CONTENTS

I. Introduction

II. Multistage Interconnection Networks

III. Permutations

IV. Reliability Analyses of Regular & Irregular Networks

V. Conclusions & Further Scope

REFERENCES

PAPER COMMUNICATED

CHAPTER-1

INTRODUCTION

1.1 MULTIPROCESSORS

Multiprocessor is a computer system characterized by the presence of several CPUs or, processing elements (PEs), which cooperate on common or shared computational tasks. Multiprocessors are distinguished from multi-computer systems and computer networks, which are systems with multiple PEs operating independently on separate tasks. The various PEs making up a multiprocessor typically share such resources as communication facilities, I/O devices, program libraries, databases etc., and are controlled by a common operating system.

The two main reasons for including multiple PEs in a single computer system are to improve performance and to increase reliability. Performance (throughput) improvement is obtained either by allowing many PEs to share the computation load associated with a single large task, or else by allowing many smaller tasks to be performed in parallel in separate PEs. System reliability is improved by the fact that the failure of one CPU need not cause the entire system to fail. The functions of the faulty processor (and its local resources) can be taken over by the other resources; consequently, multiprocessors allow fault-tolerance to be incorporated into the system.

1.2 MULTIPROCESSOR TYPES

Multiprocessors can be classified by the organization of their main memory systems. This give rise to the following classifications:

1.2.1 SHARED MEMORY COMPUTER

If main memory can be directly accessed by all the PEs of a multiprocessor then the system is termed as a shared-memory computer and the shared portion of main memory is called global memory. Information can therefore be shared among the processors simply by placing it in the global memory as shown in the Figure 1.1

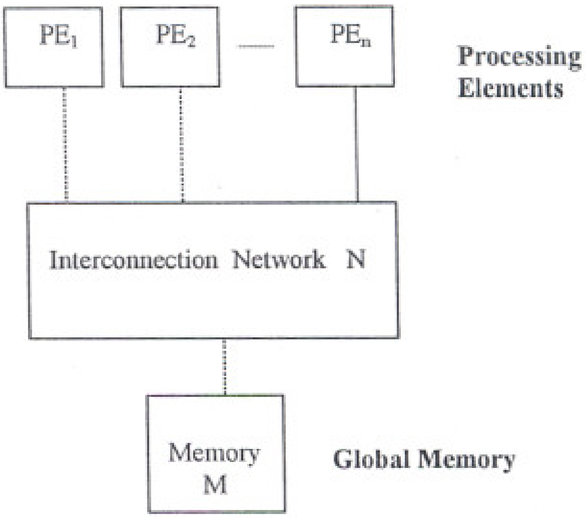


Fig 1.1: Shared-Memory systems

1.2.2 DISTRIBUTED MEMORY COMPUTER

Distributed memory computers i.e. P_1, P_2 ……P_n have no global memory. Instead each Processing Element (PE) has its own private or local main memory M_1, M_2 …… M_n shown in the figure 1.2

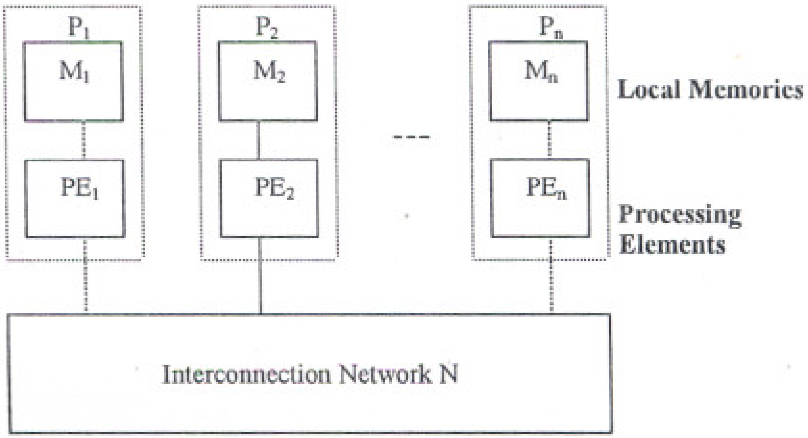


Fig 1.2: Distributed Memory Systems

Shared-memory and distributed-memory multiprocessors are also called tightly coupled and loosely coupled, respectively. Multiprocessors are sometimes classified by the interconnection structures employed to support processor-processor and processor-memory communication as shown in the Figure 1.3.

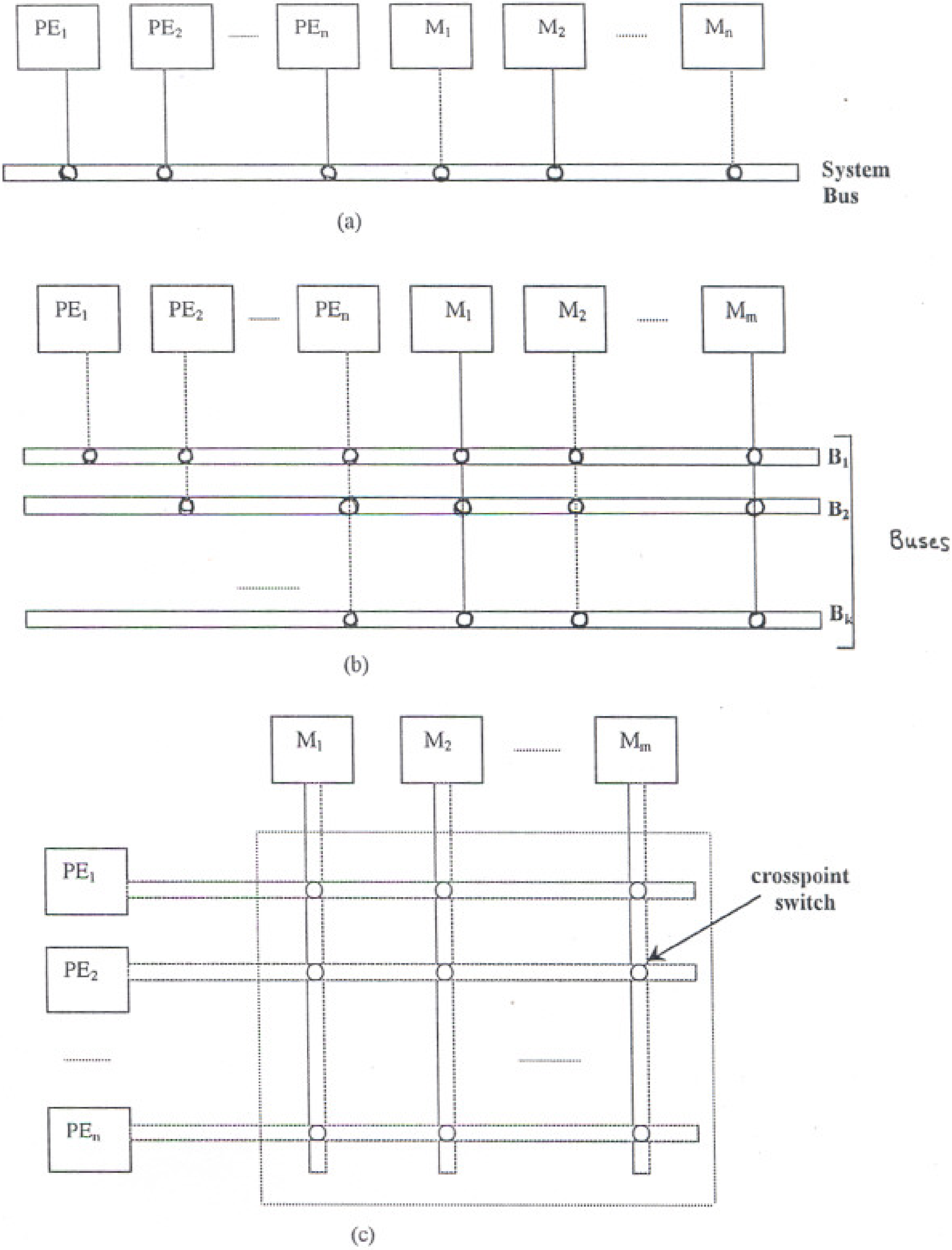


Fig 1.3: Interconnection network structures a) Single bus; b) Multiple buses c) Crossbar Network

For shared memory multiprocessors, the single system bus, as shown in Figure 1.3 (a), typically used to connect the processors to the memory and system performance can be increased considerably by increasing the capacity of the bus. One way of increasing the bus capacity, and also the system's reliability and fault-tolerance, is to increase the number of buses. In multiple bus system several processors may attempt to access the shared memory simultaneously.

1.3 INTERCONNECTION NETWORKS (INs)

An interconnection network (IN) is a complex connection of switches and links permitting processors in a multiprocessor system to communicate among themselves or with memory modules. The main criterion for designing an IN is that it should be able to provide fast, reliable, efficient and fault tolerant communication at a reasonable cost. Any IN has three operational characteristics. These are based on its timing, switching and control.

The timing control of an IN can be either synchronous or asynchronous. Synchronous systems are characterized by a central global clock that broadcasts the clock signal to all devices on the IN so that they operate in a lockstep fashion. Asynchronous systems on the other hand support independent operation of the devices without a global clock.

The IN transfers data using either circuit switching or packet switching. In circuit switching, once a device is granted a path in the IN it will occupy that path for the duration of the data transfer. In packet switching, the information is broken into small packets that individually compete for a path in the IN.

Based on the overall control of the network, an IN may be classified as centralized or decentralized. In centralized control, a global controller receives all requests and transmits the messages in the IN. In a decentralized system, requests are handled independently by different devices in the IN.

A complete interconnection such as provided by crossbar network, as shown in Figure 1.3(c), may become cost prohibitive for larger systems.

In crossbar based multiprocessors, contention may occur only when more than one processor attempts to access the same memory at the same time. The crossbar scheme thus allows high degree of parallelism between unrelated tasks, but memory contention is likely if inter-process and inter-processor synchronization are based on shared memory. A fully connected crossbar networks have a cost $O(N^2)$, which may be prohibitive for large N. An MxN Crossbar allows upto a maximum {M,N} transactions simultaneously, while a single time shared bus, as shown in Figure 1.3 (a) offers minimum bandwidth (one request per unit time). It also leads to bus contention problem.

As a compromise between these two extremes (time-shared and crossbar network), multistage interconnection networks (MINs) were proposed. MINs consist of more than one stages of small interconnection networks called switching elements (SEs). A SE is mxm crossbar switch (i.e. m number of inputs and m number of outputs) as shown in Figure 1.4 with four interconnection states: straight, exchange, upper broadcast and lower broadcast.

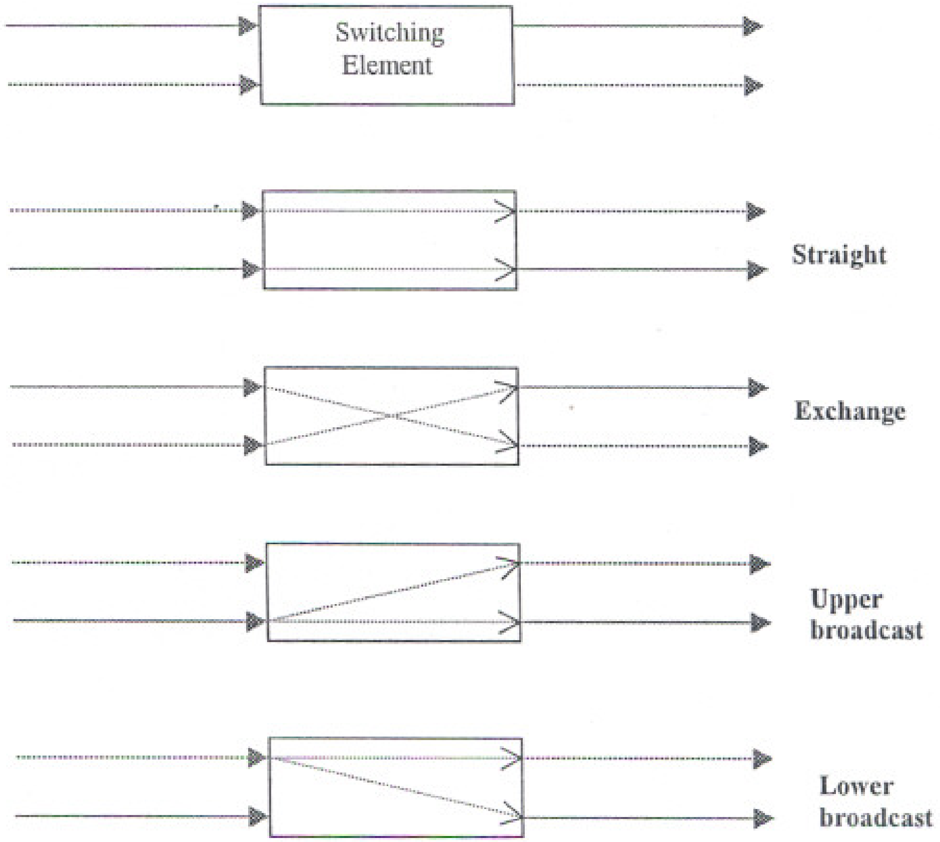


Fig 1.4 A 2x2 Switching Element with four interconnection states

A fundamental requirement of a MIN is that it should be able to connect every source to every destination by at least one of its configurations. This termed as Full Access property. MINs can be categorized as Flip controlled or Distributed controlled.

Flip controlled MINs have a common control signal for setting various switching elements in a given stage, while distributed controlled networks have got a separate control signal for every SE. Flip control is less complicated due to lesser number of control signals, but has a lesser bandwidth. Distributed control is quite complex but bandwidth is high due to flexibility in the selection of source-destination pair at a given time.

1.4 MULTISTAGE INTERCONNECTION NETWORKS (MINs)

MIN can be broadly classified as:

Unique Path MINs
Multi Path MINs

Unique path MINs provides a unique path between every source and destination. Failure of any SE along this path disconnects some of the source-destination pairs, thereby affecting the full access property of the network. These networks are not even a single fault-tolerant and hence not much reliable for large multiprocessor system.

Multi-path MINs provides more than one path for a given source-destination pair. In case of failure of a SE in one path, request gets honoured through another, thereby maintaining full access property of the network. Fault-tolerant nature of these networks make them suitable for large multiprocessor systems. Assuring high reliability is a significant task for such systems.

A Unique path MIN turns to be a multi-path MIN with the incorporation of some sort of redundancy into it. Redundancy is generally added in the form of extra SEs / links / stages / subnetworks and their combinations.

Multi path MINs can either be Static or Dynamic. For static type of networks, once a fault is encountered in the way, data has to backtrack to the source or to some other fixed point in the network for selecting an alternate path.

The implementation of backtracking is relatively expensive in terms of hardware. For dynamic networks, once a fault is encountered in a particular stage, a SE in the preceding stage will reroute the data to an available alternate path. Thus a kind of fork exists at every stage to move the data to an alternate path without resorting to backtracking.

Multi path MINs can be further classified as Regular or Irregular. Regular MINs have equal number of SEs per stage as a result of which they offer equal time delay to all the requests passing through them. Irregular MINs have unequal number of SEs per stage, and are inherently multi-path. For a given source-destination pair, multiple paths of different path lengths are available. These networks offer a minimum path length (equal to 2) for favourite source-destination pair irrespective of the size of the network.

1.5 PERMUTATION AND RELIABILITY MEASURES

The problem of layout of printed circuits and large-scale integrated chips are very complex and are, therefore, usually approached by heuristic approach i.e. permutation. To minimize the communication overhead when routing, a permutation (when every source node S needs to send a single message to destination node D) has to be scheduled. Scheduling a permutation consists of partitioning the set of nodes into m subsets E_1, E_2, E_3............. E_m for some m, such that for every i = 1,2,3.....m, the paths originating from the nodes in E_i do not conflict over links i.e. they can be established simultaneously and their corresponding messages can be delivered in parallel. E_i represents the set of source nodes that can send data to their destinations at time i, i = 1,2,3,.....m.

Reliability of a system is the probability that it will perform its intended function satisfactorily for a given time under stated operating conditions. Let T be the time till the failure of a network occurs then the probability that it will not fail before time t is given by

$$R(t) = P(T>t)$$

Mean Time to Failure (MTTF) may be defined as the expected time elapsed before some source is disconnected from some destination. Mathematically it is given as the area under the reliability curve of the concerned network system.

$$\text{MTTF} = \int_{0}^{\infty} R(t)\, dt$$

In the present analyses reliability expressions are derived based upon the Series-Parallel reliability model of the concerned network, and reliability is measured and compared in terms of MTTF.

1.5.1 Series Reliability model

A system consisting of n components is said to be of series type, if the functional diagram of the system suggests that the successful operation depends upon the proper operation of all the n components as shown in figure 1.5. Failure of even a single component will make the system useless. An important characteristic of a series system is that its reliability is always worse than the poorest component in it. Reliability of a series type of system is given by the product of reliabilities of individual components in it.

$$R_{series}(t) = p_1(t) \,.\, p_2(t) \,.\, p_3(t) \,\ldots\ldots\ldots\, p_n(t) = \prod_{i=1}^{n} p_i(t)$$

P_i being the reliability of i_{th} component

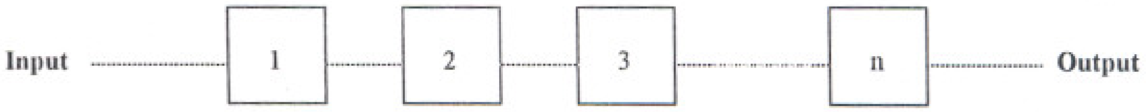


Figure 1.5: Block diagram of a series system

1.5.2 Parallel Reliability Model

A system with n components is said to be of parallel type if and only if the successful functioning of any one of the component leads to the success of the system.

System fails only when all the components fail. The reliability of a parallel type of system is as shown in figure 1.6.

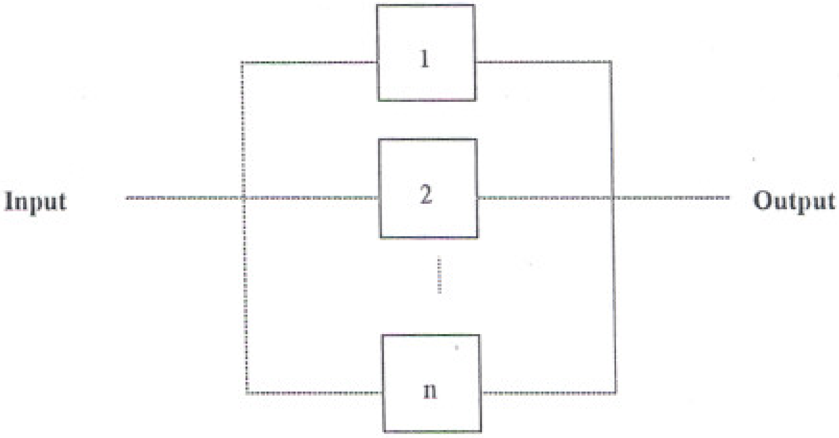


Figure 1.6: Block diagram of a parallel system

In general, if the time to failure of components is exponentially distributed then:

$$P_i\,(t) = e^{-\lambda_i t}$$

λ_i being the failure rate of i_{th} component

$$R_{series}\,(t) = \prod_{i=1}^{n} [\, e^{-\lambda_i t} \,]$$

and

$$R_{parallel}\,(t) = 1 - \prod_{i=1}^{n} [\, 1 - e^{-\lambda_i t} \,]$$

1.6 FORMULATION OF THE PROBLEM

Reliability and permutation are two major design parameters of a network, which needs to be maximised at a reasonable cost. Based upon the measures discussed above reliability is available for most of the regular and irregular networks. While, there is hardly any analyses work available in literature for permutations for regular and irregular networks. A comprehensive study of irregular networks shows them to be favourable candidate for a reasonably good performance in terms of permutations and reliability.

For Example:

1. The concept of favourite source-destination pair is completely absent in regular networks. For them the path length (a measure of communication time delay) between any source-destination is same and increases with the size of the network (typical value being Log_2N). While for favourite source-destination pair path length=2 irrespective of the size of the network.
2. For a given sized network number of SEs are comparatively lesser for irregular networks. This feature gets itself reflected in the lower cost of irregular networks in comparison to regular MINs. This feature combined with the type of network design can result in much improved reliability at a reasonable cost.
3. Irregular MINs are inherently multipath, therefore chances of blocking of a request are reduced. As a result irregular networks should tend to be more efficient and reliable. Keeping in view the above mentioned features of regular and irregular networks in the present dissertation, reliability and permutation analyses have been carried out for regular and irregular type of MINs, such as:

- Augmented Shuffle Exchange Network (ASEN)
- Fault-tolerant Double Tree Network (FDOT)
- Four Tree Network (FT)
- Proposed Smart Four Tree Network (SFT)

In this thesis, a new network has also been proposed named Smart Four Tree Network (SFT), which is designed from the combination of regular-ASEN-2 and irregular-FT networks. Permutations have been analyzed under single switch fault per stage (non-critical case) and single loop fault per stage (critical case) for regular – ASEN-2 and irregular networks – FDOT, FT, SFT of size varying from n = 4,5,6,7,……,10 and it has been found that the proposed SFT network is better in terms of permutation passable as compared to irregular FT network.

1.7 ORGANIZATION OF THE THESIS

The thesis is organized as follows:

Chapter 2 provides a brief survey of Regular/Irregular MINs with their characteristic features.

Chapter 3 analyses the performance of irregular networks in terms of permutation passable (possible number of requests served simultaneously). Results are compared with regular MINs. Also proposed a new irregular network named Smart FT (SFT) which is better in terms of permutation passable of FT network if it is single switch fault per stage or single loop fault per stage.

Chapter 4 analyses the Reliability aspects of Regular & Irregular MINs. This includes both optimistic and pessimistic analyses of the network. Also analyses the cost-effectiveness for Regular and Irregular networks.

Chapter 5 concludes the work presented in this Dissertation along with the suggestions for further scope in the field of Irregular MINs.

CHAPTER-2

MULTISTAGE INTERCONNECTION NETWORKS

The means for communication among processors, memory modules, and other devices is the Interconnection Network (IN). Many interconnection networks for large-scale multiprocessor computer systems have been proposed and analyzed by various researchers. Of these, Multistage Interconnection Networks (MINs) are well suited to communication among tightly coupled system components, and offer a good balance between cost and performance.

Primarily MINs differ from one another in terms of their constructional features like size of SEs, number of stages, number of links, control strategy and routing algorithm. Based upon the design, MINs offer varying degrees of reliability, efficiency, fault-tolerance and cost. MINs can be classified as shown in the Figure 2.1

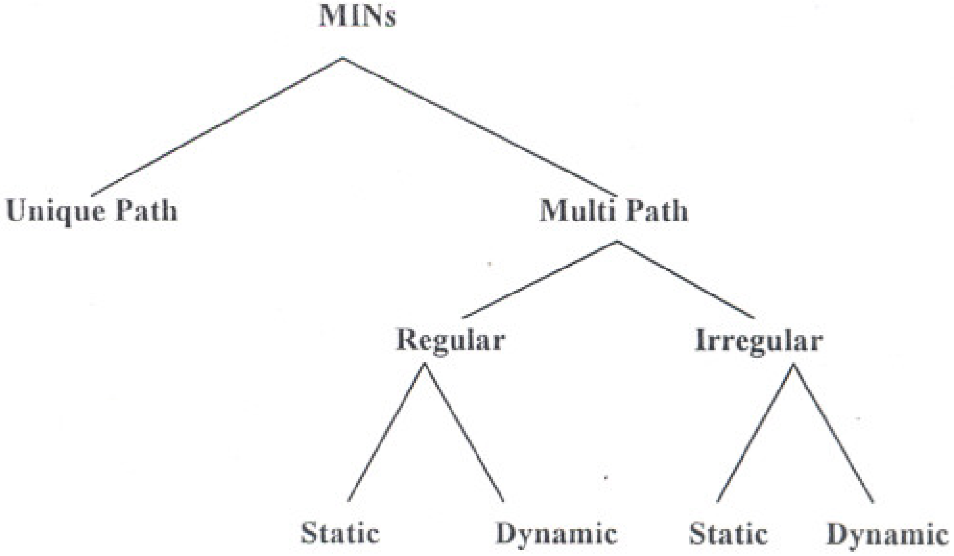


Fig. 2.1: Classification of MINs

2.1 UNIQUE PATH MINs

Unique path MINs are characterized by the presence of the single unique path between any given source-destination pair. Distinct source-destination paths may have a common link as well. As a result of which, Unique path MINs have two drawbacks:

1. In case of failure of a switch, along the desired path of a request, no alternate path is available; as a result source-destination pairs using that faulty switch will get disconnected from the network. Unique path MINs are, therefore, termed as non fault-tolerant and are least reliable.
2. A given link can be shared by more than one source-destination pairs, so in case of blocking of a particular link (due to preoccupation by another request or some other reason) no other request can get through the network. As a result in a random access environment, performance of unique path MINs tend to be poorer. This reduction in performance and reliability become more severe as the size of the network increases.

The examples of unique path MINs are Generalized Cube, Omega and Delta.

2.1.1 GENERALIZED CUBE

The generalized cube network of size 8x8 is shown in Figure 2.2. The network is a NxN MIN with Log_2N stages and each stage consisting of N links connected to N/2 switches. Each switch, called an interchange box, is 2x2, individually controlled, and can be set to one of four states (straight, exchange, lower broadcast and upper broadcast).

2.2 MULTI-PATH REGULAR STATIC MINs

In these types of networks data has to backtrack to the source for an alternate path.

The examples of multi-path regular static MINs are Extra Stage Cube (ESC), Extra Group Network (EGN), 3-Replicated network (3-rep) and INDRA network.

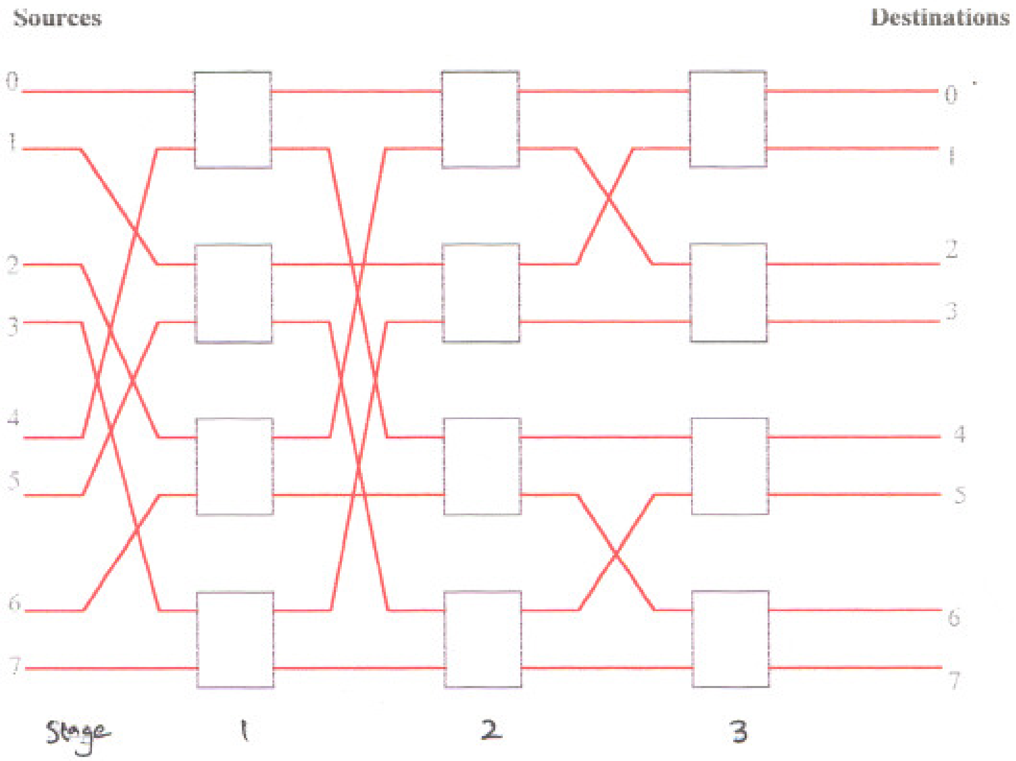


Figure 2.2 Generalized Cube Network of size N=8

2.2.1 EXTRA STAGE CUBE NETWORK

Extra –Stage Cube network of size 8x8 is shown in Figure 2.3. ESC is NxN MIN with $Log_2N + 1$ stages and each stage consisting of N links connected to N/2 switches. ESC is formed from the generalized cube network by adding an extra stage to the input side of the network along with multiplexers and demultiplexers at the input and output stage respectively. In case of a fault in a given stage, the extra stage is enabled and the request gets routed through the fault free SEs of this extra stage but in the absence of faults, extra stage remains disabled. ESC is single switch fault-tolerant and robust in the presence of multiple faults. Due to the availability of the extra stage, both performance and reliability improves, but at the expense of increase in network cost.

2.2.2 INDRA Network

INDRA network is shown in Figure 2.4. INDRA networks have $N = 2^m$ inputs and outputs with $Log_RN + 1$ stages of R x R switches. Stages are numbered 0 through n from input to output, where $n = Log_RN$, and each stage has N switches. They are connected by links in an R * (N/R) – shuffle. INDRA networks are made up as the union of R parallel networks each with Log_RN stages of R x R switches, with an initial distribution stage at the input. In INDRA network, in case of a stage 0 switch fault, data can be routed through one of the alternate links from the input port. Other switch faults can be avoided by taking one of the R^2 paths that exist between any source and destination that does not include the faulty switch. The INDRA network is R – 1 fault-tolerant and is quite reliable and efficient in general.

2.3 MULTI-PATH REGULAR DYNAMIC MINs

In this type of MINs, a fork exists at every point in a stage, which can reroute the data to an alternate available path without resorting to backtracking.

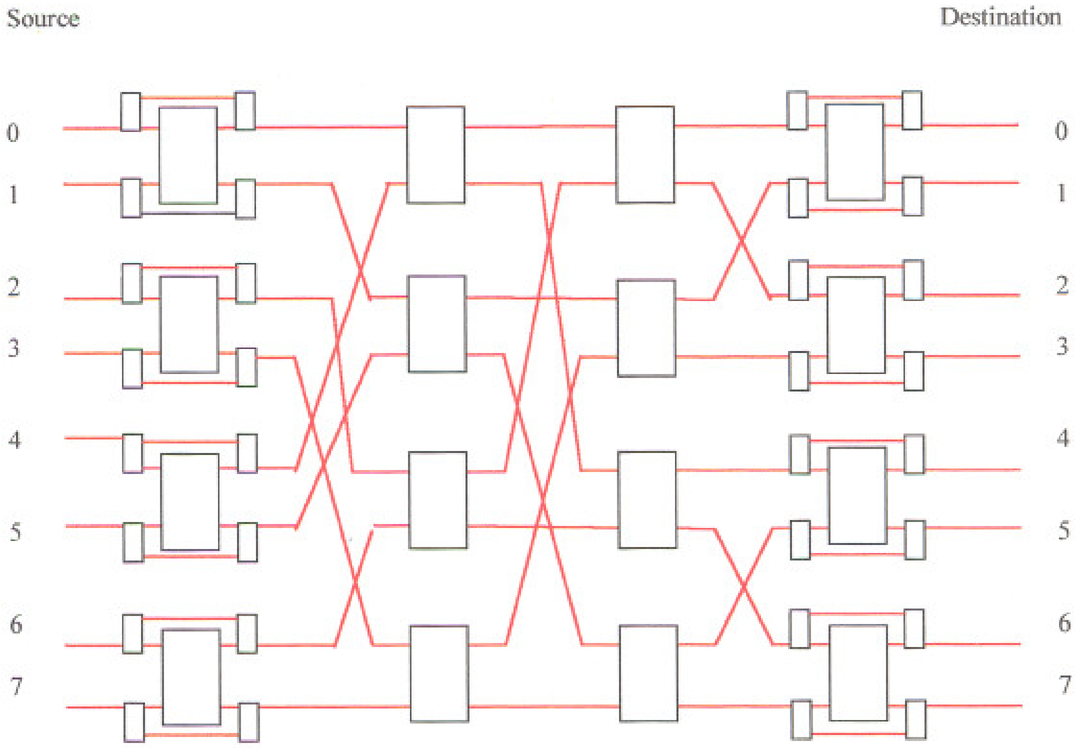


Figure 2.3 ESC Network for N=8

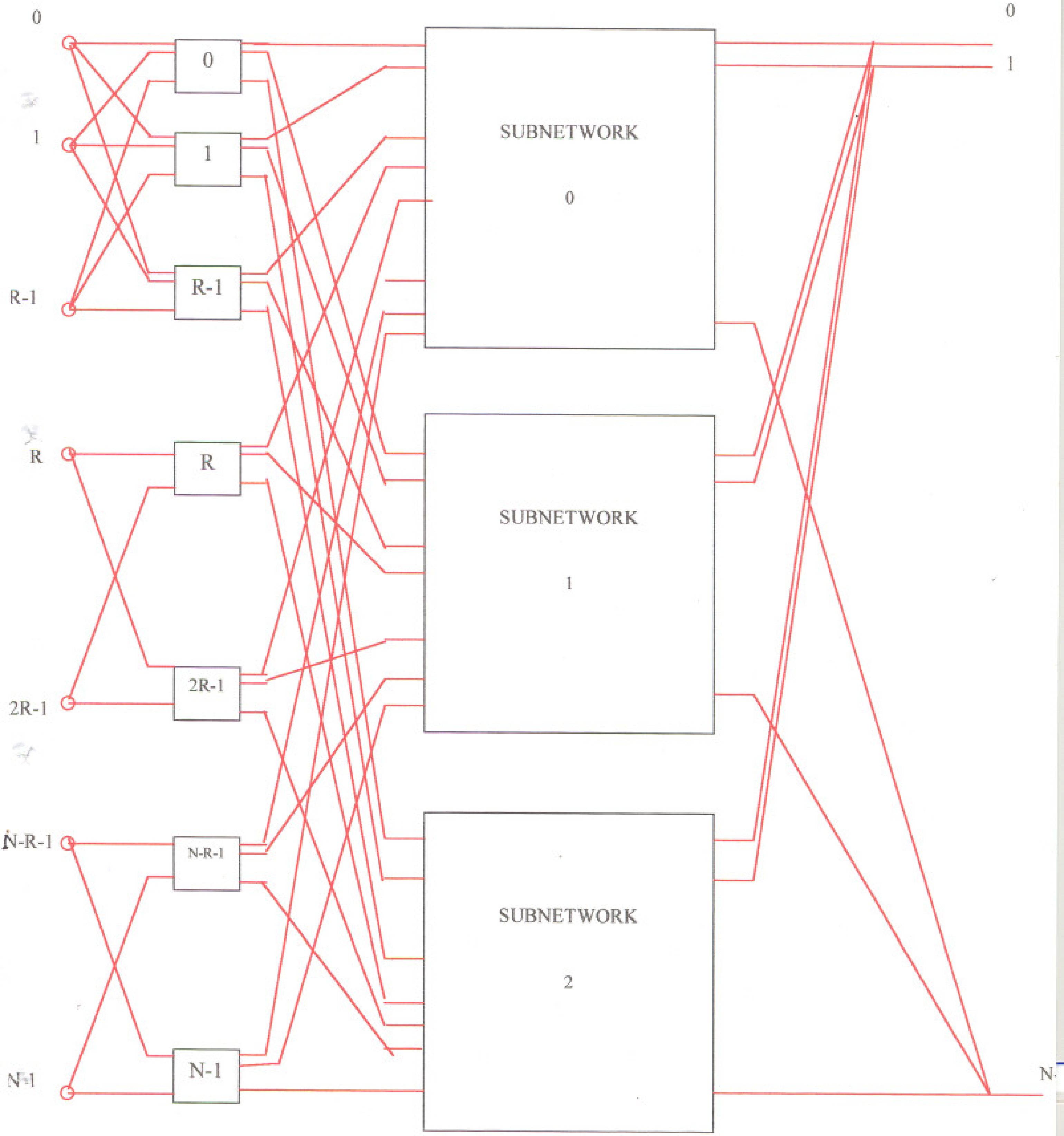


Figure 2.4 INDRA Network as the union of R subnetworks.

It also reduces the unnecessary time delay and is generally achieved by providing extra links in the network.

The examples of multi-path regular dynamic MINs are F-network, Inverse Augmented Data Manipulator (IADM), Augmented Shuffle Exchange Network (ASEN) and Augmented Baseline Network (ABN).

2.3.1 F-NETWORK

F-network of size 8x8 is shown in Figure 2.5. F-network is 2^n x 2^n with Log_2N+1 stages of N (= 2^n) switches each, that are, in general, 4 x 4 selectors. F-network basically adds links to the generalized cube network. At each stage except the output stage, two different switches can be selected while maintaining the same destination. F-network assumes 1) only internal switches fail 2) failed switches are unusable and 3) faults occur independently. F-network is single switch fault-tolerant and robust in the presence of multiple faults.

2.3.2 INVERSE AUGMENTED DATA MANIPULATOR (IADM)

IADM network of size 8x8 is shown in Figure 2.6. IADM consists of Log_2N stages each with N switches and 3N links, plus one column of output switches that are the network output ports. Each switch is 3 x 3 selector. To improve upon the reliability and performance aspects of unique path MINs extra intrastage links are provided. Unlike the generalized cube, delta, and omega networks, the IADM can tolerate some faults because of multiple paths between a source S and destination D, if $S \neq D$.

2.3.3 AUGMENTED SHUFFLE EXCHANGE NETWORK (ASEN-2)

ASEN-2 of size 16x16 is shown in Figure 2.7. ASENs have $Log_2N - 1$ stages (($Log_2N - 2$) stages consist of N/2 SEs per stage of size 3 x 3 and one stage of N/2 SEs of size 2 x 2), plus N 2x1 multiplexers and N 1x2 demultiplexers.

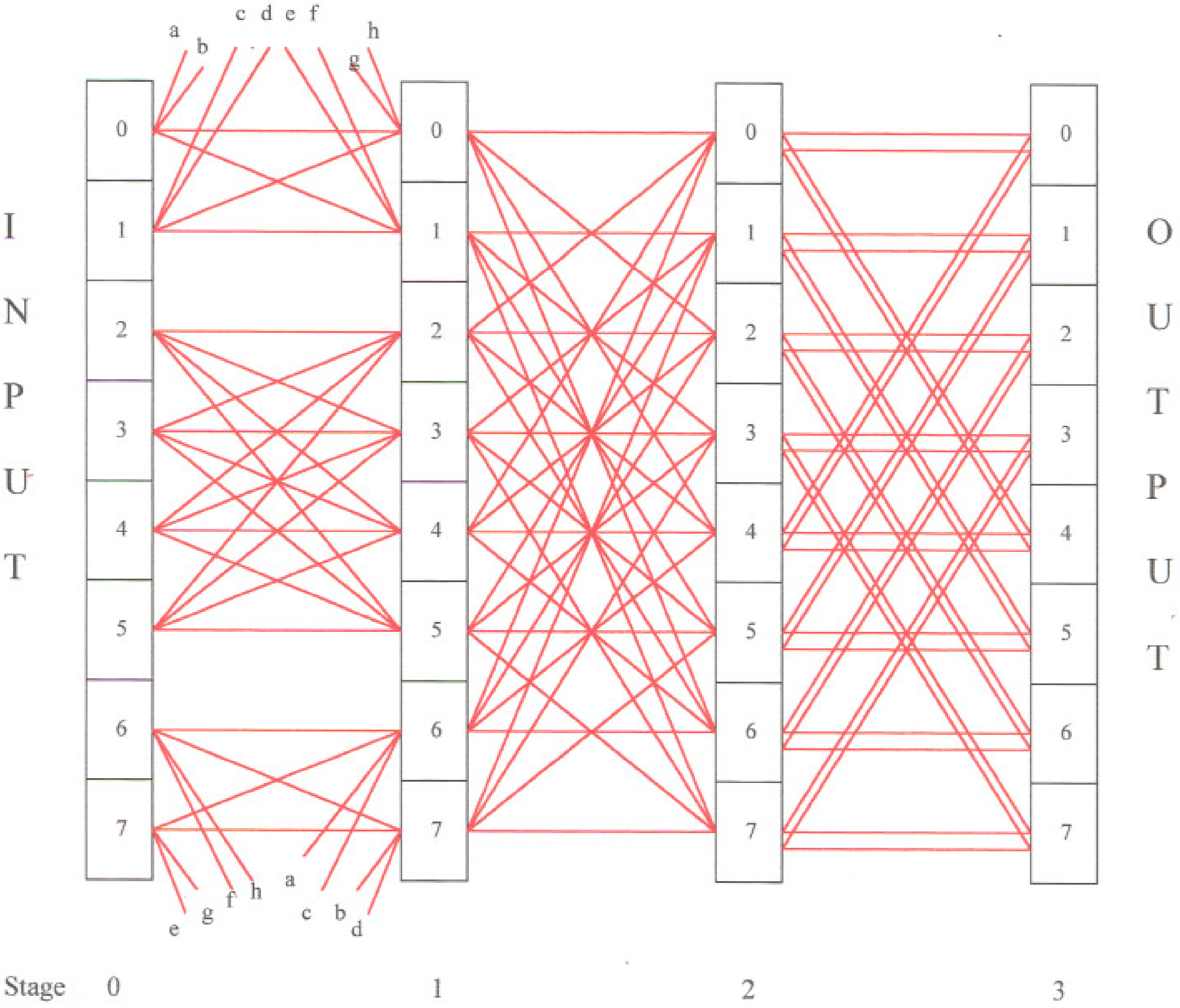


Figure 2.5 F – network of size N=8

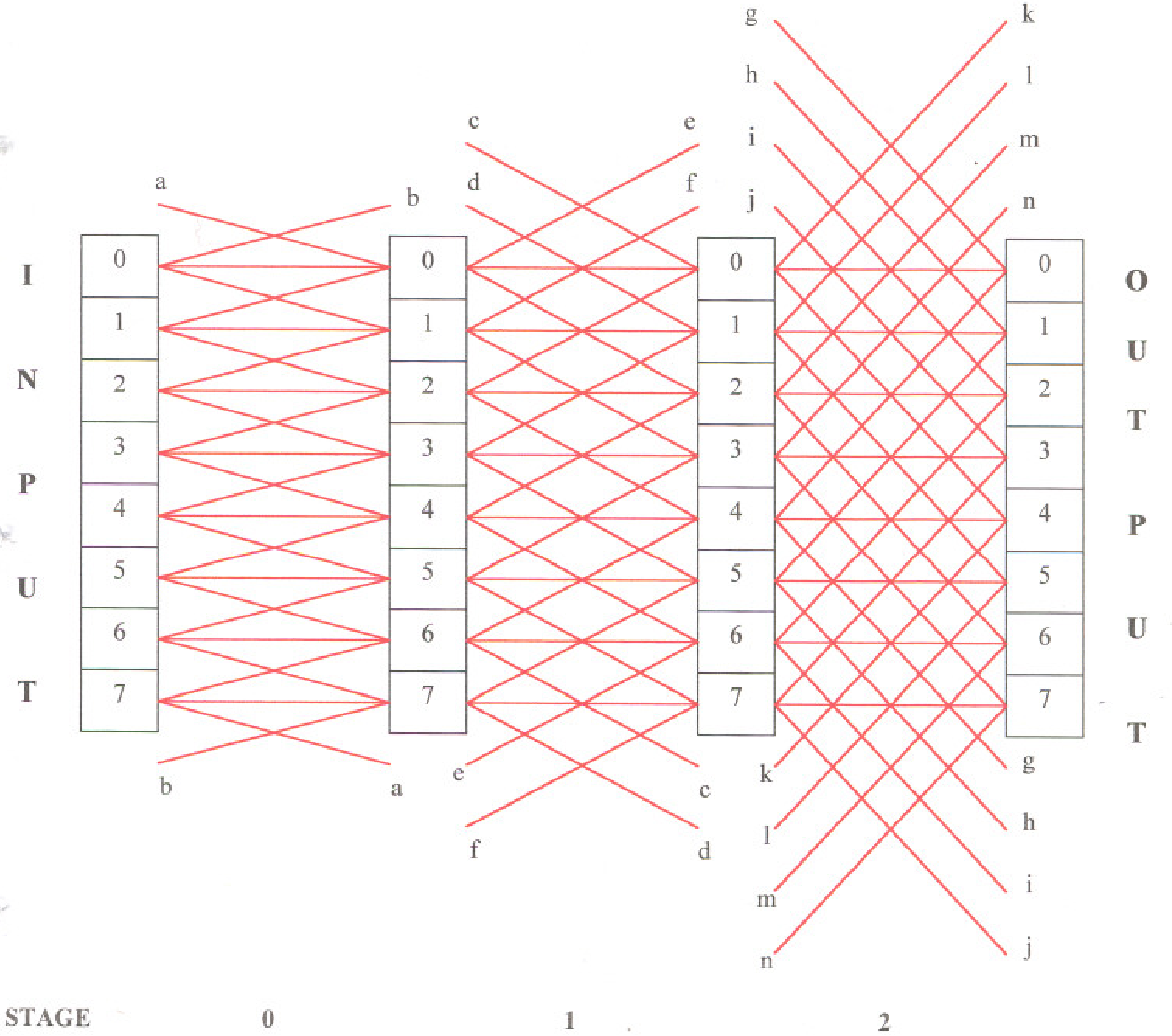


Figure 2.6 IADM network of size N=8

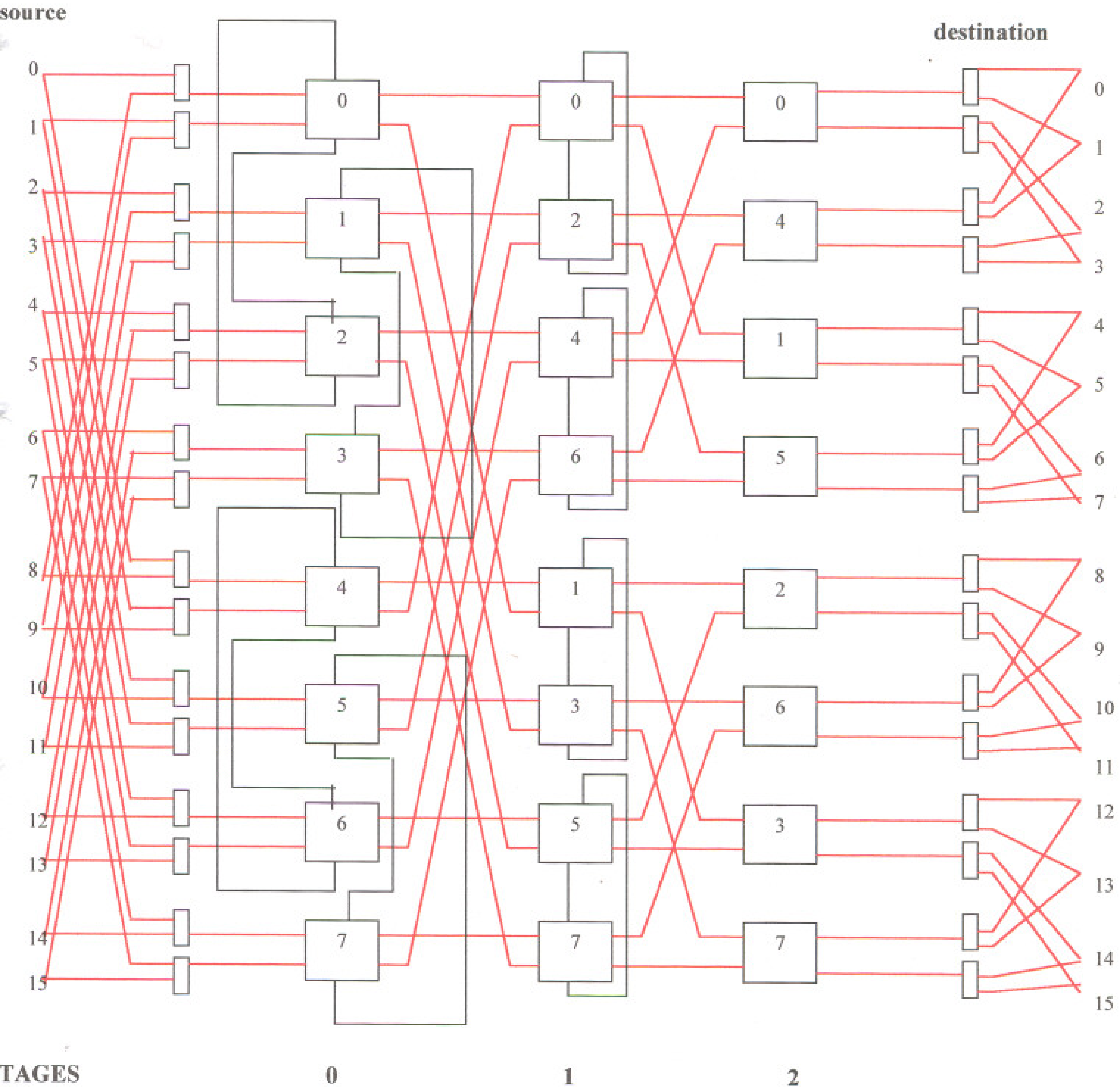


Figure 2.7 ASEN-2 Network for N=16

In ASEN-2, the loops formed in all stages except the final stage are such that for every loop there exists another loop which is connected to the same set of switches in the next stage. Such pairs of loops are called conjugate loops. In the final stage, since a loop consists of only one switch, there are in effect conjugate pairs of switches connected to the same set of destinations. A general principle to be followed in forming the loops is to ensure that if two switches are part of a loop, their conjugate switches will also be part of a loop.

2.3.4 AUGMENTED BASELINE NETWORK (ABN)

ABN network of size 16x16 is shown in Figure 2.8. ABNs have Log_2N-2 stages. The switches in the last stage are of size 2 x 2 and the remaining switches in stages 1 through $Log_2N - 3$ are of size 3 x 3. There is one 4 x 1 MUX for each input link of a switch in stage 1 and one 1 x 2 DEMUX for each input link of a switch in stage $n - 2$. This network provides two disjoint paths for each source-destination pair, so ABNs are single switch fault-tolerant. When both the switches to which a source or a destination is connected becomes faulty, then that source or destination is disconnected from the rest of the network. Since one of the input stage has been replaced by 4:1 type of multiplexers, having low hardware complexity as compared to 3 x 3 SEs, this also reduces the cost of the network and thereby improving ratio of MTTF to cost. Reliability and performance of ABN is better than other regular MINs.

2.4 IRREGULAR MINs

Irregular MINs are classified as either unique path and multi path.

2.4.1 UNIQUE PATH MINs

The example of unique-path irregular MIN is MDOT network.

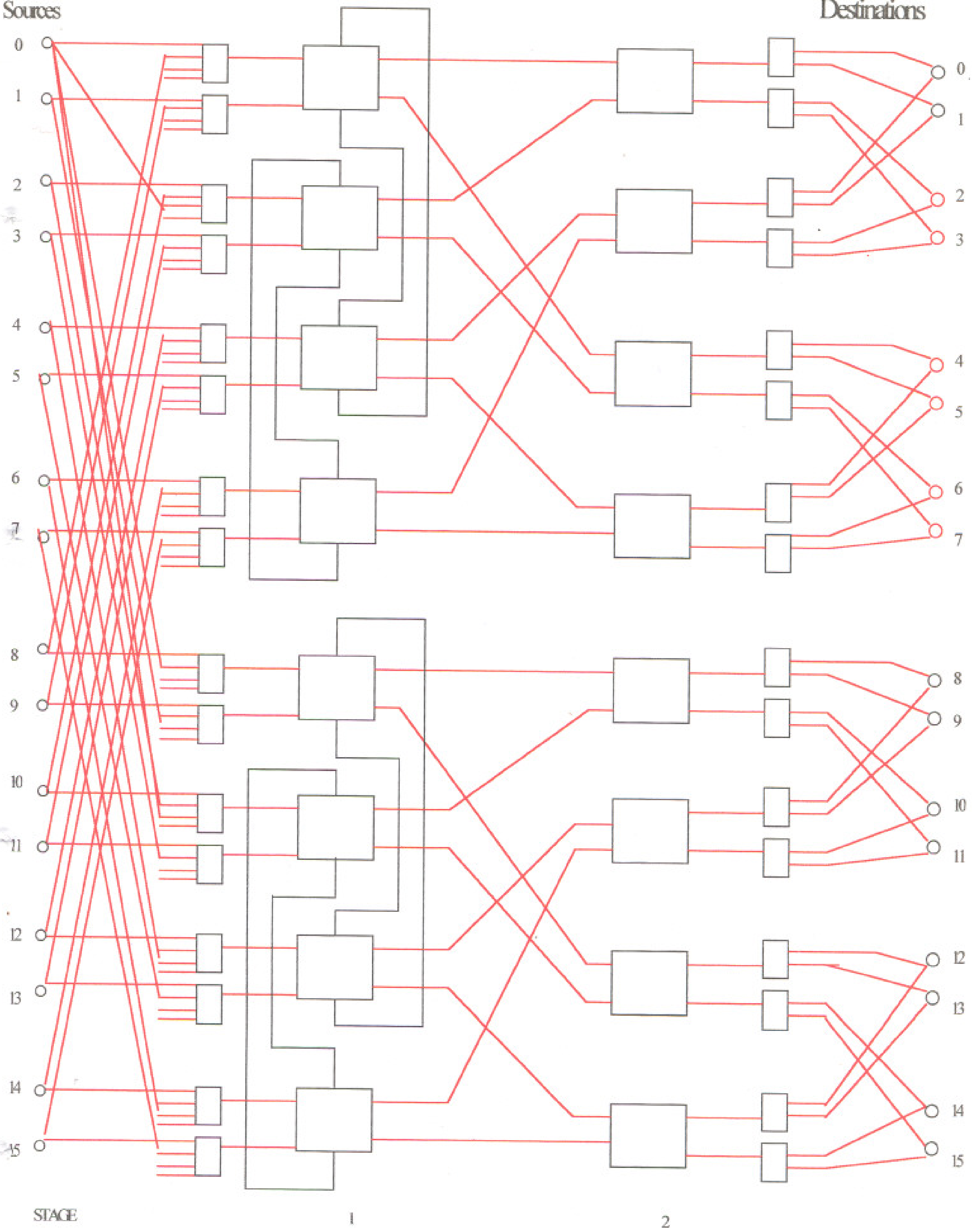


Figure 2.8 ABN Network of size N=16

3.3 CONCLUSION

From the above analyses based on the results in the Tables 6 & 8, it has been observed that the permutations are better in a proposed irregular SFT network than the existing irregular FT network.

It has been further analyzed that the permutations of FDOT irregular network (given in Tables 3 & 4) are better than the existing irregular FT network (given in Tables 5 & 6).

It has been concluded that permutations of irregular fault-tolerant MINs are better than the regular MINs.

3.2.3 IDENTITY & INCREMENTAL PERMUTATION LAYOUTS FOR SFT (N=16)

Identity and incremental permutation layouts are given in Table 7 & 8 respectively.

Faults →	MUX		S0		S1		S2		S3		S4		DEMUX	
↓ PL	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
2	.50	.50	.43	.37	.50	.50	.50	.50	.50	.50	.43	.37	.50	.50
5	.43	.37	.43	.37	.37	.25	.37	.25	.37	.25	.43	.37	.43	.37

Table 7: Identity permutation layout for SFT

Faults →	MUX		S0		S1		**S2**		S3		S4		DEMUX	
↓ PL	X	Y	X	Y	X	Y	**X**	**Y**	X	Y	X	Y	X	Y
2	0	0	0	0	0	0	**0**	**0**	0	0	0	0	0	0
5	.50	.50	.43	.37	.37	.25	**.37**	**.25**	.37	.25	.43	.37	.50	.50

Table 8: Incremental permutation layout for SFT

For Example: In case of incremental permutation layout of FT, as given in Table 6, the permutations are zero if switch2 gets faulty in a loop stage i.e. in critical case.

But in incremental permutation layout of proposed SFT network, as given in Table 8, the percentage of requests served is 0.25.

Faults →	MUX		S0		S1		S2		S3		S4		DEMUX	
↓ PL	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	.37	.37	.37	.25	.31	.18	.25	.25	.31	.18	.37	.25	.37	.37

Table 4: Incremental permutation layout for FDOT of size N=16

3.2.2 IDENTITY & INCREMENTAL PERMUTATION LAYOUTS FOR FT (N=16)

Identity and incremental permutation layouts are given in Table 5 & 6 respectively.

Faults →	MUX		S0		S1		S2		S3		S4		DEMUX	
↓ PL	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
2	.50	.50	.43	.37	.50	.50	.50	.50	.50	.50	.43	.37	.50	.50
4	.25	.25	.25	.25	.18	.12	.25	.25	.18	.18	.25	.25	.25	.25
5	.18	.12	.18	.12	.18	.12	.12	0	.18	.18	.18	.12	.18	.12

Table 5: Identity permutation layout for FT

Faults →	MUX		S0		S1		**S2**		S3		S4		DEMUX	
↓ PL	X	Y	X	Y	X	Y	**X**	**Y**	X	Y	X	Y	X	Y
2	0	0	0	0	0	0	**0**	**0**	0	0	0	0	0	0
4	0	0	0	0	0	0	**0**	**0**	0	0	0	0	0	0
5	.25	.25	.25	.25	.18	.12	**.12**	**0**	.18	.12	.25	.25	.25	.25

Table 6: Incremental permutation layout for FT

2.4.1.1 MODIFIED DOUBLE TREE NETWORK (MDOT)

MDOT of size 8x8 is shown in Figure 2.9. MDOT network of size 2^n x 2^n has 2^n source and 2^n destination terminals and 2m-1 stages. It has $2^{m+1} - 3$ switches. An ith and (2m-i)th stage has 2^{m-i} switches of size 2 x 2 for i = 1,2,3,….,n. This network has the advantage of flip control, (i.e. individual stage control) which effects the performance and reliability of the entire system, and distributed control (i.e. individual switch control). MDOT is also used in the construction of FT network.

2.4.2 MULTI PATH MINs

Irregular Networks, though inherently multi-path, may or may not be fault-tolerant. The examples of multi-path MINs are FDOT, FT and proposed SFT network.

2.4.2.1 MULTI PATH IRREGULAR STATIC MINs

The example of multi-path irregular static MIN is Fault-tolerant Double Tree Network (FDOT).

2.4.2.1.1 FAULT-TOLERANT DOUBLE TREE NETWORK (FDOT)

FDOT network of size 16x16 is shown in Figure 2.10. FDOT network of size 2^n x 2^n has 2^n sources and 2^n destinations and 2m-1 stages. Further it has $3.(2^{m+1} - 3)$ switches and 3.N/2 multiplexers or demultiplexers. FDOT-k network is an irregular type of network, with k independent sub-networks of size (N/K x N/k) and an extra one. The extra sub-network helps to enhance fault-tolerant capability and to keep a desired level of performance even in the presence of faults. FDOT-k network is k fault-tolerant. Depending on the source-destination pair, an irregular network supports multiple paths of different path lengths. Thus chances of blocking of a request are further reduced. As a result of which, performance of FDOT-k network tends to be better than other MINs.

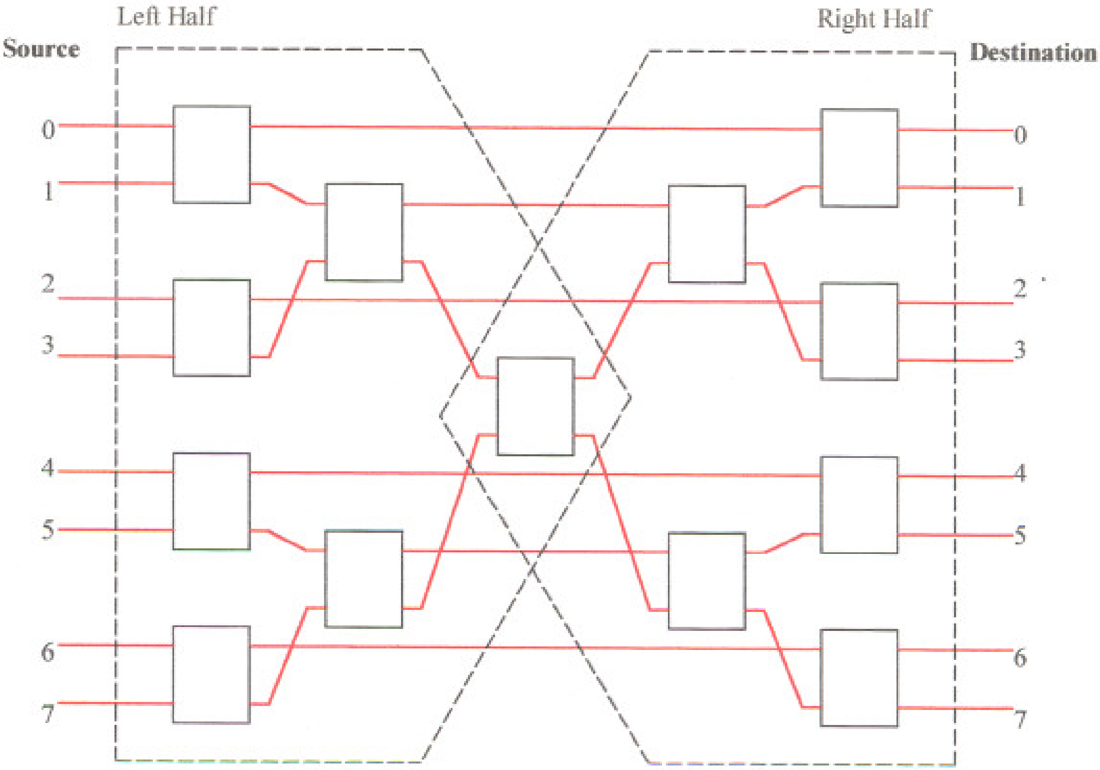


Figure 2.9 MDOT MIN of size N=8

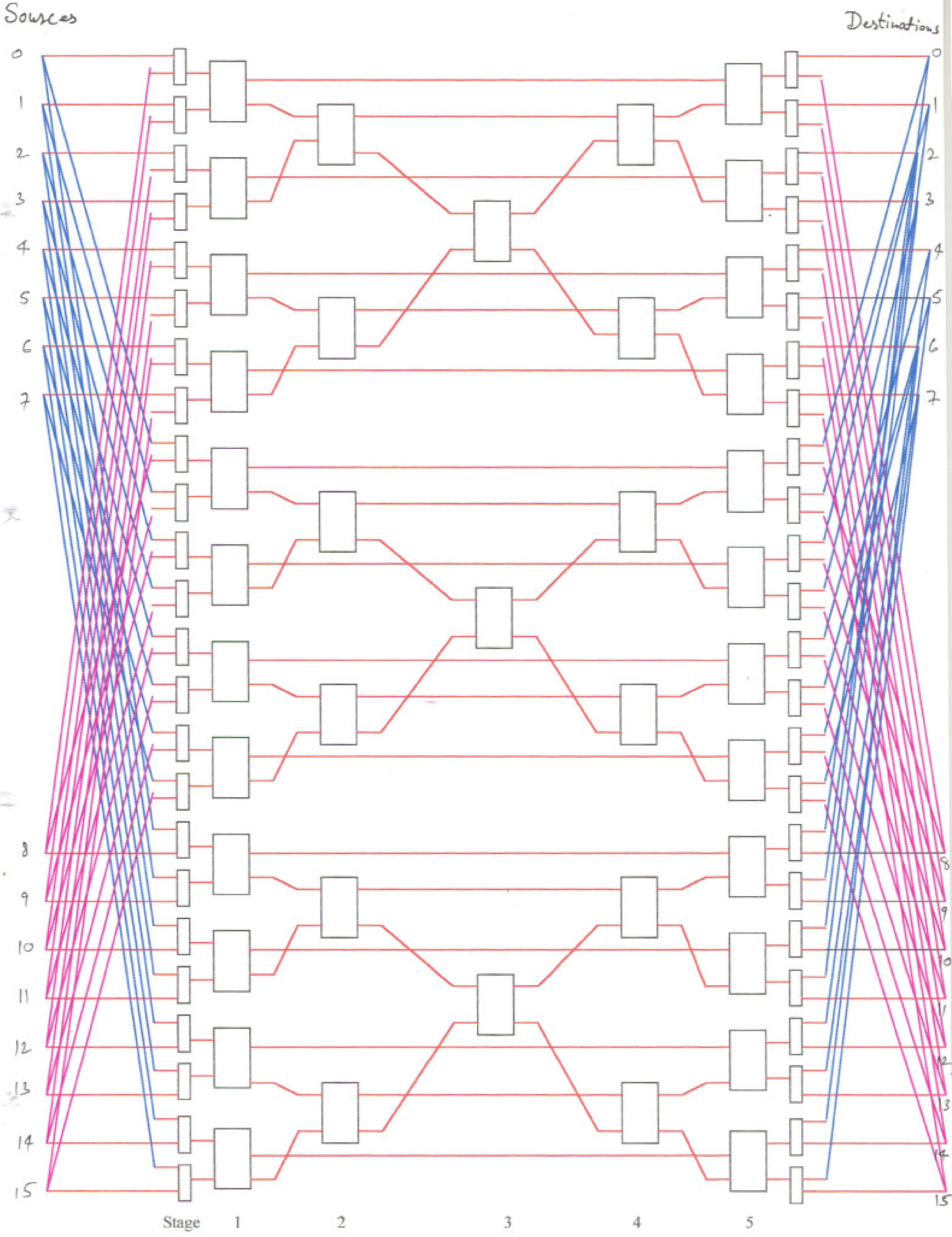


figure 2.10 FDOT-2 MIN for N=16

2.4.2.2 MULTI PATH IRREGULAR DYNAMIC MINs

The examples of multi-path irregular dynamic MINs are Four Tree network (FT) and proposed Smart Four Tree network (SFT).

2.4.2.2.1 FOUR TREE NETWORK (FT)

FT network of size 16x16 is shown in Figure 2.11. A FT network of size 2^n x 2^n consists of (2m-1) stages and $(2^{m+2} - 6)$ switches, out of which 2^{n-1} are of size 2 x 2 and the rest are of size 3 x 3. There are 2^n 2x1 multiplexers and an equal number of 1x2 demultiplexers. Both stage i and stage (2m-i) has exactly 2^{n-i} switches where i=1,2,3,….,N-1. FT network being an irregular network supports multiple paths of different path lengths.

This network is constructed with the help of two identical groups, each consisting of MDOT network of size 2^{n-1} x 2^{n-1}, 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 3x3 SE in a stage forms a loop with the corresponding numbered 3x3 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. In case the primary path is busy or faulty requests will be routed through secondary path in the sub-network. Thus a fork exists at every point in s stage except the last which makes alternate routing feasible. FT network is single switch fault-tolerant. If both switches in a loop are simultaneously faulty then clearly some sources are disconnected from some destinations. FT network is more cost-effective than other multi-path irregular MIN with high reliability and performance.

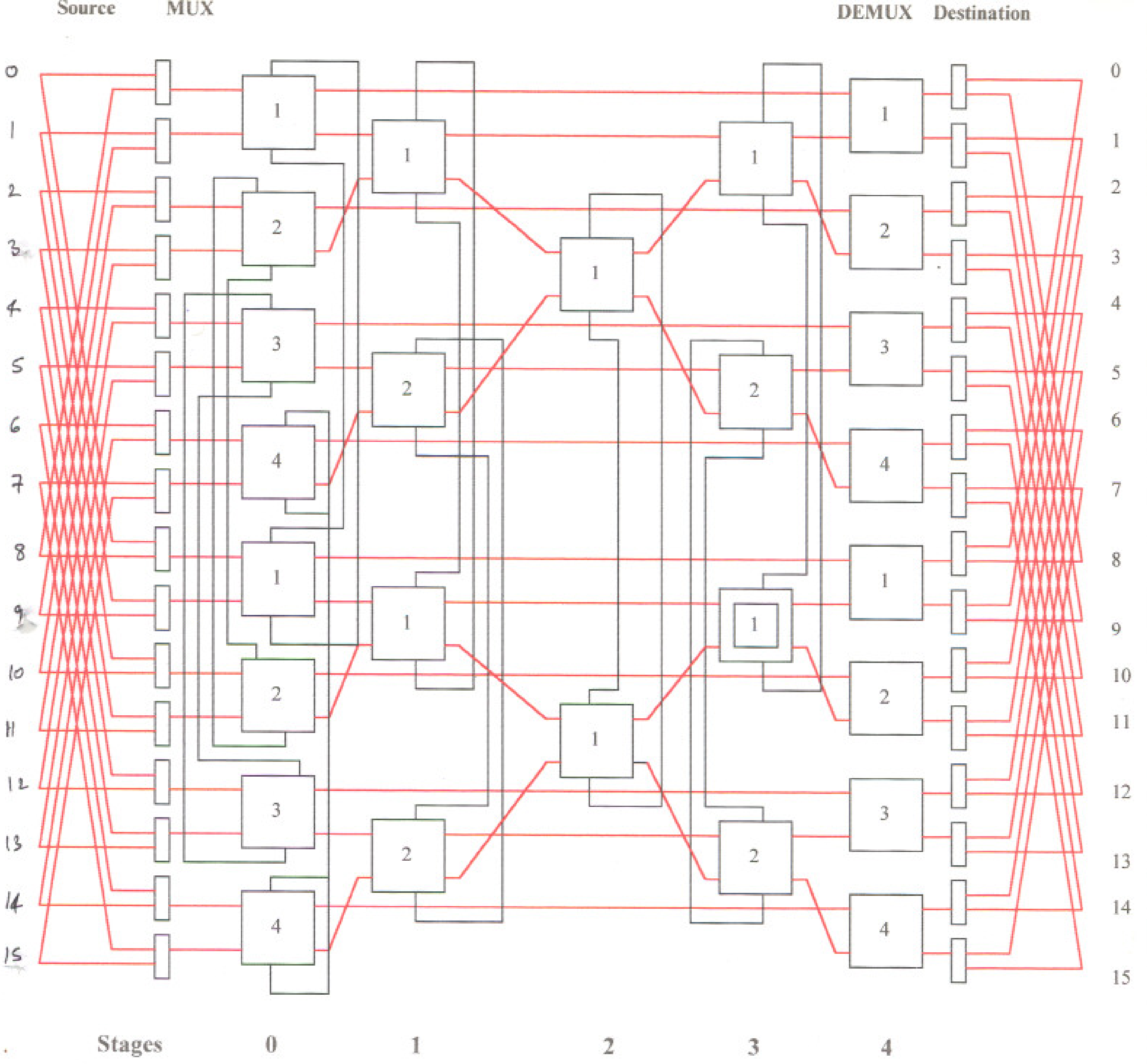


Figure 2.11 : FT MIN of size N=16

2.4.2.2.2 PROPOSED SMART FOUR TREE NETWORK (SFT)

SFT network of size 16x16 is shown in Figure 2.12. A SFT network of size 2^n x 2^n (n=Log_2N) consists of (2m-1) stages and (2^{m+2} – 4) switches, out of which 2^{n-1} are of size 2 x 2 and the rest are of size 3 x 3. There are 2^n 2x1 multiplexers and an equal number of 1x2 demultiplexers. Both stage i and stage (2m-i) has exactly 2^{n-i} switches.

A SFT network being an irregular network supports multiple paths of different path lengths. It also inherits the property of regular networks since the number of switches are same in all the stages except the last and first stage of a network. Every 3x3 SE in a stage forms a loop with the corresponding numbered 3x3 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. The advantage of this network is - if both switches in a loop are simultaneously faulty then even some sources are connected to the destinations. FT network is more cost-effective than SFT but the performance (in terms of permutation passable) of SFT is better than FT. The performance features of regular and irregular networks are discussed in the following chapters.

2.5 CONCLUSION

In this chapter unique and multi-path, regular and irregular networks have been discussed. This is useful for designer's perspective. It not only provides the present state of the art technology in multiprocessor interconnection networks but also gives an insight on the various steps/stages, to ultimate reach the present status. It also helps to know the various openings in the field, in which further work can be carried out.

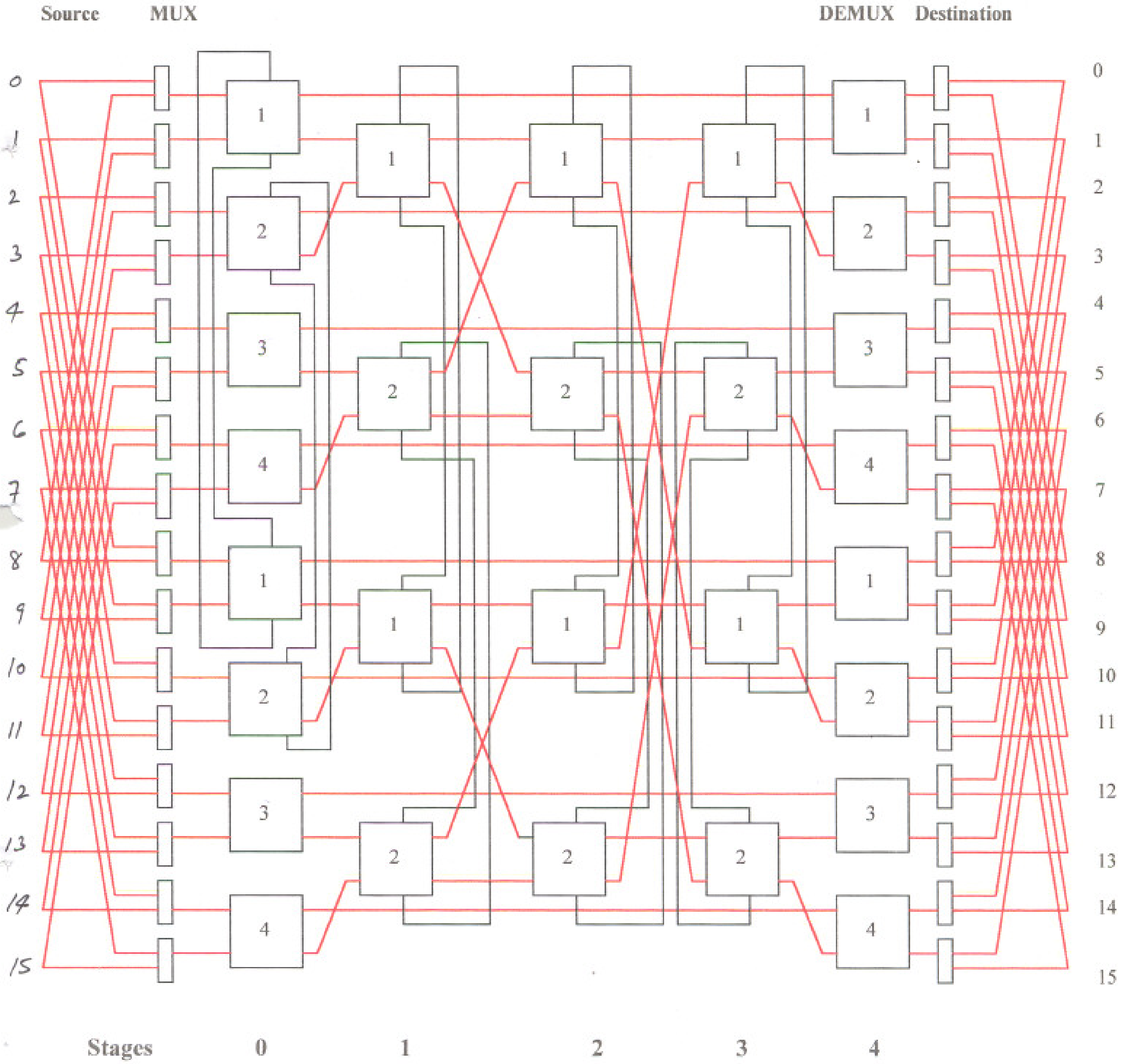


Figure 2.12 SFT MIN of size N=16

CHAPTER-3

PERMUTATIONS

A one-to-one correspondence between source and destinations are called Permutations. To find out the permutations that exists for a network, it is assumed that S_i (i=0,1,2,...N-1) represents the sources and D_i(i=0,1,2,3....N-1) represents the destinations of a network. The two possible permutation layouts are as follows:

Identity permutations

Identity permutation layout means that the sources S_i (i=0,1,2,....N-1) have one-to-one connection with the destinations D_i (i=0,1,2,...N) network and is expressed by:

$$S_i = D_i \text{ for } i=0,1,2,3,\ldots\ldots N-1$$

For Example: A network of size N=16, i.e. 16 sources and 16 destinations, is represented in Figure 3.1.

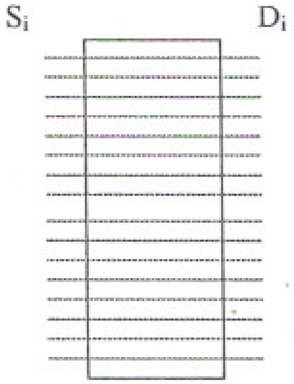
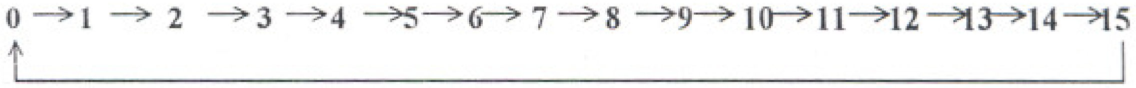


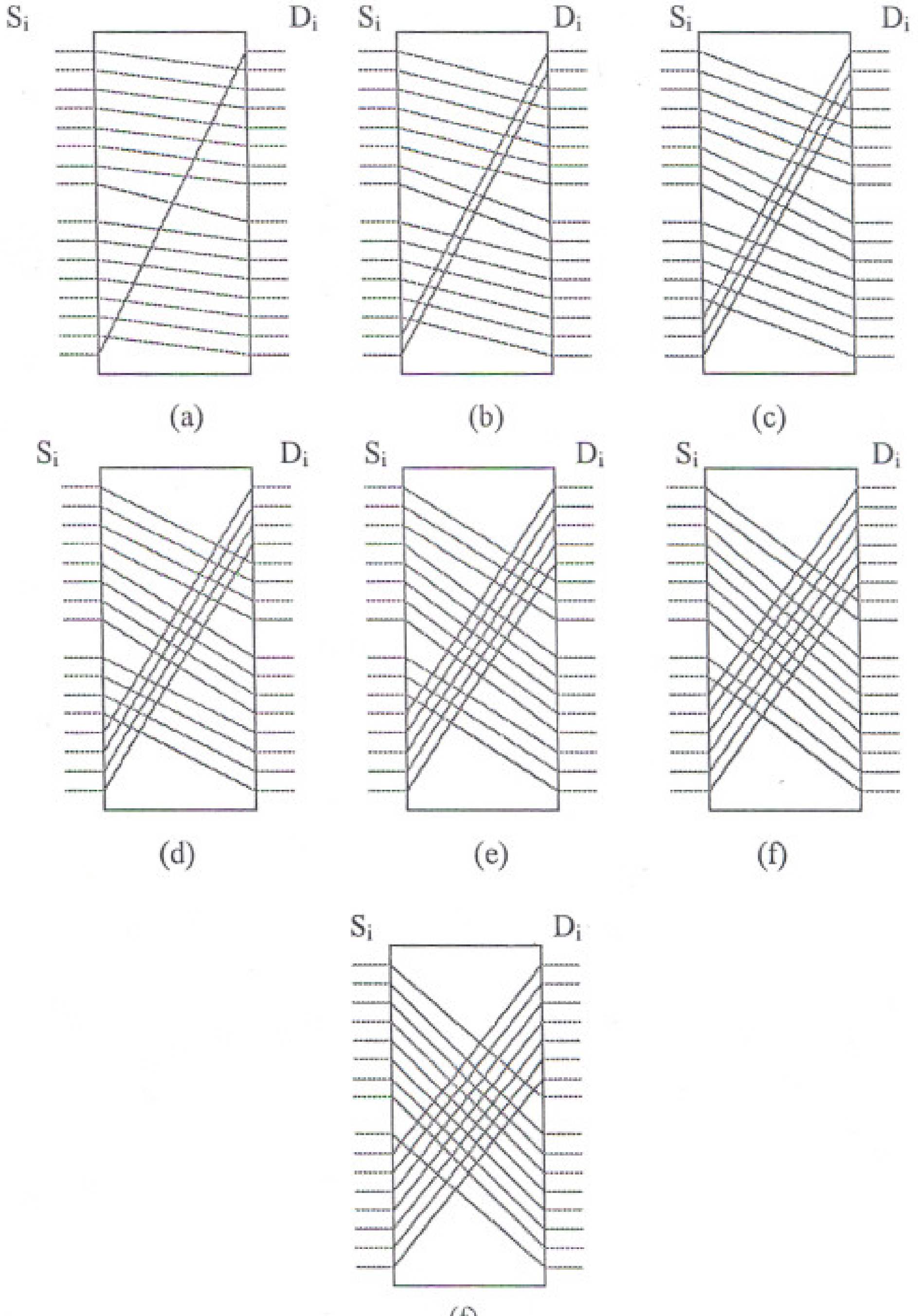
Fig. 3.1: Identity permutation layout for network of size N=16

Incremental permutations

Incremental layout permutation means that the sources S(ca[i]) where i=0,1,2,....N-1 are connected to destinations D(ca[j]) where j=1,2,.....N-1. The sources and destinations of a network are arranged in a circular array i.e. ca[i] where i=0,1,2,...N-1 which is represented as:



The incremental permutation layouts for network of size N=16 is given in Figure 3.2



(a) (b) (c)

(d) (e) (f)

(f)

Fig. 3.2: Incremental permutation layouts for network of size N=16

The incremental permutation layouts for network of size N=16 are shown in example 1 (as depicted in figure 3.2 (a)) and example 2 (as depicted in figure 3.2(b)).

1. *Source:* (0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15)
Destination: (1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 0)

2. *Source:* (0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15)
Destination: (2 3 4 5 6 7 8 9 10 11 12 13 14 15 0 1)

To find out the permutations, it is assumed that X denotes the switches in a non-critical case (if fault is present in a single switch) and Y denotes the switches in a critical case (if the switches are faulty in a loop).

PL denotes the path length(s) available for a network and probability of issuing a request is 1.0. SEs at the ith stage are represented as S0, S1, S1...... Sn ($n=Log_2N$) and multiplexers and demultiplexers are considered as MUX & DEMUX respectively.

It is further assumed that, to find out the values for incremental permutation layouts (depicted in figure 3.2(d)), the link to be taken between source and destination is:

Source: (0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15)
Destination: (4 5 6 7 8 9 10 11 12 13 14 15 0 1 2 3)

3.1 PERMUTATIONS FOR REGULAR MINs

3.1.1 IDENTITY & INCREMENTAL PERMUTATION LAYOUTS FOR ASEN-2 (N=16)

Identity and incremental permutation layouts are given in Table 1 & 2 respectively. The table shows the total number of requests served simultaneously (in percentage) depending on the single switch fault or single loop fault per stage.

Faults → ↓ PL	MUX		S0		S1		S2		S3		S4		DEMUX	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
3	.93	.87	.87	.75	.87	.75	.87	.75	.87	.75	.87	.75	.93	.87

Table 1: Identity permutation layout for ASEN-2

Faults → ↓ PL	MUX		S0		S1		S2		S3		S4		DEMUX	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
3	.93	.87	.87	.75	.87	.75	.87	.75	.87	.75	.87	.75	.87	.87

Table 2: Incremental permutation layout for ASEN-2

3.2 PERMUTATIONS FOR IRREGULAR MINs

3.2.1 IDENTITY & INCREMENTAL PERMUTATION LAYOUTS FOR FDOT (N=16)

Identity and incremental permutation layouts are given in Table 3 & 4 respectively.

Faults → ↓ PL	MUX		S0		S1		S2		S3		S4		DEMUX	
	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
2	.75	.75	.68	.57	.75	.75	.75	.75	.75	.75	.68	.57	.75	.75
4	.25	.12	.32	.18	.25	.18	.25	.25	.25	.18	.32	.18	.25	.12
5	0	0	0	0	0	.06	0	0	0	.06	0	0	0	0

Table 3: Identity permutation layout for FDOT of size N=16

CHAPTER-4

RELIABILITY ANALYSES OF REGULAR AND IRREGULAR NETWORKS

This chapter provides reliability analyses of regular and irregular MINs. The networks considered in regular MINs are Augmented Shuffle Exchange Network (ASEN-2), Augmented Baseline Network (ABN), and irregular MINs are Fault-tolerant Double Tree (FDOT), Four Tree (FT) and proposed Smart Four Tree (SFT). The analyses presented in this chapter shows that the reliability of irregular MINs is better than all the existing regular MINs.

Assumptions

The following assumptions are made during the analyses, which are identical to those made for regular and irregular networks.

1. All the SEs are independent of one another. Failure of one does not affect the reliability of other.
2. The switch failures occur independently in a network with a failure rate of λ (=10^6 per hour) per unit time. Based on the gate count failure rate for a 2 x 2 SE, $\lambda_2 = \lambda$ and for a 3 x 3 SE $\lambda_3 = 2.25\ \lambda$ for m:1 multiplexer $\lambda_m = \lambda_d = \lambda / 4 * m$ = failure rate of 1:m demultiplexer. (where λ_m - failure rate of multiplexer, λ_d - failure rate of demultiplexer, λ_2 - failure rate of 2x2 SE, λ_3 – failure rate of 3x3 SE)

4.1 RELIABILITY ANALYSES OF REGULAR NETWORKS

The general goals for the design of fault-tolerant MINs are high reliability, better performance even in the presence of faults, low cost, and simple control.

• term to be added for path length > 2
•• term to be added for path length > 4
⊗ term to be added for maximum path length

4.1.1 AUGMENTED SHUFFLE EXCHANGE NETWORK (ASEN-2)

ASENs allows two ways of routing a connection request in every stage except the final stage. If the destinations are assumed to be fault-free, it follows that all single-switch faults can be tolerated. ASENs are single switch fault-tolerant. We can characterize the faults that are tolerated and then derive probabilistic measures of the effect of faults, given that each switching element has a certain reliability. The MTTF of ASEN-2 is at least an order of magnitude higher than that of unique-path MINs, for small sizes, and the difference increases rapidly as the size of the network increases. Both pessimistic and optimistic analyses have been done for the exact value of MTTF.

4.1.1.1 Pessimistic (Lower Bound) Analyses of ASEN-2

For the pessimistic analyses, it has been assumed that an ASEN-2 becomes faulty (i.e. some source is disconnected from some destination) whenever more than one loop is affected by faults in a conjugate pair of loops. We evaluate the MTTF using a series-parallel reliability model: the two loops in a conjugate pair are in parallel and all the conjugate pair of loops are in series. The pessimistic reliability model for ASEN-2 is shown in Figure 4.1 and the reliability expression is given by:

$$R_{pas,\ ASEN}(t) = [1-(1-e^{\frac{-\lambda t}{3m}})^2]^{N/4} \cdot [1-(1-e^{\frac{-\lambda t}{3}})^2]^{N/4 * (n-3)^{\bullet}} \cdot [1-(1-e^{\frac{-\lambda t}{2m}})^2]^{N/4}$$

&

$$MTTF_{ASEN}(t) = \int_0^{\infty} R_{pas,\ ASEN}(t)\, dt$$

- • term to be added for path length > 2
- •• term to be added for path length > 4
- ⊗ term to be added for maximum path length

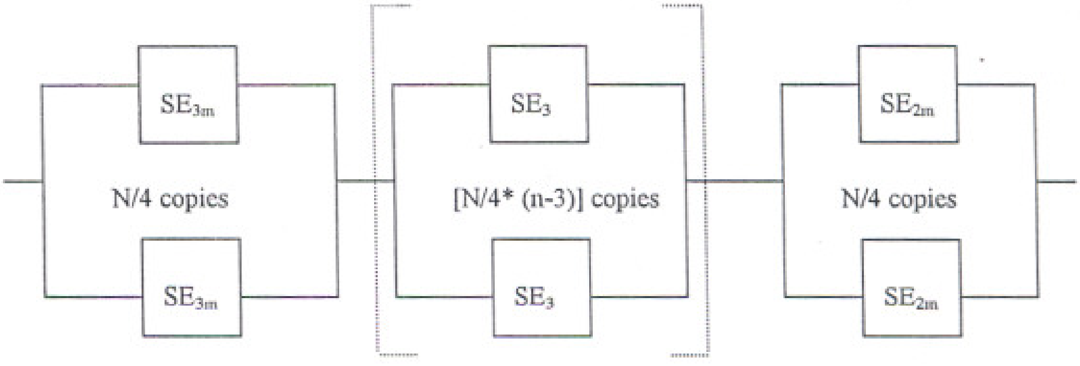


Fig 4.1: Lower Bound reliability block diagram of ASEN-2 network

4.1.1.2 Optimistic (Upper Bound) Analyses of ASEN-2

To obtain an optimistic estimate of MTTF for ASEN-2, we find that in each stage there are two "gateway" switches that receive connections from the previous stage. These two form a conjugate pair of switches. Thus, if both the switches in a conjugate pair fail, then clearly some source is disconnected from some destination. For optimistic analyses it has been assumed that an ASEN-2 becomes faulty only when a conjugate pair of switches is present in a fault pattern. This case also has a series-parallel model, with the switches in a conjugate pair in parallel and all the conjugate pairs in series. The optimistic reliability model for ASEN-2 is shown in Figure 4.2 and the reliability expression is given by:

$$\mathbf{R_{opt,\,ASEN}\,(t) = [1-(1-e^{-\lambda_m t})^2\,]^{N/2} \,.\, [1-(1-e^{-\lambda_3 t})^2\,]^{N/4+N/4\,*\,(n-3)^{\bullet}} \,.\, [1-(1-e^{-\lambda_{2m} t})^2\,]^{N/4}}$$

&

$$\mathbf{MTTF_{ASEN}\,(t) = \int_0^{\infty} R_{opt,\,ASEN}\,(t)\,dt}$$

$\bullet$ term to be added for path length > 2

$\bullet\bullet$ term to be added for path length > 4

⊗ term to be added for maximum path length

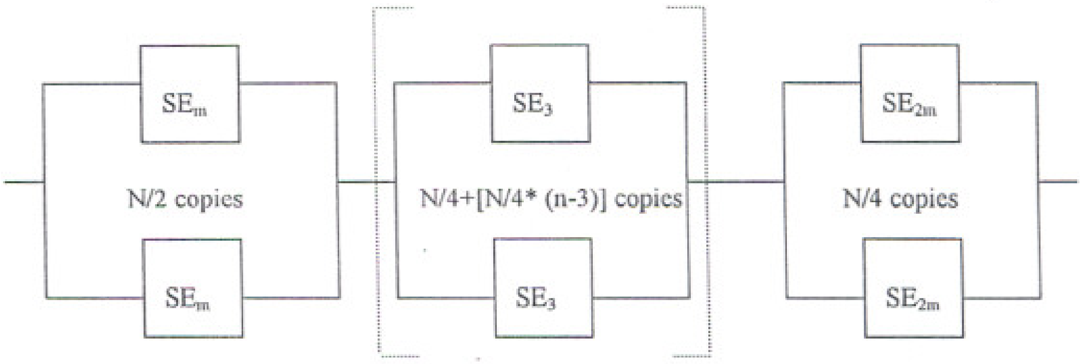


Fig 4.2: Upper Bound reliability block diagram of ASEN-2 network

4.1.2 AUGMENTED BASELINE NETWORK (ABN)

ABNs provides two disjoint paths for each source-destination pair, so ABNs are single switch fault-tolerant. The reliability of ABNs in terms of MTTF is analyzed with the following assumptions:

1. Switch failures occur independently in a network with a failure of λ for 2 x 2 crossbar switches.
2. Failure of multiplexers and demultiplexers also occur independently with failure rates of λ_m and λ_d respectively, which can be different from λ.

To simplify the analyses, it has been assumed that $\lambda_m = m\ \lambda/4$ for a m x 1 MUX or λ_d ($=\lambda_m$) for a 1 x m DEMUX.

4.1.2.1 Pessimistic (Lower Bound) Analyses of ABN

At the input side of the ABN, the routing scheme does not consider the multiplexers to be an integral part of a 3 x 3 switch.

- • term to be added for path length > 2
- •• term to be added for path length > 4
- ⊗ term to be added for maximum path length

To obtain a pessimistic (lower) bound on the reliability of ABN, 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 4.3. The expression for lower bound MTTF of the ABN is:

$$\mathbf{R_{pas,\ ABN}\ (t) = [1-(1-e^{-2\lambda_{3m} t})^2]^{N/8} \cdot [1-(1-e^{-2\lambda_3 t})^2]^{N/8 * (n-4)^{\bullet}} \cdot [1-(1-e^{-\lambda_{2d} t})^2]^{N/4}}$$

&

$$\mathbf{MTTF_{ABN}\ (t) = \int_0^{\infty} R_{pas,\ ABN}\ (t)\ dt}$$

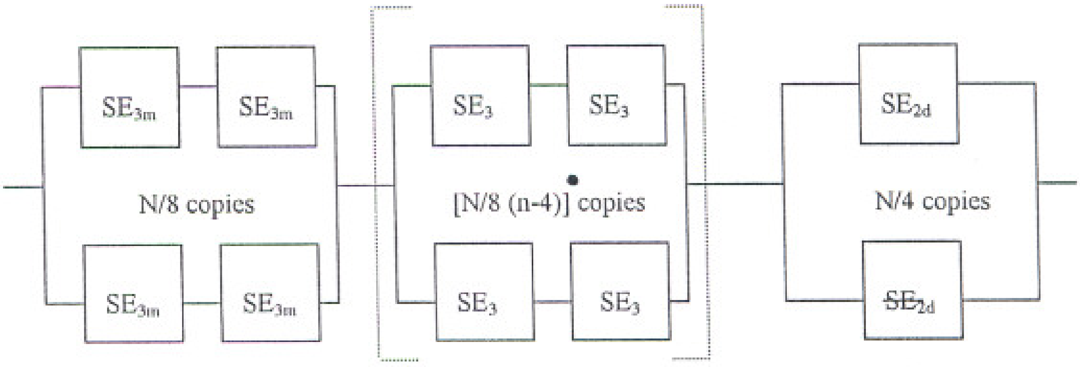


Fig. 4.3: Reliability block diagram of ABN for MTTF lower bound

4.1.2.1 Optimistic (upper bound) Analyses of ABN

To obtain an upper bound for the ABN, it has been observed that each source is connected to two multiplexers in each sub-network, and each switch has a conjugate. We assume that the ABN is operational as long as one of the two multiplexers attached to a source is operational and as long as conjugate pair (loop or switch) is faulty.The reliability block diagram of the optimistic (upper) bound is as shown in Figure 4.4 and the expression for the upper bound of the ABN reliability is:

$\bullet$ term to be added for path length > 2

$\bullet\bullet$ term to be added for path length > 4

$\otimes$ term to be added for maximum path length

$$R_{opt,\,ABN}\,(t) = [1-(1-e^{-\lambda_m t})^2]^{N/2} \cdot [1-(1-e^{-\lambda_3 t})^2]^{N/4 * (n-3)^{\bullet}} \cdot [1-(1-e^{-\lambda_{2m} t})^2]^{N/4}$$

&

$$MTTF_{ABN}\,(t) = \int_{0}^{\infty} R_{opt,\,ABN}\,(t)\,dt$$

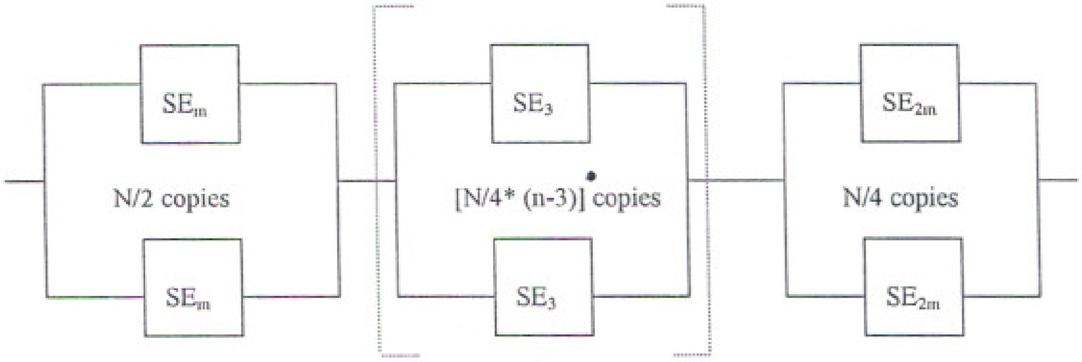


Fig. 4.4: Reliability block diagram of ABN for MTTF upper bound

4.2 RELIABILITY ANALYSES OF IRREGULAR NETWORKS

4.2.1 FAULT-TOLERANT DOUBLE TREE NETWORK (FDOT)

FDOT-k network owes its fault-tolerant property to the presence of k redundant sub-networks. It provides full access property even in the presence of k faults. Since in FDOT redundancy has been added into the system by replicating the entire network (of size N/2 x N/2), so it will be more reliable than MDOT.

But since redundancy added is at the unit level rather than at the component level. Reliability of FDOT at higher network sizes will not be better than FT network.

- ● term to be added for path length > 2
- ●● term to be added for path length > 4
- ⊗ term to be added for maximum path length

4.2.1.1 Pessimistic (lower bound) analyses of FDOT-2

FDOT network is a statically re-routable network. It backtracks to the source for selecting an alternate path. It has been assumed that input stage SE and their associated multiplexers are an integral part to find out the pessimistic reliability of FDOT. The reliability block diagram is as shown in Figure 4.5 and the corresponding reliability expression is given by:

$$\mathbf{R_{pas,\ FDOT}\ (t) = 1-(1-e^{(-N+N/\overset{\bullet}{4}+N/\overset{\bullet\bullet}{8}+\ldots\ldots+\overset{\otimes}{1})\lambda_3 t})^3}$$

&

$$\mathbf{MTTF_{FDOT}\ (t) = \int_0^\infty R_{pas,\ FDOT}\ (t)\ dt}$$

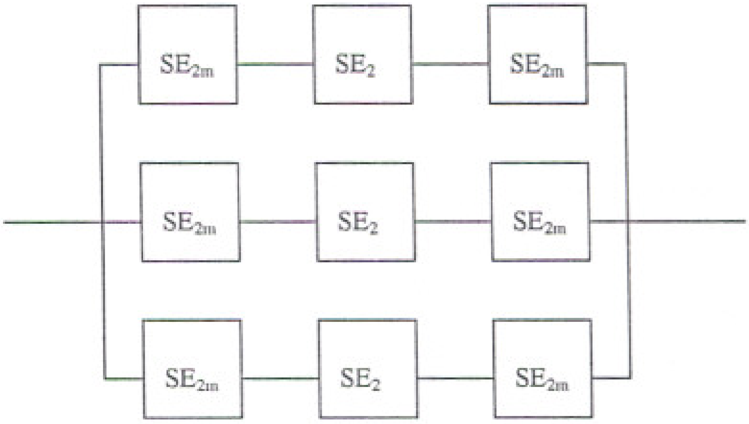


Fig. 4.5: Lower bound reliability block diagram of FDOT-2

4.2.1.2 Optimistic (upper bound) analyses of FDOT-2

Under the worst case, FDOT-k network is k-fault-tolerant. For example, the distribution of k faulty components in (k+1) sub-networks is such that the position of each faulty component is same in all the sub-networks.

$\bullet$ term to be added for path length > 2

$\bullet\bullet$ term to be added for path length > 4

$\otimes$ term to be added for maximum path length

But, if such a critical fault distribution is not there some instances of even more than k faults can be tolerated. For example, if the faulty switches are confined to k or lesser sub-networks. Under such an optimistic approach, FDOT-k network can afford complete failure of as much as k sub-networks in it. The optimistic reliability block diagram is as shown in Figure 4.6. The reliability expression is given by:

$$\mathbf{R_{opt,\,FDOT}(t) = [1-(1-e^{-\lambda_m t})^3]^{N/2} . [1-(1-e^{-\lambda_2 t})^3]^{N/4+N/\overset{\bullet}{4}+N/\overset{\bullet\bullet}{8}+\ldots+1^{\otimes}} . [1-(1-e^{-\lambda_{2m} t})^3]^{N/4}}$$

&

$$\mathbf{MTTF_{FDOT}(t) = \int_0^{\infty} R_{opt,\,FDOT}(t)\,dt}$$

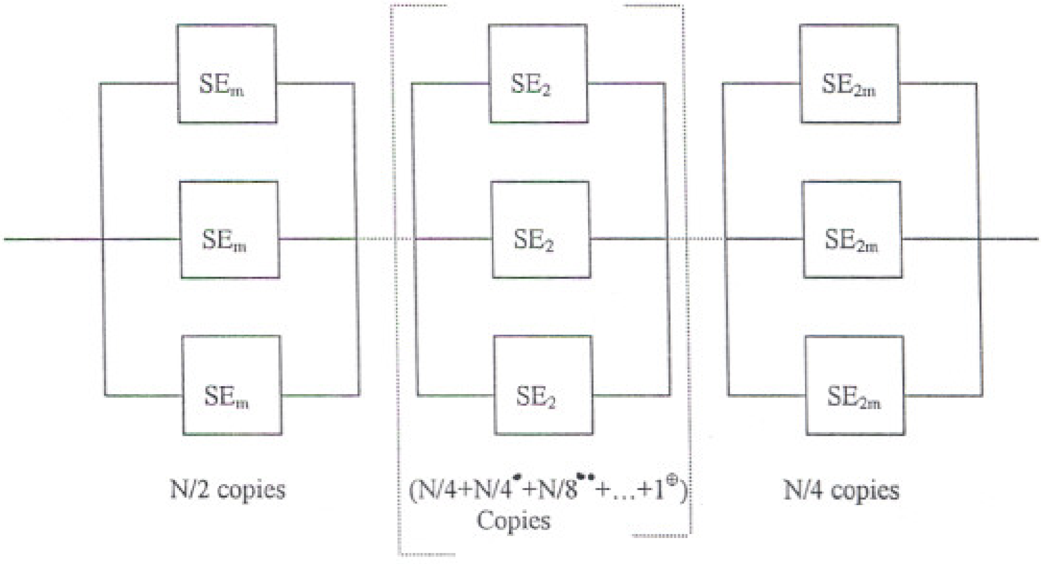


Fig. 4.6: Upper bound reliability block diagram of FDOT-2

4.2.2 FOUR TREE NETWORK (FT)

FT network owes its fault-tolerant property to the presence of a fork at every stage except the last in the network. As a result there exists two ways of routing an input request to the output stage.

$\bullet$ term to be added for path length > 2

$\bullet\bullet$ term to be added for path length > 4

$\otimes$ term to be added for maximum path length

In case of FT network redundancy has been added into the system via intrastage auxiliary links of 3 x 3 SEs. This redundancy has been introduced at the component level, which is generally more effective than the unit level redundancy. So the reliability of FT network is expected to be better.

4.2.2.1 Pessimistic (lower bound) analyses of FT

To obtain the pessimistic reliability of FT, it has been assumed that the multiplexers attached with each input side of the SE is taken as a series system. The failure of any one of the two multiplexers attached to a given SE will be considered as a failure of that particular SE. The pessimistic reliability block diagram is as shown in Figure 4.7 and the corresponding expression is given by:

$$R_{pas,\,FT}(t) = [1-(1-e^{-\lambda_{3m} t})^2]^{N/4} . \{[1-(1-e^{-\lambda_{3} t})^2]^{N/\overset{\bullet}{4}+N/\overset{\bullet\bullet}{8}+.....1}\}^{\otimes} . [1-(1-e^{-\lambda_{2m} t})^2]^{N/4}$$

&

$$MTTF_{FT}(t) = \int_{0}^{\infty} R_{pas,\,FT}(t)\, dt$$

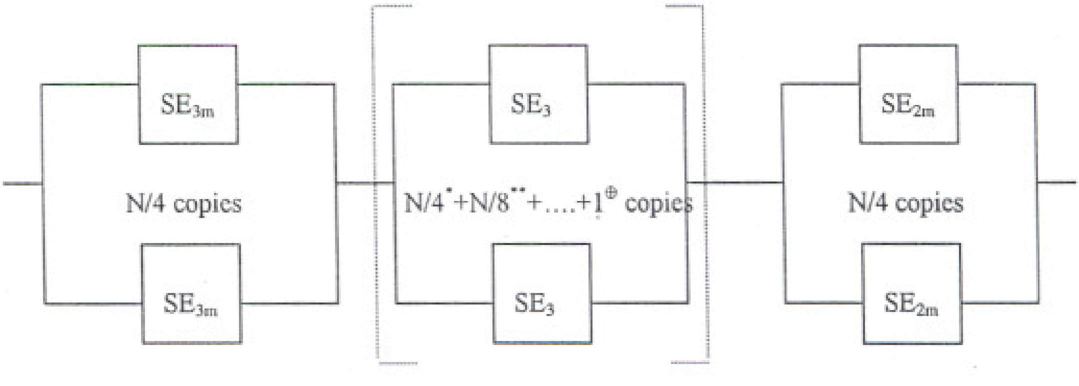


Fig. 4.7: Lower bound reliability block diagram of FT

- $\bullet$ term to be added for path length > 2
- $\bullet\bullet$ term to be added for path length > 4
- $\otimes$ term to be added for maximum path length

4.2.2.2 Optimistic (upper bound) analyses of FT

For the optimistic reliability analyses, it has been assumed that a SE remains operational as long as any one of the two multiplexers attached to it remains operational at the input side. The optimistic reliability model for FT is as shown in Figure 4.8 and the reliability expression is given by:

$$R_{opt,\,FT}(t) = [1-(1-e^{-\lambda_m t})^2]^{N/2} \cdot [1-(1-e^{-\lambda_3 t})^2]^{N/4+N/\overset{\bullet}{4}+N/\overset{\bullet\bullet}{8}+\dots+\overset{\otimes}{1}} \cdot [1-(1-e^{-\lambda_{2m} t})^2]^{N/4}$$

&

$$\mathbf{MTTF_{FT}(t) = \int_0^\infty R_{opt,\,FT}(t)\,dt}$$

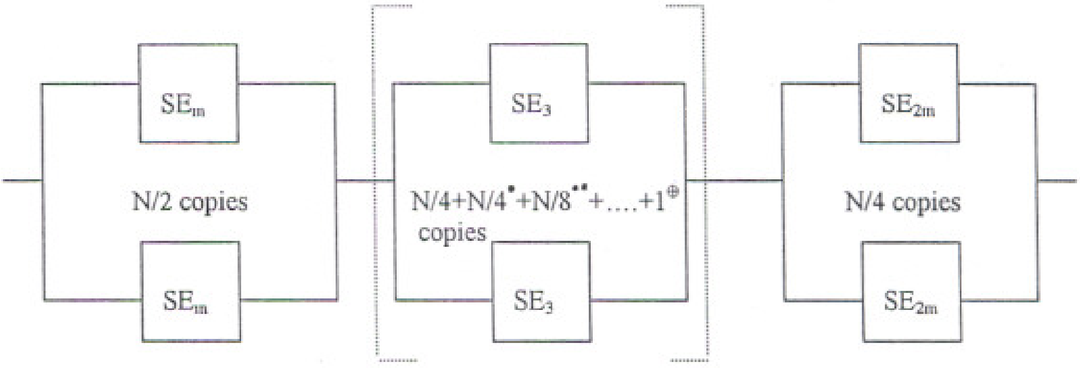


Fig. 4.8: Upper bound reliability block diagram of FT

4.2.3 PROPOSED SMART FOUR TREE NETWORK (SFT)

SFT network also owes its fault-tolerant property to the presence of a fork at every stage except the last in the network. As a result there exists two ways of routing an input request to the output stage. The reliability of SFT is expected to be lower than that of FT network as the number of switches remains same in the inner stages.

● term to be added for path length > 2
●● term to be added for path length > 4
⊗ term to be added for maximum path length

4.2.3.1 Pessimistic (lower bound) analyses of SFT

To obtain the pessimistic reliability of SFT, it has been assumed that the multiplexers attached with each input side of the SE is taken as a series system. The failure of any one of the two multiplexers attached to a given SE will be considered as a failure of that particular SE. The reliability expression for SFT is shown in Figure 4.9 and is given by the following expression:

$$R_{pas,\,SFT}\,(t) = [1-(1-e^{-\lambda_{3m} t})^2]^{N/4} \cdot [1-(1-e^{-\lambda t})^2]^{\overset{\bullet}{N/4 * (2n-5)}} \cdot [1-(1-e^{-\lambda_{2m} t})^2]^{N/4}$$

&

$$\mathrm{MTTF}_{SFT}\,(t) = \int_0^{\infty} R_{pas,\,SFT}\,(t)\,dt$$

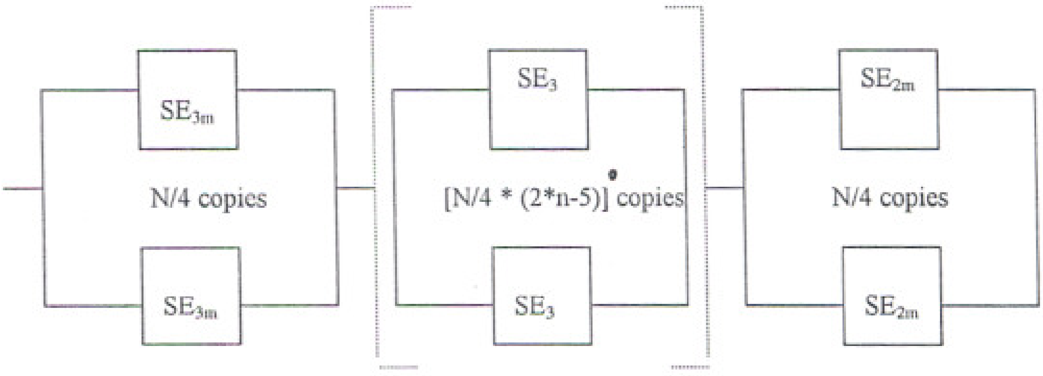


Fig. 4.9: Lower bound reliability block diagram of SFT

4.2.3.2 Optimistic (upper bound) analyses of SFT

For the optimistic reliability analyses, it has been assumed that a SE remains operational as long as any one of the two multiplexers attached to it remains operational at the input side. The reliability expression for SFT is shown in Figure 4.10 and is given by the following expression:

● term to be added for path length > 2
●● term to be added for path length > 4
⊗ term to be added for maximum path length

$$R_{opt,\,SFT}(t) = [1-(1-e^{-\lambda_m t})^2]^{N/2} \cdot [1-(1-e^{-\lambda_3 t})^2]^{N/4+N/4*(2n-5)^{\bullet}} \cdot [1-(1-e^{-\lambda_{2m} t})^2]^{N/4}$$

&

$$MTTF_{SFT}(t) = \int_0^{\infty} R_{opt,\,SFT}(t)\,dt$$

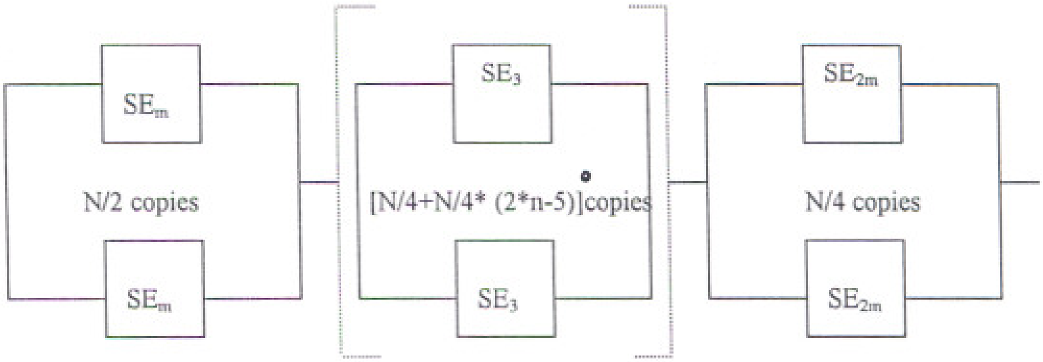


Fig. 4.10: Upper bound reliability block diagram of SFT

4.3 COST-EFFECTIVENESS

A 2 x 2 switch has four units of hardware cost, where as a 3 x 3 switch has nine units. For the multiplexers and demultiplexers, we assume that each of m x 1 multiplexers or 1 x m demultiplexers has m units of cost.

The reliability of Regular and Irregular MINs is shown in Figure 4.11. The results shows that reliability of Irregular MINs is better than Regular MINs. A measure of cost-effectiveness for reliability can be given by comparing MTTF and the cost of the network. Cost-effectiveness of FT & SFT relative to that of ASEN-2 are evaluated and compared and results are shown in Figure 4.12. From the results, we can observe that FT network is more cost-effective than the other regular network i.e. ASEN-2 but almost comparable to that of proposed SFT network for a shorter network.

● term to be added for path length > 2
●● term to be added for path length > 4
⊗ term to be added for maximum path length

4.4 CONCLUSION

Results of the reliability of ASEN-2 – regular MIN and FDOT, FT & proposed SFT-irregular MINs are presented in terms of MTTF for all possible path lengths and for all network sizes.

It is observed that FT network is more reliable than FDOT and almost comparable to that of proposed SFT network.

There is wide variation in optimistic and pessimistic results in network. Optimistic results prove it to be the most reliable of all. While pessimistic results contradict this fact. The reason is that under pessimistic assumptions every SE has got 2 redundant counterparts which can replace it in case of a failure and the probability of failure of all the three SEs in a group is very low.

Optimistic results of FT network are not as good as that of FDOT, but probability of FT network under worst fault distribution is highest among all the irregular networks i.e. FDOT and SFT. The results show that redundancy added through intra-stage link is more effective than added through extra sub-networks under worst-case fault distribution.

A look at all the tables further reveals that, after a certain path length, MTTF degrades only slightly at higher path lengths. The reason is that the number of extra SEs which come into picture at higher path lengths keep on decreasing, as the path length increases further and further.

• term to be added for path length > 2
•• term to be added for path length > 4
⊗ term to be added for maximum path length

Lower Bound MTTF Vs. Size for ASEN-2

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
3	118383	-	-	-	-	-	-
4	-	69950	-	-	-	-	-
5	-	-	43375	-	-	-	-
6	-	-	-	27700	-	-	-
7	-	-	-	-	18035	-	-
8	-	-	-	-	-	11900	-
9	-	-	-	-	-	-	7928

Upper Bound MTTF Vs. Size for ASEN-2

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
3	134935	-	-	-	-	-	-
4	-	77685	-	-	-	-	-
5	-	-	47339	-	-	-	-
6	-	-	-	29855	-	-	-
7	-	-	-	-	19255	-	-
8	-	-	-	-	-	12611	-
9	-	-	-	-	-	-	8353

Lower Bound MTTF Vs. Size for ABN

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
2	94872	-	-	-	-	-	-
3	-	53944	-	-	-	-	-
4	-	-	32667	-	-	-	-
5	-	-	-	20546	-	-	-
6	-	-	-	-	13241	-	-
7	-	-	-	-	-	8676	-
8	-	-	-	-	-	-	5752

Upper Bound MTTF Vs. Size for ABN

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
2	171627	-	-	-	-	-	-
3	-	91329	-	-	-	-	-
4	-	-	53434	-	-	-	-
5	-	-	-	32884	-	-	-
6	-	-	-	-	20867	-	-
7	-	-	-	-	-	13511	-
8	-	-	-	-	-	-	8872

Lower Bound MTTF Vs. Size for FDOT-2

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
2	114583	57291	28645	14322	7161	3580	1790
4	91666	45833	22916	11458	5729	2864	1432
5	83333	-	-	-	-	-	-
6	-	41666	20833	10416	5208	2604	1302
7	-	39855	-	-	-	-	-
8	-	-	19927	9963	4981	2490	1245
9	-	-	19503	-	-	-	-
10	-	-	-	9649	4824	2412	1206
11	-	-	-	9598	-	-	-
12	-	-	-	-	4799	2399	1199
13	-	-	-	-	4786	-	-
14	-	-	-	-	-	2390	1195
15	-	-	-	-	-	2388	-
16	-	-	-	-	-	-	1194
17	-	-	-	-	-	-	1193

Upper Bound MTTF Vs. Size for FDOT-2

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
2	374266	275514	206299	156759	120075	92807	72163
4	350031	259732	195700	149216	114815	88964	69304
5	339963	-	-	-	-	-	-
6	-	253076	191156	146020	112510	87267	68035
7	-	249992	-	-	-	-	-
8	-	-	189037	144521	111425	86466	67434
9	-	-	188012	-	-	-	-
10	-	-	-	143437	110638	85884	66997
11	-	-	-	143260	-	-	-
12	-	-	-	-	110509	85789	66926
13	-	-	-	-	110445	-	-
14	-	-	-	-	-	85718	66872
15	-	-	-	-	-	85706	-
16	-	-	-	-	-	-	66863
17	-	-	-	-	-	-	66859

Lower Bound MTTF Vs. Size for FT

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	16x16	32x32	64x64	128x128	256x256	512x512	1024x1024
2	142743	95166	64500	44244	30613	21315	14907
4	118383	79767	54483	37582	26109	18231	12777
5	110006	-	-	-	-	-	-
6	-	74387	50941	35205	24491	17118	12005
7	-	72068	-	-	-	-	-
8	-	-	49407	34173	23786	16632	11667
9	-	-	48690	-	-	-	-
10	-	-	-	33455	23296	16294	11432
11	-	-	-	33390	-	-	-
12	-	-	-	-	23217	16239	11394
13	-	-	-	-	23178	-	-
14	-	-	-	-	-	16199	11366
15	-	-	-	-	-	16192	-
16	-	-	-	-	-	-	11362
17	-	-	-	-	-	-	11359

Upper Bound MTTF Vs. Size for FT

(failure rate $\lambda = 10^{-6}$ per hour)

N/w P. L.	**16x16**	**32x32**	**64x64**	**128x128**	**256x256**	**512x512**	**1024x1024**
2	171627	115276	78546	54084	37525	26178	18334
4	134935	91329	62585	43274	30115	21054	14768
5	123345	-	-	-	-	-	-
6	-	83717	57485	39807	27731	19403	13617
7	-	80538	-	-	-	-	-
8	-	-	55352	38354	26732	26732	13134
9	-	-	54368	-	-	-	-
10	-	-	-	37361	26048	18235	12803
11	-	-	-	37203	-	-	-
12	-	-	-	-	25939	18159	12750
13	-	-	-	-	25885	-	-
14	-	-	-	-	-	18103	12711
15	-	-	-	-	-	18094	-
16	-	-	-	-	-	-	12704
17	-	-	-	-	-	-	12701

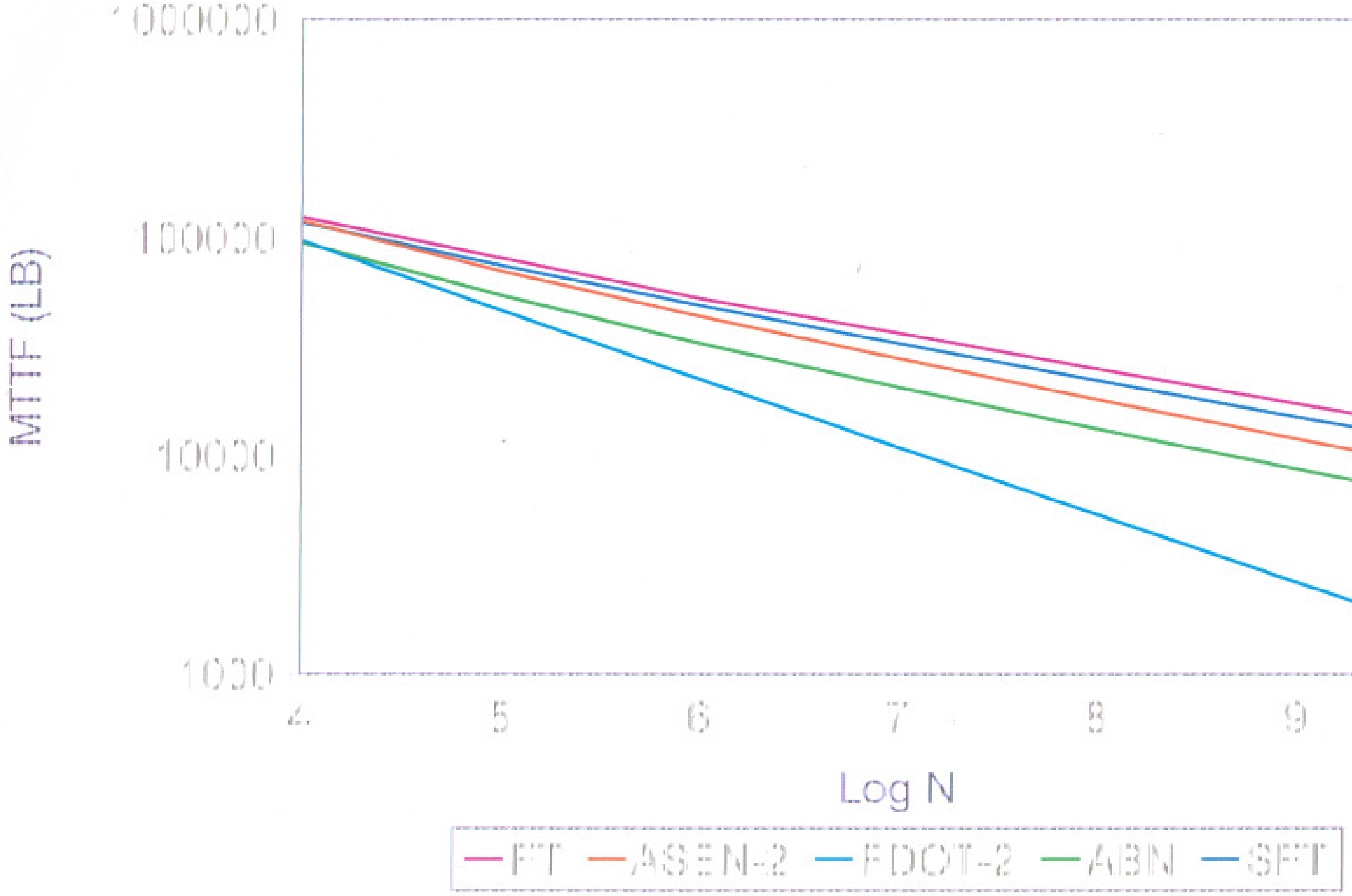
MTTF Comparisons for ASEN-2, FT and Smart FT

Size		N=16	N=32	N=64	N=128	N=256	N=512	N=1024
ASEN-2	**(UB)**	134935	77685	47339	29855	19255	12611	8353
	(LB)	118383	69950	43375	27700	18035	11900	7928
FT	**(UB)**	143302	92715	61667	41680	28567	20744	13724
	(LB)	123710	80347	53604	36341	24955	17277	12113
SFT	**(UB)**	136078	86187	56566	37909	25753	17658	12187
	(LB)	117616	74392	48691	32540	22049	15083	10390

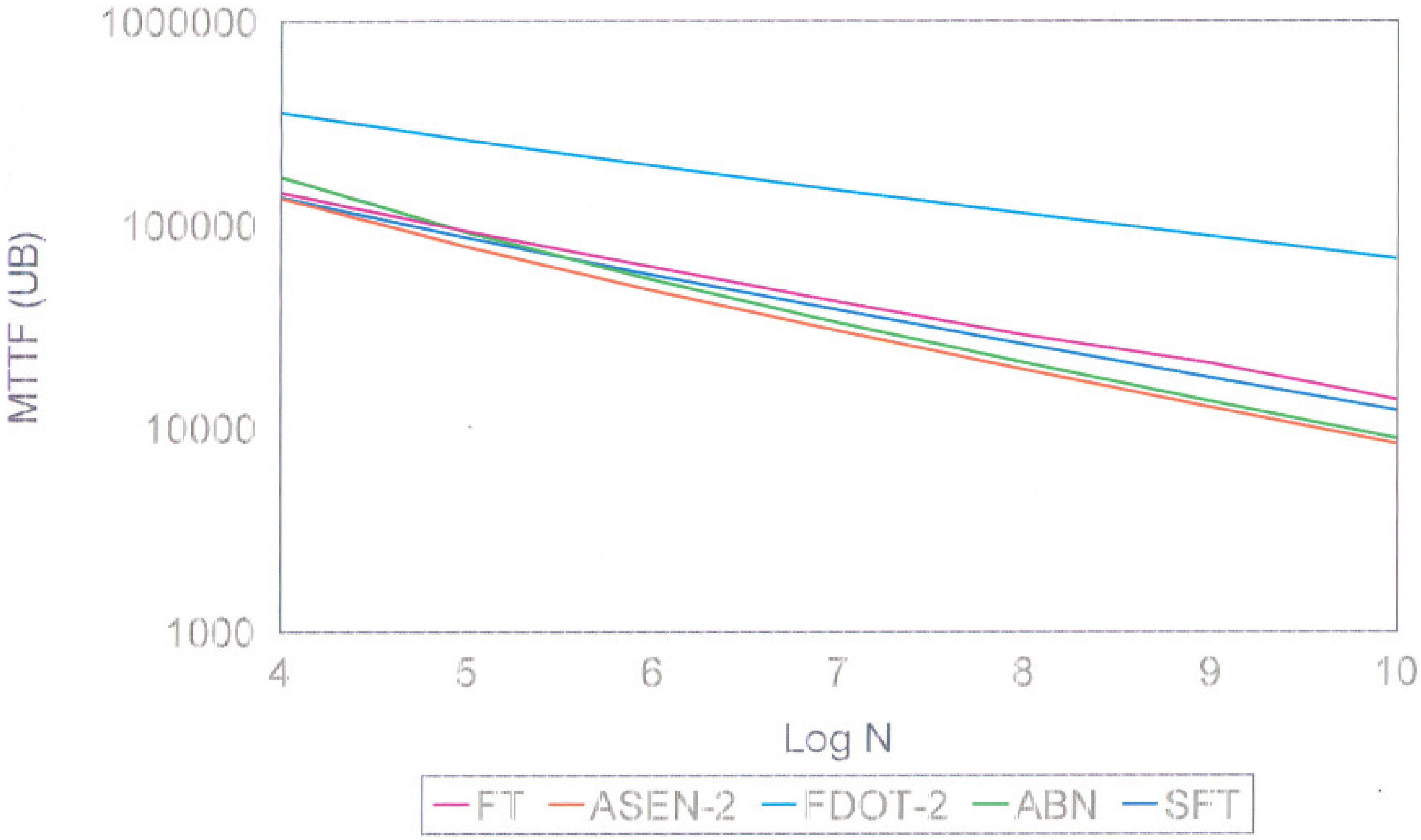
UB – Upper Bound MTTF

LB – Lower Bound MTTF

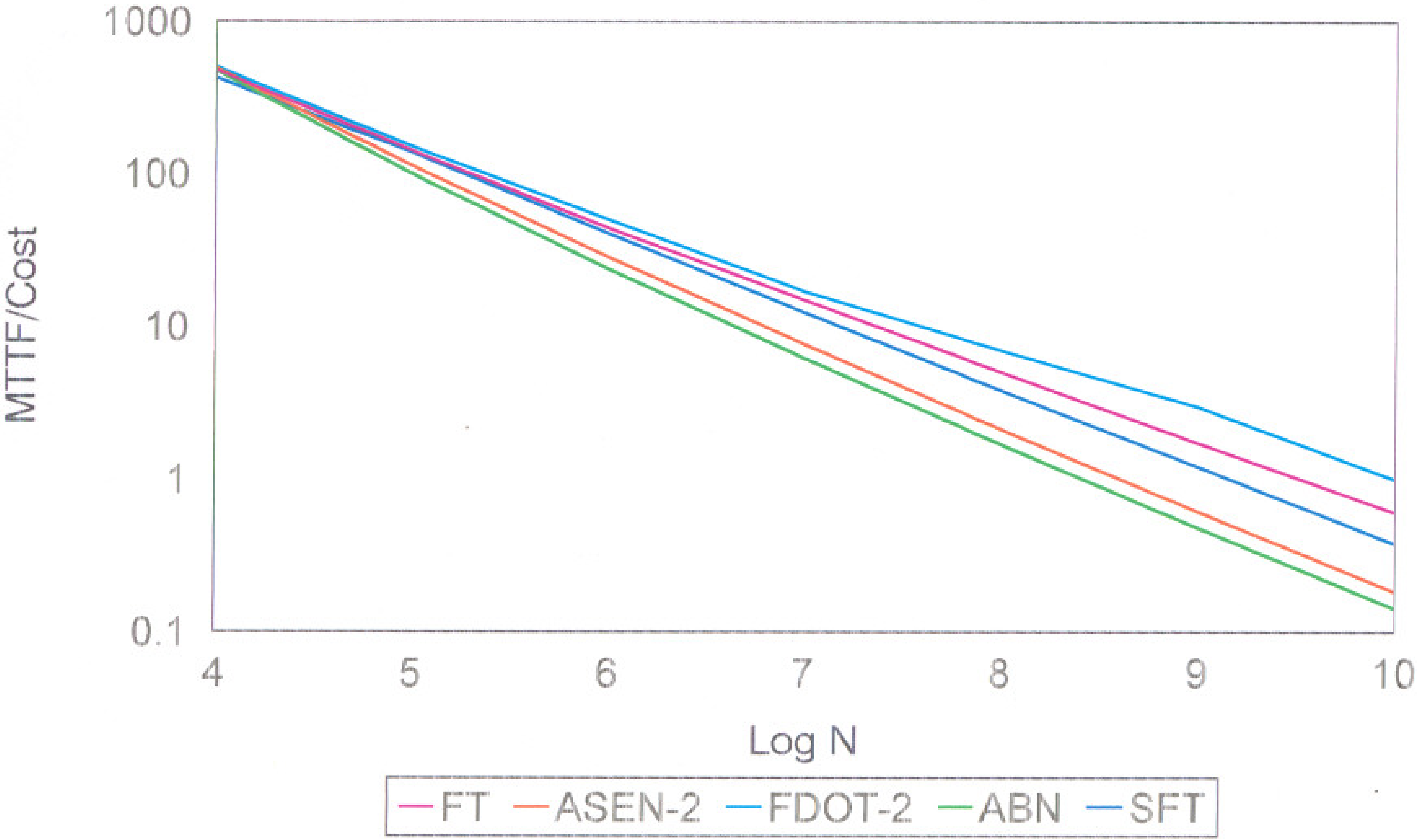
MTTF (LB) Comparison for Regular & Irregular Networks



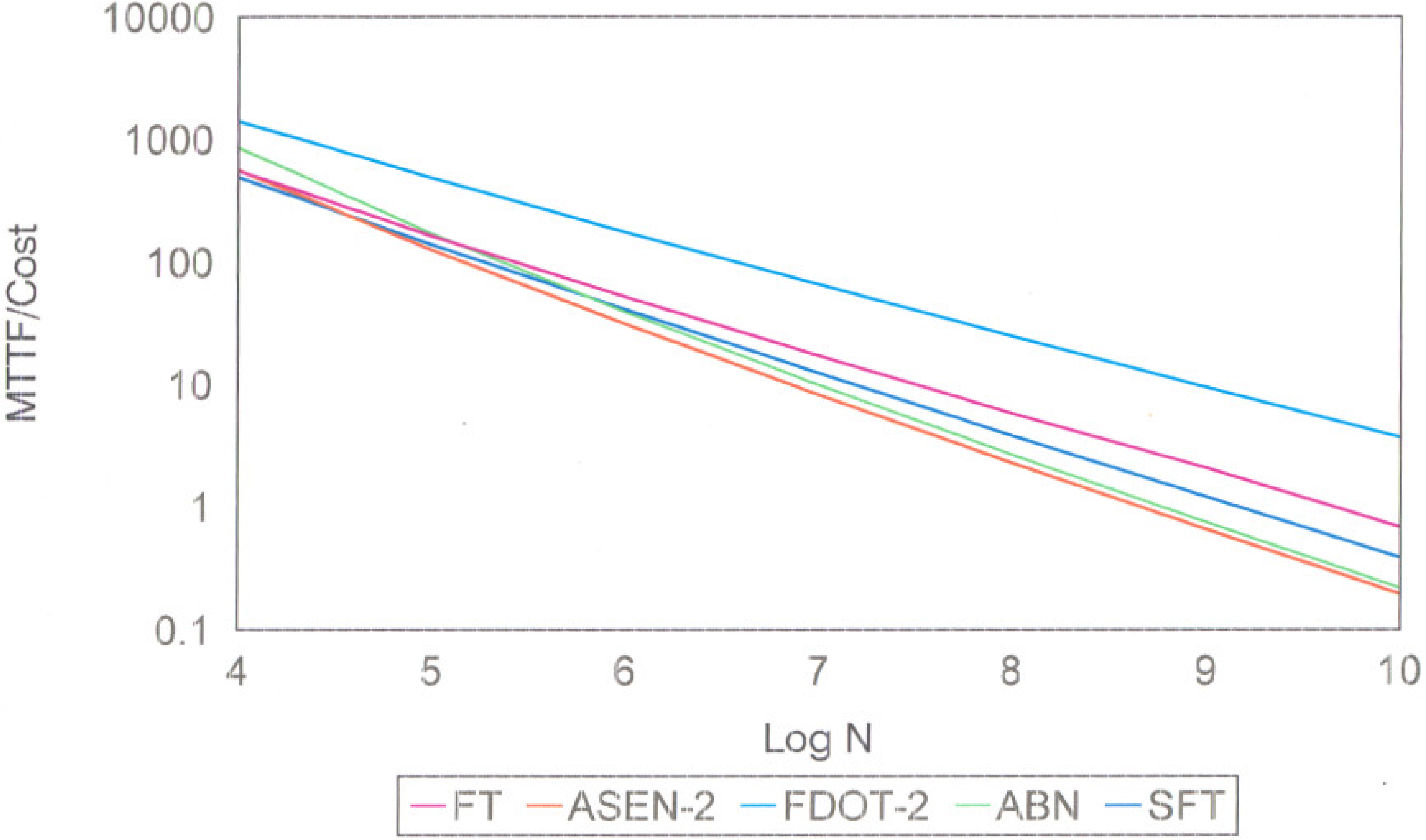
MTTF (UB) Comparison for Regular & Irregular Networks



MTTF(LB) / COST Comparison for Regular & Irregular Networks



MTTF(UB) / COST Comparison for Regular & Irregular Networks



CHAPTER-5

CONCLUSIONS & FURTHER SCOPE

5.1 CONCLUSIONS

The thesis analyses Regular and Irregular Networks in terms of Reliablility, Permutation Passability and Cost-effectiveness. The work presented here can be summarized as under:

- ❑ Survey of existing Regular and Irregular MINs has been done.
- ❑ A new network, called SFT, has been proposed. This has been designed modifying the existing FT network. It involves lesser number of stages but with marginally increased number of switches as compared to FT network. The analyses of the proposed SFT establishes that it is better than FT network in terms of Permutation Passable and is as good as the FT network as far as the Reliability and Cost is concerned.
- ❑ The reliability of various Regular (ASEN, ABN) and Irregular (MDOT, FT, FDOT-2, SFT) has been worked out in terms of MTTF. It follows from the analyses that FDOT-2 MIN is better than FT, SFT and ASEN, but under optimistic conditions only. While analyses of FT network shows it to be better than the existing regular network i.e. ASEN-2 & ABN networks, especially at higher network sizes. Also it can be deduced from this analyses that MTTF/Cost ratio is greater for Irregular MINs than Regular MINs.
- ❑ The Permutation Passable analyses, in terms of Identity and Incremental Permutation Layouts, of various Regular and Irregular MINs has been done considering a number of sample cases. This analyses shows that, in general, proposed SFT provides better performance than the various Irregular MINs discussed.

From the analyses, we can safely conclude that under non-uniform traffic input, where most of the time a source (processor) communicates with its favourite destination (processor/memory module), irregular networks offer a good substitute to regular networks in a most efficient, reliable and cost-effective manner.

5.2 FURTHER SCOPE

The field of irregular networks can be further explored in the light of the following suggestions.

- The designing of more irregular networks having better reliability and permutations can be explored.
- The networks can be applied for ATM routing.
- The detailed probability of acceptance and bandwidth analyses can be done for all the Irregular MINs.

REFERENCES

[1] Adams III G.B.; Aggarwal D.P.; Siegel H.J.; "A survey and comparison of fault-tolerant multistage interconnection network", IEEE Computers, June 1987, pp. 14-27.

[2] Algirdas Avizenis, "Towards Systematic Design of Fault-Tolerant Systems", IEEE Computer, Vol. 30, No. 4, April 1997, pp. 51-58.

[3] Bansal, P.K.; Singh, K.; Joshi, R.C.; and Siroha, G.P.; "Fault tolerant double tree network", Proceedings of International Conference IEEE INFOCOM, April 1991, pp. 462-468.

[4] Bansal, P.K.; Singh, K.; Joshi, R.C.; "Routing and path length algorithm for a cost-effective four-tree multistage interconnection network", International Journal of Electronics, Vol. 73, 1992, pp. 107-115.

[5] Bansal Savina; "Performance and Reliability Aspects of Multiprocessor Multistage Irregular Interconnection Networks", ME Dissertation, TIET, Patiala, December 1993.

[6] Bhuyan Laxmi N.; Qing Yang; Agrawal Dharma P.; "Performance of Multiprocessor Interconnection Networks", IEEE Computer, Vol. 22, Feb 1989, pp. 25-37.

[7] Binh Vien Dao; Jose Duato; Sudhakar Yalamanchili, " Architectural Support for reducing Communication Overheads Multiprocessor Interconnection Networks", Proc. Third International Symposium on High Performance Computer Architecture, Feb. 1-5, 1997, San Antonio, Texas, USA.

[8] Blake J.T.; Trivedi K.S.; "Reliabilities of two fault tolerant interconnection networks", IEEE Computers, 1988, pp. 300-305.

[9] Cutler M.; Shiloach Y., "Permutation Layout", IEEE on Networks, Vol. 8, 1978, pp. 253-278.

[10] Decegama A.L.; "Parallel processing architectures and VLSI hardware", Vol. I, PHI 1989.

[11] Gupta JyotsnaSen; Bansal P.K.; " Fault-Tolerant Routing in Irregular MINs", Proceedings of IEEE TENCON'98, New Delhi, 17-19 December 1998, pp. 638-641.

[12] Hwang K.; Brigges F.A.; Computer Architecture and Parallel Processing, McGraw Hill, Computer Organization and Architecture, 1984.

[13] Isao Shirakawa,"Some comments on Permutation Layout", IEEE on Networks, Vol. 10, 1980, pp. 179-182.

[14] Jose Duato; Sudhakar Yalamanchili; Lionel Ni; "Interconnection Networks", an Engineering Approach, IEEE Computer Society, 1997.

[15] Kruskal C.P.; Snir M.; "Performance of Multistage Interconnection networks for multiprocessors", IEEE Transactions on Computers, Vol. C-32, Dec. 1983, pp. 1091-1098.

[16] Kumar V.P.; Reddy S.M.; "Augmented Shuffle Exchange Multistage Interconnection Network", IEEE Computers, June 1987, pp. 30-40.

[17] Kumar V.P. ; Reddy S.M.; " On a class of fault-tolerant multistage interconnection networks", AT & T Bell Labs, Tech Memo, Oct. 1986.

[18] Kumar V.P. ; Reddy S.M.; "Design and analyzes of fault-tolerant multistage interconnection networks with low link complexity", Proceedings of 12th Annual Int. Symp. On Computer Architecture, Computer Society Press, Washington D.C., 1985, pp. 376-386.

[19] Kumar V.P.; Reibman A.L.; " Failure dependent performance analyzes of fault-tolerant MIN", IEEE Trans. On Copmuters, Dec. 1989, pp. 1703-1713.

[20] Mudge T.N.; Hayes J.P.; Winsor D.C.; " Multiple Bus Architecture" IEEE Computers, June 1987, pp. 42-48.

[21] Padmanabhan K.; Lawrie D.H.; "Fault-tolerant scheme in Shuffle/Exchange Type Interconnection Networks", Proc. International Conference on Parallel Processing, August 1983, pp. 153-164.

[22] Patel J.H.; " Processor – Memory Interconnections for multiprocessors", IEEE Computers 1979, pp. 168-177.

[23] Shen J.P.; Hayes J.P.; "Fault-Tolerance of Dynamic full access Interconnection Networks", IEEE Transactions on Computers, Vol. C-33, March 1984, pp. 241-248.

[24] Chalasani Suresh; Raghavendra C.S.; Varma Anujan; "Fault-tolerant Routing in MIN-Based Supercomputers", Journal of Parallel and Distributing Computing, Vol. 22, 1994, pp. 154-167.

[25] Wei S.; Lee G.; " Extra Group Network: A cost effective fault-tolerant MIN", Proceedings Annual Symp. On Computer Architecture, June 1988, pp. 108-115.

PAPER COMMUNICATED

The following paper out of this thesis have been communicated:

"Permutation & Reliability Measures of Regular & Irregular MINs."

"International Conference IEEE TENCON'2000" to be held in Kuala Lumpur, Malaysia in September'2000.