On Analytic Bounds of Regular and Irregular Fault-tolerant Multi-stage Interconnection Networks

Nitin
Department of CSE/IT
Jaypee University of Information Technology
Waknaghat, Solan 173215, INDIA

Abstract - *Two, ASENs and ZTNs, fault-tolerant multistage interconnection networks of different class are compared, based on their analytic bounds of MTTF, cost, and cost-effectiveness. The full-access condition and dead-fault model are used for the MTTF analysis. The simulation study reveals that ZTNs are better in terms of cost and cost-effectiveness while both networks are equally reliable in terms of upper bounds of MTTF.*

Keywords: Multi-stage interconnection networks, fault-tolerance, reliability, MTTF, full access, dead-fault model.

1 Introduction

As the acceptance and subsequent use of multiprocessor systems increase, the reliability and performance characteristics of the networks that interconnect processors to processors and processors to memories are receiving increased attention. A brief survey of interconnection networks (INs) is found in [1], and a survey of the fault-tolerant attributes of multistage interconnection networks (MINs) is found in [2]. A MIN in particular is an IN but consists of cascade of switching stages.

Today, INs are used in a variety of applications such as switch and router fabrics, processor-memory interconnect, I/O interconnect, and on-chip networks, to name a few. The design of an interconnection network has three aspects—the topology, the routing algorithm used, and the flow control mechanism employed. The topology is chosen to exploit the characteristics of the available packaging technology to meet the requirements (bandwidth, latency, scalability, etc.) of the application, at a minimum cost. Once the topology of the network is fixed, so are the bounds on its performance. For instance, the topology determines the maximum throughput (in bits/s) and zero-load latency (in hops) of the network. The routing algorithm and flow control must then strive to achieve these performance bounds [3].

In this paper, we are going to examine the analytic bounds of the MINs; specifically for the different class of fault-tolerant networks named as Zeta Networks (ZTNs) and Augmented Shuffle-Exchange Networks (ASENs) [4,5]. ZTNs, topology, routing algorithm, and performance issues are already proposed and discussed in [6]. Both the MINs are used as running examples throughout the paper and are compared based on three different metrics:

1. The mean time to failure (MTTF),

2. Cost and Cost-effectiveness.

The simulation corroborates all three metrics used for the analysis of the example MINs. The rest of the paper is organised as follows: Section 2 provides the network definition of ASENs, and ZTNs. Analytic bounds are analyzed in section 3. The cost and cost-effectiveness are discussed in Section 4 is followed by the conclusion.

2 Regular and Irregular MINs

This section provides brief description of two fault-tolerant MINs based on their regular and irregular topologies.

2.1 Network definition of Augmented Shuffle-Exchange Network

ASEN is a regular MIN, has the same number of switching elements (SEs) in each stage. Starting from a shuffle-exchange MIN [7,8], an Augmented Shuffle-Exchange Network (ASEN) [5] is constructed by adding a stage of 2 x 1 multiplexers (MUX) switches at the input side (for making multiple connections from the processing elements (PEs) to the MIN), replacing the last stage switches by 1 x 2 demultiplexers (DEMUX) switches (for providing multiple connections from the MIN to each PE), and by adding links to connect certain groups of switches within each stage in loops (to provide an alternate way of routing in each stage). There are several versions of the ASEN, depending upon the number of switches included in each loop. In this paper, we limit our discussion to ASEN-2, in which the loops contain exactly two switches. A 16 x 16 ASEN-2 is shown in Figure (1).

2.2 Network definition of Zeta Network

Hybrid ZTN is an irregular MIN, has different number of SEs in each stage are different while hybrid network, builds using different types of SEs in different stages i.e. 2 x 2 and 3 x 3 etc. A ZTN of size 2^n x 2^n (where 2^n are Source (S), 2^n are Destination (D), n = $\log_2 N$, and m = $\log_2(N/2)$ is constructed with the help of two identical groups G^N, [where (N = 0,1)], each consisting of a DOT Network [9] 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. Thus, half of the source-destination terminals with MSB 0 falls into the G^0 group and the others having MSB 1 fall into G^1. Each source and destination is connected to both groups with the help of MUX and DEMUX. A 16 x 16 multipath ZTN (Figure (2)), has extra SEs in the intermediate stages with additional express chaining links, provide better fault-tolerance to the network.

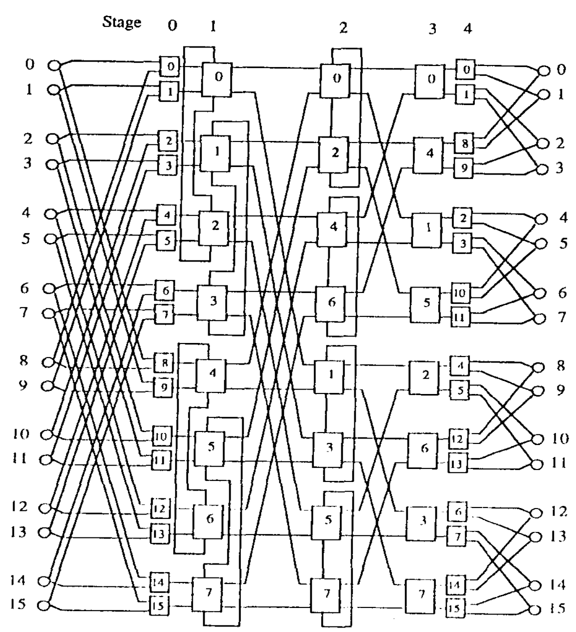

Figure 1. ASEN-2 of size 16 x 16

3 Reliability of Zeta Networks - Mean Time to Failure

In this section, we tried to analyze the reliability of ZTNs in terms of MTTF [10,11] under the criterion of full access [11]. Under this criterion, the ability to reach any output from any input in exactly one pass is preserved in the presence of some faults. To make MTTF analysis more

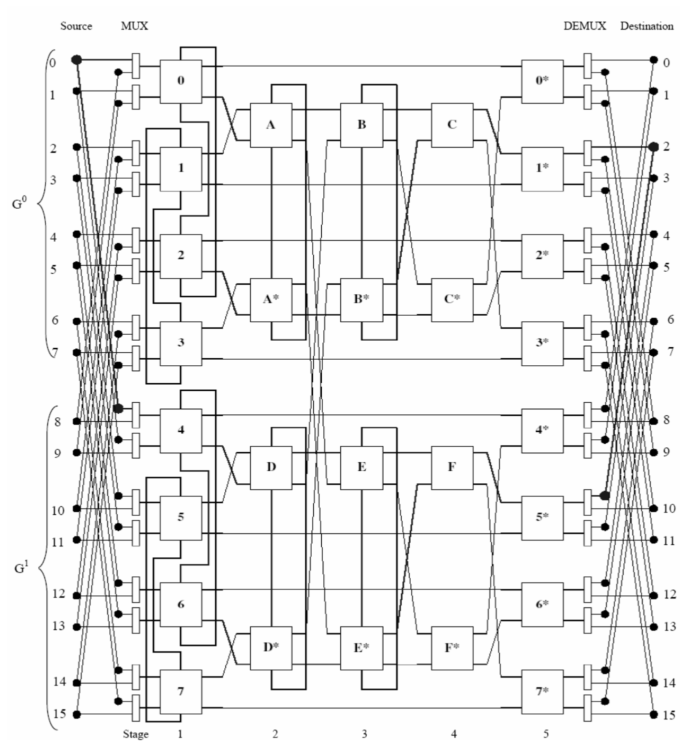

Figure 2. A ZTN of size 16 x 16, highlighting the multiple paths between S-D pair

tractable, we need to have some assumptions. We use the assumptions similar to the ones that have been made previously in the other studies of fault tolerant networks [2,4,6,8,10,11,12]. Under the common assumptions, we tried to compare the reliability of ZTNs with that of ASENs. Our comparisons are based on the upper and lower bounds on the reliability of the networks. For the analysis of MTTF, two assumptions on the failure rates of components are made below.

1. Switch failures occur independently in a network with a Poisson failure rate λ for 2 x 2 crossbar switches,

2. Failures of MUX and DEMUX also occur independently with a Poisson failure rate λ', which can be different from λ. In general, components that are more complicated lead to higher failure rate. Assuming that the hardware complexity of a component is directly proportional to the gate counts of it, one can derive a failure rate of the component. Let γ be the gate failure rate. Assuming that the failure rate of each gate inside the component is equal and failures occur independently, the failure rate of a component containing k gates is $\lambda_k=1-(1-\gamma)^k$. Also the failure rate of a component containing l x k gates is $\lambda_{l \times k}=1-(1-\gamma)^{t \times k}$. We approximate $\lambda_{l \times k}$ with l x λ_k, because λ_k is believed to be small in practice.

From the basic logic design of MUX and DEMUX, we can say that the number of gates in a 2m x l MUX or a 1 x 2m DEMUX is roughly double of that in a m x l MUX or a 1 x m DEMUX. Based on the gate counts of crossbar switches [8], the number of gates in a 2 x 2 crossbar switch is approximately equal to that in a 2 x 1 MUX or a 1 x 2 DEMUX. Thus to simplify the analysis we can assume that $\lambda_m = m\lambda/2$ for a m x 1 MUX, where λ_m failure rate of MUX or λ_d (= λ_m) for 1 x m DEMUX, where λ_d failure rate of DEMUX. The adaptive routing scheme of ZTN considers a 2 x 2 switch in the last stage and its associated DEMUX as a series system, so we consider these three elements as single component (SE_{2d}), and based on a gate count, a failure rate of $\lambda_{2d} = 2\lambda$ can be assigned to this group of elements. Also let λ_2 and λ_3 be the failure rate for the 2 x 2 (SE_2) and the 3 x 3 switch (SE_3), then based on gate count, $\lambda_2 = \lambda$ and $\lambda_3 = 2.25\lambda$ and $\lambda_{3m} = 4.25\lambda$.

3.1 Upper Bound of MTTF

The MTTF of ZTNs can be analyzed by defining a "critical set" of components. A critical set of components is defined as the set of m+1 switching components, each from different groups, such that a network failure will occur if all the m+1 components become faulty simultaneously. Although one can figure out the exact configurations of all the possible critical sets for the analysis of MTTF, we consider only two bounds on the MTTF to simplify the analysis. We first consider an upper bound (UB). For the upper bound, we consider each stage of the network separately; a critical set consists of m+l switching components on the same stage of each group. By considering only the critical sets of each stage independently, an upper bound of MTTF of a ZTN can be obtained, because critical sets can be of components from the different stages of the groups and we ignore the possible failures of such critical sets that lead to a failure of the network. Let $R_s(t)$, the reliability the reliability function of a component, be defined as the probability that a failure does not occur in a time period (0, t). Then

$$R_S(t) = \begin{cases} e^{-\lambda t} & \text{for a } 2 \text{ x } 2 \text{ crossbar switch,} \\ e^{-m\lambda t/4} & \text{for a } m \text{ x } 1 \text{ multiplexer or a} \\ & 1 \text{ x } m \text{ demultiplexer.} \end{cases} \quad (1)$$

Since a critical set of components is faulty if and only if all the components in the set are faulty, the probability, $R_{cs}(t)$, that a critical set is not fault in a time period (0, t) is

$$R_{CS}(t) = \begin{cases} 1-(1-e^{-\lambda t})^{m+1} & \text{for a critical set} \\ & \text{of } 2x2 \text{ switches,} \\ 1-(1-e^{-m\lambda t/4})^{m+1} & \text{for a critical set of} \\ & mx1 \text{ multiplexers or} \\ & 1xm \text{ demultiplexers.} \end{cases} \quad (2)$$

Since a ZTN becomes faulty if any one critical set becomes faulty, Equations (3) and (4) gives the probability that a ZTN is not faulty while Figure (3(a)) gives the combination of series-parallel model, describing the upper bound of ZTN.

$$R_{ZTN-UB}(t) = [1-(1-e^{-\lambda_m t})^2]^{N/2} \\ [1-(1-e^{-\lambda_3 t})^2(1-e^{-\lambda_2 t})^2]^{(N/4+N(2n-5)/4)-1+N/4} \\ [1-(1-e^{-\lambda_{2d} t})^2]^{N/4} \quad (3)$$

$$^{*}MTTF_{ZTN-UB} = \int_0^{\infty} R_{ZTN-UB}(t)dt \quad (4)$$

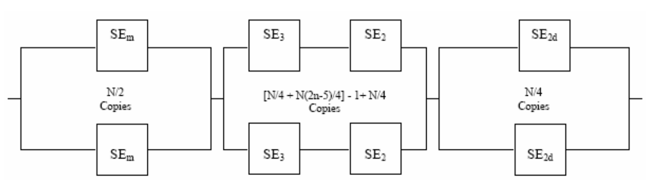

Figure 3(a). Reliability Block Diagram of ZTN for the Evaluation of MTTF in terms of Upper Bound

3.1.1 Augmented Shuffle-Exchange Networks-Upper Bound

Following the similar, approach as mentioned in previous section the MTTF for the Augmented Baseline network in terms of upper bound is:

$$R_{ASEN-2-UB}(t) = [1-(1-e^{-\lambda_m t})^2]^{N/2} \\ [1-(1-e^{-\lambda_2 t})^2]^{N/4+N(n-3)/4} \\ [1-(1-e^{-\lambda_{2d} t})^2]^{N/4} \quad (5)$$

$$^{*}MTTF_{ASEN-2-UB} = \int_0^{\infty} R_{ASEN-2-UB}(t)dt \quad (6)$$

3.2 Lower Bound of MTTF

For lower bound (LB), each group is considered independently and is assumed faulty if there is any single fault in it. Since at the input side of ZTN, routing scheme does not consider the MUX to be the integral part of the 2 x 2 switch. Hence, if both MUX are grouped with each switch in the input side and regarded, as a series system, then we will have the conservative estimate of the reliability of these components. To obtain the pessimistic (lower) bound of the reliability of ZTN here we assume that the network is failed whenever more than one conjugate loop has a faulty conjugate or more than one conjugate fails in the last stage. Following the same way used for obtaining the upper bound, the probability that a ZTN is operative for a time period (0, t) is given by Equations (7) and (8) while Figure (3(b)) gives the combination of series-parallel model, describing the lower bound of ZTN.

$$R_{ZTN-LB}(t) = [1-(1-e^{-\lambda_{3m}t})^2]^{N/4} \\ [1-(1-e^{-\lambda_3 t})^2]^{N(2n-5)/4}[1-(1-e^{-\lambda_{2d}t})^2]^{N/4} \tag{7}$$

$$^*MTTF_{ZTN-LB} = \int_0^\infty R_{ZTN-LB}(t)dt \tag{8}$$

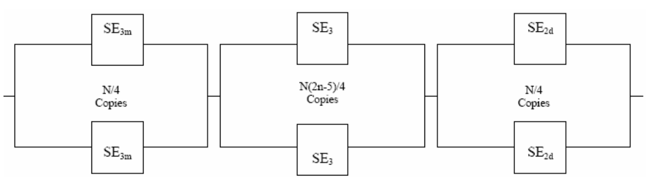

Figure 3(b). Reliability Block Diagram of ZTN for the Evaluation of MTTF in terms of Lower Bound

3.2.1 Augmented Shuffle-Exchange Networks-Lower Bound

Following the similar, approach as mentioned in previous section the MTTF for the Augmented Baseline network in terms of lower bound is :

$$R_{ASEN-2-LB}(t) = [1-(1-e^{-\lambda_{2m}t})^2]^{N/8} \\ [1-(1-e^{-\lambda_2 t})^2]^{N(n-3)/4}[1-(1-e^{-\lambda_{2d}t})^2]^{N/4} \tag{9}$$

$$^*MTTF_{ASEN-2-LB} = \int_0^\infty R_{ASEN-2-LB}(t)dt \tag{10}$$

Relative variations in upper and lower bounds of MTTF of ASEN-2 and ZTNs, are shown in Figure (4) and the results are tabulated in Table-1, from here it can be depicted that the difference in reliability of the two MINs indicates that for large network sizes (N>=64), upper bound of MTTF of ZTNs is almost equal to the upper and lower bounds of MTTF of other fault-tolerant network while on lower bound of MTTF ZTNs are less reliable in comparison to the ASEN-2.

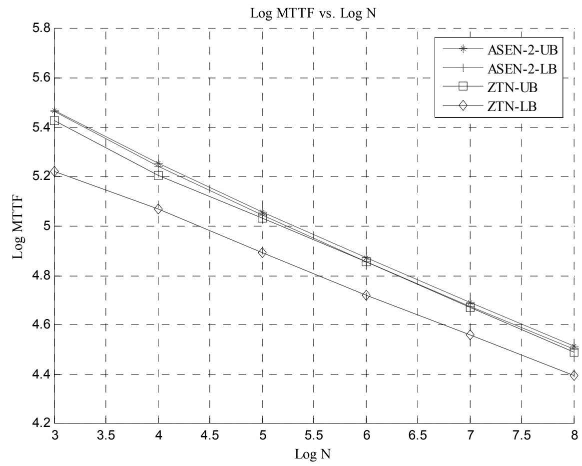

Figure 4. Comparative Upper and Lower Bounds MTTF of related MINs

Table 1
Values of Upper and Lower Bound of MTTF for different Sizes

Networks	Log MTTF vs. Size N					
	N=8	N=16	N=32	N=64	N=128	N=256
ASEN-2-UB	5.47	5.25	5.05	4.87	4.69	4.51
ASEN-2-LB	5.46	5.24	5.04	4.85	4.67	4.50
ZTN-UB	5.42	5.20	5.03	4.85	4.67	4.49
ZTN-LB	5.22	5.06	4.89	4.72	4.56	4.39

3.3 Mean Time to Failure under Repair

In real system utilizing single fault tolerant networks, it is anticipated that the detection of a fault in a switching element initiates the repair of the fault, to guard against the occurrence of a second, potentially catastrophic, fault. There is a conservative approximation of the MTTF of single fault tolerant MINs assuming repair of faults. Let the constant failure rate of individual switches be λ and the constant repair rate be μ. Now suppose we have a MIN with M switches. In a single-fault tolerant network, the Markov chain model shown in Figure (5) where there are three states can conservatively represent the whole system: State 0 is the no-fault state; State 1 is the single-fault state, while State 2 is the two-fault state. It is assumed that if the MIN reaches State 2, it has failed. Since the schemes presented in this paper can tolerate more than one faulty switch in many cases, this model should give a lower bound for the MTTF of the system.

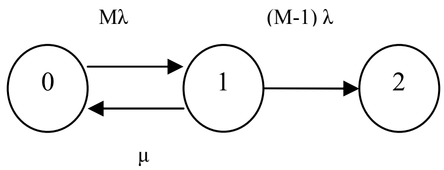

Figure 5. Markov Chain Model of Systems with Repair

The MTTF of a system such as the one in Figure (4) is in the literature [12]. For this system –

$$^{\Phi}MTTF\ with\ repair = \frac{1}{(M-1)\lambda} + \frac{(M-1)\lambda + \mu}{(M-1)N\lambda^2} \qquad (11)$$

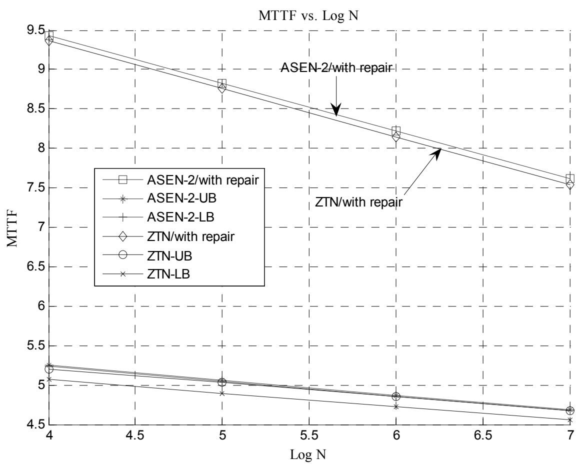

Figure 6. Lower Bound of the MTTF of ASEN-2 and ZTN with Repair

Table 2
Values of MTTF/with repair of different Networks for different Sizes

Networks	Log MTTF vs. Size N			
	N=16	N=32	N=64	N=128
ASEN-2	9.4343215	8.82309512	8.21678640	7.61313985
ZTN	9.3647029	8.75486500	8.14925639	7.54600687

Let $\lambda=10^{-6}$/hour, and $\mu=10^{-1}$/hour (which are typical values). In Figure (6) the MTTF/with repair for ASEN-2 and ZTN is drawn for sizes from 16 to 128 and values are tabulated in Table-2. The figure seems to indicate that with increasing size of the MINs, the MTTF improvement factor is actually decreasing. This is due to the conservative assumption that exactly one fault is tolerated. In reality, with increasing MIN size the average number of faults tolerated increases. However, Figure (6) does show that the actual improvement obtained by any of the fault tolerant MINs discussed in this paper is much better than the constant factor improvement suggested by Figure (4).

All *MTTF Equations [4, 6, 8, and 10,] are integrated in the interval t = 0.001 to 1 and are simulated with the help of MATLAB version 7.0.1. The values for upper and lower bound of MTTF for different class of MINs are provided in Table-1. The $^{\Phi}$ MTTF with repair for ASEN-2 and ZTN from size N=16 to 128 are calculated using the same simulation software and values are tabulated in Table-2.

4 Cost and Cost-effectiveness

To estimate the hardware cost of a network the following assumptions are considered:

1. First, measure the switch complexity with an assumption that the cost of a switch is proportional to the number of gates involved, which is roughly proportional to the number of 'crosspoints' within a switch. For example, a 4 x 4 switch has 16 units of hardware cost; a 3 x 3 switch has 9 units of hardware cost, whereas a 2 x 2 switch has 4 units,

2. For the MUX and DEMUX, it is roughly assumed that each of K x 1 MUX or 1 x K DEMUX has K units of cost.

In this way ASEN-2 and ZTNs have cost of $3N(1.5\log_2 N-1)$ and $N/8(56+24\log_2(N/2))$ respectively. The values of cost function for both the MINs are provided in Table-3. From Figure (7(a)), and Table-3, it is clear that ZTNs are more cost-effective than ASEN-2. This advantage becomes more significant as the network (N>=64) size increases.

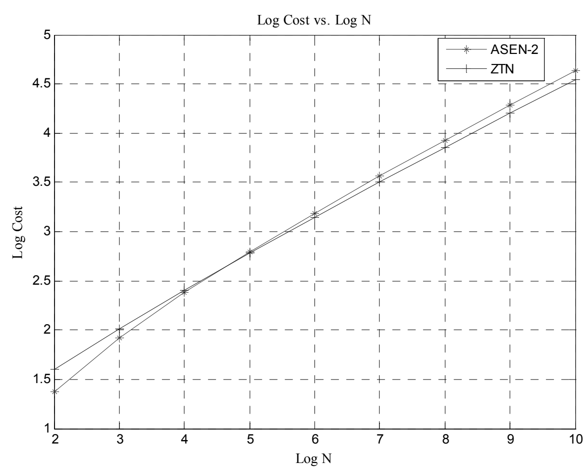

Figure 7(a). Comparative cost of related MINs

Table 3
Cost values of different Networks for different Sizes

Networks	Cost vs. Size N				
	N=4	N=16	N=64	N=256	N=1024
ASEN-2	24	240	1536	8448	43008
ZTN	40	256	1408	7168	34816

Now, a simple measure of the cost-effectiveness for reliability can be given by comparing MTTF and the cost of the network. Let

$$^{\S}Cost - effectiveness(\eta) = \frac{MTTF}{Cost} \quad (12)$$

The cost-effectiveness of ASEN-2 and ZTNs are compared, and the improvements in results are shown in Figure (7(b)) and tabulated in Table-4. From the results, one can observe that ZTNs are more cost-effective than other fault-tolerant regular MIN.

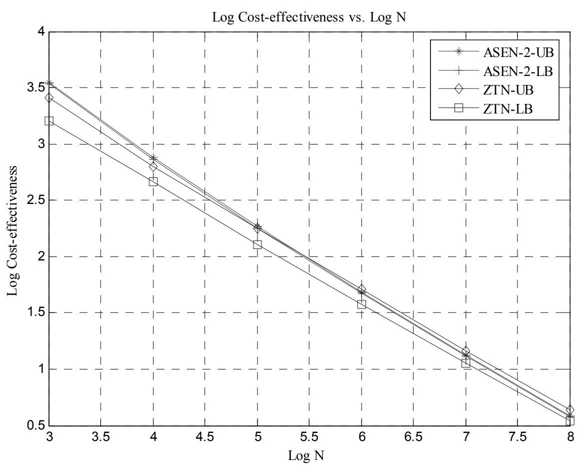

Figure 7(b). Comparative cost-effectiveness of related MINs

Table 4
Values of Cost-effectiveness of different Networks for different Sizes

Networks	Log η vs. Size N					
	N=8	N=16	N=32	N=64	N=128	N=256
ASEN-2-UB	3.54	2.87	2.26	1.68	1.12	0.58
ASEN-2-LB	3.53	2.86	2.24	1.66	1.11	0.57
ZTN-UB	3.41	2.79	2.24	1.70	1.16	0.63
ZTN-LB	3.20	2.66	2.10	1.57	1.05	0.54

The Cost-functions and §Cost-effectiveness of ASEN-2 and ZTN are evaluated and simulated using the MATLAB 7.0.1 and the values are provided in Table-3 and Table-4. All Figures [4, 6, 7(a)-(b)] are resulted from the values tabulated in Tables-[1-4].

5 Conclusions

The full access criterion and the dead-fault model were used to analyze the reliability of ZTNs. In our analysis, any switch, any multiplexer, and any demultiplexer in ZTNs are assumed to have a possibility to fail. The analysis of the upper bound of MTTF showed that ZTNs perform, in general, is equally reliable than other regular fault-tolerant networks i.e. ZTNs and ASEN-2 are equally fault tolerant and robust in the presence of faults. However, if such high reliability comes at the expense of high cost, it may have little value in practice. Our analysis on the cost of the networks shows that ZTNs are, in general, more cost-effective than other fault-tolerant multi-stage interconnection networks.

6 References

[1] Feng, T. Y., "A survey of interconnection networks," Computer, Vol. 14, pp. 12-27, December 1981.

[2] Adams III, G. B., Agrawal, D. P., and Siegel, H. J., "A survey and comparison of fault-tolerant multi-stage interconnection networks," IEEE Computer, Vol. 20, pp. 14-27, June 1987.

[3] Singh, A., "The Load-Balanced Routing in Interconnection Networks," Ph.D. Dissertation, Stanford University, March 2005.

[4] Kumar, V. P. and Reddy, S. M., "Design and analysis of fault-tolerant multistage interconnection networks with low link complexity," Proc.12 lnt 'I Symp. Computer Architecture, pp. 376-386, June 1985.

[5] Kumar, V. P. and Reddy, S. M., "Augmented shuffle-exchange multistage interconnection networks," Computer, pp. 30-40, June 1987.

[6] Subramanian, A. and Nitin, "On a performance of multi-stage interconnection network," IEEE ADCOM, pp. 73-79, December 2004.

[7] Lawrie, D. H., "Access and alignment of data in an array processor," IEEE Trans. Computers, Vol. C-24, pp. 1145-1155, December 1975.

[8] Patel, J. H., "Performance of processor-memory interconnection for multiprocessors," IEEE Transaction on Computers, Vol. 30, pp. 771-780, 1981.

[9] Bansal, P. K., Singh K., and Joshi, R. C., "Routing and path length algorithm for a cost-effective four-tree multi-stage interconnection network," International Journal of Electronics, Vol. 73, No. 1, pp. 107-115, 1992.

[10] Shooman, M. L., "Reliability of computer systems and networks: fault-tolerance, analysis, and design," John Wiley & Sons, Inc., New York, 2002.

[11] Duato, J., Yalamanchili, S., and Ni., L., "Interconnection Networks An Engineering Approach," IEEE Press, 1997.

[12] Smith, D. J., "Reliability Engineering," Barnes & Noble Books, 1972.