ISSN (Online): 2230-7893
www.IJCEM.org

Development of a New Class of Cost Effective and Fault Tolerant Multi Stage Interconnection Network for Parallel Computing System

Shama Pathania[1], Jasbir Singh Saini[2]

[1,2]Guru Nanak Dev Engineering College, Ludhiana

shama1.dec@gmail.com, sainijasbir@yahoo.com

Abstract

In this paper, a review of various regular and irregular routing algorithms of parallel computing networks is done. It can be deduced from the analysis that irregular Parallel computing networks performs better than regular ones because irregular networks are usually less costly and multipath in nature. A major problem in designing a large-scale parallel and distributed system was the construction of an Interconnection Network (IN) to provide inter-processor communication. One of the biggest issues in the development of such a system was the development of an effective architecture that have high reliability, give good performance (even in the presence of faults), low cost higher number of passes of request and a simple control.

In this paper, a new class of irregular fault-tolerant multistage interconnection network named Cost Effective and Fault Tolerant Multi stage Interconnection Network(CEFT) is proposed and analyzed. The CEFT MIN can achieve significant tolerance to faults and good performance with relatively low costs and a simple control scheme. The construction procedure of the CEFT network, algorithm for allocation of path length, routing along with the routing procedure, fault-tolerance aspect is described too.

Keywords: *Cost-effective, Fault Tolerance, Irregular Network, Multistage Interconnection Network, Routing.*

I. INTRODUCTION

In the present era of technology and development, it is very much possible to design and develops multiprocessing system with many of multi-processors[4][10]. Multistage Interconnection network (MIN) play an important role in these systems, which enables processors to communicate with themselves and with memory modules. Multistage Interconnection network consists of more than one stage of switching elements, links that interconnect them[12].

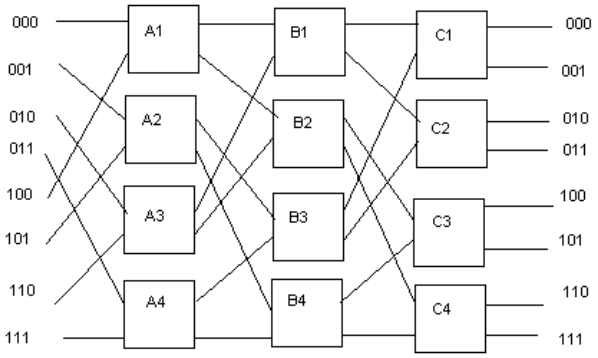

Figure 1: A 3 Stage Omega Network

An Omega network with 3 stages is illustrated in figure 1. In this paper, a new MIN is designed which is irregular and more cost-effective and simple in a way. Regular network has equal number of switching elements per stage so they impose equal time delay to all requests passing through them. Irregular network has unequal number of switching elements per stage, so for a given source to destination pair, there are different path lengths available in this network.

Different networks are available offering varying degrees of reliability, efficiency, and cost and fault tolerance. The flip network, omega network, indirect binary n-cube network, and regular SW banyan network (S = F = 2) are topologically equivalent. CROSSBAR, OMEGA[13][14] are the examples of regular MIN. MDOT, RMDOT, FT, FTD, ESC are the examples of irregular MIN[10][12].

II. CONSTRUCTION METHOD OF CEFT Network

A typical CEFT network is an Irregular Multistage Interconnection Network[1][2][10] of size $2^n x 2^n$ is constructed with the help of multiplexers, demultiplexers, switches.

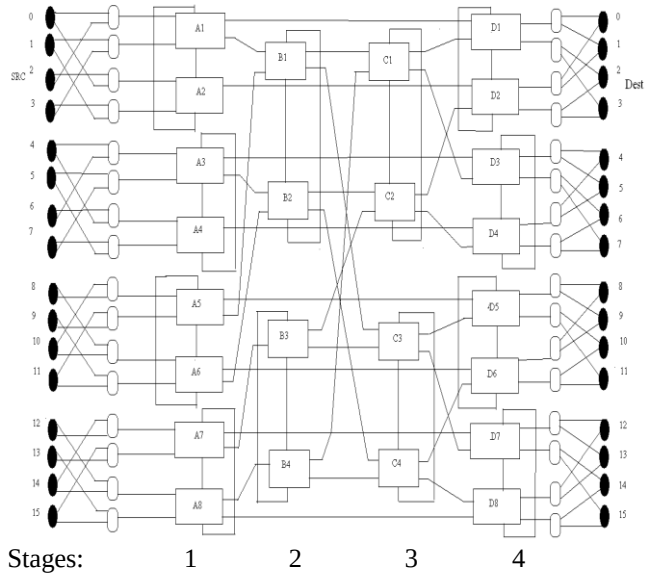

Figure 2: CEFT MIN of size N=16

The network consists of four stages. It consists of 2^{n-1} switches(where $N=2^n$ of N x N network) at stage 1 and stage 4, whereas central stages: stage 2 and stage 3 have exactly 2^{n-2} switches. Each source is connected to two different switches with the help of multiplexer and each destination is connected with demultiplexer. In CEFT MIN, more than one path from a source to destination is available. In case the shortest available route is busy or faulty, request will be routed through alternate available path.

III CEFT NETWORK ALGORITHM FOR PARALLEL COMPUTING SYSTEMS

In a CEFT network as it is very clear from the figure 2, each source and destination is connected with multiplexers, demultiplexers. A CEFT network of size $2^n * 2^n$ consists of 2^n multiplexers and 2^n demultiplexers. Each MUX and DEMUX is of size 2*1 and 1*2 respectively. In CEFT network, we are using switches of size 3*3 so that if primary switch is faulty then the request can be routed to the conjugate switch connected to this due to which less faults are occurred and request still can be matured.

IV DESIGN ISSUES OF ROUTING ALGORITHM FOR CEFT NETWORK

- Multiplexers and demultiplexers are simple with no routing capability.
- Switches have routing capability based on destination tags.
- More than one source destination pair cannot have the same values.
- No looping takes place between conjugate pair of switches.
- Path matrix is used for storing the paths of respective pairs.
- FCFS algorithm is used for serving the requests.

V ROUTING ALGORITHM

In the proposed multi stage interconnection network, there are multiple paths available for a given source to destination, minimum path length is checked if available then routed through it else other route is taken.

Let S be the source and D be the destination of the network and

S=SnSn-1_ _ _ _ _ _S2S1S0
D=DnDn-1_ _ _ _ _ D2D1D0

Check if the minimum path is available for each S to D. if there is a minimum path exists between each S to D then Set Checkflagfor_minim_path to 0 which signifies that there a minimum path exists between S to D and other possible path is also available. In other case, if this flag values not equal to zero then it means that no any path length with minimum route is available but the path with maximum length is available (length can maximum be 10). Now, for routing to take place if switches in the route are not faulty then routing is done through it. But in other case if switches in the routing path are faulty then request is routed to the conjugate switch connected to it.

The algorithm for CEFT network uses the following arrays and some of these arrays are allocated memory dynamically:

Pptr_to_sour_dest_array: for storing the source-destination pair value.

Pptr_to_path_len_array: for storing path length of each source- destination pair value.

Checkflagfor_minim_path: for storing 0 if minimum path exists.

Pptr_to_dest_bit_array: for storing the binary equivalent of destination values.

Ppath_matrix: for storing the path traversed from source to destination.

Step1: initialize the graph matrix representing value 1 if path exists from a given source to destination. Otherwise initialize it with the value 0 if no direct path exists between said sources to destination pair.

Step2: initialize the 3-dimensional array with 0.

Step3: input the number of source to destination pairs.

Generate values of source to destination either automatically or manually.

Step4: check if minimum path exists between each source to destination or not. If there is a minimum path between i^{th} sources to destination than set Checkflagfor_minim_path to 0 otherwise set it to 1.

Step5: fill the Pptr_to_dest_bit_array array according to the respective destination values in consultation with Checkflagfor_minim_path.

Step6: manipulate the numbering of destination values.

Step7: get the value of start by checking the source value as follows:

If source is 0,1 then start=1

If source is 8,9 then start=5

If source is 2,3 then start=2

If source is 10,11 then start=6

If source is 4,5 then start=3

If source is 12,13 then start=7

If source is 6,7 then start=4

If source is 14,15 then start=8

Step8: initialize row=start, time=1**,** pointer to Pptr_to_dest_bit_array =-1.

Step9: assign previous_row=row.

Step10: store the previous row in Ppath_matrix array.

Step11: go to the row equivalent to the previous row in graph matrix.

Step12: scan this row from backward and note that the two column numbers having value 1 in the variables column 1, column 2.

Step13: check the value at index location in Pptr_to_dest_bit_array.

If this bit is 1 select maximum (column 1, column 2)

If this bit is 0 select minimum (column 1, column 2)

Step14: assign row=selected column.

Step15:

- If Path from previous_row is busy at current time

Then

If

No loop forms

Then

Reroute request from conjugate switch of Previous row

Else Drop request

Else Goto conjugate of previous_row

Previous_row=conjugate (previous_row)

Increment time t by 1 Goto step 10.

Else Increment pointer of Pptr_to_dest_bit_array

Goto step 9 and repeat till end of Pptr_to_dest_bit_array

Step16: calculate path length and store in path length array.

Step17: display time wise route of request from source to destination.

Step18: display average path length and success rate.

VI. PERFORMANCE ALYSIS OF VARIOUS REGULAR AND IRREGULAR MINs

MINs	No. OF STAGES	No. OF SWITCHES PER STAGE	TOTAL No OF SWITCHES/ SWITCHING ELEMENTS	No OF MUX	No. OF DEMUX
Crossbar	N	N	N*N		
Omega	Log_2N	N/2	N/2. log_2N		
Generalized cube	Log_2N	N/2	N/2. log_2N		
Extra Stage Cube	Log_2N+1	N/2	N/2. (log_2N+1)		
M(n = 2log_2N)DOT	2n-1	-	$2(^{n+1)}{}_3$		
FT	2 $log_2N/2-1$	$2.^{n-i-1}$	-	N	N
CEFT	N	$2^{n-1},2^{n-2}$	N/2*n-1	N	N

Table I: Characteristics of various MINs

VII. PERFORMANCE EVALUATION OF CEFT ALGORITHM

The simulation of permutation passability behavior of a network will generate the number of requests appearing on source side at a particular instant of time and out of these total requests, how many of them are getting matured and the length of the path taken by the requests is calculated. These are the basic parameters for the simulation.

Simulation Environment:

- For the simulation, we are assuming that there are 16 nodes on each side i.e. there is 16 sources and 16 destinations for each network.
- The values are inputted using the algorithm of each network.
- This simulation analysis[5] of permutation passability behavior is checked for 50 times. For the simulation, different values are inputed for 50 times for each network and thus the result is calculated.

Results of Simulation: Figure 3 shows the average path length of various networks. Path length is calculated on the basis of total switches a request has to go through in the path to reach to its destination.

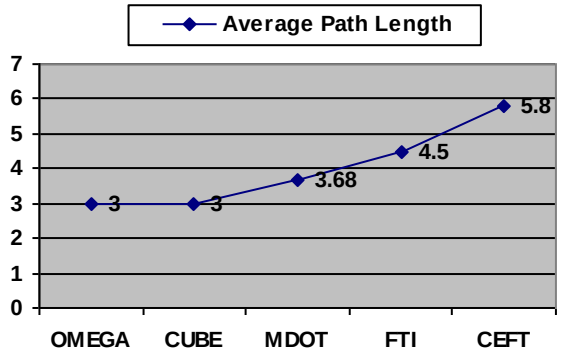

Figure 3: Comparison of Average Path Length of various Networks

The results of simulation show that Average path length of regular networks like OMEGA, CUBE[3][14] is fixed at 3 because here number of switches in each stage is equal while it varies for the irregular networks. In irregular network like MDOT, FTI network, because here numbers of switches vary for each stage so path length also varies. In our CEFT network, path length is more, as a request has to pass through the conjugate switch if the switches in the shortest path are faulty or busy. It results a long path and maturing maximum requests.

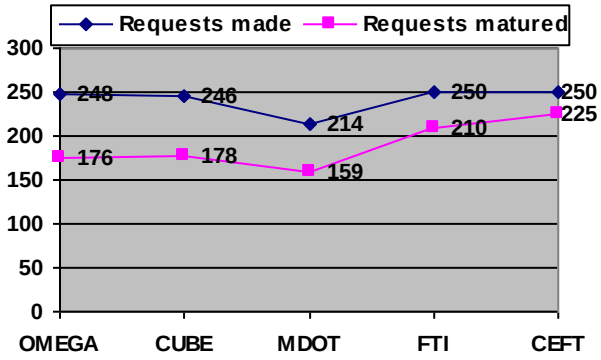

Figure 4: Comparison of Requests made and Requests matured for various Networks

The figure 4 tell about the requests sent and requests matured. The results of simulation shows that in our CEFT network total requests matured are increased as compared to the other networks[14]. In omega network, total requests made are 240 amongst them 176 requests are matured, in cube network out of total 246 requests 178 are matured , in FTI network out of total 250 requests 210 are matured and in our CEFT network out of total 250 requests 225 are matured at the time. It is also seen that if

destination of more than one requests appears at the same block then only one request gets matured. But in our proposed network, this problem is eliminated. In this way, performance is improved.

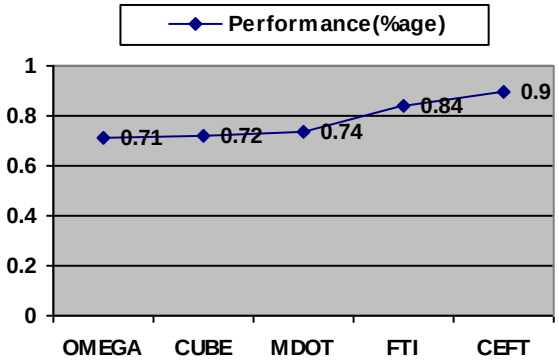

Figure 5: Comparison of performance of various networks

Figure 5 helps in showing the performance improved using our algorithm. Omega network is providing 0.71% performance while cube's performance is 0.72% and irregular network MDOT's performance is somewhat better than these regular ones. MDOT provides 0.74% performance and FTI[11][13][14] network provides 0.84% performance. In our CEFT network, performance is improved. CEFT provides the 0.90 % performance. It improves the performance from regular ones by approximately 20 % and from other irregular MDOT network by 16 %.

VIII CONCLUSION

In this study a new class of irregular fault-tolerant multistage interconnection network named Cost Effective and Fault Tolerant Multi stage Interconnection Network(CEFT) is proposed and analyzed. An irregular MIN is more cost effective and efficient than a regular MIN also they are as reliable and fault tolerant as other similar regular ones. It has been observed from the analysis that the permutation passable of CEFT network is better than existing regular and irregular networks like OMEGA, CUBE[3], MDOT and FTI networks. It also costs lesser in comparison to existing other fault tolerant irregular MINs.

We have calculated the number of requests successfully maturing and it has been found that if more number of requests are having the same destination or more number of requests are having the destination values that are lying in the same block then it resulted in more the number of clashes but CEFT achieves more fault tolerance as sources and destinations are connected to MUX, DEMUX directly so fault is tolerated at both the ends even when the request has to be routed on the alternate path.

References

[1] Yuanyuan Yang, Jianchao Wang and Yi Pan," Permutation Capability of Optical Multistage Interconnection Networks", *Journal of Parallel and Distributed Computing*, Volume 60, Issue 1, January 2000, page(s) 72-91.

[2] Mahgoub, Imad, Huang, Chien-Jen," A novel scheme to improve fault-tolerant capabilities of multistage interconnection networks", *Telecommunication Systems*, Volume 10, Numbers 1-2, October 1998, page(s): 45-66.

[3] Siegel, H.J. Nation, W.G. Kruskal, C.P. Napolitano, L.M., "Using the multistage cube network topology in parallel supercomputers", *IEEE Computer society*, Dec 1989, Volume77, Issue 12, page(s): 1932-1953.

[4] Chuan-Lin Wu, Tse-Yun Feng," On a Class of Multistage Interconnection Networks", *Transactions on Computers*, august2006, Volume: C-29, Issue: 8, page(s): 694- 702.

[5] Blake, J.T. Trivedi, K.S, "Multistage interconnection network reliability", *Transactions on Computers*, Nov 1989, Volume: 38, Issue: 11, page(s): 1600-1604.

[6] Chuan li wu, manjai lee," Performance Analysis of Multistage Interconnection Network Configurations and Operations", *IEEE Transactions on Computers*, 1992, Volume 41, Issue 1, page(s): 18 - 27.

[7] Dong Li; Mei Ming; Bo Fu,"New multistage interconnection network for multicast", The 9th Asia- Pacific Conference on Communications, 2003, *APCC 2003*, Volume 3, Issue 21, Sept. 2003, page(s): 993 – 997.

[8] López de Buen, Víctor," Multistage interconnection networks in multiprocessor systems. A simulation study", *Qüestiió*, 1987, volume11, Issue 3, page(s): 73-86.

[9] Siegel, H.J.," Interconnection Networks for Parallel and Distributed Processing: An Overview", *Transactions on Computers*, Apr 1981, Volume: C-30, Issue 4, page(s): 245- 246.

[10] Bhuyan Laxmi N., Qing Yang and P. Agrawal Dharma, 1989. Performance of multiprocessor interconnection networks. IEEE Trans. Comput., 22: 25-37. Doi: 10.1109/2.19830.

[11] Kruskal, C.P. and M. Snir, 1983. Performance of multistage interconnection networks for multiprocessors. IEEE Trans. Comput., C-32: 1091-1098. Doi: 10.1109/TC.1983.1676169.

[12] Bansal, P.K., R.C. Joshi and Kuldip Singh, 1994. On a fault tolerant multistage interconnection network. Comput. Elect. Eng., 20: 205-208. Doi: 10.1016/0045-7906(94)90047-7.

[13] Sandeep Sharma and P.K. Bansal, 2002. A new fault tolerant multistage interconnection network,TENCON '02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications,Control and Power Engineering, vol. 1,347-350, Doi: 10.1142/S0219265904001064.

[14] Yogesh Chaba, Ranjana Gulati. Development of Irregular Routing Algorithms for Parallel Computing Environment. International Journal of Computer Science and Security, Volume(1): Issue(3) 14.

[15] Park, J.H., 2006. Two-dimensional ring-Banyan network A high-performance fault-tolerant switching network. Elect. Lett., IEEE 2006, vol. 42,249-251: Doi: 10.1049/el:20062728.

[16] Bansal, P.K., K. Singh and R.C. Joshi, 1992. Quadb tree: A cost-effective fault-tolerant multistage interconnection network, Proceedings of the eleventh annual joint conference of the IEEE computer and communications societies on One world through communications, Vol. 2, 860- 866,http://portal.acm.org.

[17] Jacques Lenfant and Serge Tahé, "Permuting data with the Omega network", *Acta Informatica* ,Volume 21, Number 6,Nov 2004, page(s) 629-641.

[18] Nasser S. Fard, Indra Gunawan," Reliability Bounds for Large Multistage Interconnection Networks", *Lecture Notes in Computer Science*, 2002,Volume 2367,page 762.

[19] Aude, J.S.; Young, M.T.; Bronstein, G.," A high-performance switching element for a multistage interconnection network", *Integrated Circuit Design*, 1998.Volume 23, Issue 9, 1998, page(s): 154 –157.