International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075, Volume-8 Issue-7S2, May 2019

Design and Implementation of a New Routing Algorithm for Fault Tolerance in Network

P. Padmaja, G.V. Marutheswar

***Abstract*: The possibility to integrate more and more cores on the same chip puts severe constraints on the reliability, to which it is important to provide correct services in the presence of faults. Many fault tolerant routing algorithms are used to overcome the faults in Network on chip. However, these routing algorithms, suffer from another problem like the congestion. In this work, a novel approach inspired by Catnap is proposed for NoCs using Local and Global congestion detection mechanisms with hierarchical sub-networks architecture. With the help of these two techniques, the NoC becomes fault tolerant and is able to efficiently utilize the throughput.After simulation results shows that the proposed algorithm gives a better performance by reducing the latency and increase the reliability of the network. In addition, the algorithm has another advantage : it reduces the congestion which is considered as a temporary fault. Simulations show that our proposed algorithm reduces the latency more than 15% and throughput is improved by 20% compared to the PDA- FTR routing**

***Index Terms*:. :** *Network on Chip, Fault Tolerance, Congestion, Reliability, Sub-Network, Routing Algorithm*

I. INTRODUCTION

Network on Chip (NoC)is a technic designed for perfect communication in a system on chip (SoC). This intensification of communications leads to important questions such as performance and energy consumption. Decreasing the transistor size has made semiconductors more sensitive to faults. Thus the challenge is, to maintain the system functionality during its operational lifetime and ensure that the system performance is preserved. For this reasons, researchers have attached a great deal of importance to the reliability in networks on chip.

The fault tolerance routing algorithm is the process of finding a new path to steer packets from sources to destinations in a faulty network, by choosing an optimal path the routing algorithm can efficiently increase the performance of the network. Congestion is another key factor which leads to increase the transmission delay and power consumption. For this, routing algorithms can improve performance by re-routing packets through less congested regions and distributing traffic over the network. Finally, failures and congestions should be managed in an effective way to ensure availability and robustness into the network on the chip.

Revised Manuscript Received on May 28, 2019.

Dr.P.Padmaja, ECE, Vignan Institute of Technology and Science, Hyderabad, INDIA.
Dr.G.V.Marutheswar[2] [2]Professor, Dept of EEE, S.V.U.College of Engineering, Tirupati, Andhra Pradesh,India

II. BACKGROUND

The main challenge to increase the reliability and provide a good performance is to deal with the principal problems in routing algorithms such as deadlock, congestion and failures. In the literature, there are many various approaches which have been suggested to solve this issue. Figure 1 summarizes some significant works about fault-tolerant techniques in NoC.

The Gradient (Pratomo & Pillement.2012) suggested an adaptive routing algorithm to tolerate faulty node. The proposed routing algorithm divides the entire system in eight zones by gradient line. The main disadvantage of these approaches is that they do not foresee an efficient mechanism to control the congestion problem, in order to make a good decision routing for complex traffic condition. To avoid congestion and failure at the same time, the solution proposed by (Ebrahimi et al.2012) and, (Ebrahimi et al.2013) adopts a minimal path to reduce the congestion caused by the presence of faults.

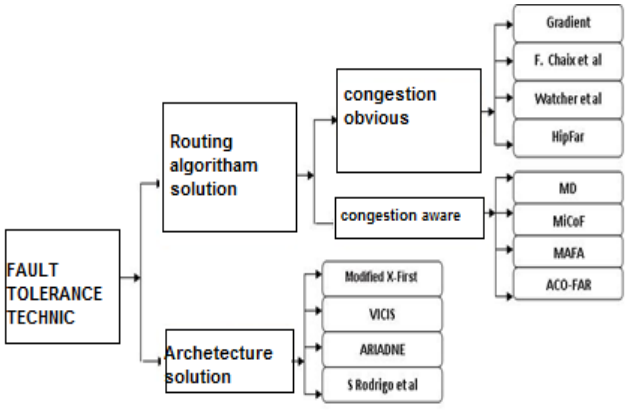

Fig1: Some related works on fault-tolerant techniques in NoC

These algorithms are able to select the shortest path to route packets as long as a path exists. (Chang and al.2013) proposed another routing algorithm which is inspired from ACO called ACO-FAR to perform a load balancing with low latency and a high throughput. The three

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

algorithms mentioned above use minimal path, therefore they are all live lock free.

Table1: Most known Fault-Tolerant Routing Algorithms

FT Routing Algorithm	Years	Components	Type	Congestion
Zhang et al.2008	2008	Up One Switch Failure	Permanent	No
Chaix et al.2010	2010	Many Faults Switch	Permanent	No
Deorio et al.2011	2011	Links And Switches	Permanent	No
Deorio et al.2012	2012	Links And Switches Failures	Permanent	No
Xiangming.2012	2012	Links And Switches Failures	Transient Faults	No
Vitskovkiyet al.2013	2013	Links	Permanent	No
Ebrahimi et al.2013	2013	Up To 6 Faulty Links	Permanent	Yes
Pratomo & Pillement.2012	2013	Switch Failures	Permanent	No
Ebrahimi et al.2013	2013	Up To 6 Faulty Nodes	Permanent	Yes
Ebrahimi et al.2013	2013	Up To 2 Faults Link	Permanent	Yes
Chen And al.2013	2013	Any Number Of Faults Links	Permanent	No
Wachter et al.2013	2013	Switch	Permanent	No
Chang And al.2013	2014	Up To 4 Faulty Switches	Permanent	Yes
Fu et al.2014	2014	Many Faulty Switches	Permanent	No
Dimopouloset al.2014	2014	Links And Switches	Permanent And Transient	Yes
Behrouz et al.2014	2014	All Faulty Links/Nodes	Permanent	No
Charif et al.2015	2016	Links And Switches	Permanent	Yes
Chen et al.2017	2017	Switches	Permanent	Yes

Table.1 recaps the most popular fault tolerance routing algorithms. The all approaches mentioned above have the following benefits and weaknesses. (a) under high traffic conditions and (b) the performance in term of throughput and latency degrade in case of important rate of fault in links and routers. Finally, all of them deal with permanent faults. Hence the aim is: to achieve a robust routing algorithm with these required functions as

(a) the ability to avoid the congested nodes for having a balanced traffic,

(b) tolerate high rate of faults which may affect links and routers and,

(c) tolerate transient error which appears in communication process.

All the techniques described above bring their approaches to the fault-tolerance problem, but they have a cost in terms of performance to be considered in the area of routing algorithms design. This is evaluated by a set of metrics, such as: latency, throughput, network congestion and energy consumption. To satisfy all these requirements at the same time is impossible. Thus, the challenge for the researchers is to find a good compromise between these costs and the reliability.

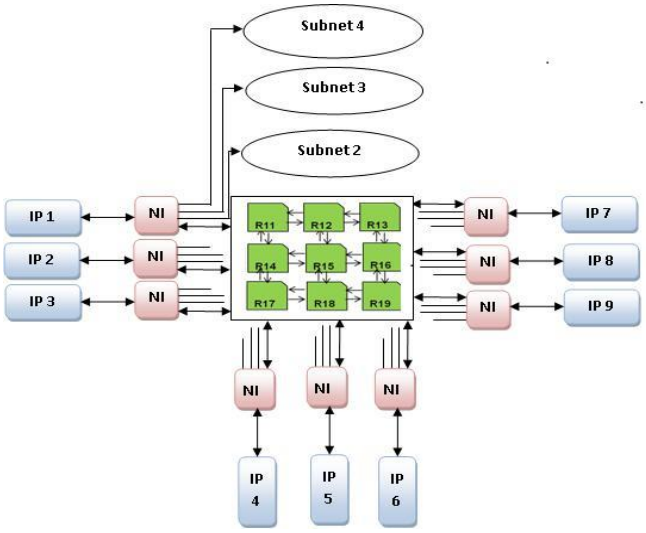

Fig2: The global architecture NI: Network Interface, R: Router, IP: Intellectual Property

III. NETWORK ARCHITECTURE

Networks-on-chip have been a very active research field since their appearance in early 2000s. Since then, many architectures have been proposed in the literature. Routers, Network Interface (NI), IPs and links are the main elements in the NoC. In this section, we introduce the proposed architecture and its main components. Before giving the details, it is necessary to specify that we have adopted two important assumptions that Intellectual Property (IP) are connected to the input and output Network Interface (NI) by the links are always non- faulty. Another assumption is that at least on path between souce and destination.

Figure 2 shows the global architecture. The network on chip is divided into subnets. Every IP is connected to four routers, each one belonging to a disjoint subnet. Thus, every switch belongs only to a single subnet. For example, we build 4 subnets with 6*6 switches, each one can communicate only with the others of the same subnet. This method represents an innovative alternative to get an efficient reliability. When a subnet is congested or faulty, we should disable the entire subnet and isolate it. When focusing on subnets, we can see that they have the same connectivity pattern.

The three IPs located in the east of the NoC are connected to the three switches of the same subnet. The same topology is applied for the west and the south IPs. We have implemented an algorithm (Section 4) to select the subnet. We have added 2 bits to the header flit in NI for distinguishing each subnet

from the others. The key advantage of this architecture is that the connectivity pattern has several alternative paths who can be used for increasing the fault tolerance, as it will be shown in section 4. When the first subnet is inoperable, we activate the second subnet, and when the second subnet is defective, we active the third subnet, and so on. The worst case scenario is when the top-subnet is faulty or congested. To provide a higher fault-tolerance, the new architecture requires an additional links. We solve this by a short connection between end nodes and IPs.The router architectured in such a way that an entity that facilitates communication between IP cores in Networks on chip. We introduce in this part the proposed router architecture and its main components.

IV. PROPOSED ROUTING ALGORITHM

This section focuses on the routing algorithm proposed in this paper. To deal with all the requirements of the fault tolerance, in the initial stage of the process, the proposed solution can provide an online detection of permanent and transient faults. Once the detection has been done, we can do the isolation of faulty components.

A. Congestion detection

Wormhole routing requires less memory than the virtual cut through or store and forward routing strategies. This may cause a congestion state because the buffer requirements may vary based on the application. The increasing number of cores will contribute to an additional traffic which will increase the congestion as well. So, there is a need to decrease the congestion and enhance the performance at the same time for greater size NoCs.These are of two types.They are local Congestion detection and Regional Congestion detection

B. Fault Detection

For on-chip network to handle the faults in routers, links and network interfaces, BIST (Fochi et al.2015) is the traditional solution. This method is a fault detection scheme. But this technique requires external circuits to perform error detection, and time to test which can reduce the performance of the NoC in term of latency. To develop our approach, we have opted for the Cyclic Redundancy Cycle (CRC), a solution that can be used to perform error detection by comparing an input and output of each router to detect wrong packets and faulty components in a NoC. This approach is inspired by (Fochi et al.2015), with some modification. Each router is equipped with a Test Module (TM). The goal of this mechanism is to enable online fault detection. The CRC polynomial used is g = 1 + x + x4, so we add 4 bits to the header of each packet. So, when an error is detected anywhere in the router or link, we know the error exists but we can't correct it with CRC decoding.

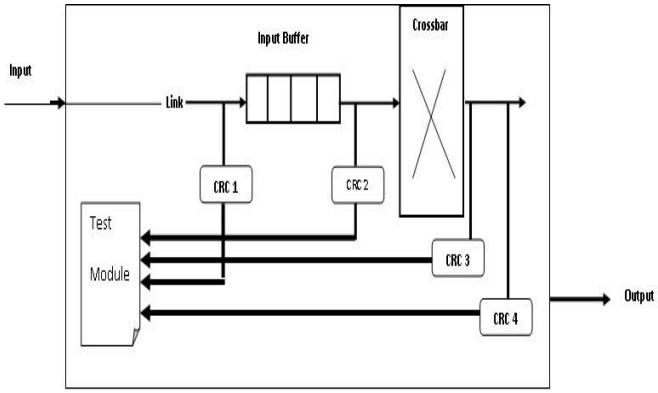

Fig3: Internal architecture of the router

To test and diagnose the communication infrastructure in the NoC. First of all, it is necessary to detect the fault, secondly how to identify the faulty components: routers, links or cores. Finally, we analyse how to use efficiently the remaining components that are fault free. In the router, the fault can occur in all the components of the router (crossbar, buffers and others). At this high core density in networks on chip, considering faults only in the routers or links do not provide the optimal safety. Other components such as input-buffers and crossbar should be given prominent attention to ensure the fault tolerance and improve the system reliability.

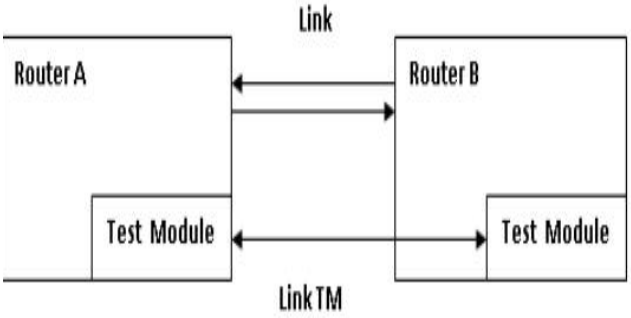

Fig4: Communication between two neighboring routers

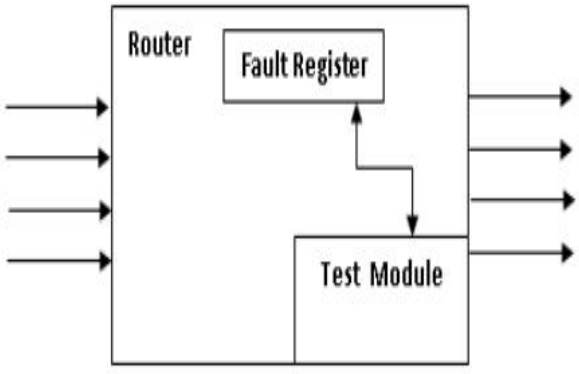

Fig5: Communication between the Fault Register and the Test Module

To achieve this goal, we

present in this section the Test Module (TM) added in our router (see figure 5). In this subsection to describe the main functionalities of the Test -Module (TM) and its important role in orchestrating the different process inside the router. Test Module detects and locates the faulty routers or links. Therefore, in the same subnet, only adjacent routers can communicate this information.

In order to keep the fault occurrence in the input-buffers and the crossbar and header flits, this information is always sent to the FR (Fault Register). The aim is to be used during the selection of the next port. Based on a Cyclic-Redundancy-Check (CRC) in each output component of the router that reads the incoming flit to detect any error. Depending on these states, our algorithm (DINRA) next nodes or links fault status received from the TM and checks the number of possible safety directions. For example, when a fault is detected in the buffer, a signal is sent to inform the TM module about the fault presence. The same case is applied for the other components.

Finally, to keep the faults information in routers or links, the TM interacts with the FR unit to exchange fault information and control signals, as shown in Fig. 5.

Table2: Codification of different states of links and routers

State	Value	Description
1	00	Port East unsafe
1	01	Port North unsafe
1	10	Port South unsafe
1	11	Fail Router
0	00	Port East safe
0	01	Port North safe
0	10	Port South safe
0	11	Safe Router

C. *Routing Algorithm*

The packet routing is one of the significant factors in the design of NoC architecture. To achieve the fault tolerance, the routing algorithm needs multiple paths to route the packets for each pair of source destination. First, the proposed routing algorithm must be adaptive to choose between them. Second, the topology must provide an alternative route. Complex routing algorithms can introduce extra area and an energy overhead. In our proposed solution, the main goal is to maximize the performance in terms of latency and reliability by having a good successful packet delivery rate.

Algorithm 1: path computation

1 State sub s0 = Active & RCS =0;
2 For (s=0, s = 3, s++) /*– Source S (x, y, z), Destination D (x', y', z')
3 If state_subnet_s = 00 /*– — normal state
4 If RCS = 0 then
5 For each S, D
6 North Last /*————-Routing Algorithm by default
7 if (link state && router) == unsafe
8 South last /——-Routing Algorithm
9 else / —— Activate s+1
10State_sub = Faulty; s=s+1;
11Else s+1 / ———— Next Subnet
12state sub = Congested; s=s+1;
13state_sub_s= Active /———active next subnet
14End loop
15End loop

Description of algoritham is NI connected to a router first inject a packets into the subnet-0 (line 2), and all the flits of a packet travel in the same subnet until they reach the destination. If the subnet-0 is congested or faulty, a higher subnet (s+1) can be activated more quickly in time to avoid a performance loss (line 11) and same case for others subnets. To route packets in same subnet we use the Last North routing algorithm, and if can't we use the North Last to avoid faulty components (Links or Routers-line 7). And, if can't route the packet in second case we active the higher subnet (s+1) and we set state-subnet equal faulty. The same state if the current subnet it going congested-line 12.Fault-tolerance in DINRA has all information about the fault status of next links and nodes. It has three possible directions for routing the flits. DINRA -FTNoC adopts Wormhole-switching.

D. *Deadlock*

The deadlock problem may arise with the adaptive routing. Most of them use Virtual-channels (VCs) or Turn Model to the routing selection to avoid the deadloc.

E. *Adaptability*

One of the key properties of the DINRA architecture is its capability to be adapted for providing several routes to transfer packets for each source destination node. This increases significantly the reliability of the entire NoC. DINRA proved its yield by ensuring both fault tolerance and congestion aware.

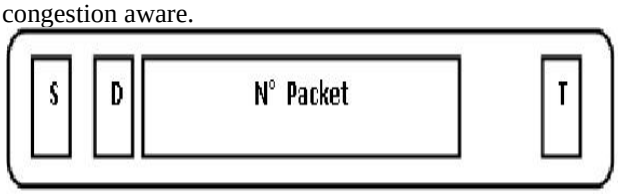

Fig.6: The notification message formatN ,S: Source, D: Destination, T: Type of Packets

F. *Retransmission of pockets*

When packets are stucked in buffer,stand the output link or the router becomes suddenly defective during the packets transit. This changes the validity of the path (dynamic fault), which later will be proved as faulty. This corruption creates sub-packets which cannot arrive at their destination. A copy of the Header of this packet will be stored in a special buffer.

.

There are two cases where a packet is considered lost and retransmitted another time by the source

- A packet fragment arrives in a faulty subnet: for example, a router without healthy output or a router or the packet is blocked by routing restrictions to its reachable destination.
- The second case, destination not attainable when a packet fragment cannot reach the destination node before a" Timeout" T (an empirical value that varies the size of NoC).

In this section to assess and analyze the performance of the proposed system and discuss the performance variation. selected two traffic patterns.One is random and be shuffled evaluate the performance of the proposed system. The first sets of analysis investigated are Latency and the Reliability of the proposed routing under each of the aforementioned patterns traffic. To evaluate our algorithm, used a simulator: Noxim.

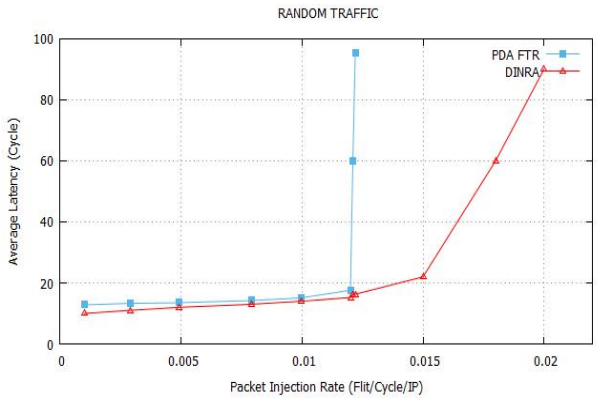

Fig7: The average latency of DINRA and PDA-FTR in 8*8 with Random traffic, in four faulty routers

V. EVALUATION

In evalution the following the following components are updated

A. *Latency Evaluation*

In this party to evaluate our approach, we vary the rate of injection of faults and measure the average latency. First,evaluated the communication latency of DINRA by calculating the average latency/ flit.

B. *Reliability Evaluation*

By performing the same series of experiments but this time taking the model of router faults (Fig 13). Again, we saw an increase in average latency based on the number of faults and the packet injection rate. It should be noted that the average latency increases here faster than in the previous case or only the links were considered defective. A router fault is a router with all its faulty links. For the DINRA algorithm, the faulty routers there are, the harder it is for a packet to find a path to its destination. Nevertheless, with more than 30% of the routers failing, the service is still rendered but with a lower packet injection rate depending on the type of traffic, compared to a fault-free situation.

$$Successratio = \frac{Total.arrivedpackets}{Totalinjectedpackets} \times 100$$

Table3: Comparison of Success ratio % with PDA-FAR Routing Algorithm

No of faults	Gradient	PDA-FTR	DINRA
2	2.8%	0.07%	00
4	3.2%	0.5%	00

DINRA makes it possible to guarantee a high level of reliability. In the case of a 12x12 NoC size and can deliver more than 95% of messages when 10% of links are defective. When we have 40% of the links defective, the delivery rate successfully varies from 50 and 60% for" Random Traffic". When considering node faults.

When considering node faults (router), the rate of packets delivered successfully decreases, as one might expect given the greater impact of a router fault. Nevertheless, even in the scenario where 40% of routers are defective, the delivery rate varies from 0 and 10% reaches 10% for "Uniform Random" traffic (Figure 13). With the addition of packet forwarding and congestion-aware techniques, DINRA is clearly improving performance and reliability over PDA FTRs. fault injection affects average network Latency and Throughput, which increases with the number of link faults, routers and the number of packets injected into the network.

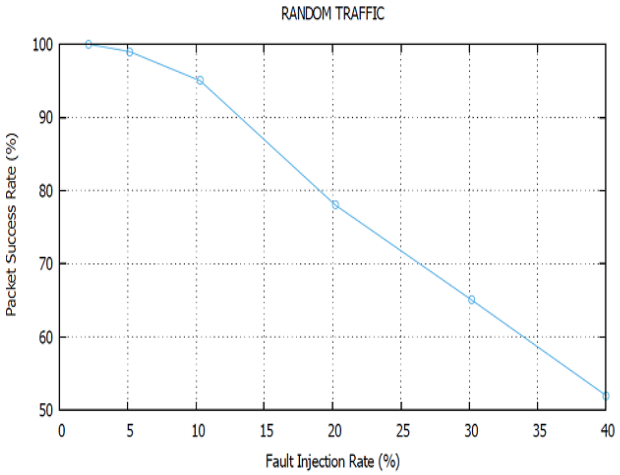

Fig8: Rate of successfully delivered packets with different fault injection ratios (Link fault) under "Random" traffic for a 12x12 NoC.

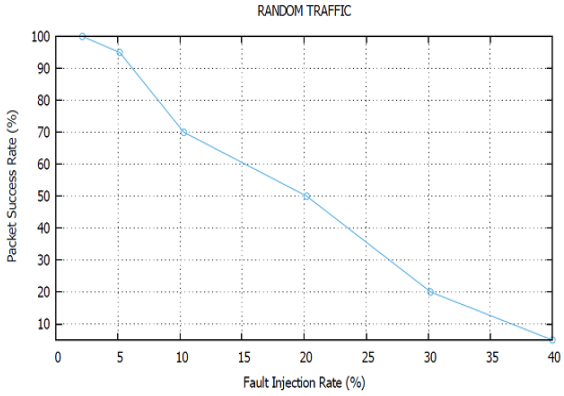

Fig9: Rate of successfully delivered packets with different fault injection ratios (router fault) under "Random" traffic for a 12x12 NoC.

5.3 Throughput

We have evaluated the throughput, which is defined as the accepted traffic of the network at a given latency. The throughput of network is presented by the formula below:

$$NetworkThroughput = SaturationThroughput \times No:ofnodes(2)$$

We then wanted to determine the impact of the Throughput and the size of the network. For this we measured the Throughput for four NoC sizes and for one single failure situation. (Fig.14) summarizes the simulations performed for various sizes. We can see as expected that our experiments confirm that DINRA is suitable for both small and large NoC systems.

VI. CONCLUSION

In this paper, presented an innovative fault- tolerant routing algorithm, called DINRA, for the Networks-on-Chip (NoC). The algorithm operates with a proposed architecture which aims to avoid a heavy sys-tem failure when some faulty components (routers and links) are observed in a NoC. The proposed solution preserves the network performance. The evaluation done proves that DINRA performs better than PDA-FTR algorithm, even at 40% fault-rate in links and routers. Despite the good results obtained, our work clearly has some limitations which should be fixed to enhance its performance and reliability. The first one is that need to analyze the best techniques which can be implemented for fault-detection. Second, it is also interesting to study if the time required to handle a fault-detection can be improved. We are currently studying the power consumption and hardware complexity of the proposed architecture. In a future work, we plan to focus on these parameters.

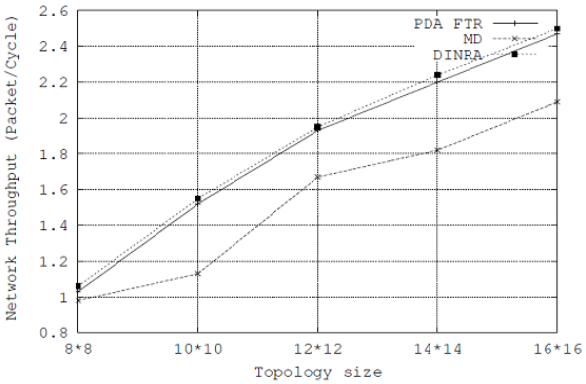

Fig10: Throughput of DINRA under different topology sizes and in single fault

REFERENCES

1. Behrouz,R.J. Modarressi, M. Sarbazi-Azad,H.(2014). Fault-tolerant routing algorithms in networks on-chip, in: Routing Algorithms in Networks on Chip, Springer, 193-210
2. W. Charif,A. Zergainoh,N.E. Nicolaidis.M. (2015). MUGEN: A high-performance fault-tolerant routing algorithm for unreliable Networks-on-Chip. IOLTS, 71-76.
3. Chaix, F. Avresky, N.E.Zergainoh, D. and Nicolaidis,M. (2010). Fault-tolerant deadlock-free adaptive routing for any set of link and node failures in multi-cores systems. In Proceedings 9th IEEE International Symposium on Network Computing and Applications, NCA, 52-59.
4. Chang,E-J .Lin,C-A. Hsin,H-K.,., Wu A.-Y.(2013) ACO-based fault-aware routing algorithm for network-on-chip systems, in: Signal Processing Systems (SiPS), Workshop on, IEEE, 342–347.
5. Chen,Y.Y. Chang,E.J. Hsin,H.K.(2017). Path-Diversity-Aware Fault-Tolerant Routing Algorithm for Network-on-Chip Systems IEEE Transactions on parallel and distributed systems, Vol. 28, NO. 3, 838–849.
6. Chen,C. Cotofana,S.D.(2013).An effective routing algorithm to avoid unnecessary link abandon in 2D mesh NoCs, IEEE, 311-318.
7. Das,R. Narayanasamy,S. Satpathy,S.K. and Dreslinski.R. (2013).Catnap: Energy Proportional Multiple Network-on-Chip, In proceedings of the 40th International Symposium on Computer Architecture, Tel Aviv, Israel ISCA.
8. DeOrio,A. Peh,L.-S and Bertacco,V.(2011). ARIADNE: Agnostic reconfiguration in a disconnected network environment in International Conference on Parallel Architectures and Compilation Techniques (PACT), 298–309.
9. DeOrio,A. Fick,D. Bertacco,V. Sylvester,D. Blaauw,D. Hu,J. and Chen,G..(2012) A reliable routing architecture and algorithm for NoCs, IEEE Trans. Comput. Aided Design Integrated Circuits Syst. 31 (5), 726-739.
10. Dimopoulos,M. Gang,Y. Anghel,L. Benabdenbi,M. Zergainoh,N. . Nicolaidis.M (2014).Fault-tolerant adaptive routing under an unconstrained set of node and link failures for manycore systems-on-chip, Microprocessors Microsystems, v.38 n.6, .620-635.
11. Ebrahimi,M. Daneshtalab,M. Plosila,J. and Tenhunen,H.(2012). MAFA: Adaptive Fault-Tolerant Routing Algorithm for Networks-on-Chip, In Proc. of 15th Euromicro Conference on Digital System Design, Sep, 201–207.
12. Ebrahimi,M. Daneshtalab,M. Plosila,J. and Mehdipour,F.(2013). MD: Minimal path-based Fault-Tolerant Routing in On-Chip Networks, In Proc. of the 18th Asia and South Pacific Design Automation Conference, Jan, 35–40.

AUTHORS PROFILE

Dr.P.Padmaja currently working as Assoc.Professor in electronics and communication engineering, Vignan Institute of technoly and science. She completed her Ph.d in March 2019 from SVU. She completed M.TECH in Digital Electronics And Communication Systems, 2008 . Her area of research is Wireless Sensor Networks. she published her papers 14 journals and 8 conferences. She guided 5 M.tech students. She did Advanced Diploma in Systems Maintanance Engineering from CEDTI,Chandigarh. she recived 'BEST FACULTY' award from VITS in the year 2011. She attended 9 workships.she has Membership in Professional Bodies of IEEE,AMIETE, IAENG.She received Diploma in electronics and Instrumentation from Govt.Polytechnic ,srikakulam.She completed school education in Jawahar navodaya vidyalaya,pedavegi,west Godavari.She participated in various state level sports and games .Conducted work shop on Ns-2 for students during summer vacation.

DR.G.VENKATA MARUTHESWAR currenly working as Professor, HOD in EEE Department, SRI VENKATESWARA UNIVERSITY, TIRUPATI. He completed Ph.d in June 2009,with specialization of Design and Evaluation of Performance of fuzzy and Integrated Fuzzy controllers for speed control of switched reluctance motor Drive. He did M.Tech with the specialization of Instrumentation &Control systems in 1990. B.Tech in EEE from SVUCE,in 1985.He published papers in 40 journals and 23 conference .He guided more than 40 students in M.tech and 10 research scholars..He conducted many workshops to enrich students knowledge in motr drives.