2021 Sixth International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET) | 978-1-6654-4086-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/WiSPNET51692.2021.9419420

A Comparative Analysis of Fault Tolerance Methods in 3D-NoC

Keerthi Somraj[1], K. Kishore Kumar[2] and B. Naresh Kumar Reddy*[3]

[1,2]Department of ECE, IcfaiTech, ICFAI Foundation for Higher Education, Hyderabad, India.
[3]Department of Electrical Engineering, IIT Delhi, India.

[1]keerthisomraj@gmail.com, [2]kkishore@ifheindia.org, [3]naresh.nitg@gmail.com

***Abstract*—Networks-on-Chips (NoCs) have found wide acceptance in many-core systems for on-chip communication due to their ease in addressing communication problems in integrated systems. Due to the incompetence of the 2D NoC structures, academia and industry have moved towards the 3D-NoC. The 3D -NoC can be manufactured with the help of Through Silicon Vias (TSV) as an interconnecting mechanism between the horizontal layers. These TSV links exhibit high vulnerability to the faults in the system. Significant research has been done to reduce this vulnerability by employing a few algorithms and schemes. This paper provides a survey of the different methods applied and their further advancements.**

***Index Terms*—NoC, 3D-NoC, Through Silicon Via (TSV).**

I. Introduction

Interconnect scaling can be attributed as one of the major factors which limit the performance scaling. The key metrics which contribute towards this are the interconnect delay and the energy dissipation [1]. Two technologies have been developed to improve the performance and scalability of the chip. One of them is the use of 3D ICs which alleviate the interconnect latency bottlenecks. It also helps to improve the I/O bandwidth with the help of the third axis to provide more connectivity. Among the viable options for the construction of 3D IC's, we have Through Silicon Vias (TSVs) which stand as a pillar between the different layers. For inter-die connectivity, TSV's are manufactured across thinned silicon substrates after die bonding [2]. On the other end, we have Network-on-chip (NoC) based architectures which have obtained acclamation due to the limitations in 2D-ICs.One of these is the limited floor planning choices of 2D ICs which restricts the increased productivity of the NoC architectures [3]. These 3D ICs have significant potential for improving the performance of a system without the need for scaling even if the packet density increases, power due to shorter wires is reduced and has higher immunity to noise as discussed in [4], [5], [6], [7].

All these features of the 3D-ICs combined with NoCs lead to the construction of 3D-NoCs. The most widely used pattern in the 2D architecture i.e. 2D mesh is extended into the 3D architecture with the help of TSV as vertical interconnects connecting the different layers of 2D mesh which leads to the formation of 3D NoCs [8]. But these TSVs have relatively low yield. To enhance the performance of these chips partially connected TSV's are also being used to reduce the complexity of the network. A huge amount of research has been done to improve this yield by increasing the fault tolerance of these chips and to obtain a deadlock-free network.

The faults present in any network can be divided as either permanent faults or transient faults. The flow-control-based methods use the error control code in conjunction with the re-transmission mechanism for transient fault tolerance. Fault tolerance is achieved by utilizing inherent redundancy present in the structure to route packets around the permanent faulty links or routers. A fault-tolerant algorithm is said to be good if it always ensures that there is no packet loss in any faulty pattern as long as a path is present to transmit the packet [9]. Mesh topology is considered to be the most basic topology in most of these schemes and applications due to its simplicity, and reliability of the system based on this topology.

One of the major concerns regarding the usage of TSVs is its large area overhead and lower yield but even with all these failures TSVs are promising enough. Hence considerations are made to enhance this fault tolerance rather than discard these chips. This paper discusses the various schemes and algorithms which try to improve the fault tolerance of these TSVs with a few compromises on the area overhead, increase in run-time, and power consumption of the device. Adaptive routing is applied in most of these algorithms rather than the deterministic model for fault tolerance routing due to

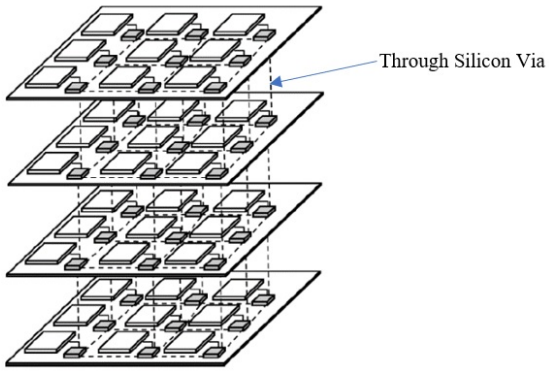

Fig. 1. 3D NoC Mesh Topology

its implicit path redundancy and as it is capable of being deadlock-free [10].

This paper provides a review of the few effective fault tolerant algorithms used to eliminate these faults.

II. LITERATURE SURVEY

V. Pasca et. al. [11] proposed a method where vertical links are used for inter-die communications with the help of serial communication and signal retransmission on non-faulty wires. These links have a degradation factor i.e., the serialization rate above which the link stops functioning. Once a fault is detected the system transmits the data in a serial manner and also remaps the remaining fault-free links for the process of data transfer. But if the serialization rate is too high then the performance of the system degrades which affects the performance and also increases the area overhead.

C. Feng et. al. [12] proposed an algorithm that utilizes a routing table proposed for 2D mesh along with 2 TSVs state vectors for the routing of the entire system rather than using global routing in the event of abstaining from the faults occurring in the horizontal links and TSVs. Since it requires a routing table of size n*4 at each router from all the 4 output ports this algorithm suffered from poor scalability.

In order to improve the yield further, C. Hernandez et. al. [13] proposed a link design which improves the yield of the above said TSVs by using an additional Ω network on either side of the faulty boundary of TSV networks where the link design depends on the condition that the probability of having faulty TSVs more than N/2 is negligible. The function of this Ω network is that at the input of the vertical link it is used to deconcentrate half of the link into N/2 fault-free TSV's and the $\Omega - 1$ network is used at the output to transmit back the data corresponding to the half-link. This design suffers from an increase in the area overhead of the chip due to the presence of the additional omega networks and due to the presence of these reconfigurable Ω networks, a large number of TSVs of a link remain unused.

K. Hsieh et. al. [14] proposed a fault tolerant mesh structure for 3-D NoC which improves both the yield and reliability of the chip with backup routing paths which are used to bypass the interconnect defect by finding a replacement path for the faulty path with the help of k- shortest path algorithm. This mesh is constructed with the help of various phases where in the first phase a 2-D NoC is optimized to 3-D NoC with constraints on area, routing length, temperature, and performance, and then the I/O optimizations. In the next phase, the routing scheme is applied with the help of few routers and then the search for the faulty paths and their replacements are done in the next phase and the final phase the fault tolerant NoC is verified.

A novel routing scheme was proposed by S. Pasricha et. al. [15] named 4NP-First for 3D NoCs using hybrid turn models where the 4N- first turn model is used first and then the 4P- first model is used, and by combining the above two models we get the 4NP First routing algorithm which achieves a higher packet rate compared to 4N and 4P first algorithms when they are used individually. A higher priority is given to the minimal paths that have higher probabilities even if faults are encountered in the path. In order to create fault tolerant NoCs, the above-said approach can be combined with error-correcting codes (ECC) techniques for transient fault resiliency and optimizations. In spite of achieving higher packet rates, with the presence of an increased number of fault rates the energy consumption of the 4NP-First becomes larger than that of the other hybrid models as it uses a higher number of non-minimal paths as the faults increase to see that the packets arrive at the destination properly since it takes too many hops to reach the destination.

A.B. Ahmed et. al. [16] proposed a new deadlock recovery technique named the Random-Access-Buffer (RAB) which can detect the occurrence of a deadlock as soon as it occurs and then blocks the further packets thereby trying to disrupt the dependency cycle between the neighboring nodes. The RAB is very much similar to the Virtual Channel (VC) but it is much simpler and has less complexity compared to the VC. In order to avoid the presence of this VC, only S. Pasricha et.al has proposed the 4NP-First algorithm. This scheme is useful when we are trying to eradicate the deadlock that occurred at high fault rates at the cost of power overhead.

S. Akbari et. al. proposed [17] a routing method that uses the ZXY when there are no faults and the XZXY routing techniques when faults are present. It uses a deterministic model rather than an adaptive model as it is simpler and doesnt need to split the virtual channel (VC). An interesting property of AFRA is that whenever all the faults are present in one direction it is deadlock-free. Hence it doesnt require the presence of any VC for the deadlock avoidance which implies that these VCs can be used by the incoming channel packets in the virtual channel flow control. The main disadvantage of using AFRA is that it can tolerate only single faults on the TSV link and each router should have the information about vertical links of all the routers in the same row.

X. Jiang and T. Watanabe [18] proposed an algorithm that can overcome the problems faced in traditional routing and improve the performance and reliability of the chip. The routing algorithm is based on assumptions that the links connecting the router to the PE are always faulty and that even if one of the directions of a bidirectional link is faulty then the entire link is considered to be faulty. The two main important considerations are that the algorithm uses a vertical node i.e., the node which has the TSV and assigns it for the inter layer routing, and secondly it applies the 2D routing algorithm for the intralayer routing. The packets are first routed towards the destination layer using the interlayer routing and the transmission within the 2D layers is done by using the intralayer routing. The energy consumption of this algorithm is very large as it uses routing computation and fault detection systems which are highly complex.

M. Ebrahimi et. al. [19] proposed a method that uses an adaptive routing algorithm called 3D-FAR where the packet can take any shortest distance from the source to destination utilizing any of the 2,2,4 virtual channels along with the X, Y, Z directions. Another algorithm called 3D-FT was also proposed for tolerating faulty routers and faulty links. Only the shortest paths are taken into consideration for the faulty

routers, whereas the non-minimal paths are considered for faulty links when the source and destination are in the same dimension. The presence of virtual channels in the X, Y, Z channels leads to a large overhead in the area of the chip.

M. Ebrahimi et. al., [20] proposed a routing algorithm that tolerates the faulty links and doesn't require any virtual channels. It uses the Hamiltonian path strategy when it has to deliver a packet to several destinations. This algorithm doesnt need any routing table thereby it reduces the area overhead and also doesnt need any extra information to be present in the header flit. This strategy eliminates deadlock conditions and remains deadlock-free but it can't tolerate too many faulty links in both the horizontal and vertical directions.

R. Alizadeh et. al. [21] proposed a circular routing algorithm for the faults present in both the TSVs as well as in the horizontal links. The algorithm is used for both 2D as well as for 3D mesh topologies where the packets are routed through a circle to reach the destination. Here the number of nodes in any network can be used to determine the number of circles in that particular network. It works based on the knowledge that the algorithm can tolerate the faults both n the horizontal and vertical directions but cannot tolerate too many faults.

J. Zhou et. al. [22] proposed an algorithm named HARS whose main idea to accomplish data transmission is to find the mid node of each layer to ensure that the packets bypass the invalid vertical links. This is done with the help of DyADM routing by extending it from 2D to 3D in the presence of faults on the vertical links. To tolerate both horizontal and vertical link faults 2D routing along with routing scheme is used. HARS is free of deadlock conditions and can detect both single and multiple faults without the need for virtual channels (VC). Reliable results are provided only when 10% of the links are faulty.

Another approach that has gained popularity with the TSVs is the partial routing of these TSVs. This is a common routing approach used for partially connected 3D-NoC to connect them vertically. Packets are first routed in a horizontal plane towards the vertical links and then again routed vertically to reach their destinations. These vertical links are called elevators.

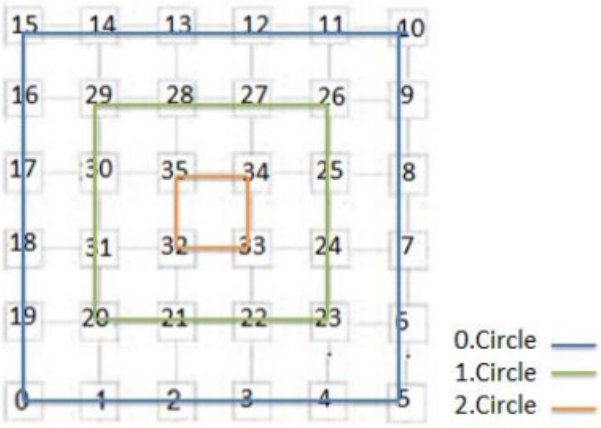

Fig. 2. 6*6 Mesh with physical network assignment

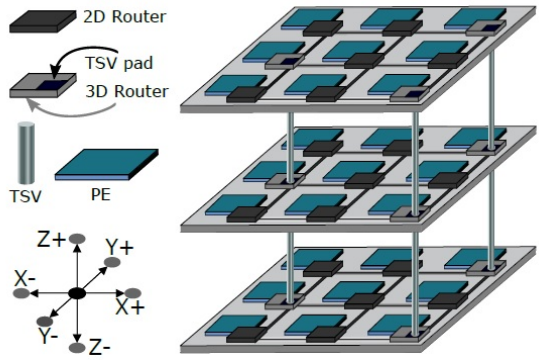

Fig. 3. Partially connected 3D-NoC

F. Dubois et. al. [23] proposed an algorithm named Elevator-First for vertically partially connected 3D-NoCs to avoid the occurrence of deadlock where two VCs per physical link only in the X, Y directions are used. The Z-direction is divided into Z+ used for ascending packets and Z- networks used for descending packets which are virtual in nature. It is a deterministic model where an elevator is attached to each packet and if an elevator is faulty then the packet assigned to it is blocked. Here a new header is added to each packet containing the address of the elevator which increases both the area and time overhead and lack of path diversity can lead to redundancies in the routing which further affects its fault tolerance capability.

J. Lee et. al. [24] proposed an algorithm called Redelf which ensures that the system is deadlock-free without using any virtual channels. In the elevator first algorithm mentioned above, there is no rule which says that a packet has to be assigned to an elevator the moment it traverses into the layers whereas in the case of Redelf it is necessary that a packet is assigned to an elevator. Similar to the elevator first algorithm the Redelf is also a deterministic algorithm because of which a lot of congestion takes place in the layers.

To overcome this problem R. Salamat et. al. [25] proposed a fault resilient routing algorithm that is both lightweight and adaptive. Here the virtual channel is present only in the Y direction. Since the vertically partially connected links cannot use the algorithms used in the XYZ directions to eliminate the problem of deadlock the authors have proposed an East Then West (ETW) algorithm which can function with only one TSV link which is present on the eastmost column. Unlike the Elevator-First algorithm, this ETW algorithm uses only one VC. A set of elevators are present at each router for source to destination transmission for the packets to choose from. This makes the algorithm fault tolerable and also reduces the area overhead. This algorithm is fault resilient as long as there is at most one TSV on the east side.

M. M. Hafizur Rahman et. al. [26] proposed a deadlock-free routing algorithm named Hierarchical 3D-Mesh (H3DM) which can be applied to a hierarchical mesh network using 2 virtual channels. Here dimension-order routing is used to evaluate the networks dynamic communication performance under the uniform traffic pattern. The minimum number of

virtual channels to be present for such a network is two. However, the number of virtual channels increases the area overhead increases.

M. Fathi et. al. [27] proposed a novel algorithm that is minimally fully adaptive (MFA) which can tolerate the permanent faults occurring in the NoC based on the Hamilton path strategy which is repeated till the packets reach the destination from the source. Since the network is subdivided into two subnets which are acyclic this algorithm is deadlock-free. As the Hamilton path is repeated for each of the switch from source to destination to find the shortest path there is a huge time overhead for this algorithm.

B. Niazmand et. al. [28] proposed a fault tolerant algorithm, Logic-Based Distributed Routing for 3D NoCs (LBDR3D), which is both deadlock and live lock-free. If the packets destination node is not in the same layer as the source node then this fault tolerant implementation utilizes only two VCs for partially connected vertical links. If a packet with vertical links in a layer wanders between different nodes, it leads to live-lock and affects performance. Since the algorithm uses the Manhattan distance to select the nearest elevator, if this distance is the same for 2 or more elevators then the algorithm inherently leads to a deadlock condition. Unlike the ETW algorithm, this algorithm can also start at the west end first, and to locate a healthy elevator only a fixed number of bits can be stored within each and every router.

Jixi Su et. al. [29] a routing algorithm with the combination of a one-hop look ahead strategy for both horizontal and vertical links. Here the odd-even (OE) turn model is used in the vertical links to avoid the deadlock condition and improve the adaptiveness. This model can identify single and horizontal faults only but not applicable for vertical faults.

R. Salamat et. al. [30] proposed an adaptive algorithm for the column-based routing named CoBRA for fault-free communication between any two nodes even when there are failures in the TSVs. This algorithm follows a dynamic elevator selection principle and the routing knowledge is based on the router having partial knowledge about the faults in the TSV links and has complete knowledge about the healthy TSV in its column. To avoid deadlock two virtual channels are used to divide the Y dimension and for this reason, the network is divided into two disjointed subnetworks: subnetwork 1 and subnetwork 2. The packets can use either of the subnetworks and can also move from one subnetwork to the other but in order to avoid deadlock, switching is allowed only in one direction. Unfortunately, CoBRA poses some problems for the location and selection of the elevator and it requires at least one healthy elevator on both the east end and west end. the packets are routed first through the east end and if there are faulty TSV's here then the packet is routed through the west end and during this reconfiguration, some of the packets are dropped until the network is back to its stable condition which leads to the loss of data.

A. Charif et. al. [31] proposed an algorithm Rout3D, which is lightweight, resilient, and highly adaptive for partially connected 3D NoCs where the number of TSVs is less. This algorithm requires very few VCs, such that it has at least one healthy TSV available to connect all the layers regardless of its position to see that there is no deadlock in the system. In order to improve the performance of this algorithm another variant of this algorithm has been proposed which is called the Rout3d+. The difference between this algorithm and the previous algorithm is that in the previous case 2 VC were present in the east and north direction and one VC in all the other directions whereas, the Rout3d+ algorithm has one additional VCs for all the vertical ports to decrease the pressure on the vertical links. Since the algorithm has more than XY routing it is much more complex than the elevator -first algorithm and as it is fully adaptive it also has a large area overhead.

E. Taheri et. al. [32] proposed an algorithm Advertiser Elevator, to overcome the failures in the partially connected NoCs. In case of any failure in the vertical link, this algorithm finds an alternate link to reroute the packets with the help of indexing each vertical link and sharing its details with the nearest links, but this increases the runtime which severely affects the deadlock condition. Alternatively, the number of elevators can be increased wherein a router can have more than one elevator in its routing table or logic. The detection of faults leads to rerouting again which increases the runtime as well as an increase in the area overhead which increases the hardware cost and limits the scalability.

A. Ben Abdallah et. al. [34] proposed a fault tolerant technique wherein a scalable TSV utilization architecture is used to overcome the interlayer reliability issues by reducing the cluster defects. Here a router is used to correct the defective TSV link by using any one of the neighbors link which is present in the same layer. For avoiding the timing violation issues the authors have placed the TSVs of two nearby routers in between them and a cluster is only shared between two neighboring routers. For TSV clusters with open defect faults K. N. Dang et. al.[34] proposed a scalable design for an online algorithm to work around new defects without the use of any redundancies. Cross talk is an issue here as the faulty TSVs affect the healthy TSVs.

R. Salamat et. al. [35] proposed an adaptive routing algorithm for partially connected 3D- NoC named Longitudinal Exclusively Adaptive or Deterministic and abbreviated as LEAD. In this algorithm an analytical model is used which is based on the queueing theory, to accommodate the adaptivity of the routing algorithm. This model distributes the traffic in the network in a uniform manner which is useful for keeping the temperature low as it doesnt route in the directions where there is more heat thereby trying to keep the temperature of the chip optimum. To be adaptive and deadlock-free this algorithm requires extra VC along the X and Y directions. Here the TSVs should be able to connect all the layers and the vertical links are to function as pillars.

Y. Cai et. al. [36] proposed an improvement to the 2D repetitive turn model for 3D structures with new rules for the vertical turns and repetitive prohibited turns along the horizontal column and row without the need for virtual channels. Deadlock free status is achieved as long as there are no 180 degree turns such that there are no cycles in the same plane or the span of two or more planes. For intra-layer routing, if the conditions observed in 2D RTM are observed then the network can be deadlock-free.

TABLE I
VARIOUS METHODOLOGIES DISCUSSED WITH RESPECT TO THE 3D-NOC AND THE THROUGH SILICON VIAS (TSVS)

Name of the author	Methodology	Limitations
Pasca et. al.	Vertical links are used for inter-die communication	Links have degradation factor called the serialization factor which if increases affect performance and increases area overhead.
C. Feng et. al.	Algorithm utilizing a routing table proposed for 2D meshes along with TSVs for entire system rather than global routing	Algorithm suffered from poor scalability
C. Hernandez et. al.	Link design for improving yield of TSVs by using additional networks	Increase in area overhead as because of using the networks a large number of TSV links remained unused
S. Pasricha et. al.	Novel 4NP-First Routing algorithm	Higher energy consumption compared to other hybrid models
S. Akbari et. al.	ZXY Routing Scheme when no faults and XZXY routing method when faults are present	Can only tolerate single faults on TSV and each router in a row should have information about other routers in that row
X. Jiang and T. Watanabe	Algorithm with two main considerations 1. Vertical node for interlayer routing and 2. 2D routing algorithm for intra layer routing	High energy consumption
M. Ebrahimi et. al.	Adaptive routing algorithm 3D-FAR where packet can utilize any virtual channel in X, Y, Z directions	Large area overhead due to virtual channels in all X, Y, Z directions.
M. Ebrahimi et. al.	Routing algorithm using Hamiltonian path strategy to deliver packets to several destinations	Cannot tolerate too many faulty links in both horizontal and vertical directions.
R. Alizadeh et. al.	Circular routing algorithm for faults on both horizontal and vertical links	Cannot tolerate too many faults.
J. Zhou et. al.	HARS algorithm which finds mid node of each layer to ensure packets bypass invalid vertical links	Results are reliable only when 10% of links are faulty.
F. Dubois et. al.	Elevator First algorithm for partially connected 3D-NoCs. Deterministic modelling is used	New header containing information address of elevator increases both area and time diversity.
J. Lee et. al.	Algorithm named Redelf where packet is assigned to elevator the moment it traverses into layers	Since it is a deterministic algorithm congestion takes place in the layers
R. Salamat et. al.	Light weight and adaptive fault resilient routing algorithm named East Then West	Algorithm is fault resilient as long as there is at most one TSV on east side.
M.M. Hafizur Rahman et. al.	Deadlock free routing algorithm named Hierarchical 3D-Mesh (H3DM) which uses 2 virtual channels	The number of virtual channels increase the area overhead.
M. Fathi et. al.	Novel algorithm that is minimally fully adaptive which can tolerate permanent faults	To find shortest path to the destination Hamilton Path strategy is used which increases the time overhead.
Jixi Su et. al.	Routing algorithm with combination of one-hop look ahead strategy for both horizontal and vertical links	This model is not applicable for vertical faults.
R. Salamat et. al.	Adaptive algorithm for column -based routing named CoBRA which follows dynamic elevator selection principle	Has issues with the location and selection of elevator which sometimes leads to loss of data.
A. Charif et. al.	Adaptive algorithms for fault resilience called Rout3D, Rout 3D+ which use virtual channels for routing in partially connected 3D NoCs	Since the algorithms are adaptive there is a large area overhead.
E. Taheri et. al.	Algorithm named Advertiser Elevator where if there is a failure in vertical link algorithm reroutes the packets to nearest fault free link	Detection of faults leading to rerouting increases runtime as well as area overhead which increases cost and limits scalability
A. Coelho. et.al	Runtime fault tolerant routing algorithms have been proposed which use the first- last routing method and one flit dedicated VC have been used as escape buffers if failures in TSVs are introduced	Hardware complexity leads to area overhead and increases power consumption
Y. Fu et. al.	This scheme has an assignment criterion based on congestion aware dynamic elevator assignment	This assignment criterion increases the complexity of the network

A. Charif et. al. [37] proposed an algorithm First-Last for targeting vertically, partially connected 3D-NoCs. This algorithm to achieve the deadlock-free situation requires a very small number of virtual channels. It has 2 VCs in the East and North directions and 1 each in the West, South, and Local port directions, where this unique structuring is given so that it can support any topology. An improvement over this algorithm Enhanced First-Last is also proposed which can improve the performance of the partially connected NoC even with a lesser number of TSVs.

An extension of the above work can be seen in [38], where a resilient routing scheme, Runtime fault-tolerant 3D Networks-on-Chip Scheme (RuNS) which can tolerate both manufacturing and runtime faults. In order to be both adaptive and deadlock-free this algorithm requires one additional virtual channel along the north and east direction. Here a set of 8 bits are present per router. Along with the TSVs status propagation, a fault tolerating routing algorithm is also used to search and identify a healthy elevator. To avoid the loss of packet date RuNS uses a rerouting mechanism and a buffer inside each router. The authors further proposed a novel lightweight and efficient fault tolerating algorithm by combining the above two algorithms First-Last Runtime and Resilient 3D Networks-on-Chip Scheme (FL-RuNS) [39] which combines the advantage of high performance observed in the First-Last algorithm and to repress the failures in the vertical links a routing scheme that is fault-tolerant is applied. Furthermore, this algorithm uses the concept of vertical link announcement to inform

the nodes about the healthy links in the network because of which the entire network can be re-configurable by the help of which there will not be any packet loss. One flit dedicated VCs which are used as an escape buffer in case the of failures in TSVs are introduced in this algorithm. It also requires very few VC's to achieve a deadlock-free network and reachability. The hardware complexity leads to area overhead and an increase in power consumption.

Y. Fu et al. [40] proposed an algorithm that is used for the assignment of elevators when there is congestion in partially connected 3D-NoCs, which takes the NoCs performance and the traffic load of the vertical links into consideration. This scheme has an assignment criterion which is based on the congestion-aware dynamic elevator assignment whose value is given by the weighted sum of the router delay and the square of the buffer utilization for the path from the source layer to the destination layer for the assigned elevator. Among all the paths obtained the scheme considers both the distance factor and the congestion information of the network, which are used to choose the path that has the least value, and the elevator along that path is assigned to the packet. Since the scheme itself is aware of congestion there is little to no chance for the presence of deadlock. The assignment criterion is dependent on many factors which leads to an increase in the complexity of the network.

III. DISCUSSION

A detailed summary of the various methodologies discussed above with respect to the 3D-NoC and their effects are given in Table 1 which is provided after the references.

IV. CONCLUSION AND FUTURE SCOPE

An attempt has been made to compare the various fault tolerant mechanisms applied onto the 3D- NoC and their performance attributes to reduce the vulnerability of these chips. It can be concluded that adaptive algorithms using virtual channels can enhance the performance of the chip and due to this a trade-off between the area overhead, run-time for evaluation, and the power consumed by the chip exists where these increased overheads lead to increased hardware cost and complexity which limit the scalability of the chip. Recently these fault tolerance improvements on the 3D-NoC are being carried out with the help of various machine learning algorithms that are computationally powerful as well as energy-efficient. In the future, we would like to propose an algorithm for improving the fault-tolerant routing by using optimization techniques and evaluate their performance metrics.

REFERENCES

[1] NKR Becchu et al., System level fault-tolerance core mapping and FPGA-based verification of NoC, Microelectronics Journal, vol 70, pp. 16–26, 2017.

[2] B. Naresh Kumar Reddy, M.H.Vasantha, Y.B.Nithin Kumar and Dheeraj Sharma, Communication Energy Constrained Spare Core on NoC, 6th International Conference on Computing, Communication and Networking Technologies (ICCCNT), PP. 1–4, 2015.

[3] B. Naresh Kumar Reddy, M.H.Vasantha, Y.B.Nithin Kumar and Dheeraj Sharma, A Fine Grained Position for Modular Core on NoC IEEE International Conference on Computer, Communication and Control, PP. 1–4, 2015.

[4] BNK Reddy et al., An energy-efficient fault-aware core mapping in mesh-based network on chip systems, Journal of Network and Computer Applications, vol 105, pp. 79–87, 2017.

[5] NKR Beechu et al., High-performance and energy-efficient fault-tolerance core mapping in NoC, Sustainable Computing: Informatics and Systems, vol 16, pp. 1–10, 2017.

[6] BNK Reddy et al., Energy-Aware and Reliability-Aware Mapping for NoC-Based Architectures, Wireless Personal Communications 100 (2), 213–225, 2018.

[7] NKR Beechu, VM Harishchandra, NKY Balachandra, Hardware implementation of fault tolerance NoC core mapping Telecommunication Systems 68 (4), 621–630, 2018.

[8] B. Naresh Kumar Reddy, and B. Sireesha, An Efficient Core Mapping Algorithm on Network on Chip, 22 nd International Symposium on VLSI Design and Test (VDAT), 2018

[9] B. Naresh Kumar Reddy, Dharavath Kishan and B. Veena VaniPerformance constrained multi-application network on chip core mapping," International Journal of Speech Technology , vol 22(4), pp. 927–936, 2019.

[10] AS Kumar, TVKH Rao, BNK Reddy, Exact Formulas for Fault Aware Core Mapping on NoC Reliability, 2020 IEEE 17th India Council International Conference (INDICON), 1–5, 2020.

[11] V. Pasca, L. Anghel, C. Rusu, and M. Benabdenbi, "Configurable serial fault-tolerant link for communication in 3D integrated systems," 2010 IEEE 16th International On-Line Testing Symposium, Corfu, 2010, pp. 115–120.

[12] C. Feng, M. Zhang, J. Li, J. Jiang, Z. Lu, and A. Jantsch, "A Low-Overhead Fault-Aware Deflection Routing Algorithm for 3D Network-on-Chip," 2011 IEEE Computer Society Annual Symposium on VLSI, Chennai, 2011, pp. 19–24.

[13] C. Hernandez, A. Roca, J. Flich, F. Silla, and J. Duato, "Fault-Tolerant Vertical Link Design for Effective 3D Stacking," in IEEE Computer Architecture Letters, vol. 10, no. 2, pp. 41–44, 2011.

[14] K. Hsieh, B. Cheng, R. Gu and K. S. Li, "Fault-tolerant mesh for 3D network on chip," 2011 6th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), Taipei, 2011, pp. 202–205.

[15] S. Pasricha and Y. Zou, "A low overhead fault tolerant routing scheme for 3D Networks-on-Chip," 2011 12th International Symposium on Quality Electronic Design, Santa Clara, CA, 2011, pp. 1–8.

[16] A. B. Ahmed, A. B. Ahmed, and A. B. Abdallah, "Deadlock-Recovery Support for Fault-tolerant Routing Algorithms in 3D-NoC Architectures," 2013 IEEE 7th International Symposium on Embedded Multicore Socs, Tokyo, 2013, pp. 67–72.

[17] S. Akbari, A. Shafiee, M. Fathy and R. Berangi, "AFRA: A low cost high performance reliable routing for 3D mesh NoCs," 2012 Design, Automation & Test in Europe Conference & Exhibition (DATE), Dresden, 2012, pp. 332–337.

[18] X. Jiang and T. Watanabe, "A novel fully adaptive fault-tolerant routing algorithm for 3D Network-on-Chip," 2013 IEEE International Conference of IEEE Region 10 (TENCON 2013), Xi'an, 2013, pp. 1–4.

[19] M. Ebrahimi, M. Daneshtalab, P. Liljeberg, and H. Tenhunen, "Fault-tolerant method with distributed monitoring and management technique for 3D stacked meshes," The 17th CSI International Symposium on Computer Architecture & Digital Systems (CADS 2013), Tehran, 2013, pp. 93–98.

[20] M. Ebrahimi, M. Daneshtalab, and J. Plosila, "Fault-tolerant routing algorithm for 3D NoC using hamiltonian path strategy," 2013 Design, Automation & Test in Europe Conference & Exhibition (DATE), Grenoble, France, 2013, pp. 1601–1604.

[21] R. Alizadeh, M. Saneei, and M. Ebrahimi, "Fault-tolerant circular routing algorithm for 3D-NoC," 2014 International Congress on Technology, Communication and Knowledge (ICTCK), Mashhad, 2014, pp. 1–7.

[22] J. Zhou, H. Li, Y. Fang, T. Wang, Y. Cheng, and X. Li, "HARS: A High-Performance Reliable Routing Scheme for 3D NoCs," 2014 IEEE Computer Society Annual Symposium on VLSI, Tampa, FL, 2014, pp. 392–397.

[23] F. Dubois, A. Sheibanyrad, F. Pétrot, and M. Bahmani, "Elevator-First: A Deadlock-Free Distributed Routing Algorithm for Vertically Partially Connected 3D-NoCs," in IEEE Transactions on Computers, vol. 62, no. 3, pp. 609–615, 2013.

[24] J. Lee and K. Choi, "A deadlock-free routing algorithm requiring no virtual channel on 3D-NoCs with partial vertical connections," 2013 Seventh IEEE/ACM International Symposium on Networks-on-Chip (NoCS), Tempe, AZ, 2013, pp. 1–2.

[25] R. Salamat, M. Khayambashi, M. Ebrahimi, and N. Bagherzadeh, "A Resilient Routing Algorithm with Formal Reliability Analysis for

Partially Connected 3D-NoCs," in IEEE Transactions on Computers, vol. 65, no. 11, pp. 3265–3279, 1 2016.

[26] M. M. Hafizur Rahman, A. Shah, and Y. Inoguchi, "A deadlock-free dimension order routing for hierarchical 3D-Mesh network," 2012 International Conference on Computer & Information Science (ICCIS), Kuala Lumpur, 2012, pp. 563–568.

[27] M. Fathi, S. Ebrahimi, and H. Pedram, "A fault-tolerant routing algorithm in 3D topology manycore processors," 2015 2nd International Conference on Knowledge-Based Engineering and Innovation (KBEI), Tehran, 2015, pp. 217–222.

[28] B. Niazmand, S. P. Azad, J. Flich, J. Raik, G. Jervan, and T. Hollstein, "Logic-based implementation of fault-tolerant routing in 3D network-on-chips," 2016 Tenth IEEE/ACM International Symposium on Networks-on-Chip (NOCS), Nara, 2016, pp. 1–8.

[29] Jixi Su, Changchun Chai, Xiangyang Lei, and YinTang Yang, "Vertical-mesh-conscious-dynamic routing algorithm for fault tolerant 3D NoC," 2016 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, 2016, pp. 2004–2008.

[30] R. Salamat, M. Ebrahimi, N. Bagherzadeh, and F. Verbeek, "CoBRA: Low-cost compensation of TSV failures in 3D-NoC," 2016 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT), Storrs, CT, 2016, pp. 115–120.

[31] A. Charif, N. Zergainoh, A. Coelho, and M. Nicolaidis, "Rout3D: A lightweight adaptive routing algorithm for tolerating faulty vertical links in 3D-NoCs," 2017 22nd IEEE European Test Symposium (ETS), Limassol, 2017, pp. 1–6.

[32] E. Taheri, M. Isakov, A. Patooghy, and M. A. Kinsy, "Advertiser elevator: A fault tolerant routing algorithm for partially connected 3D Network-on-Chips," 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, 2017, pp. 136–139.

[33] A. Ben Abdallah, K. N. Dang, and Y. Okuyama, "A low-overhead fault tolerant technique for TSV-based interconnects in 3D-IC systems," 2017 18th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA), Monastir, 2017, pp. 179–184.

[34] K. N. Dang, A. Ben Ahmed, Y. Okuyama, and A. Ben Abdallah, "Scalable design methodology and online algorithm for TSV-cluster defects recovery in highly reliable 3D-NoC systems," in IEEE Transactions on Emerging Topics in Computing.

[35] R. Salamat, M. Khayambashi, M. Ebrahimi, and N. Bagherzadeh, "LEAD: An Adaptive 3D-NoC Routing Algorithm with Queuing-Theory Based Analytical Verification," in IEEE Transactions on Computers, vol. 67, no. 8, pp. 1153–1166, 2018.

[36] Y. Cai, D. Xiang, and X. Ji, "Deadlock-Free Adaptive Routing Based on the Repetitive Turn Model for 3D Network-on-Chip," 2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/SustainCom), Melbourne, Australia, 2018, pp. 722–728.

[37] A. Charif, A. Coelho, M. Ebrahimi, N. Bagherzadeh, and N. Zergainoh, "First-Last: A Cost-Effective Adaptive Routing Solution for TSV-Based Three-Dimensional Networks-on-Chip," in IEEE Transactions on Computers, vol. 67, no. 10, pp. 1430–1444, 2018.

[38] A. Coelho, A. Charif, N. Zergainoh, and R. Velazco, "A Runtime Fault-Tolerant Routing Scheme for Partially Connected 3D Networks-on-Chip," 2018 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT), Chicago, IL, 2018, pp. 1–6.

[39] A. Coelho, A. Charif, N. Zergainoh, and R. Velazco, "FL-RuNS: A High-Performance and Runtime Reconfigurable Fault-Tolerant Routing Scheme for Partially Connected Three-Dimensional Networks on Chip," in IEEE Transactions on Nanotechnology, vol. 18, pp. 806–818, 2019.

[40] Y. Fu et al., "Congestion-Aware Dynamic Elevator Assignment for Partially Connected 3D-NoCs," 2019 IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 2019, pp. 1–5.