AFRA: A Low Cost High Performance Reliable Routing for 3D Mesh NoCs

Sara Akbari*, Ali Shafiee†, Mahmoud Fathy*, and Reza Berangi*
*School of Computer Engineering University of Science and Technology, Iran
†School of Computer Engineering Sharif University of Technology, Iran
saraakbari@comp.iust.ac.ir a_shafiei@ce.sharif.edu mahfathy@iust.ac.ir rberangi@iust.ac.ir

[1]

***Abstract*—Three-dimensional network-on-chips are suitable communication fabrics for high-density 3D many-core ICs. Such networks have shorter communication hop count, compared to 2D NoCs, and enjoy fast and power efficient TSV wires in vertical links. Unfortunately, the fabrication process of TSV connections has not matured yet, which results in poor vertical links yield. In this work, we address this challenge and introduce AFRA, a deadlock-free routing algorithm for 3D mesh-based NoCs that tolerates faults on vertical links. AFRA is designed to be simple, high performance, and robust. The simplicity is achieved by applying ZXY and XZXY routings in the absence and presence of fault, respectively. Furthermore, AFRA, as will be proved, is deadlock-free when all vertical faulty links have the same direction. This enables the routing to save virtual channels for performance rather than scarifying them for deadlock avoidance. Finally, AFRA provides robustness, which means supporting connection for all possible pairs of communicating nodes in high fault rates.**

AFRA is evaluated, though cycle accurate network simulation, and is compared with planar adaptive routing. Results reveal that AFRA significantly outperforms planar adaptive routing in both synthetic and real traffic patterns. In addition, the robustness of AFRA is calculated analytically.

I. INTRODUCTION

Three dimensional (3D) integration, i.e., stacking up multiple layers of die, is gaining popularity for its enormous benefits including heterogeneous technology integration, shorter interconnect, and better yield compared to 2D chip [1]. Such benefits motivate industry to adopt this technology, especially for systems with high number of cores. One potential bottleneck in such 3D many-core systems is inter-core communication. To tackle with this challenge, 3D Network-on-Chips (3D NoCs) are introduced that inherits both benefits of NoCs and 3D integration.

Typical 3D NoCs are composed of conventional 2D NoCs, which are vertically connected using Through-Silicon Vias (TSVs). Compared to horizontal links used in typical 2D NoCs, TSV-based links have different characteristics. More precisely, TSVs are significantly shorter than horizontal links and have larger pitch leading to fast and power efficient connections [4]. Thanks to these features and low diameter, 3D NoCs outperform 2D NoCs of the same size.

While 3D NoCs are superior in terms of performance and power, the commercial fabrication of such networks have not yet matured. The main reason accounting for this is low TSV yields, which is the result of different factors such as misalignment, failure on bonding, and random open defects [5][6]. As a result of this poor yield, a fault tolerant mechanism is necessary to realize efficient 3D NoCs. In this work we address permanent faults on vertical TSV based links and propose AFRA, a fault tolerant routing algorithm for 3D mesh topology. AFRA relies on the fact that fault rate on vertical links is much larger than conventional horizontal one. This is compatible with recently sub-100 nm fabricated 2D NoCs, which prefer XY routing, as standard circuit techniques are sufficient enough to reach high yield network [7][8]. ARFA is, therefore, designed to just tackle faults on vertical links. In addition, AFRA aims at satisfying the three following major goals:

Simplicity: This is a key factor in the design of routing as it may impact clock frequency and router complexity. This is also a widely accepted merit for industry since there are many off-chips and on-chip interconnects implemented with simple routing. AFRA realizes this goal by following ZXY routing and XZXY routing algorithms in the absence and presence of faults, respectively. As a result of these simple routings, some turns are useless hence can be removed from the crossbar. Moreover, it is possible to adopt decoupled crossbar design in order to achieve low-cost and high-performance router architecture [9]. Note that, adaptive fault tolerant routings lack this feature, as they tackle faults using adaptivity, which requires larger crossbar.

In addition to benefits from simple crossbar, AFRA is deterministic. This implies AFRA provides in-order packet arrival and, compared to adaptive routings, is better modeled analytically.

Good Performance: By means of recent advancement on resilient TVS links [5][6], most of the fabricated 3D NoCs will be fault-free. Reliable routing is just used to maximize the profit. This fact implies that proposed reliable routings are expected to compete with non-reliable high-performance routing in the absence of fault. In addition, performance metrics must not drop dramatically with a few faults, since such faulty cases cover most faulty situations.

To satisfy this goal, AFRA tries not to split virtual channels (VCs) for adaptivity or deadlock avoidance and saves them for

---

[1] Miss Akbari and Mr. Shafiee equally contributed to this paper.

978-3-9810801-8-6/DATE12/

performance improvement. As we will prove in Section III, AFRA does not require splitting into group of VCs to avoid deadlock, when all faulty links have the same direction (i.e., all faults happen exclusively on upward or downward links). The direct consequence of this fact is that, AFRA guarantees tolerating single-fault with no need to any VCs.

In addition to this interesting property, AFRA avoids adaptivity by using a routing similar to dimension ordered routing. Therefore, it prevents congestion happens due to adaptivity in the center of network in popular traffics such as uniform and bit-complement.

In order to evaluate the performance of AFRA, we compare it with planar adaptive routing, a well-known fault-tolerant adaptive routing for n-dimensional mesh and k-ary n-cube [18]. Results reveal that AFRA improves injection rate by up to 44% and 118% for synthetic traffics in the absence and presence of fault, respectively. For SPLASH-2 programs [19], AFRA outperforms planar Adaptive routing by up to 14% and 5% in faulty and fault-free situations.

Robustness: This measures the power of routing algorithm to provide connectivity for all pairs of communication nodes, even in the presence of many faulty links. As we will show in Section V, AFRA is robust enough (> 0.99) to support fault rates up to 15%.

The rest of the paper is organized as follows. In the next section we review prior work on fault-tolerant 3D NoCs and discriminate AFRA from other proposed routings. Section III explains suggested routing and the way deadlock is handled in more detail. We evaluate AFRA using network simulation in Section IV. Finally, Section V offers conclusion.

II. RELATED WORK

The problem with TSV disconnection is addressed in many prior researches. Such work dealt with this challenge by either suggesting resilient TSV links or reliable routings. Resilient TSV links are achieved by mean of redundancy, which in turns realized in different cost-reliability tradeoffs. In an extreme case, it has been suggested to duplicate every TSV [2]. This approach is highly reliable but at the cost of TSV footprint duplication. Other proposals tried achieving acceptable reliability with much less area occupation. More specifically, a few more wires, says M, are augmented to an unreliable N-wire link. The resulting link has M+N wires to support at most M faults. This link requires two circuits to distribute bits of a conveying flit over M+N wires and then to extract the N original bits out of the (M+N)-bit link. Other suggested reliable links differ on circuit used distribution and extraction power, ranging from simple multiplexers [5][6] to omega network [10].

In addition to link level mechanisms, fault tolerant routing has been suggested to tackle problems due to poor vertical yields. Rusu et al. proposed RILM [11], a reliable routing best suited for 3D NoC with heterogeneous topology in different layers. This work relies on two virtual networks for deadlock avoidance. AFRA, on the other hand, requires no extra virtual networks in many cases including single fault occurrence. Moreover, AFRA discriminates vertical from horizontal links by consider faults occurrence just on TSV wires. Rahmani et al. also proposed a method for Hybrid NoC-bus structure that primarily focuses on avoiding congestion and thermal hotspots [12]. Their method, as they mentioned, is also applicable for fault tolerance by consider faulty links as highly congested connections. The method relies on virtual channels and adaptive routing for deadlock avoidance and communication in 2D layer, respectively. Another approach, which is suitable for both 2D and 3D NoCs, is leveraging routing table to keep routing paths [13]. This method is highly resilient but suffers from poor scalability due to area required for the tables [14].

Besides the methods proposed for NoCs, there are many researches that address reliable routing in off-chip interconnection networks. Most of such ideas rely on adaptive routing to tolerate faults. For instance, planar adaptive routing routes packets using adaptivity in 2D planes to reduce the need for VCs [18]. Nortdbotten et al. also proposed a method based on adaptive routing that tackle fault by means on some intermediate nodes, which construct fault-free sub-paths suitable for adaptive routing [15]. Unfortunately, this method requires at least two VCs and tables to keep track of intermediate nodes positions. Wu, also, presented a minimal routing based on the concept of faulty block [16]. The model is then extended to 2D faulty block to scarify fewer nodes, hence achieving better performance [17]. The main problem for the last two proposals is that they are based on faulty node model (and not faulty link), which disable many working parts or nodes.

The main drawback of all interconnect networks reliable routings is that they treat all links, either vertical or horizontal, in the same way. While this is helpful for off-chip interconnects, it wastes some resources such as virtual channels, since the fault rates in vertical and horizontal links are different. We evaluate the impact of resource misuse on performance in Section IV.

III. AFRA

In this section we explain our proposed deadlock-free routing algorithm, AFRA, in more detail. As discussed in prior sections, AFRA designed to tolerate faults occurred on vertical TSV-based links. In the rest of this session, we first describe the routing algorithm (Section III-A) and then we investigate how to make the routing algorithm deadlock-free using virtual channels (Section III-B).

A. Routing Algorithm

In the absence of fault, AFRA routes packets to destination through ZXY routing algorithm. In the presence of fault on vertical links, however, AFRA switches to XZXY routing. Here, we explain this routing algorithm using an example illustrated in Figure 1. Assume that the vertical link AB is faulty and Node Src knows about this fault. This fault disconnects Src and Dst through ZXY path forcing AFRA to choose another path to resume connectivity. The new path is selected in two steps. In the first step, Src sends packets to an intermediate node (Here, Node M) known as escape

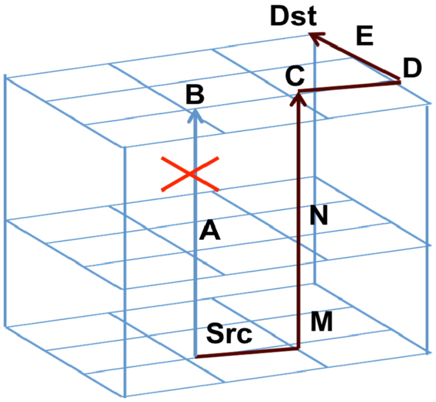


Fig. 1: *A case for XZXY routing as the result of fault on vertical link. M is the escape node from Src to Dst.*

node for communication pair (Src,Dst). In the second step, AFRA routes packet from escape node through ZXY routing to destination. Note that, packets are not ejected in escape node; hence, there is no performance penalty as the result of ejection-reinjection.

As mentioned above, AFRA relies on escape node for reliable communications. Escape node is formally defined as follows:

Definition 1: Node (x, y, z) is an escape node for communication from source Node (x_1, y_1, z_1) to destination Node (x_2, y_2, z_2) if it satisfies the two following conditions:

1) $y = y_1$ and $z = z_1$.
2) For $z_1 < z_2$ ($z_2 < z_1$), all upward (downward) links entering Node (x, y, z_0) with $z_1 < z_0 \leq z_2$ ($z_2 < z_0 \leq z_1$) must be healthy.

The first condition limits escape nodes to those reachable through X direction, hence simplifies routing algorithm. The second condition, guarantees a fault-free ZXY path from escape node to the destination. Therefore, it allows continuous traversing of Z dimension. This means that packets entering escape node are routed vertically toward destination layer with no need to dimension change. In many faulty patterns, there is more than one potential escape node for a communication pair. In such cases, AFRA selects an escape node for each pair of communication, which is selected based on the following rules:

1) Source node first tries to select the nearest escape node on the minimal path.
2) If there is no escape node on the minimal path, Source node selects the escape node with the smallest ID. (Node IDs are numbers unique to each node and ranges from 0 to $N-1$ where N is network size.)

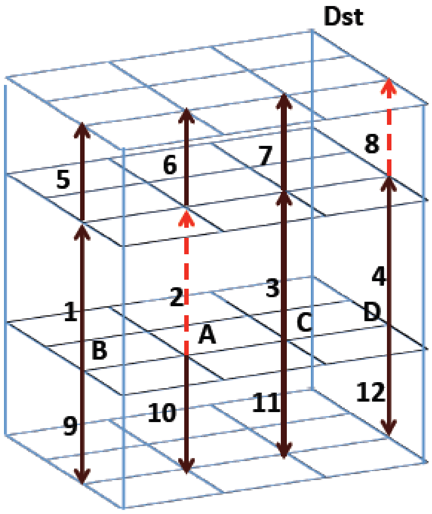


Fig. 2: *For source node A, B and C are potential escape nodes. Note that D is not an escape node as Link 8 is faulty.*

Regarding the definition of AFRA, every node has to keep some information to detect faulty ZXY path and to select an appropriate escape node for each pair of communicating nodes. This is realized by associating one bit that indicates faulty/working situation of vertical links potentially reachable through AFRA routing. Figure 2 shows such vertical links (links 1 to 12) for Node A. By mean of information recorded for these links, Node A (source node) is now able to detect fault on ZXY path for packets destined to Dst. Furthermore, this node finds Node B and C as escape nodes for this destination (Node Dst). In general, assuming $m \times n \times p$ mesh, for a node located at (x, y, z) coordination, AFRA keeps information of upward links entered nodes (x_1, y_1, z_1) with $z_1 > z$, $y_1 = y$, and $0 \leq x_1 < m$. In addition, it stores bits for downward links entered nodes (x_2, y_2, z_2) such that $z_2 < z$, $y_2 = y$, and $0 \leq x_2 < m$. This implies that AFRA records information for $m \times (p-1)$ vertical links per router. For example, in $8 \times 8 \times 4$ 3D mesh, the information overhead of AFRA per router is only $8 \times (4-1) = 24$ bits.

1) Deadlock-freedom: In this session, we study different fault patterns and their potential deadlock situations for wormhole flow control. Then we propose VC-based approach to avoid such cases. As mentioned in Section III-A, AFRA relies on XZXY routing to resume connection between two nodes when their corresponding ZXY path is faulty. An interesting feature of AFRA is that it is deadlock-free when all faults have the same direction (i.e., all faults happen on upward links or downward links). Therefore, in such cases, AFRA needs no VC for deadlock avoidance and each incoming packet can use any of VCs in virtual channel flow control. Here, we justify this feature. To simplify the proof, we first define the following notations for dependency turn.

Definition 2: We refer to turn ABC (i.e., dependency among link AB to link BC) as x^+z^- turn iff AB is in X dimension with increasing direction and BC is in Z dimension with decreasing direction. Similar to x^+z^+, we can define other notations such as x^-y^-, z^+x^+, etc., for other turns. In our notation, we also consider x as either x^+ or x^-. For example, we define turn xz^+ as either x^+z^+ or x^-z^+ turn. This definition is extensible for other dimensions. With the

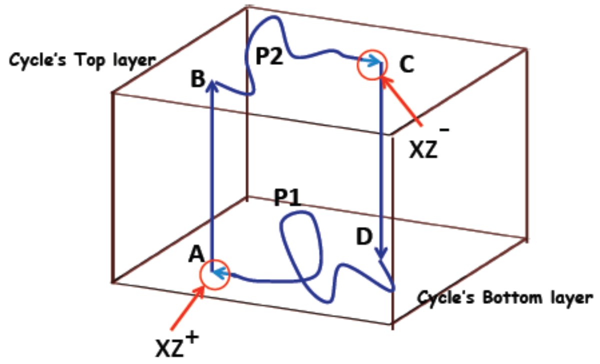


Fig. 3: *A presumed dependency cycle with its track on the top and the bottom layer of the cycle.*

notations defined above, we can now prove the deadlock-freedom of AFRA when all faults have the same directions.

Theorem 1: In 3D mesh NoC using wormhole switching, AFRA is deadlock-free, when all faults have the same direction (i.e., all faulty links are upward or downward).

Proof: Without loss of generality, we assume that all faulty links are upward. In the absence of fault, the only permitted turns are zx and xy as AFRA route packet using ZXY routing. In the presence of fault on upward links, however, AFRA further allows xz^+ turn in escape nodes (e.g., turn SrcMN in Figure 1). Therefore AFRA limits permitted turns to zx, xy, and xz^+.

Now consider that there exists a dependency cycle similar to that shown in Figure 3. Note that such cycles span to multiple layers (i.e., XY plane). This is because AFRA follows XY routing rules in each layer; hence, it is not possible to have a dependency cycle with all its nodes in just a single layer. Therefore, there are two layers as the top and the bottom layer in any such dependency cycle. Figure 3 illustrates a cycle with Path P1 and P2 as its intersection with the bottom and the top layers, respectively. In the node that this cycle leaves from the top layer (i.e., Node C), an xz or yz turn is needed. However, as mentioned before, these turns are not allowed which is a contradiction.

The same argument applies, when faults happen on downward links but for turns in the nodes leaving the bottom layer (i.e., Node A). Corollary 1: AFRA wont scarify VCs for deadlock avoidance in the case of single fault and in 50% of double fault cases, assuming uniform distribution of faults on all vertical links. Unfortunately, when there are at least two faults on different directions, there are some potential deadlock situations. To tackle such cases, AFRA relies on at least two virtual channels to construct two separate virtual networks. The first virtual network is for packets destined upward or for packet of even layer with source and destination in the same layer. Other packets are conveyed using the second network.

IV. METHODOLOGY AND RESULTS

In this section we evaluate AFRA in terms of performance and robustness. First, AFRA is compared with planar adaptive routing, in Section IV-A. Then, in Section IV-B, we analytically calculate its robustness.

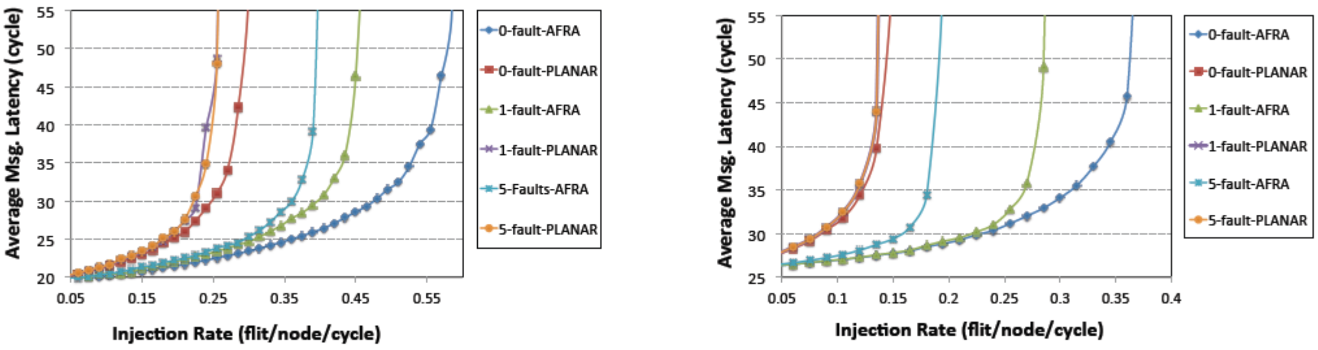
A. *Performance Evaluation*

We run cycle-accurate network simulations using Booksim [20], for both synthetic and real traffic patterns to compare AFRA and planar adaptive routing. In our simulations, we consider a 4-ary 3-mesh with three 5-flit VCs per input channel. Our experiment is based on 5-flit packets injected according to poison model (synthetic traffic) or 3-flit and single flit packets read from trace files of SPLASH-2 programs (real traffic) [19].

Figure 4 show average message latency of both AFRA and planar adaptive in three different situations (i.e., fault free, single fault, and five faults) and under uniform and bit complement patterns. According to this figure, AFRA outperforms planar adaptive routing in all cases. This is the result of using VCs to avoid head of line blocking rather than scarifying them for deadlock avoidance. In addition, adaptivity does not necessary leads to performance improvement, especially in the case of NoCs with few VCs per input channels. As a result, AFRA improves saturation injection rate by 70% and 54.1% for uniform traffic, and by 207% and 44% for bit complement traffic, under single and five faults respectively. Additionally, in the absence of fault, AFRA shows 1.8x and 3x increase in saturation injection rates for uniform and bit complement traffic patterns, respectively. We also compared AFRA and planar adaptive under real traffic patterns of SPLASH-2 programs. In our primary comparison, we observe that, with our router configuration, planar adaptive is not able to support acceptable average message latency and the latency exceeds to more than 200 cycles. AFRA, on the other hand, provide latency of less than 20 cycle as it uses all its VC for performance and follows dimension-order routing, which nicely fit with unifrom-like of traffic patterns of SPLASH-2 traces. We then double inter packet latency of trace files to reach the half rate of the original traces. For these calibrated traces, we run 30 single fault simulation scenarios. In each simulation, we randomly disable one vertical links. For these 30 samples, we evaluate minimum, maximum, and the average of improvement achieved by AFRA compared to baseline planar adaptive routing. Figure 5 illustrates these data. According to the figure, AFRA outperforms planar adaptive routing by more than 14%, even in half the real traffic rates.

B. *Robustness Evaluation*

In this section we calculate the probability of network disconnection, assuming vertical links with fault probability p. A network is disconnected if there is at least one source that finds no path to its destination using AFRA. If source node is below (above) the destination node we call it, upward (downward) disconnection. The calculation of disconnection probability relies on the following lemmas. In the interest of



(a) uniform

(b) bit complement

Fig. 4: *Averge message latency for AFRA and planar adaptive routing (PLANAR) in different fault scenarios (i.e. no fault, single fault, 5-fault case) under uniform and bit complement patterns.*

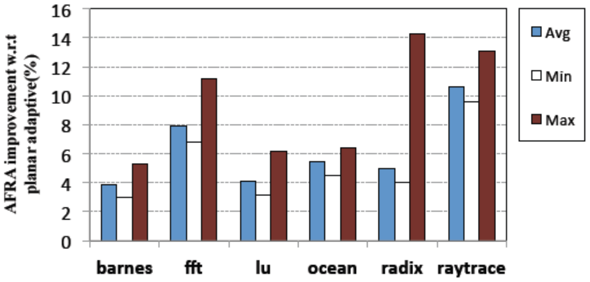


Fig. 5: *The average message latency comparison AFRA and planar adaptive for SPLASH-2 traffics.*

space, these lemmas consider n-ary 3-mesh as the network topology with faults just on upward links. The proof for downward links is similar.

Lemma 1: If Node (x_1, y_1, z_1) is upward disconnected from Node (x_2, y_2, z_2) $(z_1 < z_2)$, then it is disconnected from Node (x_2, y_1, z_2) as well.

Proof: Assume that Node (x_1, y_1, z_1) is connected to Node (x_2, y_1, z_2). Since we consider horizontal links are fault free, Node (x_2, y_2, z_2) is reachable from Node (x_2, y_1, z_2) using links on Y direction. If Node (x_1, y_1, z_1) could communicate with Node (x_2, y_1, z_2), which is achievable through XZ or XZX path, then it also reaches Node (x_2, y_2, z_2), which is a contradiction.

Lemma 2: If Node (x_1, y_1, z_1) is upward disconnected from Node (x_2, y_1, z_2), then Node $(x_1, y_1, 0)$ is also disconnected from Node $(x_2, y_1, n-1)$.

Proof: Assume that Node $(x_1, y_1, 0)$ is connected to Node $(x_2, y_1, n-1)$. Note that the path from Node $(x_1, y_1, 0)$ to Node $(x_2, y_1, n-1)$ shares some vertical links with a potential path from Node (x_1, y_1, z_1) to Node (x_2, y_1, z_2). If the former path is fault-free, then the latter one has fault-free vertical links, which means Node (x_1, y_1, z_1) is connected to Node (x_2, y_1, z_2). This contradicts with the lemmas assumption.

Lemma 3: For any $y (0 \leq y < n)$, if Node $(x_1, y, 0)$ is upward disconnected from Node $(x_2, y, n-1)$ for two specific values x_1 and x_2 $(x_1 \neq x_2)$, then Node $(x, y, 0)$ is upward disconnected from Node $(x_0, y, n-1)$ for all $0 \leq x < n$ and $0 \leq x_0 < n$ $(x \neq x_0)$.

Proof: Consider otherwise, that there are two Nodes A and B, located at $(a, y, 0)$ and $(b, y, n-1)$ respectively, that are not disconnect. Assume the connecting path starts from Node A passes through escape Node $(e, y, 0)$, going up safely to Node $(e, y, n-1)$ and finally reaches Node B. Based on this route, one can construct a fault-free XZXY path passing through Positions $(x_1, y, 0)$, $(e, y, 0)$, $(e, y, n-1)$, and $(x_2, y, n-1)$ in order to connect Node $(x_1, y, 0)$ to Node $(x_2, y, n-1)$. This contradicts our assumption that Node $(x_1, y, 0)$ is upward disconnected from Node $(x_2, y, n-1)$.

Based on Lemma 3, the probability of any disconnection in a XZ plane (i.e., all nodes with similar Y coordinates) equals the probability of disconnection between any two nodes in the top and the bottom of that plane. Now, we consider a concrete case of Node $(x_1, y_1, 0)$ in the bottom sends packets to Node $(x_2, y_1, n-1)$ in the top. According to definition of AFRA, Node $(x_1, y_1, 0)$ is upward disconnected from Node $(x_2, y_1, n-1)$ iff 1) there is a fault on one upward link entering a Node (x_1, y_1, z) with $0 < z < n$. and 2) there is no (fault free) escape Node $(x, y_1, 0)$ with $0 \leq x < n$ and $x \neq x_1$.

The probability of the first condition to happen is $1-(1-p)^{n-1}$. The probability that any Node $(x, y_1, 0)$ not be an escape node is also $1-(1-p)^{n-1}(0 \leq x < nx \neq x_1)$. Consequently, the probability of occurrence of condition 2 (i.e., all potential escape nodes fail to establish a path) is $(1-(1-p)^{n-1})^{n-1}$. Therefore, the probability that Node $(x_1, y_1, 0)$ fails to reach Node $(x_2, y_1, n-1)$ is $(1-(1-p)^{n-1})^n$. This value also represents the probability of upward disconnection for XZ plane y_1, i.e., XZ plane with Y coordinate y_1. Considering the same argument for downward links, the probability of disconnection (both upward and downward) in any XZ plane will be $(1-(1-p)^{n-1})^{2n}$. Now for whole the network to be disconnected at least one of

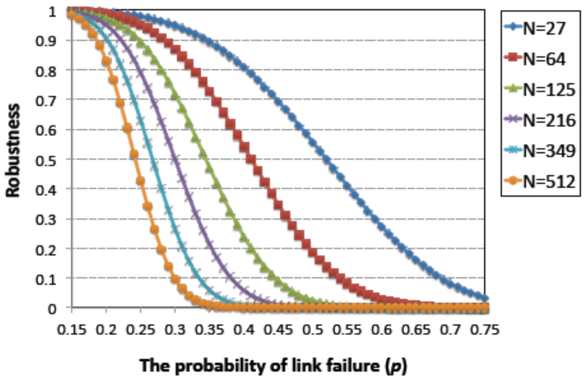


Fig. 6: *Robustness for n-ary 3-mesh of different size (N is the size of Network.*

its XZ plane must be disconnected. As a result, the probability of network disconnection is:

$$P = 1 - Pro(XZ \ \ plane\ is\ connected)^n$$
$$= 1 - (1 - (1 - (1-p)^{n-1})^{2n})^n \qquad (1)$$

In general, and for MNL 3D meshes, the robustness (i.e., the probability that all nodes of the network are connected) is:

$$Robustness = (1 - (1 - (1-p)^{L-1})^{2M})^N \qquad (2)$$

Based on the above formula, Figure 6 plots the robustness of n-ary 3-mesh topology of different size for vertical links fault probability ranging from 0 to 1. As shown in the figure, the robustness decrease by increasing the size of network. This is because larger networks have more vertical links. Furthermore, for low probability fault range (i.e., 0 to 0.15), almost all (>0.99) the fabricated networks of different size, with AFRA routing, are connected. Fortunately, this fault probability range is the practical range achievable by state-of-the-art link layer technology [5][6]. Therefore, AFRA is robust enough to cover all possible case not tolerated by link layer approaches.

V. CONCLUSION

In this paper we have proposed, AFRA, a low-cost highly efficient, reliable routing for 3D mesh NoCs. AFRA relies on this fact that reliability consideration is only necessary for vertical links in routing layer. Therefore, AFRA tries to follow simple ZXY routing when possible and just switch to XZXY routing when a fault is detected in baseline path. We also proved that AFRA successfully tolerate, with no need to VCs for deadlock avoidance, when all faulty links are exclusively upward or downward. In addition the robustness of AFRA has been analytically investigated. Based on our analysis, AFRA is highly robust even in high fault rate (15%). Finally, the performance of AFRA is compared with planar adaptive routing using cycle-accurate network simulations. According to the results, AFRA improves saturation injection rate by 70% and 54.1% for uniform traffic, and by 207% and 44% for bit complement traffic, under single and five faults respectively.

ACKNOWLEDGMENT

The authors would like to thank Mr. Mahdy Zolghadr for his invaluable comments.We would Also like to thank Caspian Company as part of this research is done there.

REFERENCES

[1] A. W. Topol, et al., Three-dimensional integrated circuits. IBM Journal of Researches and Developments .July 2006.
[2] U. Kang et al., 8 Gb 3-D DDR3 DRAM Using Through-Silicon-Via Technology. IEEE JSSC, January 2010.
[3] M. B. Healy, et al. Design and Analysis of 3D-MAPS: A Many-Core 3D Processor with Stacked Memory. In CICC, San Jose, California, September, 2010.
[4] S. Spiesshoefer, et al. Z-axis interconnects using fine pitch, nanoscale through-silicon vias. In ECTC 2004
[5] A. Hsieh, et al. TSV redundancy: architecture and design issues in 3D IC. In DATE. Leuven, Belgium, 2010.
[6] I. Loi, et al. Characterization and Implementation of Fault-Tolerant Vertical Links for 3-D Networks-on-Chip. IEEE TCAD, 2011.
[7] Y. Hoskote, et al. 2007. A 5-GHz Mesh Interconnect for a Teraflops Processor. IEEE Micro, September, 2007.
[8] J. Howard, et al. A 48-Core IA-32 message-passing processor with DVFS in 45nm CMOS ISSCC 2010.
[9] John Kim. Low-cost router microarchitecture for on-chip networks. In MICRO. ACM, New York, USA, 2009.
[10] C Hernndez, et al. Fault-Tolerant Vertical Link Design for Effective 3D Stacking. IEEE CAL,June. 2011.
[11] C. Rusu, L. Anghel, D. Avresky. RILM: Reconfigurable inter-layer routing mechanism for 3D multi-layer networks-on-chip. In IOLTS, 2010.
[12] A. Rahmani ,et al. Congestion aware, fault tolerant, and thermally efficient inter-layer communication scheme for hybrid NoC-bus 3D architectures. In NoCs, Pittsburgh, Pennsylvania, May, 2010.
[13] D. Fick, et al.A Highly Resilient Routing Algorithm for Fault-Tolerant NoCs. In DATE, 2009.
[14] A. Mejia, et al. Region-based routing: a mechanism to support efficient routing algorithms in NoCs. IEEE TVLSI, March, 2009.
[15] N. A.Nordbotten et al. A fully adaptive fault-tolerant routing methodology based on intermediate nodes. In LNCS , 2004.
[16] J. Wu. A simple fault-tolerant adaptive and minimal routing approach in 3-D meshes. In Journal of Computer Science and Technology, January 2003.
[17] D. Xiang, et al. Fault-tolerant routing in meshes/tori using planarly constructed fault blocks. In ICPP,, 2005.
[18] A. A. Chien and J. H. Kim. Planar-adaptive routing: low-cost adaptive networks for multiprocessors. In Journal of ACM, January, 1995.
[19] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, and A. Gupta. The SPLASH-2 Programs: Characterization and Methodological Considerations. In ISCA , June, 1995.
[20] https://nocs.stanford.edu/