

Non-Blocking BIST for Continuous Reliability Monitoring of Networks-on-Chip

Junshi Wang^{*†‡}, Letian Huang^{*}, Masoumeh Ebrahimi[†], Qiang Li^{*}, Guangjun Li^{*} and Axel Jantsch^{†*}

^{*}University of Electronic Science and Technology of China, Chengdu, China

Email: wangjsh@std.uestc.edu.cn, huanglt@uestc.edu.cn

[†]Technische Universität Wien, Vienna, Austria

[‡]Royal Institute of Technology, Stockholm, Sweden and University of Turku, Turku, Finland

Abstract—To achieve high reliability in on-chip networks, frequent runs of Built-in Self-Test allow the detection of and recovery from faults before they affect packets and the system functionality. However, to test routers, wrappers isolate cores from the network which leads to execution blocking and performance loss. In this paper, we propose a design-for-test reconfigurable router with two alternative bypassing channels. The router architecture allows maintaining the connection between cores and the network during the testing procedure by utilizing the bypassing channels. With the help of an adaptive routing algorithm and a testing strategy, networks can be fully tested at a high testing frequency with <15% increase of execution time.

I. INTRODUCTION

Due to aggressive technology scaling, reliability has become one of the critical issues for Network-on-Chips (NoCs). Physical failures can affect the behavior of circuits. Packets infected by faulty components may lead to misbehavior of the network and the application [1]. The *incubation period* is defined as the period between the occurrence of a physical failure and the observation of an error on the network level. If faults can be detected within the incubation period, failures can be avoided, or their impact can be limited.

Built-in self-test (BIST) can detect faults at runtime by injecting well-designed testing vectors. By increasing the frequency of testing procedure, BISTs can detect and diagnose physical failures before they lead to serious effects on the network. Since the test procedure isolates the routers under test (RUTs) by wrappers formed by multiplexers and demultiplexers, the network connectivity and integrity are disturbed, and all packets passing through the RUT or sent to or from the connected core are blocked. High testing frequency is desirable to identify faults as early as possible, but it leads to longer execution times of applications [2].

In our previous work, we proposed TARRA [3] to overcome the shortage of traditional BIST methods. As shown in Figure 1, by applying the bypassing channels, the core connected to RUTs can normally be accessed during the testing period. With the help of bypassing channels, an adaptive routing algorithm, and a proper testing strategy, the execution time only increases no more than 3% at a high testing frequency. However, the main drawback of TARRA is that the global links cannot be tested because the links must be available all the time. The global links are used to maintain the connection of the network in both the normal phase and the testing phase (Figure 1).

Also, the overhead of applying three bi-directional bypassing channels is significant.

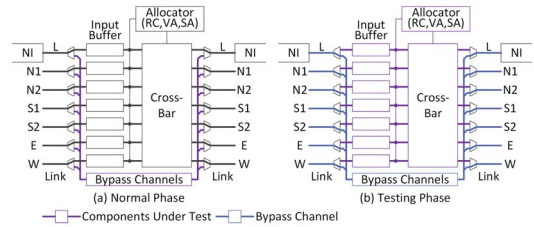


Fig. 1: Test coverage for TARRA.

To overcome the shortage of TARRA, we propose an enhanced design called **AlterTest**. AlterTest utilizes only one bypassing channel to enable the access between the core and the RUT during the testing period. The bypass channel can be selectively connected to one of the two physical ports at a time so that all the components can be tested. Moreover, the adaptive deadlock-free routing algorithm and network-level testing strategy are described in this work to fit the alternative bypass channel and to achieve better system performance.

The testing vectors for BIST are beyond the scope of this work. The paper is organized as follows. The router architecture and the routing algorithm are described in Section II and Section III. The experimental results are discussed in Section IV while the conclusion is given in the final section.

II. RECONFIGURABLE ROUTER ARCHITECTURE AND TESTING STRATEGY

A. Router Architecture

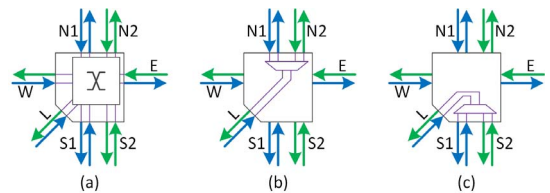


Fig. 2: Router architecture (a) normal mode using crossbar unit; (b) under test; (c) under test for top borderline routers

As shown in Figure 2, each router has seven pairs of channels, i.e. Local(L), East(E), West(W), North1(N1),

North2(N2), South1(S1), and South2(S2). Two channels along the Y dimension are used to avoid deadlock and also to provide two options to the bypass channel.

By default, the input and output channels are connected through a crossbar unit. When a router is under test, either the N1 or N2 port is connected to the local port through the bypass channel and a multiplexer (Figure 2(b)). If the router is on the top border, the bypass channel connects the local port to the south port (S1 or S2) as shown in Figure 2(c). The north or south neighbor of a RUT, which helps the packets to access the core and the network, is called *ladder router*. Therefore, at a time, one and only one port connects the local port to the ladder router. Other ports and links are under test.

B. Testing Procedure Control

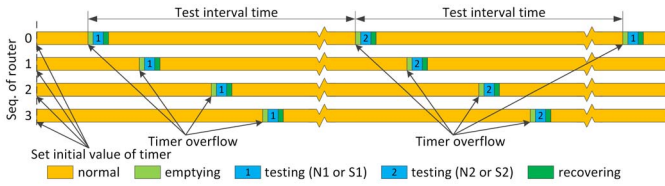


Fig. 3: Test state machine and testing control

Each router has one timer to trigger the testing procedure periodically. The overflow and initialization value determine the interval time between two procedures and the start time of first procedures (Figure 3), respectively. The first testing round connects the N1 (or S1) port to the bypass channel while the second testing round connects the local port to the N2 (or S2) port. Therefore, the router is fully tested after two rounds.

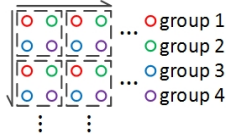


Fig. 4: Four-group testing sequence

Because packets turn around the RUTs, the neighbors and corner neighbors of RUTs suffer from increased traffic. To avoid deadlock and hotspots, RUTs should not be each others' neighbors at any time. Hence, a specific testing sequence is necessary. The routers are divided into four groups. The distance of routers in one group is larger than 1 in both the X and the Y dimension. As shown in Figure 4, groups are tested in turn, and the routers in one group are tested from top-left to bottom-right. With the increase of testing frequency, more and more routers go under test at the same time. At most, one-fourth of routers can be tested concurrently.

One testing period contains four phases, Normal, Emptying, Testing and Recovering (Figure 3). In the Normal phase, the crossbar is used to deliver packets. In the Testing mode, the router is disabled except one of the bypassing channel.

The paths of packets are different in the Normal and Testing phases. To avoid loss of packets, the RUT are temporarily emptied during the Emptying and Recovering phases. The RUT

and neighboring routers stop sending new packets but proceed to deliver all packets which have already been partially delivered. The clearing behavior in neighboring routers is triggered by ER signals from the RUT, and neighboring routers respond with EA signals. The signals to inform the status and synchronize are shown in Figure 5.

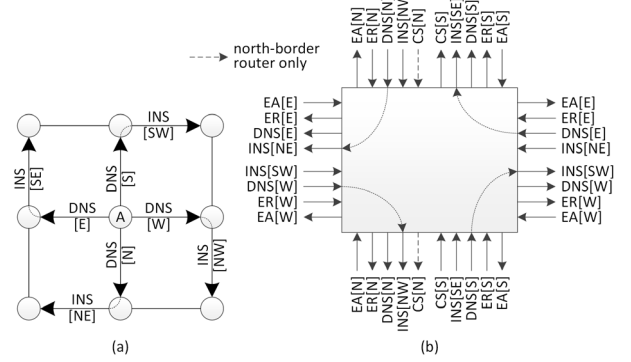


Fig. 5: (a) Propagation of synchronization signals in the network. (b) Connections of synchronization signals in a router

During the Normal and Emptying phases, the entire control path and the crossbar are active while during the Testing and Recovering phases, only the bypassing channel is active. In the Emptying, Testing, and Recovering phases, neighboring routers route packets by assuming that the RUT is disabled. Signals DNS and INS propagate the status of the RUT to a 3x3 area with the RUT being in the center. With the CS signal, the RUT informs the ladder router to switch between N1 and N2 (S1 and S2).

C. Testing coverage

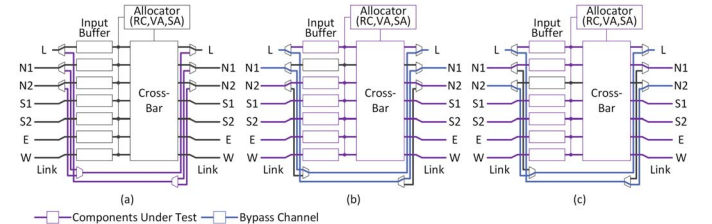


Fig. 6: Test coverage in different phases.

As shown in Figure 6, testing procedures test different parts of the router. In the Testing (N1 or S1) phase (Figure 3), all channels except N1 are tested. N1 is tested in the Testing (N2 or S2) phase. The additional components of testing such as bypassing channels, testing interval timer, and testing controller are tested in the Normal phase. After two testing procedures, all the components are tested except the connections between N1 and N2 in the crossbar. However, the connections between N1 and N2 are not used in our routing algorithm, so these connections are not necessary in the crossbar. In short, the proposed testing strategy can fully test the router.

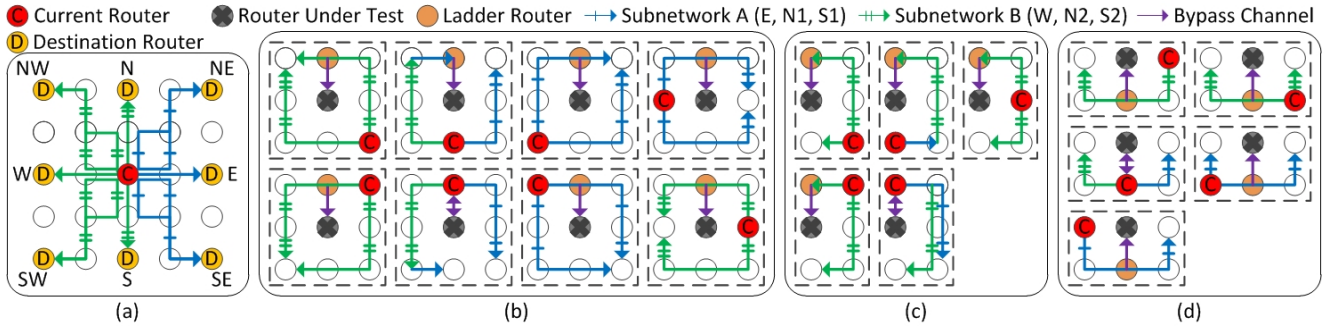


Fig. 7: Example of paths for the routing algorithm with different location of current router, destination and router under test. (a) No router under test; (b) Router under test is not on the top border or west border; (c) Router under test is on the west border; (d) Router under test is on the top border.

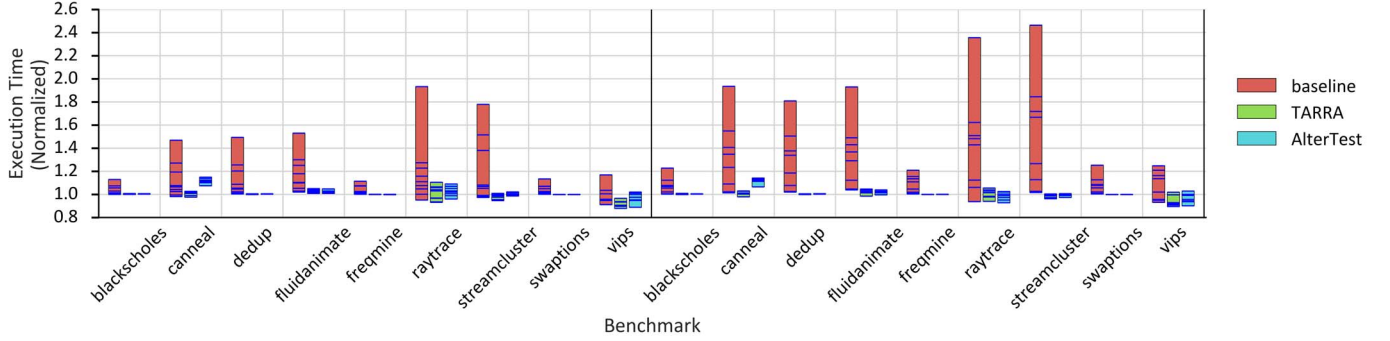


Fig. 8: Execution time of 100 million instructions of the PARSEC benchmarks. The blue short lines mark the execution time for different testing interval times. The top and bottom border of boxes denote the longest and shortest execution time for experiments with different testing interval times.

III. FULLY ADAPTIVE ROUTING ALGORITHM

To avoid deadlock, the network is divided into two subnetworks. *Subnetwork A* covers the E, N1, and S1 channels and the east (E), northeast (NE), and southeast (SE)-bound packets are routed in this subnetwork. The *Subnetwork B* covers the W, N2, and S2 channels and the west (W), northwest (NW) and southwest (SW)-bound packets are assigned to this subnetwork. The north (N) and south (S)-bound packets can be routed either in Subnetwork A using N1 and S1 channels or in subnetwork B using N2 or S2 channels. In this work, the N2 or S2 channels are used by N and S-bound packets. The subnetworks are deadlock free since the availability of three channels within each subnetwork is not sufficient to close a cycle. In addition, a cycle cannot be closed by a single transition from one subnetwork to another. We use this property to allow a transition from Subnetwork B to Subnetwork A, meaning that if packets use the channels of subnetwork A, they cannot use any channels of subnetwork B any longer. We should also note that in borderline routers, specified transition from subnetwork A to subnetwork B is possible because a cycle cannot be formed in borderlines. As shown in Figure 7(c), only the turns E-N2 and E-S2 are allowed in this network because the turns at the indirect neighboring router of the RUT could not form a cycle. We summarize the rules as:

- 1) E, NE, and SE packets should use the channels of

Subnetwork A (i.e. E, N1, and S1) for routing.

- 2) N, S, W, NW, SE packets should use the channels of Subnetwork B (i.e. W, N2, and S2) for routing.
- 3) Packets of Subnetwork B can change to use Subnetwork A, but packets of Subnetwork A cannot change to Subnetwork B.
- 4) If both the source and destination are in the left-borderline, packets can turn to N2 or S2 in Subnetwork B after using the E channel of Subnetwork A.

In addition to these rules, in order to enable the algorithm to choose from the shortest paths, the status of the RUT is delivered to eight direct and indirect neighboring routers, forming a 3x3 area.

As shown in Figure 7(a), in fault-free cases, Rule1 and Rule2 are applied to select the channels from the current router to any possible destination position. Due to the limited information about the location of faults, the NW, NE, SW and SE-bound packets are adaptively routed to the indirect neighbor of the destination router that is located on the shortest path. After reaching this position, packets are delivered to the destination router using the YX algorithm. Figure 7(b) shows all possible positions of the current router with respect to the RUT. As can be seen in this figure, Rule1, Rule2, and Rule3 are applied to route packets around the RUT toward any possible position of destination. Note that, when the N and S-bound packets turn around the RUT, they will not turn

back to the column of destination until they arrive in the same row as the destination. This is due to the fact that the transition from Subnetwork B to Subnetwork A can occur only once. In addition, bypassing channels are used to enable sending/receiving packets from/to RUTs. Bypassing channels directly connect the core of the RUT to the ladder router (i.e. immediate north or south neighbor of RUT). Figure 7(c) covers the cases where RUT is located in the left borderline. When the current router is located in the same column as the destination, Rule4 is utilized to turn around the RUT. In other cases, either Rule1 or Rule2 is applied. Finally, Figure 7(d) illustrates the conditions where the RUT is located at the top borderline. To deliver packets from the current router to the destination router, either Rule1 or Rule2 is applied.

A. Reliability analysis of multiple routers under test

As we have shown, all locations of a single router under test are supported in the network without any packet loss and by allowing all cores to function normally. If there are more than one RUT, some patterns are not supported due to deadlock or unreachable destinations. After analyzing all patterns with two RUTs, we observed that as long as the distance between RUTs in X-axis or Y-axis is larger than 1, the patterns are supported. Therefore, there is no unsupported pattern based on the 4-group testing sequence (Figure 4).

IV. EXPERIMENT AND DISCUSSION

AlterTest is compared with TARRA [3] and the baseline approach described in [2]. In the baseline method, one testing period has four phases as well. Packets have to be blocked in the neighbors until the Testing phase ends.

A. Hardware overhead

To assess the area overhead and power consumption, router architectures are synthesized using Synopsys Design Compiler with TSMC65nm technology. Because all three methods can be used to implement the same BIST methods, the overhead of the BIST circuit is not included. As Table I shows, AlterTest and TARRA cost more area and power than the baseline method due to the bypass channels and the routing algorithms. The routing algorithm of AlterTest is more complex than TARRA which is more than compensated by the smaller number of bypass channels. Thus, AlterTest has the same area but consumes less power than TARRA.

TABLE I: Power and area analysis. The area and power of BIST circuits are not considered in this table.

	AlterTest	TARRA [3]	BTest [2]
Power	8.577 μ W	12.53 μ W	6.869 μ W
Area	0.079mm ²	0.079mm ²	0.076 mm ²

B. Simulation results of execution time under PARSEC benchmarks

We measure the execution time of PARSEC benchmarks [4] for 100 million instructions to evaluate the performance of AlterTest, TARRA, and the Baseline method. The many-core platform built on SimpleScalar has 64 Alpha-21264 cores

and 4 memory controllers. The system implies a 10×8 interconnection network. Inspired from the published works [2], [5], we assume the testing period of a signal router as 500 or 1000 cycles. The testing interval time reduces from 10^6 cycles to 10000 or 20000 cycles.

Figure 8 shows the range of execution times for different testing interval times. The value has normalized to the execution time without BIST. In the baseline method, packets are blocked during the entire testing procedure, so the execution time rapidly increases as the testing interval time reduces. Especially when multiple routers are under test at the same time, the execution time can be extended to more than twice of the execution time without BIST. On the other hand, in AlterTest and TARRA, packets are only blocked during Emptying and Recovering phases which are 4 cycles on average. The execution time of both AlterTest and TARRA stays very close to the execution time without BIST except for Raytrace and Streamcluster. Because AlterTest does not provide bypass channels to cross the RUTs, the distance of paths is longer than in TARRA. Longer latency of packets leads to the increase on the execution time. The maximum increase of execution time is 15% under the Canneal benchmark.

V. CONCLUSION

BIST can improve the reliability of NoCs, but testing wrappers isolate the core connected to the RUT. AlterTest is proposed in this paper which consists of a reconfigurable router architecture, an adaptive routing algorithm, and a testing strategy. During the testing, one of the north/south channels is connected to the local port by one bypass channel while other components are under test. By changing the bypassing connections, a router can be fully tested. In summary, AlterTest exhibits lower power consumption at the expense of higher execution time in some cases.

VI. ACKNOWLEDGMENTS

This paper was supported by the National Natural Science Foundation of China under grant (NSFC) No.61176025, No.61006027, No.61534002, No.61471407, Chinese Scholarship Council and the Oversea Academic Training Funds (OATF), UESTC. This work is also supported by VINNOVA-MarieCurie within the ERoT project and Academy of Finland.

REFERENCES

- [1] A. Prodromou, A. Panteli, C. Nicopoulos, and Y. Sazeides, "Nocalert: An on-lien and real-time fault detection mechanism for network-on-chip architectures," in *45th Annual IEEE/ACM International Symposium on Microarchitecture*. ACM/IEEE, 2012, pp. 60–70.
- [2] M. Kakoei, V. Bertacco, and L. Benini, "At-speed distributed functional testing to detect logic and delay faults in nocs," *IEEE Transactions on Computers*, vol. 63, no. 3, pp. 703–717, 2014.
- [3] L. Huang, J. Wang, M. Ebrahimi, M. Daneshmand, X. Zhang, G. Li, and A. Jantsch, "Non-blocking testing for network-on-chip," *IEEE Transactions on Computers*, vol. 65, no. 3, pp. 679–692, March 2016.
- [4] *the PASRC Benchmark Suite*, <http://parsec.cs.princeton.edu/>.
- [5] A. DeOrion, D. Fick, V. Bertacco, D. Sylvester, D. Blaauw, J. Hu, and G. Chen, "A reliable routing architecture and algorithm for nocs," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 5, pp. 726–739, 2012.