TSINGHUA SCIENCE AND TECHNOLOGY
ISSN 1007-0214 08/10 pp85–94
DOI: 10.26599/TST.2019.9010039
Volume 26, Number 1, February 2021

Fault-Tolerant Design for Data Efficient Retransmission in WiNoC

Yiming Ouyang, Qi Wang*, Zhe Li, Huaguo Liang, and Jianhua Li

Abstract: There is a sharp decline in the network performance when the wireless link fails as a data path in the Wireless Network-on-Chip (WiNoC). To counteract this problem, we propose a fault-tolerance mechanism for the efficient retransmission of data in the WiNoC. When an error is detected in the data transmission process, this mechanism works to feed back the fault information to the source node in real time via fault signal lines. In the source node, the highest transmission priority is assigned to the backup retransmitted data, and the corresponding direct link is positioned to enable the data packet for its efficient retransmission to the destination node, thereby ensuring efficiency in fault tolerance. Additionally, we have improved the receiving port of the wireless router, added the corresponding redundant buffers and mux, and dynamically selected the retransmitted non-faulty data packets to be written to the local router in order to avoid the disorderly retransmission of the data packets. The evaluation results of this paper demonstrate that compared with the methods which are under different fault conditions, this fault-tolerant method drastically improves the data throughput rate, reduces the delay, effectively guarantees the reliability of the network, and improves the system performance.

Key words: wireless network-on-chip; retransmission; fault-tolerant; reliability

1 Introduction

With the continuous development of integrated circuit technology, there are a wide range of problems, such as poor scalability and low communication efficiency faced by the traditional system-on-chip system, which is based on the bus architecture[1].These drawbacks have been greatly overcome with the advent of Network-on-Chip (NoC). NoC is a new type of communication architecture, which is characterized by high bandwidth and scalability. However, the communication delay and network performance between multi-hop nodes are greatly constrained with the increase in the scale of the NoC network. Based on the abovementioned limitations, the researchers had proposed a new solution, Wireless Network-on-Chip (WiNoC)[2–5].

As a new type of interconnect technology, WiNoC can effectively solve the problem of high latency among multi-hop nodes in a network. While describing an On-Off Keying (OOK) wireless transceiver architecture, Deb et al.[6] have confirmed that zig-zag antennas and Wireless Interfaces (WIs) can be integrated on the chip, by relying on the current microelectronics technology, thereby providing a realistic basis for WiNoC research. Ganguly et al.[7] proposed a WiNoC interconnection scheme, in which a small-world communication architecture was designed, and the entire network was divided into several sub-networks. The Wireless Routers (WRs) are used for the communication of data among sub-networks to reduce the latency among multi-hop nodes. The authors in Ref. [8] proposed a multi-hop based 2-level hybrid mesh and a competition avoiding the routing algorithm. The authors in Ref. [9] designed a WiNoC architecture, mWNoC based on the millimeter wave (mm-wave) technology to improve

• Yiming Ouyang, Zhe Li, and Jianhua Li are with School of Computer and Information, Hefei University of Technology, Hefei 230009, China. E-mail: oyymbox@163.com; 2272229156@qq.com; jhli@heut.edu.cn.
• Qi Wang and Huaguo Liang are with School of Electronic Science and Applied Physics, Hefei University of Technology, Hefei 230009, China. E-mail: keywenchester@outlook.com; huagulg@hfut.edu.cn.
* To whom correspondence should be addressed.
Manuscript received: 2019-06-26; accepted: 2019-08-27

© The author(s) 2021. The articles published in this open access journal are distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/).

the network performance through efficient routing algorithms and with the optimal placement of wireless nodes.

However, the components in the router are prone to permanent or transient failure[10], which will affect the correct transmission of data and reduce the network performance due to manufacturing defects, circuit aging, process instability, electron migration, and other reasons. This means that it is necessary to consider the potential hardware failures in the design of WiNoC so as to improve the reliability of the system. The fundamental guarantee of system performance is to ensure that the routers (basic and wireless) are working properly by building data transmission links for data storage and forwarding.

We propose a WiNoC based on the efficient retransmission of data with the following contributions:

(1) Based on the real-time feedback of fault information, a network topology is proposed in which the fault information can be fed back to the source node via the fault signal line in real time.

(2) A data retransmission straight-through link is used to set the highest transmission priority of the backup retransmission data packet, which will effectively guarantee the efficiency of fault tolerance.

(3) We added redundant data buffers and mux to improve the architecture of the WR receiver, which will dynamically select the fault-free packets for transmission and prevent the retransmission of data packets from being damaged.

2 Related Work

Permanent and transient failures of the on-chip multi-core interconnect architecture are unavoidable. The authors in Ref. [11] exploited the fact that microbial communities have high fault-tolerance capabilities and proposed a robust fault-tolerant NoC based on the inherent robustness of complex network connections. However, this mechanism cannot cope with more complicated routing situations. The authors in Ref. [12] proposed a WiNoC Error Control Code (ECC), which used crossed Hamming codes in the wireless links and in joint cross-talk avoidance triple-error correction and quadruple-error detection methods in wire links. However, this kind of wireless detection and error correction method needs to wait for the complete transmission of all the data, and therefore, has a high delay and area overhead. A fault-tolerant WiNoC architecture was designed in the research work of Ref. [13], and innovative routing strategies and efficient fault-tolerant communication protocols were proposed, greatly increasing the robustness of the network. The authors in Ref. [11] designed a hybrid fault-tolerant hierarchical architecture by using a node-and-link-disjoint path communication structure between two points. The use and minimization of redundant wireless nodes improve network performance and have an acceptable area overhead. A traditional NoC structure error detection strategy between nodes was proposed in Ref. [14], which added an additional detection module between each pair of nodes and could notify the source node to retransmit data on the discovery of a fault. However, the area overhead of this strategy was particularly high, and a large number of retransmitted packets were generated in the case of multiple faults, greatly affecting the network throughput rate. The authors in Ref. [15] introduced SurfNoC: an on-chip network that significantly reduces the latency incurred by temporal partitioning. By scheduling the network into waves that flow across the interconnect, data from different domains carried by these waves are strictly non-interfering while avoiding the significant overheads associated with cycle-by-cycle time multiplexing. This strategy reduces the number of retransmitted packets, but also brings a lot of area overhead.

The objective of this paper is to propose a WiNoC based on the efficient data retransmission, which feeds back the fault information to the source node in real-time through the fault signal line. Additionally, a direct link for data retransmission based on the Manhattan path is designed to quickly release and retransmit the error information. We have also improved the wireless receiver architecture to effectively avoid problems, such as out-of-order packets.

3 WiNoC Architecture

3.1 Topology

This paper constructs a recursive topology based on the real-time feedback of fault information. As shown in Fig. 1, the topology is divided into four subnets and each subnet consists of 5×5 routers. The basic routers communicate over the wired interconnect network, and the WRs communicate over the wireless link by using the WI. Among them, the location of the WR can be obtained by the simulated annealing algorithm[16]. By considering the placement of the WR, its placement at

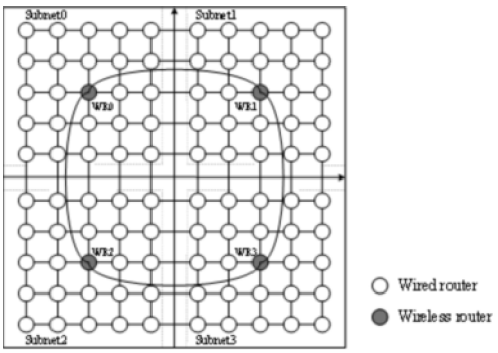


Fig. 1 Network topology of WiNoC.

the center of the subnet had no effect on the experimental results in this paper. In this paper, the description of the center origin coordinates is used in order to make the routing algorithm efficient and universal. The form of the coordinate description method is (x, y), where x and y represent the abscissa and ordinate of the router, respectively. Taking Fig. 1 as an example, the WR coordinates are $(-3, 3)$, $(3, 3)$, $(-3, 3)$, and $(3, -3)$. Recursion can drastically increase the efficiency of on-chip integration and routing algorithms.

Additionally, fault information feedback lines are set between the WRs in this paper. The destination router responds to the received data packet in the real time by sending back a fault feedback signal to the source router when it decodes and finds a transmission error. For example, if the source router WR1 sends a data packet to WR2 and WR2 detects the packet transmission error through the corresponding decoding mechanism, then WR2 sends the fault information back to WR1 in a fixed format through the fault information feedback line. The specific mechanism will be elaborated below.

3.2 Wireless router architecture

The conventional on-chip network WR consists of the basic architecture of the wired router and the WI. The specific architecture is shown in Fig. 2. It consists of six modules[17, 18]: input buffer, Route Calculation (RC), Virtual channel Allocation (VA), Switch Arbitration (SA), crossbar, and WI. Here, each input buffer contains multiple Virtual Channels (VCs). The WI is composed of a wireless transmitter (TX) and a wireless receiver (RX). The TX includes Transmitter Data Buffer (TDB), serializer, modulator, and Power Amplifier (PA). The TDB is used to store the data transmitted via the WI. The serializer is used to serialize the data in TDB. The modulator can modulate the low-frequency signal into a high-frequency signal that can be easily transmitted by the carrier. The PA can amplify the wireless signal transmission power to achieve the demand of on-chip wireless transmission. Similarly, RX consists of Low Noise Amplifier (LNA), demodulator, deserializer, and Receive Data Buffer (RDB). Wireless data transmitting and receiving tasks are performed by the on-chip antennas installed between wireless nodes. We use a sawtooth on-chip antenna, which is characterized by long transmission distance, high gain, and strong anti-interference[6].

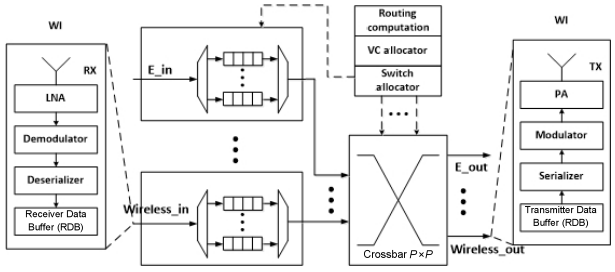


Fig. 2 Architecture of wireless router.

3.3 Problem analysis

While focusing on WiNoC data communications, the reliability of the wireless channel is considered to be critical since it affects the overall performance of the network. To resolve this issue, the common fault-tolerance scheme is divided into data redundancy and path redundancy[19]. Data redundancy refers to sending multiple copies of data along the same path or adding check bits and other different encoding schemes. However, this solution requires high error detection and correction capabilities along with a large area overhead. The path redundancy method implements the fault tolerance of faults on the on-chip network through redundant transmission links or it implements the efficient transmission of WiNoC through corresponding routing algorithms. As valuable communication resources, wireless links are prone to congestion in WiNoC and bear the task of sending and receiving wireless data. There are numerous wired links, and the communication pressure above them is far less than that on the wireless links. Therefore, we make an efficient use of the wired links and transmit the retransmitted data packets over the wired links, which greatly improve the utilization of network resources.

On the traditional on-chip router, the received packet is detected by the error detection unit in the destination router. If the data transmission error is found, the data will be retransmitted through the original path by the source router. As shown in Fig. 3, the First-In-First-Out

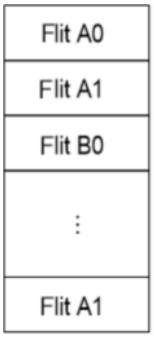


Fig. 3 Traditional retransmission write state.

(FIFO) structure is used in the data-receiving buffer of the router, and the source router is notified to retransmit the Flit A1 if it has detected an error during transmission. At the same time, the Flit B0 and other flits in the source router have also been transmitted. When Flit A1 is retransmitted to the destination router, Flit B0 has already been written to the buffer. At this point, the problem of discontinuity of data occurs due to the discontinuities present in the placements of Flits A0 and A1. At the same time, the communication efficiency is reduced, and performances, such as delay and throughput rate, are greatly affected, since the same link is used to transmit the normal and retransmission flits. This paper presents a fault-tolerant design for the efficient retransmission of data. Multipath transmission is used for different types of data in order to improve the transmission efficiency while ensuring the reliability of network.

4 Design of Efficient Data Retransmission in WiNoC

4.1 Design of retransmission architecture

Figure 4 shows the retransmission architecture between source and destination nodes. As shown in Fig. 4, the Retransmission Buffer (RB) is set in the source WR in this paper. Waiting for the data to be sent after the routing stages, such as VA and SA, the local data in the router are written to the TDB. The wireless data are simultaneously written into the RB to back up the redundant data. On the basis of the corresponding wireless transmission protocol, the wireless data in the source WR are sent to the destination WR through the antenna. The ECC unit of the destination WR decodes and detects the received wireless data packet, and if a data transmission error is detected, it will retransmit the data retransmission signal to the Retransmission Control Logic (RCL) of source WR via the error signal line. The RCL retransmits the backup information of the RB by controlling the corresponding power switch.

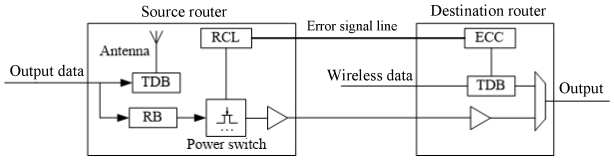


Fig. 4 Retransmission architecture.

Serving as a link for faulty transmission of information between the source WR and the destination WR, the faulty transmission line is essential for improving the system reliability. There are four WRs set in the 10×10 topology shown in Fig. 1. The signal format consists of 3 bits. The first bit is "1", which is used to represent the retransmission signal; the last 2 bits are router feature numbers. The feature numbers for WR0, WR1, WR2, and WR3 are 00, 01, 10, and 11, respectively. If the destination WR returns a 3-bit signal to the source WR, then the corresponding source WR needs to perform the retransmission of data. Based on this observation, this paper designs a fault signal transmission protocol. As shown in Fig. 5, the protocol mainly consists of a Clock Line (CL) and a Signal Line (SL). CL and SL stay high at the same time in idle. This is flagged as the start signal of the transmission when CL holds the high voltage and SL changes from the high voltage to the low voltage or flagged as the termination signal of the transmission when CL holds the high voltage and SL changes from the low voltage to the high voltage. When the clock signal is high, the data on the data line must remain stable, when it is in the transmission state. Only when the voltage on CL is low, it can change the state of the data line, otherwise, the signal transmission fails. As shown in State 0 in Fig. 5, data start to be transmitted after the appearance of the start signal. This state indicates that WR3 will perform the retransmission of data during the transmission of Signal 111.

As shown in Fig. 4, the RCL analyzes and processes the fault signal from the destination WR and turns the RB on or off via the power switch. The RCL is mainly composed of a sequence detector and a corresponding output circuit. The sequence detector in RCL is different for different WRs. Taking WR3 as an example, the

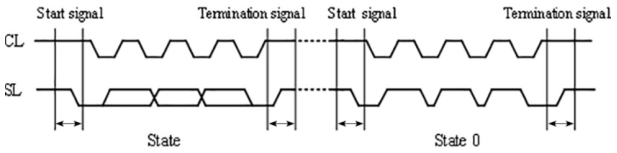


Fig. 5 Retransmission signal transmission timing diagram.

RCL structure is shown in Fig. 6, and the sequence detector consists of two JK triggers. The sequence detector outputs a high voltage when the fault signal from the ECC is a 3-bit continuous high voltage, i.e., 111. The output circuit consists of corresponding output buffers, which can greatly reduce the cross-talk caused by uncertainties and effectively ensure the stability of the output waveform.

Taking WR3 as an example, Verilog is used to describe the sequence detector. The program is shown in Algorithm 1.

4.2 Design of direct link

Data transmission delays and out-of-order phenomenon can be caused by retransmission mechanisms in the traditional on-chip routers. We propose a retransmission data pass-through link based on bypass transmission and design the corresponding data packet format. As shown in Fig. 7, the packet consists of a head flit, body flits, and a tail flit. Head flit, a total of 16 bytes, records the packet routing information. Among them, the first bit is flag bit, which is used to record whether the data packet is a retransmission data packet. If the bit is 1, then the data

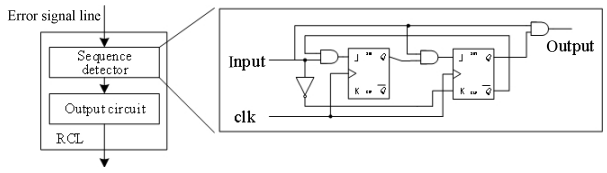


Fig. 6 Design of RCL in WR3.

Algorithm 1 Sequence detector designed by verilog

```
module Detector111 (input a, clk, reset, output w);
  parameter [1:0] s0 = 2'b00,
                  s1 = 2'b01,
                  s2 = 2'b10;
  reg [1:0] current;
  always @ (posedge clk)
    begin
      if(reset)
        current = s0;
      Else
        Case (current)
        s0: if(a) current ≤ s1;
            current≤ s0;
        s1: if(a) current ≤ s2;
            current ≤ s0;
        s2: if(a) current ≤ s2;
            current ≤ s0;
        end case
    end
    assign w = (current == s2) ? 1 : 0;
end module
```

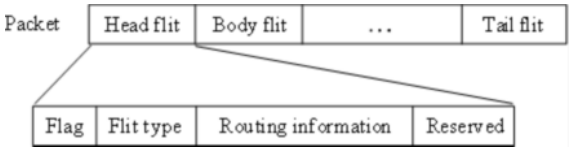


Fig. 7 Design of packet format.

packet is a retransmission data packet and if the bit is 0, then it is an ordinary data packet. The flit type occupies 1 bit and is used to represent the packet type. If the bit is 1, then the packet is a wireless data packet; and if the bit is 0, then it is a wired data packet. Routing information occupies 8 bits, containing VC number, destination address, and other packet information, and the remaining 6 bits are reserved bits.

The RCL retransmits data in the fault flit by controlling the RB. Taking the topology shown in Fig. 1 as an example: When WR2 transmits a packet to WR3, ECC feeds back the retransmission signal to RCL in WR2 through the error SL and RCL performs the corresponding analysis to determine whether the data need to be retransmitted due to incorrect transmission. In this paper, the Manhattan path between source and destination nodes is set as the data packet retransmission path so as to reduce the number of retransmitted hops and the transmission delay. Based on the topology presented in Fig. 1, the Manhattan path between WRs routes passes through several wired routers. Taking the Manhattan router between WR2 and WR3 as an example, the structure is shown in Fig. 8. We added Head Flit Recognition (HFR) and B_mux in Manhattan routers compared to non-Manhattan wired routers. The specific working mechanism of the Manhattan router is divided into two steps.

Step 1: The RB in WR2 writes the flag bit of the backup packet as 1 and transmits it. When the

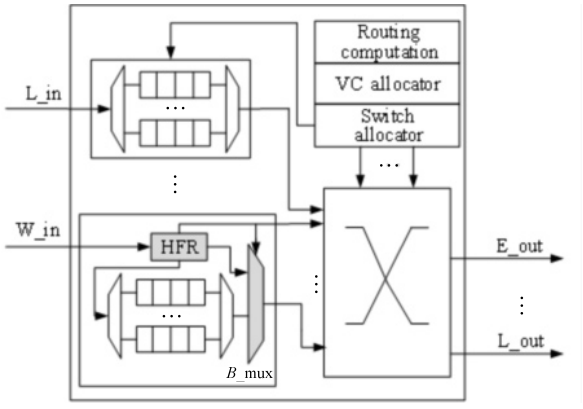


Fig. 8 Design of router architecture on Manhattan path.

retransmitted packet, using the Manhattan path as the shortest path, passes through the Manhattan router, the HFR in the router identifies and detects the flag in flit. When the bit is identified as 1, it is an indication that the flit is retransmission data. At this point, HFR directly forwards the data to B_mux and generates a corresponding enable signal to B_mux. Then B_mux preferentially transmits the retransmission packet.

Step 2: The crossbar switch receives the signal enabled by HFR. Retransmitted packets enter through the west port of the crossbar and are directly generated from the west port without going through the SA stage. Normal packet transmission in the port is resumed after the completion of retransmission of the data packet. The next level of router transmission is similar.

Ports between HFR and B_mux among different WRs are different because the Manhattan path is not the same, but the overall architecture of the Manhattan router is similar. It can be seen from the above that the retransmitted packets have been assigned the highest priority during the entire transmission process, which ensures the fault-tolerance efficiency of the system.

Figure 9 is a schematic diagram of the retransmission of data between WR2 and WR3. The red path represents the path of retransmission data and BR represents the Baseline Router. The retransmitted data in this paper, once identified as retransmission packets via HFR, can be transmitted over the direct link without RC, VA, and SA, etc. Retransmission data have been assigned the highest priority of data transmission, which greatly ensures the real-time reliability of the communication.

4.3 Architecture of wireless receiver

We have improved the RX of WRs by considering that the retransmission mechanism of traditional on-chip routers will cause data to go out of order. Figure 10 shows the architecture of the receiving end of the WR in this paper. Ordinary wireless data are written to RDB after LNA, demodulator, and deserializer. Assume that the depth of RDB is n bits. The received wireless data are first written to bit 1, because RDB is an FIFO structure. At this point, the ECC unit decodes and detects the bit. If the data packet is correct, then no signal feedback

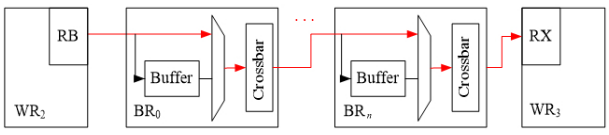


Fig. 9 Simplified schematic diagram of data retransmission links.

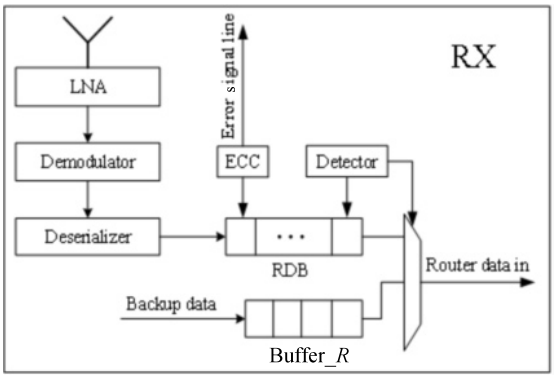


Fig. 10 Receiver architecture design of destination wireless router.

will be made. If the data packet is incorrect, then it is run as described in Section 3.1. Fault information is transmitted to the source router through the faulty transmission line. The source router retransmits the backup data through the Manhattan path until it is written into Buffer_R. At the same time, the fault data are moved to the next RDB. The fault data will be shifted into the n-th data as more wireless data are written. The n-th data are continuously detected by detector. If the data are detected as fault data, the mux is enabled, and the data of the Buffer_R that have completed retransmission data writing are transmitted. At this point, the correct retransmission data will be written to the local router and the failed packet will be discarded. The mux blocks the error data in the RDB, waits for the retransmitted data to arrive in the Buffer_R, strobes the Buffer_R, and discards the error data. This scheme effectively performs and measures the fault tolerance of wireless data and no out-of-order data phenomenon occurs.

The transmission diagram is presented in Fig. 11, where the RDB depth is 8 bits. The ECC decodes the first data written in TDB to check at the T_0 stage. During the T_1–T_6 period, as RX continues to receive new wireless data, the data are continuously shifted and written until the T_7 period, the data are transmitted to the last bit of the TDB, and the decoder detects the data. The retransmission data are written in Buffer_R to the local router by strobe mux if it is to be fault data. The correct data for retransmission are written to the local router during the T_8 period.

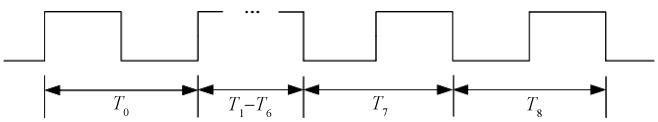


Fig. 11 Data retransmission timing diagram.

5 Experiment

Our experiments were conducted from three aspects: network performance, area, and power consumption. The experimental comparison objects are Schemes 1 and 2. Scheme 1 is WiNoC based on error control coding mechanism proposed in Ref. [12]. Scheme 2 is a fault-tolerant network-on-chip based on complex networks proposed in Ref. [11].

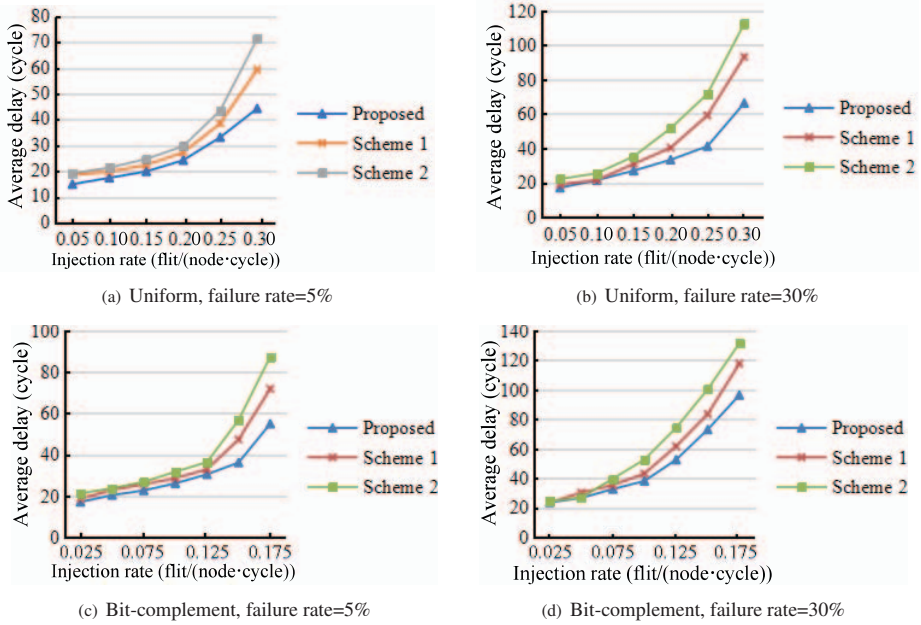
Table 1 Setting of experiment parameters.

Parameter	Setting
Network size	100
Wireless technology	mm-wave
Algorithm	XY routing algorithm
Flit size	32
Clock frequency (GHz)	1
Antenna frequency (GHz)	32
Buffer depth (flit)	20

5.1 Network performance

We simulate network performance based on Noxim[20] which is a clock-level simulation tool. We make analysis of the network's latency and throughput under a 10×10 topology. We assume that there will be no error when the packet enters the router from the local input port. The fault distribution is only for the data transmission link between the routers in the on-chip network and not for the internal transmission link of the router. The traffic mode adopts uniform and bit complement, and the sampling period is 1000 cycles. The experimental parameters, such as network scale and flow model, are the same to ensure the fairness and authenticity of the experimental results. The basic parameters of the experiment are detailed in Table 1.

Figure 12 compares the average delay at different failure rates in the uniform and bit-complement mode. Figures 12a and 12b compare the average delay of the uniform mode at the failure rates of 5% and 30%, respectively. Under the low failure rate of 5%, when the injection rate is relatively low, the advantages of the proposed scheme are not apparent compared with the comparison object. The shortcomings of the low fault-tolerance capacity of Schemes 1 and 2 and the advantages of the solution of this paper compared with the comparison scheme are gradually highlighted with the increase in injection rate. This is because the ECC of Scheme 1 needs to wait for the entire data packet to be transmitted, further causing a high communication delay. This complex situation in the actual network is not considered by Scheme 2 and there is an increase in delay with the busy network traffic. When the injection rate is relatively low, the delay of the scheme of this paper is slightly smaller than that of the comparison object, under the high failure rate of 30%. The delay of the solution in this paper is obviously lower than that of Schemes 1 and 2 with the increase of the injection rate. It can be found that the data efficient retransmission scheme adopted in this paper has a better fault tolerance. When the injection rate is 0.20 flit/(node · cycle), the proposed scheme reduces the average delay by 11.10% compared with Scheme 1 at the failure rate of 5% and decreases



(a) Uniform, failure rate=5%

(b) Uniform, failure rate=30%

(c) Bit-complement, failure rate=5%

(d) Bit-complement, failure rate=30%

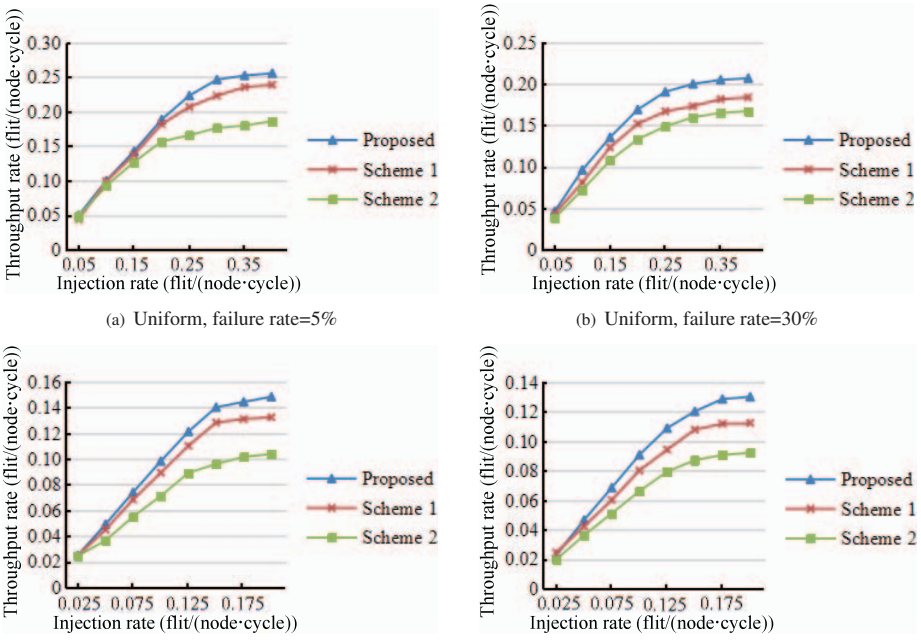
Fig. 12 Comparison of average delays of different schemes.

by 18.26% compared with Scheme 2. The scheme proposed in this paper reduces average delay by 17.36% compared with Scheme 1, and decreases average delay by 35.43% compared with Scheme 2, at a failure rate of 30%. Figures 12c and 12d are comparison of average delays at the failure rates of 5% and 30%, respectively, in the bit-complement mode. When the injection rate is 0.125 flit/(node·cycle), the proposed scheme reduces the average delay by 7.44% compared with Scheme 1 and decreases by 16.24% compared with Scheme 2 at the failure rate of 5%. Compared with Scheme 1, the average delay is reduced by 14.79%. Compared with Scheme 2, the average delay is reduced by 29.15% at a failure rate of 30%.

Figure 13 compares the throughput rates at different failure rates in uniform and bit-complement modes. It can be observed that the throughput rate of the scheme presented in this paper is observed to be better than that of the other schemes with the increase of the injection rate. The error control coding mechanism proposed in this paper is more tolerant to error than that proposed in Scheme 1. The retransmission data do not occupy the wireless link, which reduces the transmission pressure of the wireless link to some extent. Therefore, the throughput rate is the highest in this paper. The routing mechanism of Scheme 2 is relatively simple. Due to the limited resources of WRs, it is easy to cause congestion due to the inflow of a large number of packets. The throughput rate is significantly lower than what we proposed, especially when the failure rate is high. Figures 13a and 13b are comparison of throughput rates at the failure rates of 5% and 30%, respectively, in the uniform mode. When the injection rate is 0.20 flit/(node·cycle), the throughput rate of the proposed scheme is 3.74% higher than that of Scheme 1, and 20.34% higher than that of Scheme 2 at the failure rate of 5%. Our proposed scheme has an 11.32% increase in throughput rate over Scheme 1, and a 26.95% increase in throughput rate over Scheme 2 at the failure rate of 30%. Figures 13c and 13d are comparison of throughput rates at the failure rates of 5% and 30%, respectively, in the bit-complement mode. When the injection rate is 0.125 flit/(node·cycle), the throughput rate of the proposed scheme is 9.98% higher than that of Scheme 1 at the failure rate of 5%, and 35.97% higher than that of Scheme 2. Compared with Scheme 1, the throughput rate of our solution increases by 15.48%, and the throughput rate of Scheme 2 increases by 37.15% under the failure rate of 30%.

5.2 Area and power consumption

The experimental tool used in this paper is Synopsys' Design Compiler. The technology library used is the 45 nm standard cell library. The proposed area and power consumption of the router and the basic wired router, the router designed by Scheme 1, and the router proposed in Scheme 2 designed in this paper are simulated, and the depth of the RB in the comparison router is set to be the



(a) Uniform, failure rate=5%

(b) Uniform, failure rate=30%

(c) Bit-complement, failure rate=5%

(d) Bit-complement, failure rate=30%

Fig. 13 Comparison of throughput rates of different schemes.

same. The experimental results are shown in Table 2.

The data show that compared with the reference effective router, the WR has increased the area overhead due to the addition of wireless transceivers and on-chip antennas. The WR of Scheme 1 has the largest area overhead. This is because it adds a corresponding codec unit that has a large area overhead. The WR of Scheme 2 has the smallest area overhead since it has not been modified. In this paper, WRs have added redundant buffers, emitter coupled logic, and other modules. The area cost of the WR in this paper is about 5.72% lower than that of Scheme 1, and about 8.20% higher than that of Scheme 2. For the overhead of power consumption, the scheme of this paper increases about 0.62% compared with Scheme 1, and increases about 12.05% compared with Scheme 2. This is due to the addition of a corresponding fault-tolerant module by the WR used in this paper on the basis of the basic router, thereby increasing the power consumption. Considering the advantages of the scheme in terms of delay and throughput rate, the increased area overhead and power consumption are acceptable in the event of a failure.

6 Conclusion

WiNoC has gradually become the prime focus of research with the development of integrated circuit technology. The wireless link serves as the backbone of the wireless data transmission in the WiNoC. When it fails, system performance is greatly reduced. Based on this reason, we designed a fault-tolerant mechanism for the efficient retransmission of data. The fault SL is set in the network, and the fault information can be fed back to the source node in real time. By setting up the direct link, the backup data packet in the source WR has been assigned the highest transmission priority and is preferentially transmitted to the destination WR, thereby effectively ensuring the efficiency of fault tolerance. Additionally, we have improved the WR receiver architecture. By dynamically selecting the backup of the non-failure data packet, the architecture avoids out-of-order retransmission of the data packet. The experimental results demonstrate that under a slight increase of power consumption overhead, compared with other WiNoC solutions, the proposed scheme has reduced delay and area overhead, significantly increased throughput rate, and effectively improved network performance.

Table 2 Setting of experiment parameters.

Router type	Area overhead (m^2)	Power consumption (mW)
Baseline wired router	49 820	21.25
Scheme 1	82 681	25.79
Scheme 2	72 041	23.16
Proposed scheme	77 951	25.95

Acknowledgment

This research was supported in part by the National Natural Science Foundation of China (NSFC) (Nos. 61874157, 61674048, and 61874156) and the Natural Science Foundation of Anhui Province, China (No. KJ2018A0783).

References

[1] W. J. Dally and B. Towles, Route packets, not wires: On-chip interconnection networks, in *Proc. 38^{th} Design Automation Conf.*, Las Vegas, NV, USA, 2001, pp. 684–689.

[2] Y. M. Ouyang, Y. D. Zhang, H. G. Liang, Z. F. Huang, and H. Chang, Design of fault-tolerant router for 3D NoC based on virtual channel fault granularity partition, (in Chinese), *Journal of Computer Research and Development*, vol. 51, no. 9, pp. 1993–2002, 2014.

[3] X. Y. Wang, D. Xiang, and Z. G. Yu, TM: A new topology for networks-on-chip, (in Chinese), *Chinese Journal of Computers*, vol. 37, no. 11, pp. 2327–2341, 2014.

[4] H. L. Zhu, Y. X. Peng, Y. M. Yin, and S. G. Chen, A NoC router with dynamically allocated virtual-output-queueing, (in Chinese), *Journal of Computer Research and Development*, vol. 49, no. 1, pp. 183–192, 2012.

[5] B. Z. Fu, Y. H. Han, H. W. Li, and X. W. Li, Building resilient NoC with a reconfigurable routing algorithm, (in Chinese), *Journal of Computer-Aided Design & Computer Graphics*, vol. 23, no. 3, pp. 448–455, 2011.

[6] S. Deb, K. Chang, A. Ganguly, X. M. Yu, C. Teuscher, P. Pande, D. Heo, and B. Belzer, Design of an efficient NoC architecture using millimeter-wave wireless links, in *13^{th} Int. Symp. Quality Electronic Design*, Santa Clara, CA, USA, 2012, pp. 165–172.

[7] A. Ganguly, K. Chang, S. Deb, P. P. Pande, B. Belzer, and C. Teuscher, Scalable hybrid wireless network-on-chip architectures for multicore systems, *IEEE Transactions on Computers*, vol. 60, no. 10, pp. 1485–1502, 2011.

[8] C. F. Wang, W. H. Hu, and N. Bagherzadeh, A wireless network-on-chip design for multicore platforms, in *2011 19^{th} Int. Euromicro Conf. Parallel, Distributed and Network-Based Processing*, Ayia Napa, Gyprus, 2011, pp. 409–416.

[9] S. Deb, K. Chang, X. M. Yu, S. P. Sah, M. Cosic, A. Ganguly, P. P. Pande, B. Belzer, and D. Heo, Design of an energy-efficient CMOS-compatible NoC architecture with millimeter-wave wireless interconnects, *IEEE Transactions on Computers*, vol. 62, no. 12, pp. 2382–2396, 2013.

[10] A. Dehghani and K. Jamshidi, A novel approach to optimize fault-tolerant hybrid wireless network-on-chip architectures,

ACM Journal on Emerging Technologies in Computing Systems, vol. 12, no. 4, pp. 1–37, 2016.

[11] A. Ganguly, P. Wettin, K. Chang, and P. Pande, Complex network inspired fault-tolerant NoC architectures with wireless links, in *Proc. 5th ACM/IEEE Int. Symp. Networks-On-Chip*, Pittsburgh, PA, USA, 2011, pp. 169–176.

[12] A. Ganguly, P. Pande, B. Belzer, and A. Nojeh, A unified error control coding scheme to enhance the reliability of a hybrid wireless network-on-chip, presented at IEEE Int. Symp. Defect and Fault Tolerance in VLSI and Nanotechnology Systems, Vancouver, Canada, 2011, pp. 277–285.

[13] A. Dehghani and K. Jamshidi, A fault-tolerant hierarchical hybrid mesh-based wireless network-on-chip architecture for multicore platforms, *The Journal of Supercomputing*, vol. 71, no. 8, pp. 3116–3148, 2015.

[14] A. Prodromou, A. Panteli, C. Nicopoulos, and Y. Sazeides, NoCAlert: An on-line and real-time fault detection mechanism for network-on-chip architectures, in *Proc. 45th Annu. IEEE/ACM Int. Symp. Microarchitecture*, Vancouver, Canada, 2012, pp. 60–71.

[15] H. M. G. Wassel, Y. Gao, J. K. Oberg, T. Huffmire, R. Kastner, F. T. Chong, and T. Sherwood, SurfNoC: A low latency and provably non-interfering approach to secure networks-on-chip, *ACM SIGARCH Computer Architecture News*, vol. 41, no. 3, pp. 583–594, 2013.

[16] B. Bahrami, M. A. J. Jamali, and S. Saeidi, Proposing an optimal structure for the architecture of wireless networks on chip, *Telecommunication Systems*, vol. 62, no. 1, pp. 199–214, 2016.

[17] Y. M. Ouyang, Y. J. Chen, H. G. Liang, M. X. Yi, and J. H. Li, Design of a low-overhead fault channel isolated fault-tolerant router, (in Chinese), *Acta Electronica Sinica*, vol. 42, no. 11, pp. 2142–2149, 2014.

[18] Y. M. Ouyang, J. F. Yang, K. Xing, Z. F. Huang, and H. G. Liang, An improved communication scheme for non-HOL-blocking wireless NoC, *Integration*, vol. 60, pp. 240–247, 2018.

[19] M. O. Agyeman, Q. T. Vien, A. Ahmadinia, A. Yakovlev, K. F. Tong, and T. Mak, A resilient 2-D waveguide communication fabric for hybrid wired-wireless NoC design, *IEEE Transactions on Parallel & Distributed Systems*, vol. 28, no. 2, pp. 359–373, 2017.

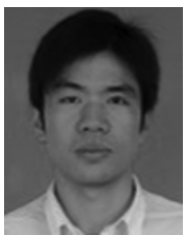
[20] V. Catania, A. Mineo, S. Monteleone, M. Palesi, and D. Patti, Noxim: An open, extensible and cycle-accurate network on chip simulator, in *Proc. 26th Int. Conf. Application-Specific Systems, Architectures and Processors*, Toronto, Canada, 2015, pp. 162–163.

Yiming Ouyang received the bachelor's, master's, and doctoral degrees from Hefei University of Technology in 1984, 1991, and 2013, respectively. He is currently a professor at Hefei University of Technology and a leading expert in research. His main research directions are NoC and on-Chip Systems (SoC), integration and testing of embedded systems, and automation of digital system design.

Qi Wang received the bachelor degree from Jilin University in 2016. During 2014–2015, he was an exchange student in computer and system science at Stockholm University. He received the bachelor's degree from Stevens Institute of Technology in 2018. Now he is a PhD student at Hefei University of Technology. His main research fields include network on chip, machine learning, and fault tolerance.

Zhe Li received the bachelor's and master's degrees from Hefei University of Technology in 2015 and 2018, respectively. His main research fields include NoC architecture and embedded system design.

Huaguo Liang received the bachelor's and master's degrees from Hefei University of Technology in 1982 and 1989, respectively. And he received the PhD degree from the University of Stuttgart in 2003. Now he is a professor and doctoral supervisor at School of Electronic Science and Applied Physics, Hefei University of Technology. His main research fields include SoC and VLSI design & test, etc.



Jianhua Li received the BS degree from Anqing Normal University, Anhui, China in 2007 and the PhD degree from University of Science & Technology of China in 2013. He is now a full-time lecturer at the School of Computer and Information, Hefei University of Technology, Anhui, China. His research interests include multi-core memory system, emerging nonvolatile memories, and on-chip networks.