



# A novel arbitration mechanism for crossbar switch in wireless network-on-chip

Farhad Rad<sup>1</sup> · Midia Reshadi<sup>1</sup> · Ahmad Khademzadeh<sup>2</sup>

Received: 25 December 2018 / Revised: 15 June 2020 / Accepted: 21 June 2020 / Published online: 12 September 2020  
© Springer Science+Business Media, LLC, part of Springer Nature 2020

## Abstract

Wireless network-on-chip (WiNoC) has been introduced as an efficient communication paradigm that can provide high bandwidth and low latency wireless links among long-distance cores. However, the increase of demand for using these shared wireless links and the presence of few numbers of these channels on a chip leads to port contention in WiNoCs. This problem increases the average packet latency as well as the network latency. Therefore, a fair arbitration mechanism is required to eliminate port contention in wireless routers (WRs) and improve the utility of the output port. In this study, we propose a new arbitration mechanism for crossbar switch that can fairly allocate the port priorities based on the current traffic load and the wireless channel bandwidth. The simulation results under synthetic traffic patterns show that the proposed arbitration scheme reduces the average latency and improves the network throughput compared to round robin-based WiNoC.

**Keywords** Network-on-chip · Arbitration mechanism · Congestion control · Wireless network-on-chip · Performance evaluation · Scheduling mechanism

## 1 Introduction

The Network-on-Chip (NoC) has sufficient capabilities to integrate high-density processing elements (PEs) on a single chip [1]. However, by increasing the network size, high latency and power consumption are the significant challenges in the NoCs. In recent years, several solutions such as 3D NoCs [2], photonic NoC [3], and wireless NoC (WiNoC) [4, 5] have been proposed to overcome these constraints. Table 1 shows the comparative analyses of these technologies. As can be seen, CMOS compatible wireless NoC is the best candidate between these alternative approaches that can change the multi-hop communication to the single-hop communication using low latency and high bandwidth wireless links.

In WiNoCs, each wireless router (WR) is equipped with a wireless interface (WI) to attain the unique features of wireless interconnects. However, with the increasing competition for using these shared wireless channels, congestion in WRs requires more attention. In general, congestion can reduce network performance. There are two types of congestion in the WiNoCs; link-level congestion and node-level congestion [6]. In wireless link-level congestion, when the wireless link is shared among several WRs, the competition to access these channels occurs. In this circumstance, interference reduces link efficiency and network performance. To solve this problem, a suitable medium access control (MAC) mechanism is needed to ensure that only one pair of WRs can share the wireless link to communicate with each other without any interference. In wireless node-level congestion, when the different input ports of the WR will compete for the shared output port, congestion can occur. Hence, the arbitration stage of crossbar switch in WRs becomes a critical stage for network performance improvement. So, some fair arbitration methods are required to eliminate congestion in WRs.

---

✉ Midia Reshadi  
reshadi@srbiau.ac.ir

<sup>1</sup> Department of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>2</sup> Education and International Scientific Cooperation Department, Iran Telecommunication Research Center, Tehran, Iran

**Table 1** Three emerging NoCs [7]

	3D NoCs	Photonic NoCs	WiNoCs
Design Requirements	Multiple layers for interconnections	Resonators (Silicon photonic components)	CNT/Graphene-based antennas
Performances	Low delay and power in vertical directions	Very high speed and power, independent from data-rate	Single-Hop capability, low latency, and high bandwidth
Reliability	Vertical links via failure	Temperature sensitivity of photonic components	Noisy wireless channel
Challenge	Heat dissipation	CMOS compatibility and Integration of photonic components	Design low area and power consumption transceivers

The Round Robin (RR) algorithm is an appropriate scheduling algorithm that provides good performance in the low traffic pattern. However, the network suffers from a poor performance on heavy traffic patterns, especially when RR is not considered the current state of the network to grant requests with different priorities. Therefore, an efficient scheduling algorithm is required in the congestion condition to allow the crowded input port to obtain the highest priority for packet transmission. In this paper, a novel arbitration mechanism is proposed for choosing an appropriate input port to access the output port. The main contributions of this paper include the following:

- We proposed a new arbitration strategy that can maintain good fairness among input requests. Our proposed scheme provides dynamic priority orders for acquiring the grant signal.
- The dynamic priority is assigned to the input ports of WR based on the network status. Three factors are considered for calculating the priority of each input port including the length of the current packet, the antenna buffer length, and the length of the last transmitted packet.
- The performance of the proposed method in various synthetic traffic patterns has been evaluated. Our simulation results show that the proposed scheme can significantly improve network performance.

The remainder of this paper is organized as follows: Sect. 2 describes a brief review of the related work. The proposed arbitration mechanism is presented in Sect. 3. Section 4 reports the performance evaluation of the proposed algorithm under various synthetic traffic patterns. Finally, Sect. 5 provides the conclusion of the paper.

## 2 Related work

In recent years, Wireless NoC architecture (WiNoC) has been suggested as an acceptable solution for multiprocessors' interconnection on the chip. Antennas and transceivers are two main blocks in WiNoCs that used for wireless

communications. Different types of antennas based on the portion of the electromagnetic spectrum have been proposed in WiNoCs. Some of them are millimeter-wave (mm-wave) antenna [8], carbon nanotube (CNT) antenna [9], and sub-THz antenna [10].

In addition to wireless antenna technologies that introduce briefly in the above mention, several hybrid wired-wireless topologies have been proposed for WiNoCs [11–15]. In two-dimensional hybrid mesh topologies, at first, the optimal number of WRs and their locations are determined, and then a wireless link is shared by a pair of WRs. However, due to the hardware complexity and power overhead of WRs, the number of WRs is limited. Also, the resources of WRs such as input buffers, the number of wireless channels, and communication bandwidth are finite. These specifications and inherent constraints will lead to traffic load exceed compare to the available resource capacity. Hence, under heavy traffic loads, the input buffers of WR are filled, and congestion occurs.

Congestion is one of the key issues in the WiNoCs that can affect all of network performance parameters. Therefore, different methods have been proposed to reduce congestion in WiNoCs. In [30], a comprehensive survey of the significant congestion control mechanisms in WiNoCs is present. The authors classified the available congestion control schemes into six categories, including hardware resources-based congestion control, congestion-aware routing algorithms, medium access control protocol, congestion-aware architectures, rate-based congestion control, and application-mapping with task-migration techniques.

The Frequency/Time Division Multiplexing (FDMA/TDMA) [7], Code Division Multiple Access (CDMA) [16], and Token-Passing (TP) protocols [17] are general mechanisms that used as medium access control strategies in WiNoCs. In [18], the authors designed and implemented a new synchronous and distributed protocol (SD-MAC) that using binary countdown-based arbitration policy within each zone to resolve the channel contention between wireless interfaces. In [19], a novel radio access control mechanism (RACM) has been presented. The authors at first, establish a ring-based topology among radio hubs to

minimize the hop count and alleviate the wiring congestion and routing algorithm problems compared to other topologies. Secondly, they proposed RACM as a token-based MAC approach. In the RACM, the unemployed clock cycles are redistributed among the radio-hubs, them that have used the radio channel completely in the previous timelines. According to the simulation results, the communication radio delay decreased by about 30%, and the energy is saved up to 25%. In [20], a broadcast, reliability, and sensing protocol (BRS-MAC) protocol has been proposed to provide the broadcast flow communications in WiNoCs. The BRS protocol is a contention-based protocol that can minimize the penalty of collisions in the wireless interfaces via three techniques: preamble transmission, collision detection, and scalable acknowledging.

On the other hand, using the deterministic routing algorithms lead to Head-of-Line (HoL) blocking in the input buffers of routers during high traffic workload [31–33]. So, some studies attempt to eliminate HoL, and congestion in WiNoC routers by proposing congestion-aware architectures [21] and congestion-aware routing algorithms [22].

Scheduling algorithms attempt to prevent node-level congestion using a fast and fair crossbar switch arbitration block. An efficient scheduling algorithm must be allocated a shared output port fairly between the input ports so that an input request is guaranteed in each arbitration cycle. Arbitration methods can be static or dynamic depending on the Quality of Service (QoS) requirements. In static priority-based arbitration, priorities are constant throughout the network lifetime and never changes at runtime. While in dynamic priority mechanisms, the priority order and priority assignment can be changed in each cycle. Based on static and dynamic priorities, an arbitration algorithm may be preemptive or non-preemptive. In recent years, many studies have been carried out on the arbitration mechanism in NoCs [23–27]. Most of the existing research works improve the performance of round-robin (RR) arbiter while other works focus on fairness enhancement. In general, the RR arbitrator's fairness is limited due to the static priority specified by the previous granting in an arbitration cycle. It can be observed that the RR arbiter cannot achieve the optimal results under heavy traffic patterns. Therefore, the design and implementation of the fairly arbitration methods for WiNoCs are required, especially when ports contention is occurring in WRs due to using the shared wireless channels. In this paper, we focus on the design and implementation of a new arbitration mechanism in WRs to reduce contention between input ports.

### 3 Wireless NoC architecture

In this paper, a hybrid WiNoC topology built on a  $10 \times 10$  two-dimensional mesh network. At first, the network is divided into four  $5 \times 5$  subnets, and in the center of each subnet, a baseline router is equipped with a wireless interface (WI). These wireless routers are responsible for transmitting packets between long-distance nodes using the wireless links. The WiNoC topology and the main components of a WR illustrated in Fig. 1. Each WR includes input ports with virtual channels (VCs), routing computing logic, virtual channel (VC) allocator, switch arbitration (SA), a crossbar switch, and a wireless interface. The baseline routers are VC compatible routers which using the buffer queues to store packets. In this work, the input port of the baseline routers contains two virtual channels to avoid head-of-line (HOL) blocking.

For routing in WiNoC, a deterministic routing algorithm such as XY routing can be used if the source and the destination nodes are in the same subnet or the adjacent subnet. Otherwise, the packet will be transmitted through wired and wireless links. In this paper, we used an adaptive routing algorithm schemes which have been proposed in [21]. Figure 2 represents the routing algorithm steps that commonly used for routing in mesh-based WiNoC. If the hop count ( $H$ ) between source and a destination node in a wireless path ( $H_{wl}$ ) is lower than the sum of hop count through the wired link ( $H_w$ ) and  $\lambda$ , which is  $H_{wl} < H_w + \lambda$ , the wireless connection is used to route packets to the destination node. The colored blocks in Fig. 2 show the routing steps in this case. Nevertheless, it is possible that in heavy traffic loads, a large number of packets to pass through WRs. In this case, if the wireless channel is occupied by the other WRs or token-based MAC protocols could not allocate the wireless channel due to the unavailability of the token, the data packets will be stored in the input ports of WRs. As a result of massive traffic emergence, WRs become the hot-spots point due to buffer occupancy and unfair distribution of the shared channels.

To effectively balance the utilization of the wired and wireless links interconnections, a suitable balance parameter called  $\lambda$  is employed to the routing decision. It can be observed, under different traffic patterns, the optimal value  $\lambda$  is different. We have shown these situations in Fig. 3. In the uniform traffic pattern, the maximum saturation throughput is at  $\lambda=3$ , while under other traffic patterns, the optimal value is 4. Therefore, in this paper based on the experiment results, we add the optimal value of 4 to the routing algorithm.

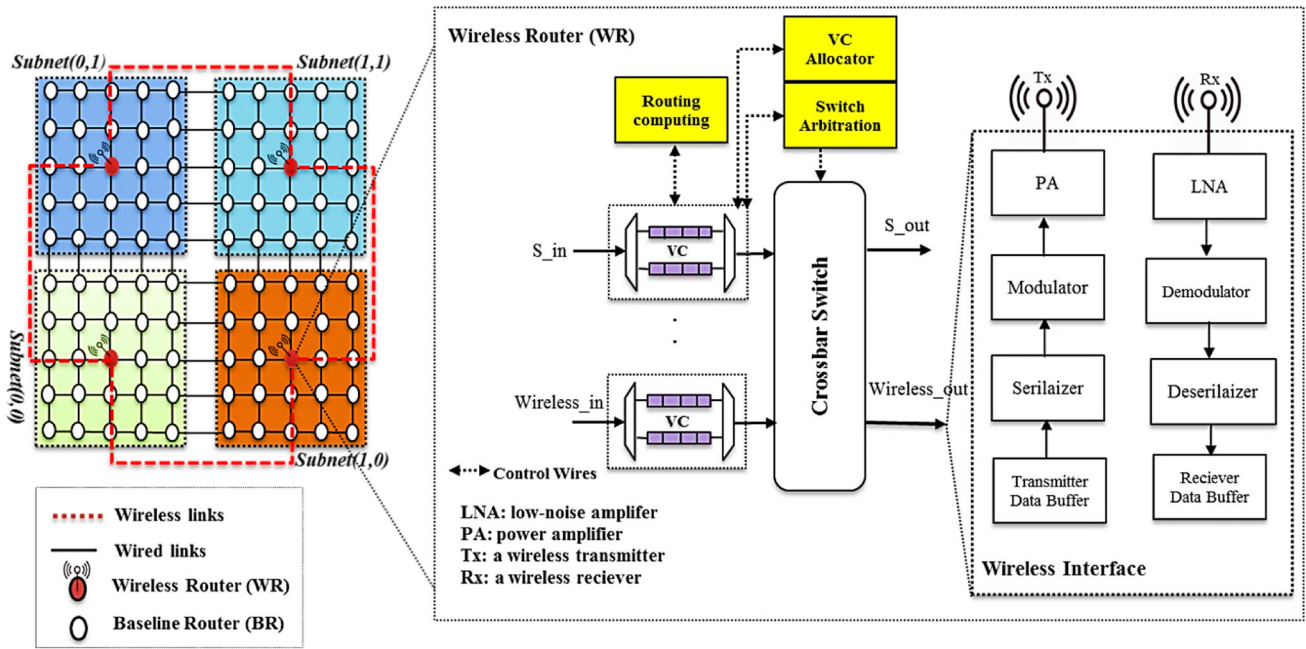


Fig. 1 Wireless NoC topology and WI structure

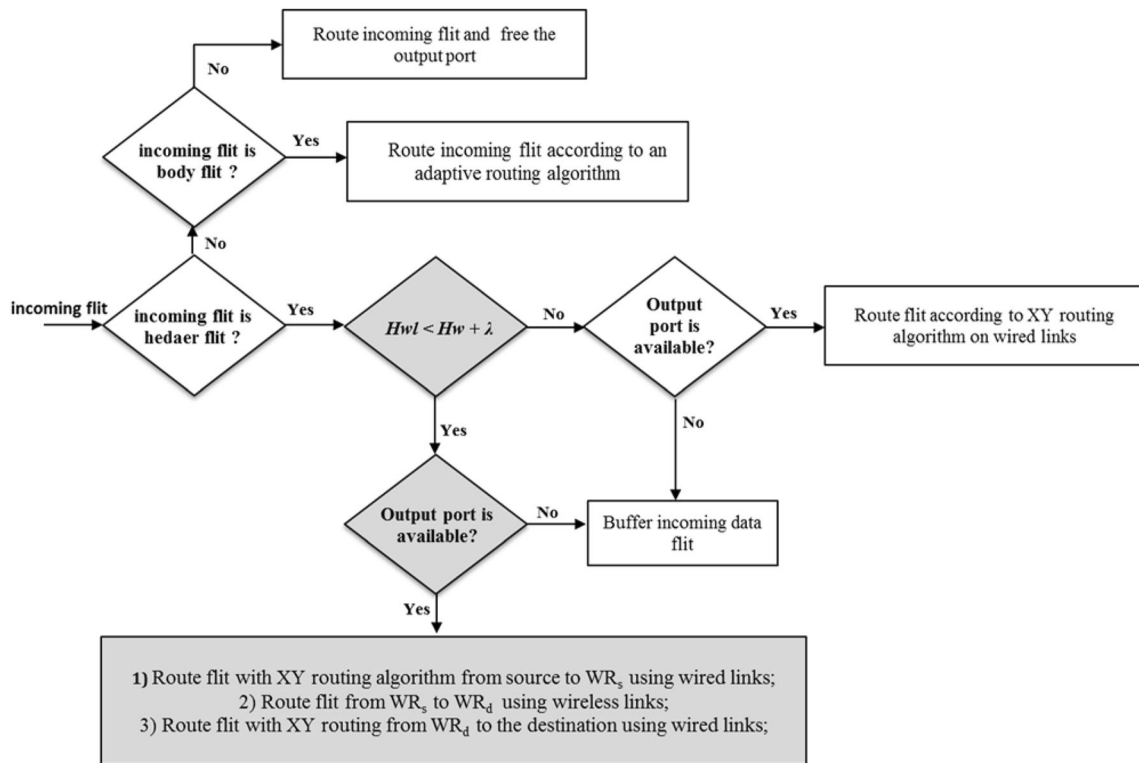


Fig. 2 Routing algorithm in WiNoC architecture

### 3.1 The proposed arbitration mechanism

In routers, the incoming packets that cannot be transmitted due to congestion or contending, store in first-in-first-out (FIFO) input buffers. Although increasing the buffer slots

of the input queues can improve the network performance on heavy traffic load, it causes more area and power consumption overhead. So, the arbitration mechanism plays an important role in crossbar switch to reduce congestion and enhance router performance. The round-robin (RR)

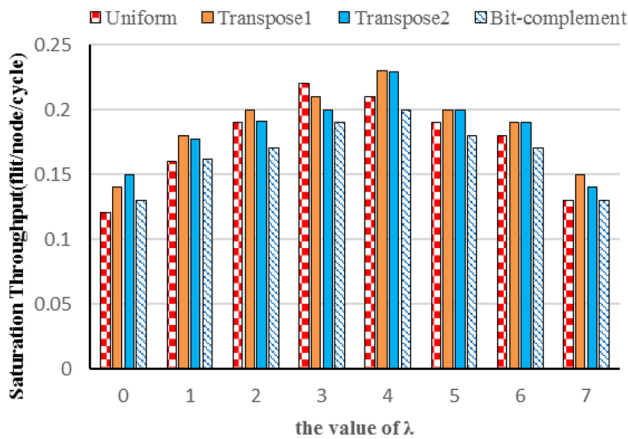


Fig. 3 Saturation throughput with different values of  $\lambda$  under traffic patterns

algorithm is a well-known arbitration mechanism that can guarantee fair scheduling. In the RR arbitration strategy, all input ports have an equal chance to own the output port. Although RR arbiter performs well under uniform traffic loads, it is not flexible for some customized applications and distribution traffic loads, especially when a prioritization policy is required. In this section, we propose a new arbitration mechanism, which can be used in the arbitration of crossbar switch in the wireless routers. Figure 4 illustrates the micro-architecture of the proposed mechanism in a wireless router.

According to our new arbitration structure, when wireless communication is required, the incoming packets are inserted into the FIFO queues of the input ports. Then, the header flits analyzed by the Header Analysis Unit (HAU), and the packet length field (*flit.sequence\_length field in the header flit*) is stored in the buffer status registers. In the next step, the Priority Calculation Unit (PCU) received this

information of buffer status registers and calculates the priority of input ports based on three factors consist of the length of the current packet, the length of the last transmitted packet which transmits through the wireless channel and the occupation length of antenna buffer (*Tx\_buffer\_occupancy*). The values of these parameters can be changed during the arbitration cycle. Finally, as shown in Fig. 4, the Decision Maker Unit (DMU) provides grant signals for crossbar switch to specifies which input ports will be allowed to obtain the output wireless channel.

### 3.2 The design of priority calculation unit

The PCU block generates the input ports priority based on the values that mention in the previous section as follows: *Packet\_size* is the packet length of the input port and equal with the number of available flits in the packet. If the wireless channel bandwidth is shown as *B* then the required time for transmitting the current packet through the wireless channel can be written as to Eq. (1).

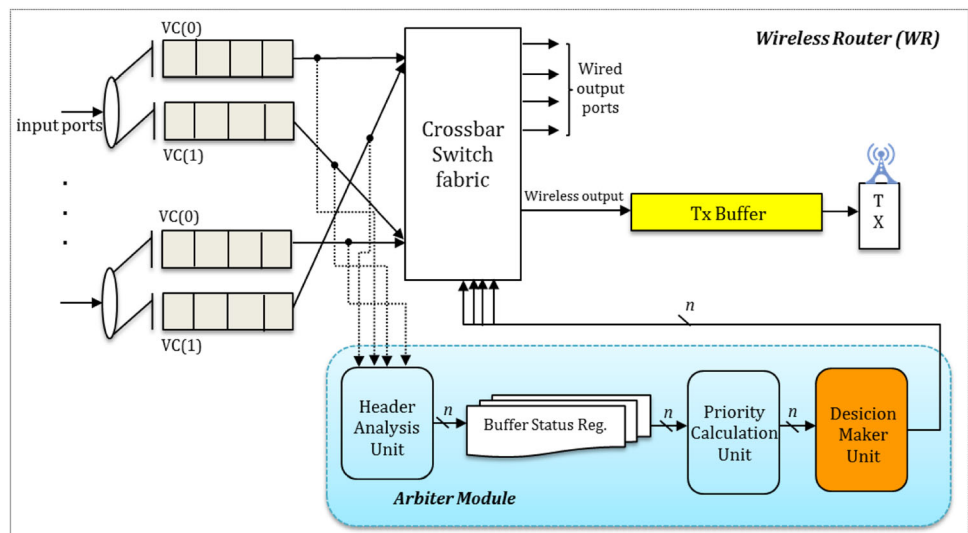
$$\Delta T = \frac{Packet\_size}{B} \tag{1}$$

The number of fits that have been stored in the antenna buffer (*Tx\_buffer\_occupancy*) is *numFlit* and can be changed at any time. So,  $\Delta T_{Tx}$  is the packet waiting time in the antenna buffer and can be defined based on Eq. (2).

$$\Delta T_{Tx} = \frac{numFlit}{B} \tag{2}$$

The length of the last data packet transmitted through the wireless channel is considered to be *LastPacket\_size*. Therefore, the transmission time of the last packet is calculated as:

Fig. 4 Schematic of the proposed arbitration





$$\Delta T_w = \frac{LastPacket_{size}}{B} \tag{3}$$

Based on Eqs. (2) and (3), the total packet waiting time in the corresponding input port for transmitting the data packet through the wireless channels is  $W_{time}$ . So, we can calculate the input port priority as Eq. (5).

$$W_{time} = (\Delta T_w + \Delta T_{Tx}) \tag{4}$$

$$priority = \frac{\Delta T + W_{time}}{\Delta T} \tag{5}$$

By substitution Eqs. (1) and (3) in Eq. (5), the input port priority in the corresponding input port can be rewritten as:

$$priority = \frac{Packet_{size} + nmuFlit + LastPacket_{size}}{Packet_{size}} \tag{6}$$

For generality, we do not consider the latency for acquiring a grant signal from the cross switch arbiter and the latency of the MAC mechanism for acquiring the wireless channel. Figure 5 shows the structure of the PCU block. This block is responsible to calculate the priority of input ports. These priorities are the fixed integer number during arbitration time in each WR. So, the input port with a higher priority has a higher integer number.

As can be seen, the ports priority calculation method based on Eq. (6) has the following advantages: 1) the priority of input ports can change dynamically by varying the packet length in each arbitration cycle. 2) although the proposed method assigns the higher priority to the shortest packets, this method can avoid the starvation problem. For example, when the transmission time of packets is the same, the input priorities values depend on the waiting time parameters. Therefore, as the input port waiting time increases, the input port priority becomes higher. Hence, low priority ports obtain the chance to transmit packets at the next time.

### 3.3 The design of decision maker unit

The DMU block is an arbiter module that receives the values of input port priorities from the PCU and

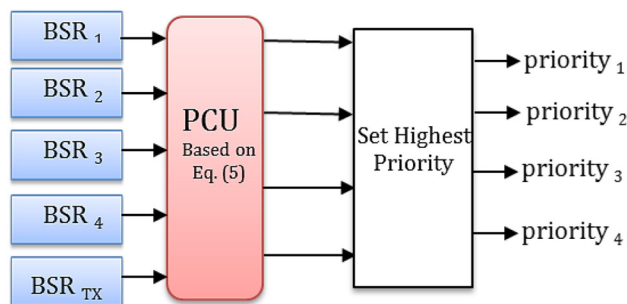


Fig. 5 The PCU scheme

dynamically adjusts the granted signal for an input port. The selected input port has a shorter packet transmission time compared to the other ports during the arbitration cycle. In this case, the other input ports will be favored in the next clock cycles. Figure 6 shows the hardware structure of the DMU block for 4-input signals that demand the output port. The following equations define the proposed module functionality.

$$\begin{aligned} Grant_i &= R_i \cdot (priority_i + C_{in_i}) \\ C_{out_i} &= \bar{R}_i \cdot (priority_i + C_{in_i}) \end{aligned} \tag{7}$$

In Eq. (7),  $R_i$  is the input port request,  $Grant_i$  indicates the input port granted signal,  $priority_i$  is the current input port priority, and  $C_{in_i}/C_{out_i}$  passes the priority among cells. According to Eq. (7), if an input port request has the highest priority ( $priority_i = 1$ ) then it obtained the guaranteed signal ( $Grant_i = 1$ ). The Pseudocode of the proposed arbitration algorithm is shown in Fig. 7.

Inline 3, the PCU block is called and the input port's priorities are calculated. In lines 8 to 11, if the current packet in the input port has not the highest priority, the input port is ignored. Otherwise, it obtains the output port. In lines 12 to 13, when a higher-priority input port arrives while the output channel is reserved for a lower-priority port, this channel be free and given to the higher-priority port. In each WR, the steps for selecting an input port that can to transmit the data packet through the wireless interface are briefly described:

when wireless communications are required, the data packets first enter to the FIFO buffer of input ports. The header flits of the data packets which located at the head of each non-empty input queues are passed to the head analysis unit. The HAU extracts the data packets length from the header flits and then sends them to PCU. The PCU determines the priority of input ports based on Eq. (6). When the priority signal of the current data packet set to the higher integer number, the DMU activates a grant signal ( $Grant$  line in Fig. 6). Hence, the selected input port is allowed to transmits the current data packet to the antenna buffer of the wireless interface ( $Tx\ buffer$ ) and the priority value decreases cyclicly from that point during the arbitration cycle.

To further clarify the functioning of the DMU arbiter, assume that in Fig. 6, the input port 0,1 and 3 ( $R_0, R_1,$  and  $R_3$ ) are input requests to obtained the wireless channel. The PCU be adjusted the priority value of input ports based on Eq. 6 with an integer number. We assume that a priority vector is defined as “2130” belongs to these values. Thus, the input port request  $R_1$  gets the highest priority and  $R_0$  receives the least priority. (the higher integer indicates a higher priority). Therefore, the DMU arbiter based on Eq. 7 selects the input port  $R_1$  in according to the higher

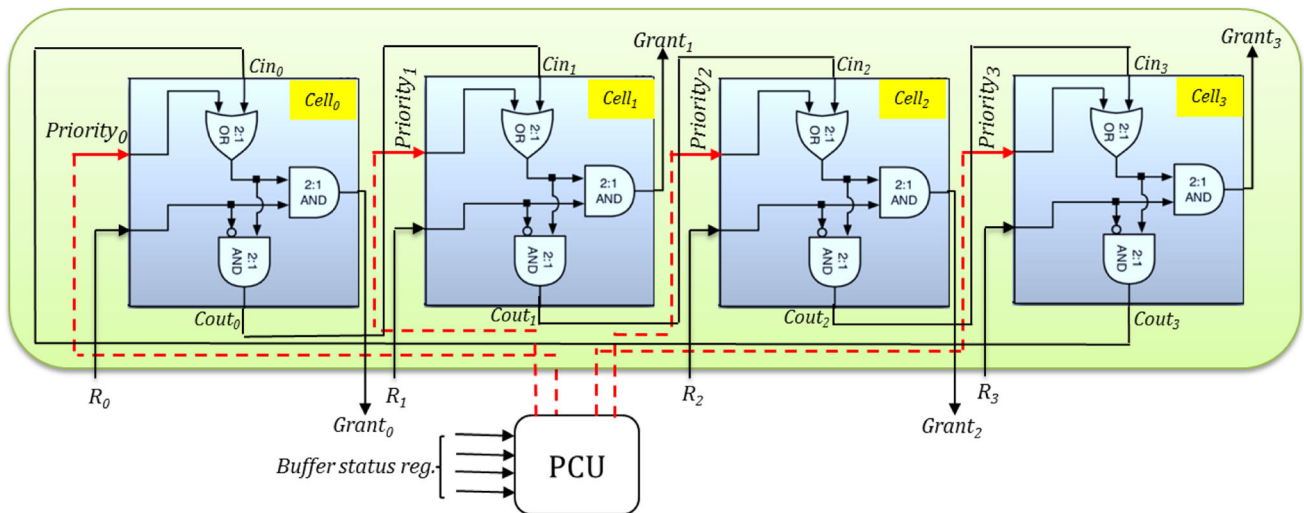


Fig. 6 The DMU architecture

```

input: inportsP
output: reservation_table()
*****
// Channels: txChannel and rxChannels;
Begin
1: for i = 1 to number of input ports do
2:   if (inportsP(i).HasAwaitingPacket()) Then
3:     inportsP(i).CurPackPri = Calculate.Priority(Packet) // Calculation of the priorities
4:   end if
5: end for
6: Sort inportsP(i).CurPackPri in descending order;
7: i=1;
8: While (!inportsP(i).CurPackPri.Empty()) do
9:   if (inportsP(i).HasAwaitingPacket()) Then
10:    if (reservation_table.isAvailable(channel)) Then
11:      reservation_table.reserve(i, channel); // Grant index is set
12:    else if (the channel is reached through the lower priority) Then
13:      Free the channel and make it available to the higher priority;
14:    end if
15:  end if
16: end if
17: i++;
18: end While
19: end function
    
```

Fig. 7 The proposed algorithm

priority, and grants the output port to  $R_1$  by setting the grant line ( $Grant_1$ ) to 1.

### 4 Experiments and simulation environment

In this section, we develop our WiNoC platform with a cycle-accurate simulator [28], and the results of our design compare with the baseline NoC and RR-based WiNoC. The simulation parameters are summarized in Table 2. The simulation environment uses mesh network topology with four and nine WRs in the center of any  $5 \times 5$  subnets. Each WR input port consists of a FIFO buffer with four flits capacity and each packet has between three to nine flits. The buffer size of the wireless antenna at each WR has 64

flits. Wormhole mechanism and an adaptive deadlock-free routing algorithm are used as the switching mechanism and routing function. The simulation runs for 10,000 cycles and with the first 1000 warmup cycles. To increase the accuracy of the experimental results, the simulations have been repeated ten times and the results were averaged.

#### 4.1 Traffic patterns

The proposed arbitration mechanism evaluated by using synthetic traffic patterns. In uniform or random traffic patterns, a source node sends flits to other nodes with the same probability while under the transpose traffic patterns, source nodes send flits to specific destination nodes. For a mesh  $n \times m$ , a node  $(i, j)$  only sends flits to a node in  $(n-1-j, m-1-i)$  position, where  $n$  is the number of columns and  $m$  is the number of rows in a mesh network. For the transpose patterns, we have applied two transpose traffic patterns, Transpose1 and Transpose2. For transpose1, we use the mention definition mechanism and for transpose2 we send the packet from the source node  $(i, j)$  to the destination node  $(j, i)$ . In the Bit-complement traffic pattern, source nodes send the data to one’s complement of its address.

#### 4.2 Evaluation metrics

One of the most important parameters to measure WiNoCs performance is the average latency. The latency of on-chip networks is due to buffering, routing calculations, switch arbitration, synchronization, and so on. The latency of each packet defined as different times between the injection of a header flits into the network and the time that a tail flit is received at the destination node. The average packet latency is defined as follows:

**Table 2** Simulation parameters

Parameter	Value
Mesh network size	10×10, 15×15
Number of radio-hubs	4,9
Switching technique	Wormhole
Packet injection rate	0.05
Traffic distribution	Uniform random, Bit-complement, Transpose1, Transpose2
Flit size [bit]	32
Radio Access Control Mechanism	Token-based
Wireless Bandwidth	16 [Gbps]
Packet length [flit]	3 to 9
Buffer depth [flit]	4
Buffer size of WI-antenna [flit]	64
Clock_Period_ps	1000
Reset_time cycles	1000
Simulation_time cycles	10,000
Stats_warm_up_time cycles	1000

$$Latency = \frac{Total\ received\ flits}{Total\ cycle \times Network\ size} \quad (8)$$

The network throughput refers to the number of successfully received packets (or total received flits) per total cycle time. Throughput in WiNoCs define as follows:

$$Throughput = \frac{Total\ received\ flits}{Total\ cycle \times Network\ size} \quad (9)$$

The power consumption is also a critical network performance metric that indicates the network energy consumption. The power consumed by WiNoC can be partitioned into two main contributions, namely, wired and wireless power consumption. Also, the wired/wireless power consumed can be classified into static and dynamic energy. Dynamic power depended on the density of the packet that to be moved without errors through the network and expressed in joules per bit. In contrast, static power is expressed in watts and depended on the essential wire/wireless elements in WiNoCs. In a WiNoC, the overall power consumption can compute as follows:

$$Power_{total}^{(WiNoC)} = Power_{total}^{(wired)} + Power_{total}^{(wireless)} \quad (10)$$

### 4.3 Experiment results

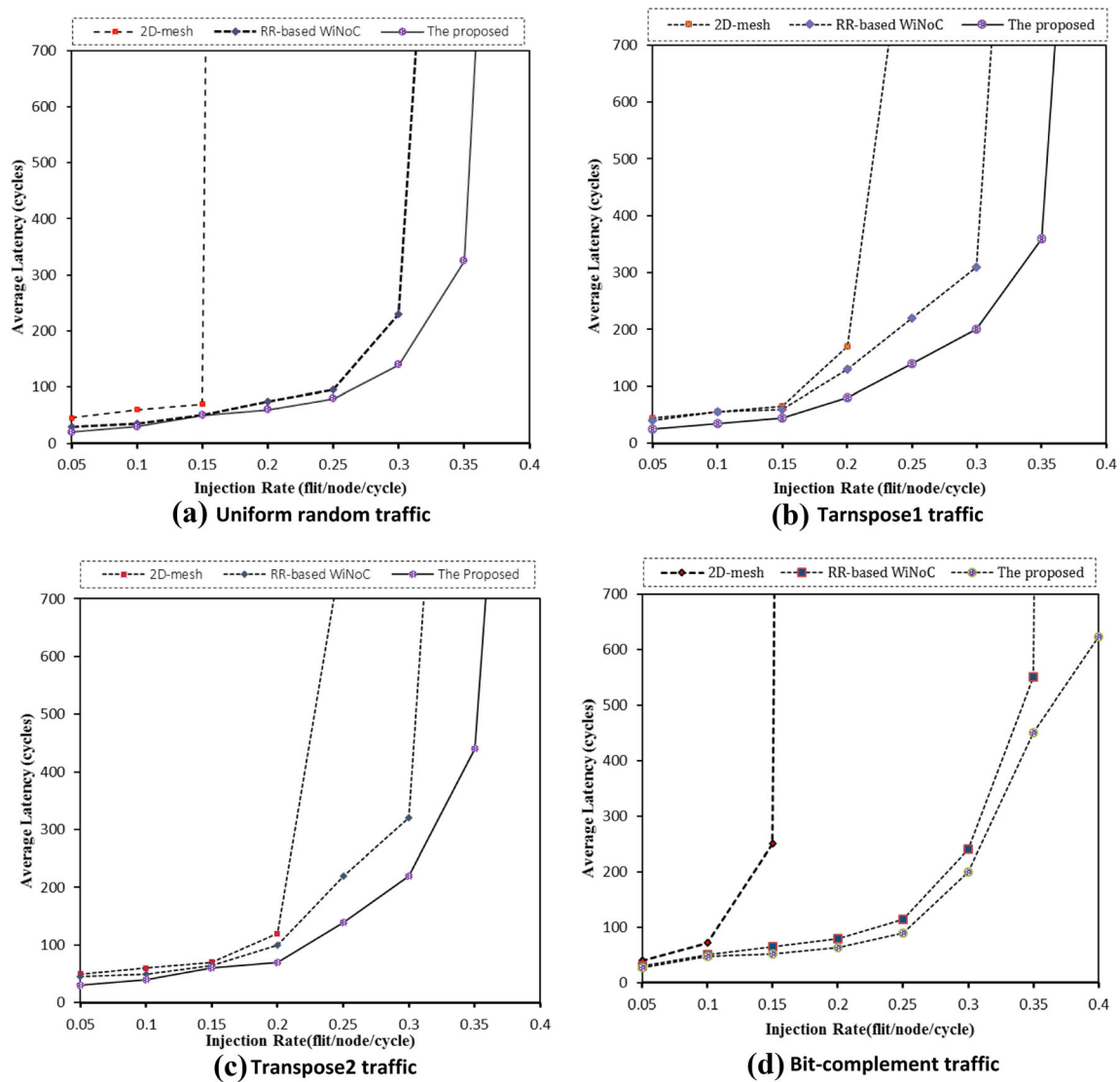
To obtain accurate results, throughput, average latency, and energy consumption have been shown under several synthetic traffic patterns with different packet injection rate.

Fig. 8a–d shows the average latency simulation results for a 10×10 mesh network with four WR. We can see that WiNoCs has a lower average delay than 2D mesh NoC. This is mainly due to decreasing the average number of hops for sending the data packets from sources to

destinations in WiNoCs. In the low packet injection rate, congestion in the network resources is low so, the average latency reduces. Using the proposed mechanism in WiNoC has a better performance than RR-based WiNoC. By increasing the packet injection rates, the competition to receive the output port (wireless link) between the input ports of WR is increased. In this situation, an efficient scheduling policy can decide which of the input ports get a higher priority to access the output port. For this purpose, the proposed mechanism shows more advantages over the static RR-based mechanism. When the higher priority is allocated to an input port, this port receiving a grant signal so, it can send more flows to the wireless interface. Hence, the input buffers of WR will be empty as soon as and congestion in WR will eliminated. The network saturation point of our mechanism under uniform random traffic pattern is 0.7 flit/nodes/cycles, which is improved by about 12% compared to RR-based WiNoC. Also, the average latency improved by 29%. Figure 8b–d shows the average latency under two transpose traffic patterns and Bit-complement traffic respectively. As the traffic load becomes heavier, the average latency of RR-based WiNoC will increase due to congestion in WRs. This is because some of the input ports have more packets than the other input ports. Therefore, these packets suffer from a longer competition delay to access to the wireless interface channel. Nevertheless, our mechanism allocates the highest priority to these ports. The average packet latency is about 14% and 12% lower than the RR-based WiNoC in transpose1 and Bit-complement traffic patterns.

Another important factor to measure WiNoCs performance is the network throughput. Figure 9a–d shows the network throughput (flits/node/cycle) of two arbitration mechanisms under synthetic traffic patterns respectively.





**Fig. 8** Average latency of the 10×10-core WiNoC with four WRs under synthetic traffic patterns

As can be observed in these figures, while the packet injection rate gradually increases, the network throughput increases. Compared to RR-based WiNoC, the proposed arbitration mechanism can dynamically adjust the priority of input ports for using the output port. Hence, the injection rate saturation point and network throughput are increased.

The simulation result of total energy consumption is shown in Fig. 10. The energy consumption of the 2D-mesh network is more than WiNoCs. In the proposed mechanism, the network latency is reduced due to the allocation of the output port to the input ports with the higher traffic flow. In other words, the proposed arbitration leads to lower contention among the input ports of WRs and reduces the dynamic power consumption.

To evaluate the area and power consumption overhead, the general RR arbiter and the proposed arbitration

mechanism implemented and synthesized in a TSMC 45nm library in CMOS technology using Synopsys Design Compiler. Results obtained with operating frequency 1 GHz are shown in Table 3. As can be seen, the area overhead of our arbitration scheme is more than the general RR arbiter. Our scheme has an additional block like a PCB, hence, the area and power consumption overhead have been increased. Figure 11 shows the area and power consumption breakdown of a wireless router with further details. Due to the low number of wireless routers in WiNoC, the area and power consumption of the proposed arbitration is negligible compared to the transceiver, router, and other components [29].

In other experiments, Figs. 12 and 13 show the latency and network throughput versus the packet injection rate for a large (i.e., 15 × 15) mesh network under synthetic traffic

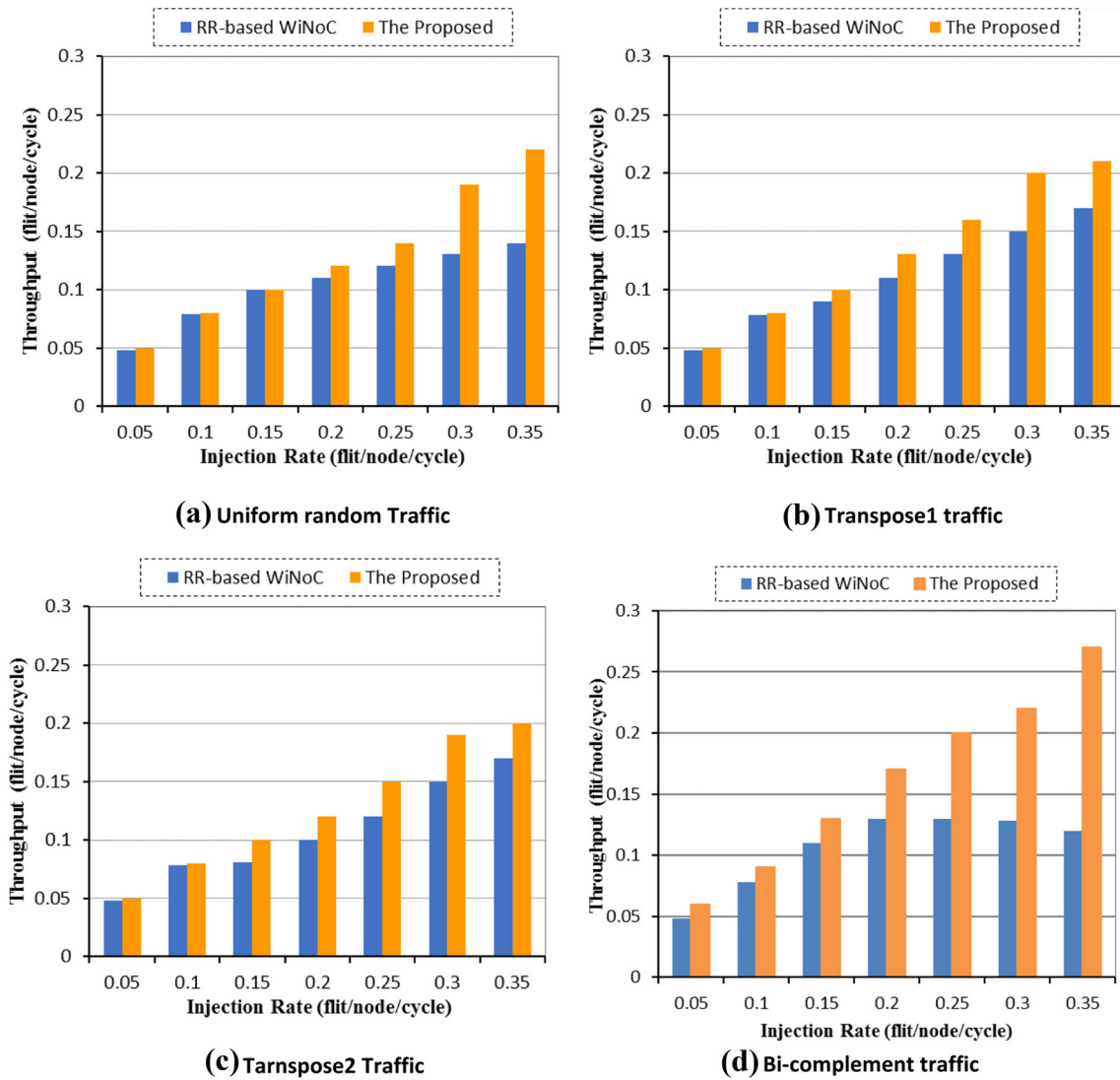


Fig. 9 Throughput of the 10×10-core WiNoC with four WRs under synthetic traffic patterns

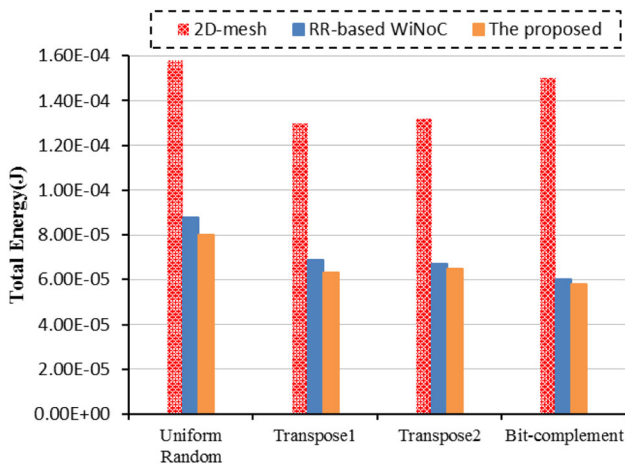


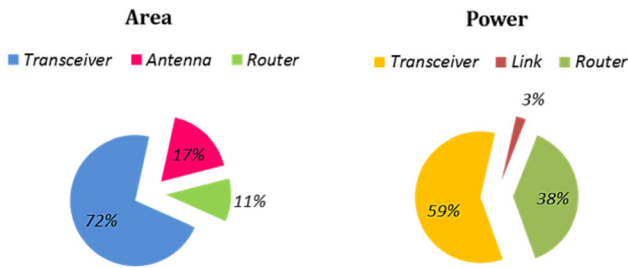
Fig. 10 Total energy of the 10×10-core WiNoC with four WRs under synthetic traffic patterns

Table 3 Area and power overheads

Block type	Area ( $\mu\text{m}^2$ )	Power (mW)
RR	650	70.15
The proposed	1231	115.3

pattern. As it can be observed, by increasing the network size, in all traffic patterns, the network latency is increased due to increasing the distribution of packets, the number of nodes, and distance among long-distance nodes.

It can be also noted that at the beginning of the simulation, because of the low packets injection rate, the average latency of networks is low. As the network traffic load becomes heavier, the average latencies increase gradually. This is because of the congestion in the BRs and



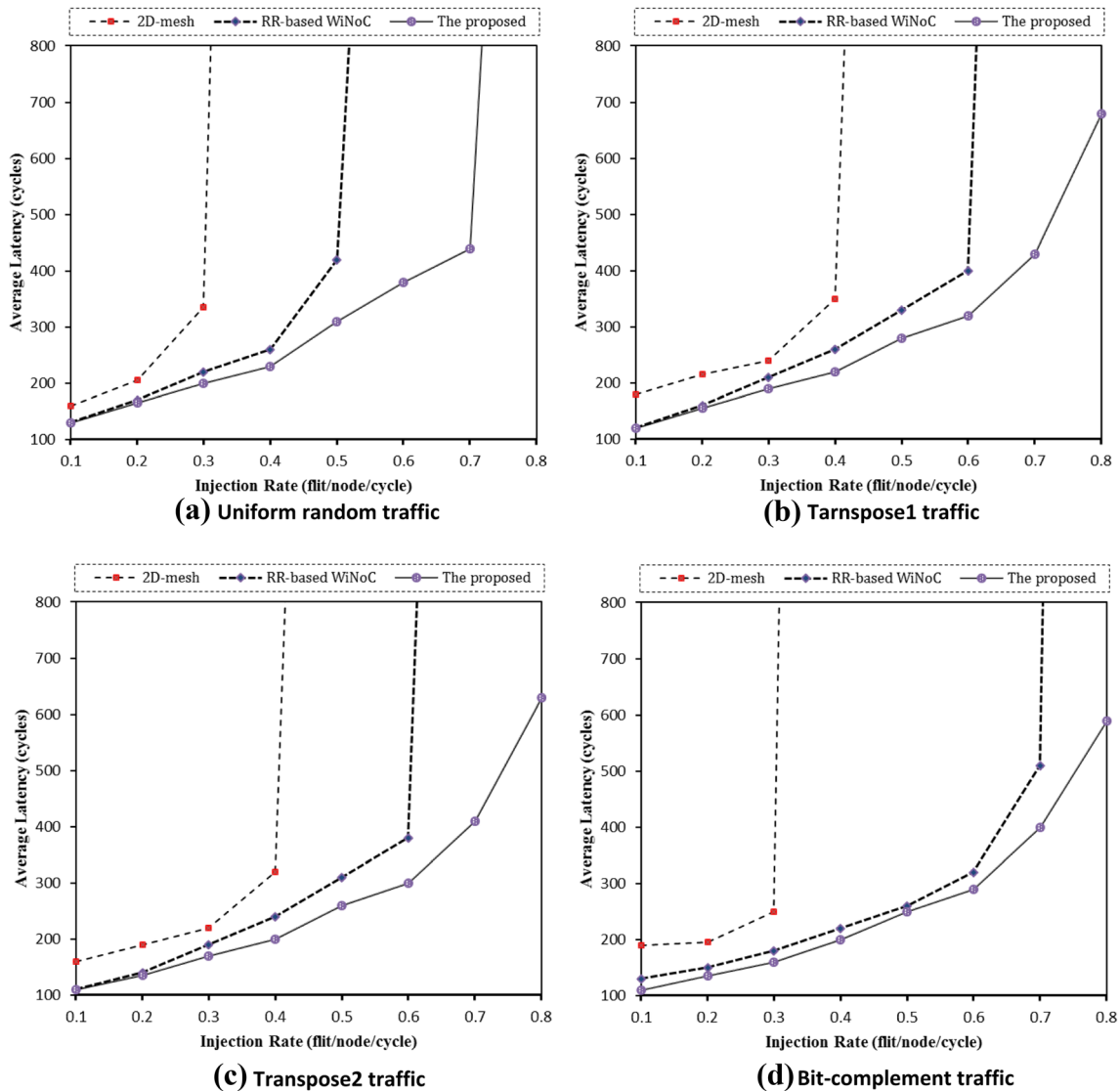
**Fig. 11** Quantities of the area and power consumption overhead in a wireless router

WRs. Thus, the links utilization could be severely decreased, and the network performance degrades. Moreover, the input ports buffers are also greatly influenced by the issue of excessive data packets. However, in all scenarios, the proposed scheme has better latency performance

compared with RR-based WiNoC. The latency improvement of the proposed mechanism under random, and transpose 1 traffic patterns are 35% and 25% respectively. The measured values are obtained at a similar saturation injection point for each traffic pattern.

The results of the network throughput of the  $15 \times 15$ -core WiNoC with nine WRs under synthetic traffic patterns are illustrated in Fig. 13. The proposed scheme under different traffic patterns has higher throughput in comparison to RR-based WiNoC. As to the saturating point of the proposed mechanism, the network throughput is increased 12%, 10%, 6%, and 23% in comparison to RR-based WiNoC under random, transpose1, transpose2, and Bit-complement traffic patterns respectively.

In Fig. 14, the energy consumption of the proposed scheme is presented. As Fig. 14 illustrates the 2-D mesh-



**Fig. 12** Average latency of the  $15 \times 15$ -core WiNoC with nine WRs under synthetic traffic patterns

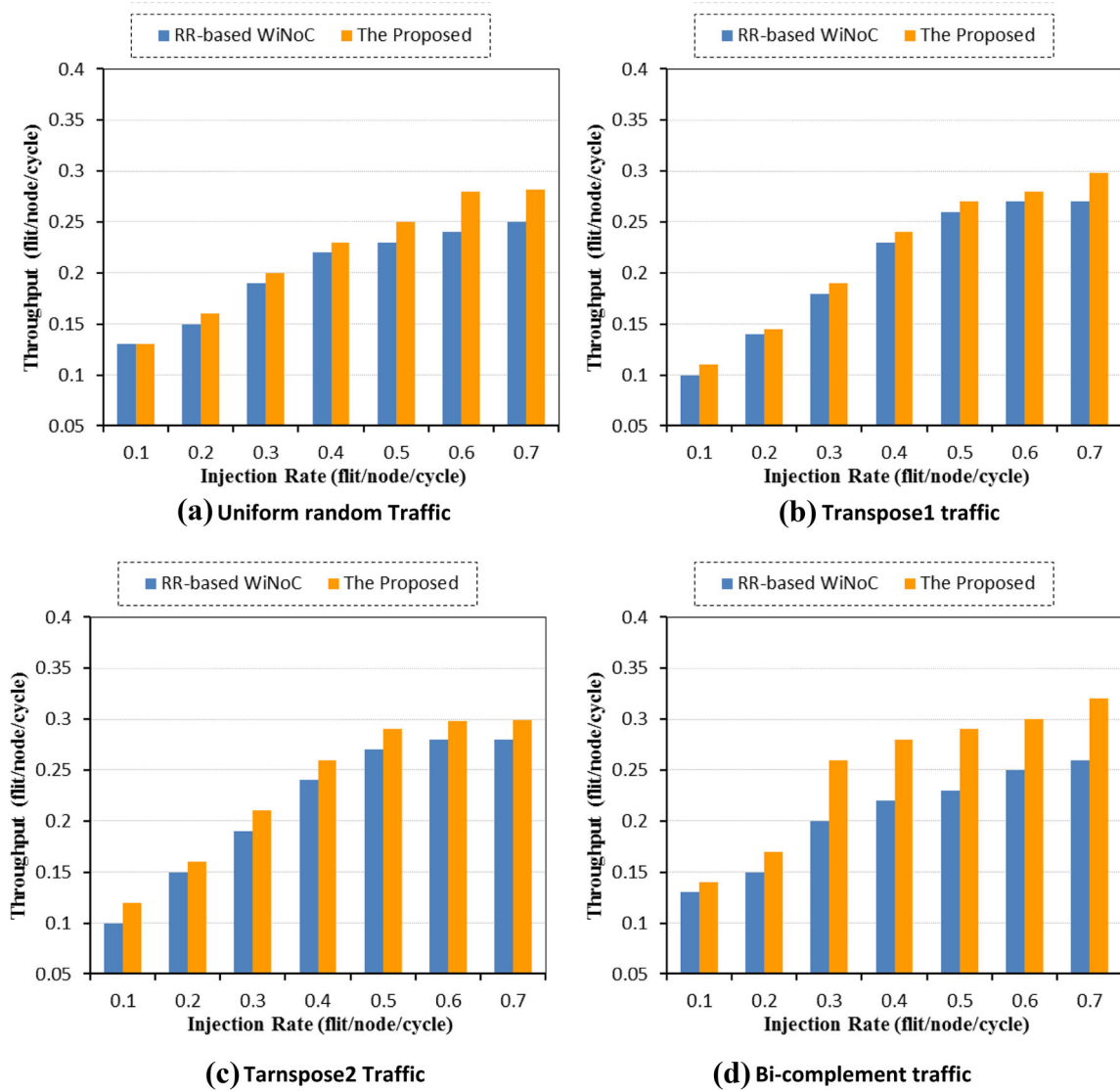


Fig. 13 Throughput of the 15×15-core WiNoC with nine WRs under synthetic traffic patterns

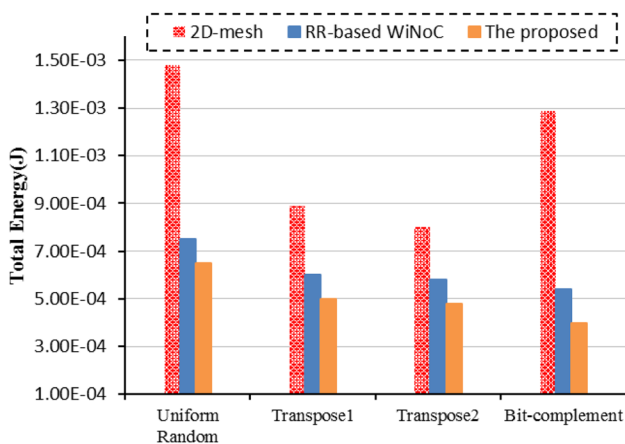


Fig. 14 Total energy of the 15×15-core WiNoC with nine WRs under synthetic traffic patterns

based NoC has more power consumption compared with WiNoCs. The proposed method is more energy-efficient than RR-based WiNoC due to the better utilization of wire/wireless links and less contention of the input ports. Under the Bit-complement traffic pattern, the energy consumption of the proposed method is much less than RR-based WiNoC because of the topology.

### 5 Conclusions

This paper proposes a novel arbitration mechanism for wireless routers in WiNoC. The basic idea is to adjust the dynamic priority for each input port of WR based on the current network status and the data packet length. Initially, the PCU block calculates the priority of input ports and

then the DMU assigns the output port to the highest priority input port. The simulation results under synthetic traffic patterns show that the proposed arbitration scheme reduces the average latency and improves the network throughput compared to the RR-based arbitration mechanism in WiNoC. Meanwhile, the area overhead of the proposed mechanism has a negligible impact on the overhead of WiNoC.

## References

1. L. Benini and G. De Micheli, “Networks on chip: a new paradigm for systems on chip design,” in *Design, Automation, and Test in Europe Conference and Exhibition, 2002. Proceedings, 2002*, pp. 418–419: IEEE.
2. Feero, B.S., Pande, P.P.: Networks-on-chip in a three-dimensional environment: A performance evaluation. *IEEE Transactions on Computers* **58**(1), 32–45 (2009)
3. Shacham, A., Bergman, K., Carloni, L.P.: Photonic networks-on-chip for future generations of chip multiprocessors. *IEEE Transactions on Computers* **57**(9), 1246–1260 (2008)
4. D. DiTomaso, A. Kodi, S. Kaya, and D. Matolak, “iWISE: Inter-router wireless scalable express channels for network-on-chips (NoCs) architecture,” in *High-Performance Interconnects (HOTI), 2011 IEEE 19th Annual Symposium on*, 2011, pp. 11–18: IEEE.
5. DiTomaso, D., Kodi, A., Matolak, D., Kaya, S., Laha, S., Rayess, W.: A-WiNoC: Adaptive wireless network-on-chip architecture for chip multiprocessors. *IEEE Transactions on Parallel and Distributed Systems* **26**(12), 3289–3302 (2015)
6. Ouyang, Y., Li, Z., Xing, K., Huang, Z., Liang, H., Li, J.: Design of Low-Power WiNoC with Congestion-Aware Wireless Node. *Journal of Circuits, Systems, and Computers* **27**(09), 1850148 (2018)
7. Deb, S., Ganguly, A., Pande, P.P., Belzer, B., Heo, D.: Wireless NoC as interconnection backbone for multicore chips: Promises and challenges. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems* **2**(2), 228–239 (2012)
8. Lin, J.-J., Wu, H.-T., Su, Y., Gao, L., Sugavanam, A., Brewer, J.E.: Communication using antennas fabricated in silicon integrated circuits. *IEEE Journal of solid-state circuits* **42**(8), 1678–1687 (2007)
9. Kempa, K., et al.: Carbon nanotubes as optical antennae. *Advanced Materials* **19**(3), 421–426 (2007)
10. S.-B. Lee *et al.*, “A scalable micro wireless interconnect structure for CMPs,” in *Proceedings of the 15th annual international conference on Mobile computing and networking*, 2009, pp. 217–228: ACM.
11. W.-H. Hu, C. Wang, and N. Bagherzadeh, “Design and analysis of a mesh-based wireless network-on-chip,” in *Parallel, Distributed and Network-Based Processing (PDP), 2012 20th Euromicro International Conference on*, 2012, pp. 483–490: IEEE.
12. Dehghani, A., RahimiZadeh, K.: Design and performance evaluation of Mesh-of-Tree-based hierarchical wireless network-on-chip for multicore systems. *Journal of Parallel and Distributed Computing* **123**, 100–117 (2019)
13. J. Murray, P. Wettin, P. P. Pande, and B. Shirazi, *Sustainable wireless network-on-chip architectures*. Morgan Kaufmann, 2016.
14. Bahrami, B., Jamali, M.A.J., Saeidi, S.: A demand-based structure for the architecture of wireless networks-on-chip. *Wireless Personal Communications* **96**(1), 455–473 (2017)
15. Chidella, K.K., Asaduzzaman, A.: A novel Wireless Network-on-Chip architecture with distributed directories for faster execution and minimal energy. *Computers & Electrical Engineering* **65**, 18–31 (2018)
16. Vijayakumar, V., Yuvaraj, M.P., Mansoor, N., Nerurkar, N., Ganguly, A., Kwasinski, A.: CDMA enabled wireless network-on-chip. *ACM Journal on Emerging Technologies in Computing Systems (JETC)* **10**(4), 28 (2014)
17. Deb, S., et al.: Design of an energy-efficient CMOS-compatible NoC architecture with millimeter-wave wireless interconnects. *IEEE Transactions on Computers* **62**(12), 2382–2396 (2013)
18. Zhao, D., Wang, Y.: SD-MAC: Design and synthesis of a hardware-efficient collision-free QoS-aware MAC protocol for wireless network-on-chip. *IEEE Transactions on Computers* **57**(9), 1230–1245 (2008)
19. Palesi, M., Collotta, M., Mineo, A., Catania, V.: An efficient radio access control mechanism for wireless network-on-chip architectures. *Journal of Low Power Electronics and Applications* **5**(2), 38–56 (2015)
20. Abadal, S., Torrellas, J., Alarcón, E., Cabellos-Aparicio, A.: OrthoNoC: A Broadcast-Oriented Dual-Plane Wireless Network-on-Chip Architecture. *IEEE Transactions on Parallel & Distributed Systems* **1**, 1–1 (2018)
21. Rezaei, A., Daneshlab, M., Zhao, D.: CAP-W: Congestion-aware platform for wireless-based network-on-chip in many-core era. *Microprocessors and Microsystems* **52**, 23–33 (2017)
22. S. M. Mamaghani and M. A. J. Jamali, “A load-balanced congestion-aware routing algorithm based on time interval in wireless network-on-chip,” *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–14, 2018.
23. Noghondar, A.F., Reshadi, M., Bagherzadeh, N.: Reducing bypass-based network-on-chip latency using priority mechanism. *IET Computers & Digital Techniques* **12**(1), 1–8 (2017)
24. R. Das, O. Mutlu, T. Moscibroda, and C. R. Das, “Application-aware prioritization mechanisms for on-chip networks,” in *Proceedings of the 42nd Annual IEEE/ACM International Symposium on Microarchitecture*, 2009, pp. 280–291: ACM.
25. Rahmati, D., Sarbazi-Azad, H.: Classified Round Robin: A Simple Prioritized Arbitration to Equip Best Effort NoCs With Effective Hard QoS. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* **37**(1), 257–269 (2018)
26. Bashizade, R., Sarbazi-Azad, H.: P2R2: Parallel pseudo-round-Robin arbiter for high performance NoCs. *Integration, the VLSI Journal* **50**, 173–182 (2015)
27. Liu, Y., Jin, J., Lai, Z.: A dynamic adaptive arbiter for Network-on-Chip. *Journal of Microelectronics, Electronic Components, and Materials* **43**(2), 111–118 (2013)
28. Catania, V., Mineo, A., Monteleone, S., Palesi, M., Patti, D.: cycle-accurate network on chip simulation with noxim. *ACM Transactions on Modeling and Computer Simulation (TOMACS)* **27**(1), 4 (2016)
29. A. Mineo, M. Palesi, G. Ascia, and V. Catania, “An adaptive transmitting power technique for energy-efficient mm-wave wireless NoCs,” in *Design, Automation, and Test in Europe Conference and Exhibition (DATE), 2014*, 2014, pp. 1–6: IEEE.
30. Farhad Rad, Midia Reshadi, Ahmad Khademzadeh, “A survey and taxonomy of congestion control mechanisms in wireless network on chip,” *Journal of Systems Architecture*, Volume 108, 2020, 101807, ISSN 1383-7621.
31. Mohtavipour, S. M., Mollajafari, M., & Naseri, A., “A novel packet exchanging strategy for preventing HoL-blocking in fat-trees,” *Cluster Computing*, 2019, pp: 1–22.



32. Mortazavi, Seyed Hassan, Reza Akbar, Farshad Safaei, and Amin Rezaei. "A fault-tolerant and congestion-aware architecture for wireless networks-on-chip," *Wireless Networks* 25, no. 6, 2019, pp: 3675-3687, <https://doi.org/10.1007/s11276-019-01962-3>.
33. Yazdanpanah, F., AfsharMazayejani, R., Alaei, M., Rezaei, A., Daneshtalab, M.: An energy-efficient partition-based XYZ-planar routing algorithm for a wireless network-on-chip. *The Journal of Supercomputing* 75(2), 837–861 (2019)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Farhad Rad** received her B.Sc. and M.Sc. degree in computer engineering from Shiraz University and Iran University and Science & Technology (IUST) respectively in 2004 and 2008. He is currently a Ph.D. student in the Department of Computer Engineering, Science and Research Branch, Islamic Azad University (SRBIAU), Tehran, Iran. Her research interests include on-chip interconnection networks, communication infrastructure of the

multi-core chip, Wireless Network-on-Chip (WiNoC), and optimization techniques for on-chip architecture and high-performance computing.



**Midia Reshadi** is currently an Assistant Professor in Computer Engineering Department at Science and Research Branch of Azad University since 2010. His research interest is Network-on-chip including performance and cost improvement in topology, routing and application-mapping design levels of various types of NoCs such as 3D, photonic and wireless. Recently, he has started carrying out research in NoC based deep neural network accelerators and

Silicon interposer based NoC with his team consists of M.Sc. and Ph.D. students.



**Ahmad Khademzadeh** was born in Mashhad, Iran, in 1943. He received a B.Sc. degree in Applied Physics from Ferdowsi University, Mashhad, Iran, in 1969 and the M.Sc., Ph.D. degrees respectively in Digital Communication and Information Theory and Error Control Coding from the University of Kent, Canterbury, UK. He is currently the Head of Education and National Scientific and Informational Scientific Cooperation Department at Iran

Telecom Research Center (ITRC). He was the head of Test Engineering Group and the director of Computer and Communication Department at ITRC. Dr. Khademzadeh has been received four distinguished national and international awards including Kharazmi International Award and has been selected as the National outstanding researcher of the Iran Ministry of Information and Communication Technology.