An efficient adaptive reconfigurable routing protocol for optimized data packet distribution in network on chips

Pavithra Goravi Sukumar¹, Modugu Krishnaiah², Rekha Velluri³, Pooja Satish⁴, Sharmila Nagaraju⁵, Nandini Gowda Puttaswamy⁶, Mallikarjunaswamy Srikanthswamy⁷

¹Department of Computer Science and Engineering, JSS Academy of Technical Education, Bangalore, India
²Department of Electronics and Communications, Vidyajyothi Institute of Technology, Hyderabad, India
³Department of Computer Science and Engineering, CHRI (Deemed to be University), Bangalore, India
⁴Department of Electronics and Communications, K.S. Institute of Technology, Bangalore, India
⁵Department of Electrical and Electronics Engineering, Sri Jayachamarajendra College of Engineering, Mysore, India
⁶Computer Science and Engineering, East Point College of Engineering and Technology, Bangalore, India
⁷Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Bangalore, India

ABSTRACT

The deadlock-free and live lock-free routing at the same time is minimized in the network on chip (NoC) using the proposed adaptive reconfigurable routing protocol (ARRP). Congestion condition emergencies are avoided using the proposed algorithm. The input packet distribution process is improved among all its shortest paths of output points. The performance analysis has been initiated by considering different configuration (N*N) mesh networks, by sending various ranges of data packets to the network on chip. The average and maximum power dissipation of XY, odd-even, Dy-XY algorithm, and proposed algorithm are determined. In this paper, an analysis of gate utilization during data packet transfer in various mesh configurations is carried out. The number of cycles required for each message injection in different mesh configurations is analyzed. The proposed routing algorithm is implemented and compared with conventional algorithms. The simulation has been carried out using reconfigurable two-dimensional mesh for the NoC. The proposed algorithm has been implemented considering array size, the routing operating frequency, link width length, value of probability, and traffic types. The proposed ARRP algorithm reduces the average latency, avoids routing congestion, and is more feasible for NoC compared to conventional methods.

Keywords:
Dynamic XY routing algorithm
Network on chip
Random traffic profile
System on chip
Transpose traffic profile

1. INTRODUCTION

In any complex system on chips (SoCs), communication between the cores is very essential to perform the proper functioning. Hence in this paper, we proposed a network on chips (NoCs) for complex SoCs. To find the best path to deliver the packets from source to destination several routing algorithms are used in NoCs. The routing algorithms used for routing are classified as adaptive and deterministic algorithms. The best route from the source to the destination in the network is determined based on the congestion of the network in the case of adaptive-type routing algorithms. In a network packet injection rate of cores must be above the threshold of NoC, then the implementation of deterministic routing algorithms is preferable [1], [2]. In the case of the adaptive routing algorithm packet passing through the congested link is reduced as compared to other routing
algorithms. Hence it is important to adopt the congestion control mechanism to the algorithm. In recent days several deadlock-free adaptive routing algorithms have been introduced. In routing algorithms, an introduced virtual channel assists the use of different network topologies to develop the network topologies [3]–[5]. In the mesh structure-based network, the proposed routing algorithms can use virtual channels to route the packet to the destination [6]. Abadal et al. [7] proposed the two-dimensional mesh XY static routing algorithm, in which each transmitted packet travels in both directions to reach the destination [8]. The algorithm is witnessed for the occurrence of no deadlock, but no possibility of providing adaptively. Another adaptive routing algorithm called turn-model is introduced in [9]. An adaptive routing algorithm called odd-even turn is proposed in [10]. This algorithm uses a mesh-based topology to avoid the deadlock. In study [11] combination of adaptive routing odd-even and static routing algorithms or fix-based algorithms called DyAD is introduced. Based on the congestion condition this algorithm uses the above-said congestion condition. The condition adopted algorithm is finally called DyXY which is proposed in [12]. The static XY algorithm is considered as a basis for the implementation of this algorithm. In the literature to avoid congestion, adaptive routing algorithms are incorporated. In networks, congestion is a common problem and it can be minimized by developing an effective routing algorithm. Hence an adaptive routing algorithm called adaptive reconfigurable routing protocols (ARRP) is proposed in this work to avoid congestion.

The structure of the paper is formulated as follows. The proposed ARRP is presented in section 2. The proposed algorithm hardware implementation is presented in section 3. Section 4 reports experimental results. Section 5 concludes the paper.

2. PROPOSED ALGORITHM FOR ADAPTIVE RECONFIGURABLE ROUTING PROTOCOL

The mesh topology-based networks have many paths from source to destination. Among these paths, some of them have minimum distance. The proposed algorithm routes the packets through the minimal path by avoiding congestion. An example of such an algorithm is the live lock-free feature which meets the requirements. Our proposed ARRP algorithm has two operational procedures to avoid congestion [13]. The ARRP in the first procedure over the network distributes the traffic just as local information. In the second case, ARRP tries to exchange information based on the situation. Every router in the network distributes uniformly all the incoming packets to the destination device through the shortest path. If the router has one incoming and outgoing port, then the packets are forwarded through that port only. On the other side, if the packet is forwarded through the two ports, then, half time through the first port and remaining through the second port. If both cases are not satisfied packet through either sub-network [15].

Figure 1 shows the network structure of sending packets from core (0, 3) to (3, 3) and core (0,0) to core (3, 1). While sending the packets from the core (0, 3) to (3, 3) it travels through the core (1, 3). Packets are always routed through the minimal path. Similarly, packets routing from (0, 0) to (3, 1) are forwarded through the core (1, 0). Also, this router (1, 0) has two options to forward the packets, half of the packets through the (1, 1) and the remaining part through the router (2, 0).

In an ARRP routing protocol, the feature of deadlock-freeness has a Y dimension with two virtual channels [14]. The network is partitioned -X and +X subnetwork each having a pair of channels. The packet is forwarded through the +X sub-network if the destination node is located right of the source node. Similarly, the packets are forwarded to the destination node through the -X sub-network located left of the source node. If both cases are not satisfied packet through either sub-network [15].
The congestion is decreased, by forwarding the incoming packets by forwarding packets among the different paths. However, the congestion in the network cannot be eliminated by following this procedure. Therefore, the ARRP routing protocol mechanism is used to detect and avoid congestion by changing the mechanism of routing. The routers in the network continuously monitor its buffet input and when a number of packets reaches above the threshold value. Each router in the network forwards packets in their pre-determined routes based on the congestion of the router [16].

The routing mechanism is changed whenever congestion occurs at the router. To control the congestion, follow the structure of Figure 2. Assume the congested dedicated link between core node (2, 2) to core node (3, 2). If the group G4 router has a packet for a member of another group member of G4 then the router has the option to forward through south or east ports. In this case, if the router chooses the south port, then it passes the congestion link, however, this is not the case if the router chooses the south port. Hence it is required to increase the probability of selection of the east port rather than the south port. The same argument is true for the other routers while forwarding their packets.

Let us consider another situation where G1 of group 1 has a packet sent to G7 of group 7. The north or east port is selected for forwarding the packets. If the router chooses the east port, then the packet does not pass near the congestion link, however, the probability of occurrence of congestion is higher if it chooses the north port. The same is a true member of the other group in the network [17].

The packets are forwarded to the group G3 members, if they pass through the G10/G8 group by choosing the north/south port then congestion may likely happen. Hence changing of routing mechanism is critical in these groups. The changes made in the routers of group G1, G3, and G5 routing mechanisms are less likely to pass the congestion link, however, even changes not made also do pass the congestion link. Hence changes made in the routing mechanism for these groups are not critical for congestion control. The overall summary above discussion about the mechanism of congestion control is given in Table 1.

<table>
<thead>
<tr>
<th>Group</th>
<th>South Port</th>
<th>North Port</th>
<th>East Port</th>
<th>West Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td>DEC</td>
<td>INC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G2</td>
<td>DEC</td>
<td>INC</td>
<td>INC</td>
<td></td>
</tr>
<tr>
<td>G3</td>
<td>INC</td>
<td>INC</td>
<td>DEC</td>
<td></td>
</tr>
<tr>
<td>G4</td>
<td>DEC</td>
<td>INC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G5</td>
<td>DEC</td>
<td>INC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2. Mesh topology-based sample NoC
As mentioned before every router in the network monitors its buffer and it is detected once it is full. The routing mechanism is adjusted as per Table 1 to find an appropriate route to forward the packets [18, n Figure 2, router (2, 2) of the east link is congested, and then the router generates 5 packets for routing five groups as given in Table 1. Here the south port is chosen for forwarding the packets to G1 and G2 groups and the west port is selected for forwarding the packets to group G3. North port is used to forward the packets belonging to groups G4 and G5. The routing mechanism is changed as given in Table 1. The ARRP algorithm is used to avoid congestion at the router by defining two threshold values at the buffer input of every router. When the buffer size of the router reaches the maximum value then it uses the critical groups by generating the routing mechanism.

3. HARDWARE IMPLEMENTATION

The hardware implementation ARRP routing protocol makes use of the advantage of the priority table. Each neighboring port of a router has two adjacent ports, for example, northwest, northeast, southwest, and south-east as per the location of the priority table. The packets can be forwarded in more than ports as indicated in the priority table. Priority table each location has a 2-bit width and indicates how many packets are sent out of 10 packets through each port. The number of packets sent through each port is tabulated in Table 2.

<table>
<thead>
<tr>
<th>Table 2. Priority-based data packets are transmitted to different ports</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of Transmitted packets out of 10</strong></td>
</tr>
<tr>
<td>North-West</td>
</tr>
<tr>
<td>------------</td>
</tr>
<tr>
<td>(West-09, North-01)</td>
</tr>
<tr>
<td>00</td>
</tr>
<tr>
<td>(West-06, North-04)</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>(West-04, North-06)</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>(West-01, North-09)</td>
</tr>
</tbody>
</table>

In the router, each port has 6-bit saturating (Saturated-10) counters, and two counter of that port is incremented whenever a packet is sent. A combination of two letters is used to identify the counters. The owning port is represented by the first letter and the neighbor port is represented by the second letter. The combination of the priority table and counters together identifies the packet destination port. Suppose the packet is forwarded in more than one direction, then the packet is routed through the counter with letters corresponding to that direction [19]–[21]. All eight counters’ initial value is zero and priority table content is set to forward the incoming traffic (01, 10, and 01) for South-East, South-West, North-West, and North-East. The priority table content is maintained and remains constant until a delivered packet, and it changed in this case as per Table 1.

4. RESULTS AND DISCUSSION

The proposed routing algorithm efficiency is computed with different parameters and prove its performance in the other three existing routing algorithms is also implemented. The others implemented algorithms such as odd-even, XY, and DyXY. They have developed a simulator for NoC which is based on C++ for calculating power consumption and average delay of the transmission. This type of simulator is effectively used to switch into a two-dimensional configuration of the mesh for the NoC [22]. The inputs included in the simulator are frequency, array size, array size, and link width length. The simulator is capable of generating different traffic profiles. Orion library functions [14] are used to calculate the power consumption of the algorithm. The simulation speed of the simulator is high since it is event-driven based. The data width for all switches is 16 bits wide and 12 flits of channel buffer size with a congestion threshold set of 25% and 75% of the capacity of the buffer.

4.1. Proposed ARRP based transpose traffic profile architecture

The transpose traffic profile is considered the first set of simulations. In this profile for a mesh network of size \( n \times n \) a processing element (PE) at position \((i,j) (i,j \in [0,n])\) which sends only data packets into another position PE \((n - 1 - i, n - j - 1)\). As shown in 1 has been utilized to compute the traffic profile score by summing the link load and the routers along the routes.
An efficient adaptive reconfigurable routing protocol for optimized ... (Pavithra Goravi Sukumar)

Figure 3 shows the proposed ARRP architecture. It consists of three extra reconfigurable additional hardware units compared to existing methods. First, the routine unit is a core part of the proposed architecture. It controls the switch operation of the crossbar [23], [24]. The entire operation routine unit is controlled by input buffers. Second, the input buffer avoids data overlapping and loss of information. The input buffer has been controlled by the port controller. Each port controller is controlled by the reconfigurable routing algorithm and is capable of handling a huge amount of data packets from peripheral devices. Third, the congestion flag is connected to each input buffer and is used to control the congestion in the link. It is capable of changing the path direction to avoid huge data traffic. The congestion flags change their condition with respect to the results and number of parallel operations [25].

\[ T_{psum} = \sum_{i=0}^{L_r} l_i + \sum_{i=0}^{S_l} \sum_{j=0}^{S_l} f_{ij} \]  

(1)

where \( L_r \) is described as set of the link load values, \( l_i \) is identified as a load of a link, \( f_{ij} \) represents the number of flits per second arriving towards the router and \( S_l \) is described as the set of the router load [26]. The load of the router is directly proportional to the load of the link arriving towards the routers as given in (2),

\[ f_{ij} = \sum_{i=0}^{L} \frac{l_i}{L} \]

(2)

where \( L \) is described as the number of the router links. In [27], a similar concept of finding transpose can be found. In the simulated results, ten flit data messages were generated by the processing elements and used in network time intervals. The array size of 8×8 and 14×14 have been considered and this leads to the traffic profile distribution of non-uniform with nodes in the mesh for heavy traffics [28]. The creation of a hotspot is not required if the injection rate of a data packet is very low. As the increases in data rates, the proposed algorithm leads to fewer delays. The corresponding analysis is shown in Figure 4. Figure 4(a) shows the performance with a load of 8×8 2D mesh and Figure 4(b) shows the performance with a load of 14×14 2D mesh.

4.2. Proposed ARRP based random traffic profile

To simulate the proposed algorithm, we considered the two array sizes of 8×8 and 14×14. In the profile of the traffic, each PE sends many messages to the destination. Each message is directed to the destination for uniform distribution. Here we set the 12-destination node for the simulation. The relationship between average message injection rate and average communication delay is plotted in Figure 5. The simulation results show that our
The proposed algorithm offers better performance as compared to existing algorithms as shown in Figure 5(a) performance with a load of 8×8 2D mesh and Figure 5(b) performance with a load of 14×14 2D mesh.

Figure 5. Performance analysis between the proposed algorithm and conventional methods with respect to random traffic model (12 destinations) under different loads in (a) performance with a load of 8×8 2D mesh and (b) performance with a load of 14×14 2D mesh

4.3. Power dissipation

An algorithm power dissipation for routing is measured and compared with the other existing routing algorithms such as odd-even, XY, and DyXY under the model of random traffic and transpose. Figure 6 shows the average power results under the traffic model of 14×14 2D-mesh and corresponding results are presented in Table 3. The average power dissipation of our proposed algorithm under the model of transpose traffic profile is 12% more than DyXY, 22% more than that of the XY algorithm, and 6.5% less than the DyXY algorithm. Figure 6(a) shows the transpose model of the average power dissipation comparison analysis of ARRP and Figure 6(b) shows the random model of the average power dissipation comparison analysis of ARRP.
The proposed algorithm’s maximum power dissipation is 35% less than XY, 24% less than odd-even, and 10% less than DyXY algorithms. The analysis is also given in Table 4. From the experiment result, the maximum and average power of our algorithm is considerably lower as compared to existing algorithms.

![Graph showing average power dissipation comparison of ARRP with existing algorithms](image)

**Figure 6.** Average power dissipation comparison of ARRP with existing algorithms (a) transpose model and (b) random model

<table>
<thead>
<tr>
<th>Power dissipation (Average)</th>
<th>DyXY</th>
<th>Odd-even</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-5%</td>
<td>11%</td>
<td>22%</td>
</tr>
<tr>
<td>Transpose</td>
<td>-6.5%</td>
<td>12%</td>
<td>24%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Power dissipation (Maximum)</th>
<th>DyXY</th>
<th>Odd-even</th>
<th>XY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>-6%</td>
<td>-20%</td>
<td>-31%</td>
</tr>
<tr>
<td>Transpose</td>
<td>-10%</td>
<td>-24%</td>
<td>-35%</td>
</tr>
</tbody>
</table>

### 4.4. Hardware overhead

The proposed algorithm, area overhead is computed with very high-speed integrated circuit (VHSIC) hardware description language (VHDL) by designing the switches programming synthesized with Leonardo-Spectrum application-specific integrated circuit (ASIC) and standard cell library of semi-conductor laboratory (SCL) 0.25 μm. The data width for all switches is set at 16 bits and 12 flits buffer size. In order to obtain better performance, implement the first in, first out (FIFO) using registers. Figure 7 shows the comparative analysis of the proposed algorithm and existing algorithms with respect to the number of gates per micrometer.

![Graph showing comparative analysis of the proposed algorithm and existing algorithms with respect to the number of gates per micrometers](image)

**Figure 7.** Comparative analysis of the proposed algorithm and existing algorithms with respect to the number of gates per micrometers

*An efficient adaptive reconfigurable routing protocol for optimized ...* (Pavithra Goravi Sukumar)
5. CONCLUSION

The proposed ARRP is capable of handling more complex operations in NoCs. The proposed algorithm improves the number of gates accumulation per micrometer compared to conventional methods such as 2%, 1%, 10%, Dy-xy, odd-even, and XY, respectively. The proposed algorithm improves the message injection rate per cycle compared to conventional methods like 4%, 2%, and 4%, Dy-xy, odd-even, and XY respectively.

The proposed algorithm reduces the maximum power dissipation compared to conventional methods like 6%, 2%, and 5%, Dy-xy, odd-even, and XY, respectively. Moreover, passing congestion information is more feasible and reliable compared to conventional methods. The proposed hardware architecture improves the switching speed and is capable of performing more parallel operations and improves the throughput efficiency of the device. The proposed algorithm can be improved for gate accumulation in nanometers. It can enhance the switching speed to operate in gigahertz frequency signals. When it is operated in 6G mode, it produces more attention, jitters, and hazards. It will reduce the original information strength.

ACKNOWLEDGEMENTS

The authors would like to thank JSS Academy of Technical Education, Bangalore, JSS Science and Technology, Mysore, KS Institute of Technology, Bangalore, Visvesvaraya Technological University (VTU), Belagavi and Vision Group on Science and Technology (VGST) Karnataka Fund for Infrastructure strengthening in Science and Technology Level–2 sponsored “Establishment of Renewable Smart Grid Laboratory” for all the support and encouragement provided by them to take up this research work and publish this paper.

REFERENCES


BIographies of Authors

Pavithra Goravi Sukumar completed her B.E. and M.Tech. degrees in computer science and engineering from Visvesvaraya Technological University, Belgaum, India in 2003 and 2009 respectively. Currently, she is working as an assistant professor in the Department of Computer Science and Engineering at JSS Academy of Technical Education, Bangalore, India. Her areas of interest include clustering techniques and routing protocols for wireless sensor networks. He can be contacted at pavi.pgs@gmail.com.

Modugu Krishnaiah is currently working as an assistant professor in the Department of Electronics and Communication Engineering, Vidyav Jyothi Institute of Technology, Hyderabad, India. He completed a B.Tech. in Electronics and Communication Engineering from KU Affiliated College, an M.Tech. in VLSI System Design from JNTUH Affiliated College, and a Ph.D. in School of Engineering (VLSI Circuit Testing) under the Faculty of Electronics and Communication Engineering from SSUTMS, Sehore. He has research and teaching experience of 15 years and has published 7 research papers in SCI/Scopus journals and reputed conferences. He can be contacted at krish.mtvlsi@gmail.com.

Rekha Velluri completed her B.E and M.Tech. in computer science and engineering from Visvesvaraya Technological University Belgaum. She has more than 16 years of teaching experience, published many papers at national and international conferences, and currently working as an assistant professor at Christ University. She can be contacted at rekha.v@christuniversity.in.
Pooja Satish completed her B.E. in ECE, M.Tech. in digital electronics and communication, and Ph.D. from Visvesvaraya Technological University, Belgaum in 2008, 2015, and 2023, respectively. She has eight years of teaching experience and is currently working as an assistant professor in the Department of Electronics and Communication Engineering, KSIT, Bangalore. She can be contacted at: poojas@ksit.edu.in.

Sharmila Nagaraju has completed her B.E in EEE at SJCE, Mysore and M.Tech. in CAID at NIE Mysore. Secured second rank in Bachelor of Engineering degree. She has eight years of experience in teaching and is currently working as an assistant professor at RNSIT, Bangalore, in electronics and its applications. She can be contacted at Sharmila.n.89@gmail.com.

Nandini Gowda Puttaswamy currently working as asst. professor in information science and engineering, Sathagiri College of Engineering, Bangalore. She completed her B.E. in CSE from VTU, her M.Tech. in Software Engineering from VTU, and is now pursuing a Ph.D. from Visvesvaraya Technological University. She has published around 4 papers at the national conference and her areas of research interest are cloud computing, fog computing, edge computing, and internet of thing, AI, ML, and big data analytics. She can be contacted at nandini.educator@gmail.com.

Mallikarjunaswamy Srikantaswamy is currently working as an associate professor in the Department of Electronics and Communication Engineering at JSS Academy of Technical Education, Bangalore. He obtained his B.E. degree in Telecommunication Engineering from Visvesvaraya Technological University Belgaum in 2008, M.Tech. degree from Visvesvaraya Technological University Belgaum in 2010 and was awarded a Ph.D. from Jain University in 2015. He has more than 11 years of teaching experience. His research work has been published in more than 68 international journals and conferences. He received funds from different funding agencies. He is currently guiding five research scholars at Visvesvaraya Technological University Belgaum. He can be contacted at mallikarjunaswamys@jssateb.ac.in.