

PARS – An Efficient Congestion-Aware Routing Method for Networks-on-Chip

Xin Chang, Masoumeh Ebrahimi, Masoud Daneshtalab, Tomi Westerlund, Juha Plosila
Department of Information Technology, University of Turku
{xinchang,masebr,massdan,tovewe,juplos}@utu.fi

Abstract—The performance of NoCs (Networks-On-Chip) highly relies on the routing algorithm. Despite the higher implementation complexity compared with deterministic routing, adaptive routing has several merits, such as lower latency, higher throughput and better fault-tolerance performance. Most of the existing adaptive routing algorithms are based on the comparison of the horizontal and vertical congestion status in the network. However the performance of adaptive routing schemes suffers from the inadequate global congestion information. To address this issue, we proposed a novel routing algorithm with a congestion aware subnetwork to obtain more accurate non-local congestion information. This subnetwork will propagate the congestion information along the paths directly towards the destination. To find a less congested path, PARS (Path-Aware Routing Scheme) uses the congestion information of paths in a straight direction towards the destination rather than local congestion information. The simulation results reveal that the new presented scheme can offer better performance under different traffic profiles with a small hardware overhead.

Keywords *Network-on-Chip, congestion awareness, highly adaptiveness, minimal routing algorithm, mesh topology, two-dimensional networks.*

I. INTRODUCTION

With the growth of the technology scale and chip integrity, interconnection network on the chip has already become the mainstream in complex digital systems. The paradigm of Network-on-Chips (NoCs) enjoys better performance, fault tolerance and lower power consumption than the other traditional on-chip interconnections such as shared bus. Parallel communication can greatly benefit from these direct networks [1][2][3].

NoC provides scalability and reusability of communication infrastructure to reduce time to design and time to market for new products. The performance and efficiency of NoC highly depend on the underlying communication infrastructure; this, in turn, relies on the performance of the on-chip routers [24]. Thus, the design of high performance on-chip routers represents a critical issue for the success of the NoC approach. The router is the main component of the interconnection architecture that has a functionality of transferring information from one of its input ports to any of its output ports. A network

consists of an interconnection of many routers to enable a large number of cores to communicate with each other.

The topology, routing algorithm and flow control are the three most significant aspects in the NoCs. Topology defines the way of connecting routers together to form a network. Because of shortening link length and facilitating the manufacture process, simple topologies are usually employed in the NoCs such as two dimensional meshes (2-D mesh) [22][23][24]. To reduce the area overhead, a router with limited virtual channels, shorter pipeline stages is preferred [5].

Routing algorithms determine a path to deliver a packet from a source node to a destination node. Routings can be classified into *deterministic and adaptive* [4]. In deterministic routing models, the path between a source and a destination is fixed and determined by the source and destination nodes positions without considering the network condition. Deterministic routings are relatively simple in implementation, however, the traffic load is not balanced which tends to cause a traffic jam in the network and degrades the performance. XY is a deterministic wormhole routing algorithm that is deadlock and livelock free. Despite the deterministic schemes, adaptive routings are based on the network condition [11]. This can help to avoid the congested links or mal-function links to reduce the latency and provide fault tolerance with increasing the implementation complexity.

There are various routing algorithms such as table-based [7], deflection routing [8], non-minimal routing [9], and minimal congestion-aware routing [15]. The table-based routing provides many advantages including lower latency and better fault tolerance characteristics but suffers from high area and power consumption which may not be well suited in NoCs. In the deflection routing, the live-lock is a difficult problem to solve which may cause more latency. Non-minimal routing may increase the latency if the congestion cannot be accurately measured in the network. Many researches addressed that minimal and adaptive congestion-aware routings are simple and efficient approaches to distribute packets through the network and decrease the network latency [5][6]. In order to have a better performance, the global congestion information of the network is required [14][19][26].

Unicast and collective communication can be considered in NoC [20][21]. In the unicast communication case, a message is sent from a source node to a single destination node, while in the collective communication case one or several nodes send a message to one or several destinations. However, in this work, we limit our considerations to unicast commutation.

In this paper, we proposed a congestion-aware adaptive routing algorithm, named Path-Aware Routing Scheme (PARS). PARS is able to provide efficient global congestion information by gathering the congestion status of highly adaptive routers on the path towards the destination.

Section II summarizes a number of common routing algorithms in the NoCs which related to this work and the motivation of the new algorithm. In the Section III, we will describe the congestion aware subnetwork and the proposed algorithm. The experiment results are represented in the Section IV and Section IV concludes.

II. BACKGROUND AND RELATED WORKS

A. XY Routing Scheme

The XY algorithm is one of the most typical deterministic routing algorithms. In this algorithm, each packet first is routed in the X dimension until it reaches the column of the destination, and then is turned in the Y dimension. This routing is unable to balance the traffic load, especially in the non-uniform traffic pattern.

B. DyXY Routing Scheme

Dynamic XY (DyXY) is proposed as an adaptive deadlock-free routing algorithm [10]. In this scheme, when a packet reaches an intermediate router, then the next router is determined based on the congestion status of the corresponding buffers at the adjacent routers. The packet tends to travel along the less congested path to reduce the latency.

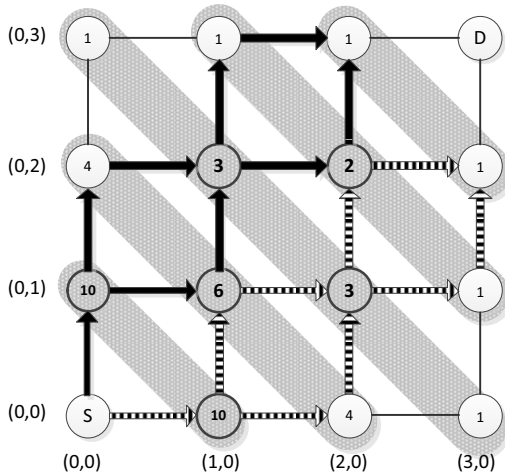


Figure 1. 4x4 2D mesh network with degree of adaptiveness

C. EDXY Routing Scheme

Enhanced Dynamic XY (EDXY) scheme is introduced in [11] where the congestion information of a router is propagated

to its row and column nodes via separate wires. Non-local information provided by these wires is used only when the packet is one row or column away from the destination. This scheme provides better performance than DyXY and can tolerate some faults in the network.

Both DyXY and EDXY routing algorithms have better performance than XY under all traffic patterns except for the uniform traffic. However, no matter in DyXY, EDXY or other routing schemes, the congestion information is mostly based on the buffer statuses at adjacent routers. These routers cannot provide the further information about the congestion status along the possible paths. This may lead the packet in a wrong direction that congestion cannot be avoided. It is insufficient in finding the less congested path.

III. PATH-AWARE ROUTING SCHEME

A. The Basic Idea of PARS

The objective of the Path-Aware Routing Scheme (PARS) is to exploit more comprehensive congestion information of the network. This is achieved by using a congestion aware subnetwork (CAS) which includes the congestion information of certain routers. These routers are selected in such a way as to favor adaptive routing algorithms. Adaptiveness helps to avoid hotspots and faulty components while reducing blocking probability of packets [12]. The degree of adaptiveness is a metric for measuring the adaptiveness of the routing algorithm [13]. It represents the number of possible minimal paths between any source and destination nodes. The degree of adaptiveness P can be defined as:

$$P = \frac{(\Delta x + \Delta y)!}{\Delta x! \Delta y!}$$

The source and destination addresses are represented by (X_s, Y_s) and (X_d, Y_d) , while Δx and Δy indicates the distance between source and destination in X and Y directions, respectively.

The degree of adaptiveness of each router is shown in Figure.1, when a packet travels between the source node $(0, 0)$ and destination node $(3, 3)$. All routers except the source and the destination nodes are divided into five groups according to their distances to the destination (e.g. 1-hop, 2-hop, 3-hop). Every router with the greatest degree of adaptiveness is highlighted in each group. These routers are more influential as many packets pass through them. By increasing the probability of passing packets through these routers, the choices of alternative paths are intensified when needed. Thus, the congestion information of the paths which pass through these routers is more valuable and can be used as a metric in the routing decision. In Fig. 1 several paths passing through these routers are marked from the source to the destination node. By comparing the congestion information of the marked paths (i.e. that is available at router $(1, 0)$ and $(0, 1)$) the routing unit can select a less congested path.

B. Congestion Aware Subnetwork

The basic idea of PARS is to route packets based on the congestion status of the potential routers with a larger degree of adaptiveness. For the source and destination nodes which are not in the same row or column, the orientations of the potential paths can be divided into Northeast, Northwest, Southeast, and Southwest. In each router, there are four two-bit registers called C-register to record the congestion status of these four directions. The congestion status is represented by the status of buffer. The first bit of the C-register is computed by using the OR and AND logic gates in each router. After computation, this bit will also be sent to the next router along the direction. The second bit travels from the previous router in the corresponding directions. Figure. 2 shows the congestion-aware subnetwork oriented to the Router 12 while Figure. 3 shows the way of recording the congestion values in the Router 12.

For example, the first bit of the register of in NW direction at Router 12 is coming from

$$\{S(11) \text{ and } E(16)\} \text{ or } \{S(17) \text{ and } E(16)\}.$$

S(11) and E(16) stands for the congestion statuses of the Router 11's south port and Router 16's east port. These two ports compose a path to the Router 16 which is two hops away from the source node and with a higher degree of adaptiveness. While S (17) and E (16) compose another path towards the same router. After ORing them together, the result can indicate the congestion information of the paths.

The congestion values propagate through the congestion aware subnetwork. Figure. 3 shows how the congestion information travels to the Router 12. The arrows stand for the propagation direction of the corresponding imports' congestion value. The black lines mean those congestion values are used for generating the first bit of the Router 12's congestion propagating value. The second bits are directly passed from the previous router in the directions towards the destination nodes. For instance, the second bit of the Router 12's congestion propagating value in the NW comes from the first bit of the Router 16's congestion propagation in the NW direction. The zebra-strip lines indicate where the second bits come from.

C. The Routing Scheme

When a new packet reaches a router, the corresponding imports' congestion statuses and the congestion propagation information of two minimal adjacent routers are compared to find a proper path. These two adjacent routers have greatest degrees of adaptiveness, so their congestion statuses become the most significant bit of congestion value. The corresponding direction's congestion propagating values of these two routers are as the less important two bits. According to the different position of the destination, there are different approaches to compare the congestion information. Figure. 4 shows four different situations have occurred. The PARS routing algorithm is shown in Figure. 5.

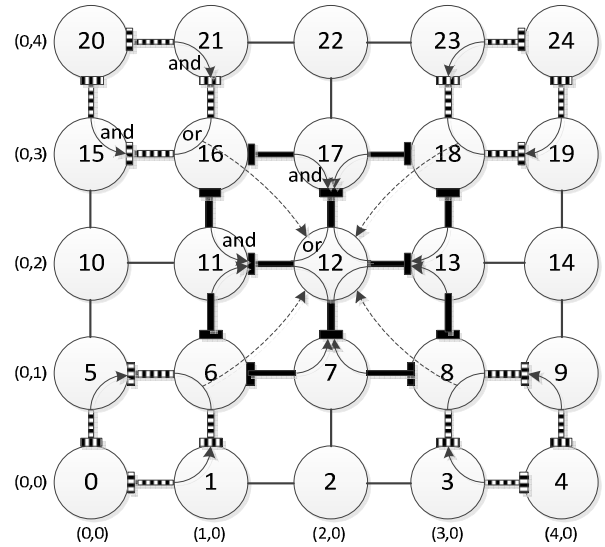


Figure 2. Congestion Aware Subnetwork(Oriented to Router 12)

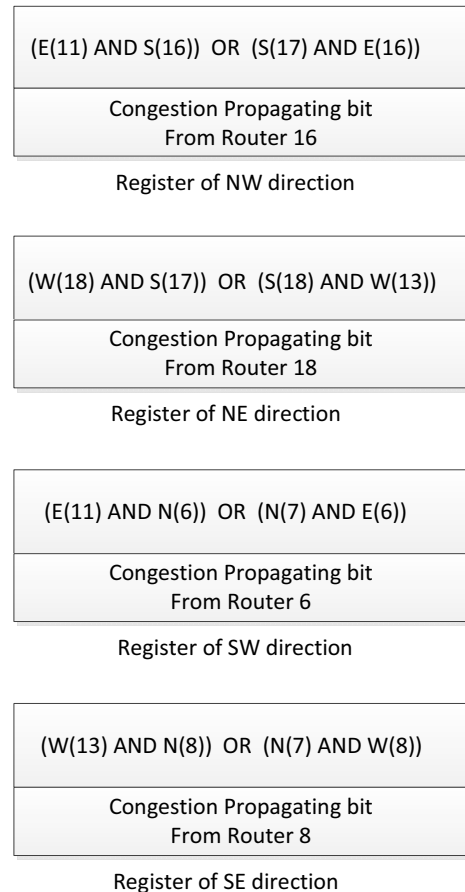


Figure 3. The Congestion Propagation Value Registers in Router 12

If the distance to the destination along each of the X (Δx) and Y direction (Δy) are greater than two hops, as presented in Figure. 4(a), 3-bit congestion value in the north direction is compared with 3-bit congestion value in the west direction. Finally, the packet is sent to a less congested path. As Figure. 4(b) shows when the distances along both the X and Y direction are equal to two, the first two bits of congestion values are compared with each other.

However, when $\Delta x \neq \Delta y$ and one of them is smaller than 3, the situation is noteworthy. Figure 4(c) and Figure 4(d) show how the situations may look like. In the direction which is closer to the destination, the number of bits used for comparison is the same as the minimum distance on X or Y dimension. The rest of the bits should be set to 1. It avoids the routing based on the congestion information of unrelated routers. In the other direction, one more bit should be considered to gain the accurate congestion information. If the number of bits is still less than three, the rest bits are also set to 1. If the corresponding bits of the C-registers are equal, PARS selects the output port according to the remaining distances along each direction. The packets are sent through a direction with greater distance (i.e. higher adaptiveness) in order of increase the number of alternative paths.

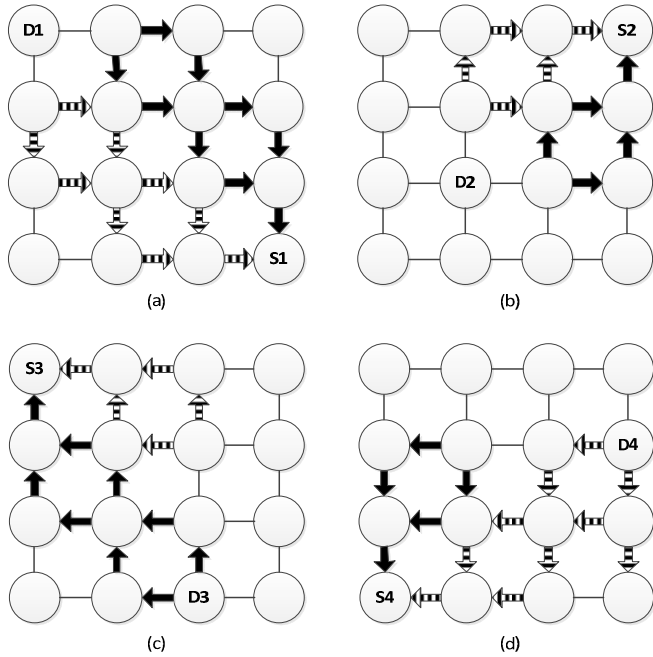


Figure 4. The converging paths in different situation (a) S1-> D1 $\Delta x = \Delta y = 3$; (b) S2-> D2 $\Delta x = \Delta y = 2$; (c) S3-> D3 $\Delta x = 2 \Delta y = 3$; (d) S4-> D4 $\Delta x = 3 \Delta y = 2$;

IV. EVALUATION

A. Methodology

This simulation is performed in the Hermes simulator [16][17]. The Hermes simulator is a platform to evaluate NoCs that are described in VHDL. It allows the

```

dX=ABS(Xcurrent - Xdestination)
dY= ABS(Ycurrent - Ydestination)
num = MIN (dX, dY)
flag = 1 when dX > dY else 0

If (dX ≥ 3 and dY ≥ 3) then
    compand_X <= congestion_X
    compand_Y <= congestion_Y
Elsif (dX = dY < 3) then
    compand_X (2 : 3 - num) <= congestion_X (2 : 3 - num)
    compand_X (3 - num : 0) <= (others=>'1')
    compand_Y (2 : 3 - num) <= congestion_Y (2 : 3 - num)
    compand_Y (3 - num : 0) <= (others=>'1')
Elsif (dX != dY and flag = 1) then
    compand_X (2 : 2 - num) <= congestion_X (2 : 2 - num)
    compand_X (2 - num : 0) <= (others=>'1')
    compand_Y (2 : 3 - num) <= congestion_Y (2 : 3 - num)
    compand_Y (3 - num : 0) <= (others=>'1')
Elsif (dX != dY and flag = 0) then
    compand_X (2 : 3 - num) <= congestion_X (2 : 3 - num)
    compand_X (3 - num : 0) <= (others=>'1')
    compand_Y (2 : 2 - num) <= congestion_Y (2 : 2 - num)
    compand_Y (2 - num : 0) <= (others=>'1')
End If
USE compand_X and compand_Y to choose the destination port

```

Figure 5. PARS Routing Algorithm

manipulation of the configuration parameters, customizing the topology of the network and the establishing routing algorithm serving our purposes extremely well.

We tested the proposed algorithm with different packet injection rates under three different traffic patterns. With the used patterns, we are able to verify the efficiency of the proposed algorithm. The efficiency of NoCs is measured by latency. Latency is the time interval between the first flit of a packet driven onto the network and the last flit of the packet coming out of the network. To improve our simulation results, we used the following technique to avoid deadlocks (as presented in [18]): if the destination router is east to the source router, the packets are obligated to send through virtual channel 1. On the opposite, if the destination router is west to the source, only virtual channel 2 is available. If the source and destination are in the same column, then the packets are allowed to travel through either channel. Thus, deadlock avoidance is guaranteed.

We simulated and analyzed our algorithm in a 8×8 2D mesh network. The configurations of the network are shown in Table. 1. The smallest unit of data that we used was in the simulation was a flit. Every packet consists of flits. In the simulation, if more than 40% of buffer is occupied, the buffer is considered as congested and the signal stands for congestion is set to 1.

TABLE I. Baseline network configuration

Characteristic	Baseline
Topology	8 x 8 2D Mesh
Routing	Minimal, fully-adaptive, reserved VC deadlock avoidance
Per-hop Latency	5 cycles
Virtual Channel	2
Flit Buffers	8
Packet Length	5
Traffic Load	Transpose, Uniform, Hotspot
Analyzed Packets	3000

B. Traffic Scenarios

We evaluation the PARS routing algorithm using three standard traffic patterns: transpose traffic, uniform traffic and hotspot traffic. Transpose traffic implies the destination of the package which generates in the node (i, j) is the node (j, i) evermore [19]. Under uniform traffic, every nodes send packets to the other nodes randomly with the same probability. Each time the destination is indiscriminately generated. Hotspot traffic is based on the uniform traffic. Some nodes are taken as the hotspots; other nodes have more probabilities to send data to hotspots compare with the normal nodes. In this simulation, nodes (3, 3), (3, 4), (4, 3) and (4, 4) which locate in the center of NoCs are taken as hotspots. The hotspot traffic is closest to the real application among these patterns. So we take two experimental groups of hotspot traffic. The probabilities of sending data to hotspot are respectively 5% and 10%.

C. Simulation Results

Results of simulation under transpose traffic are shown in the Figure. 6. Figure. 7 indicates the results under the uniform traffic. Figure. 8 and Figure. 9 show the latency vs. the packets injection rate under respectively 5% and 10% hotspot traffic. No matter in which traffic profile, PARS always represents the best performance.

The results under transpose traffic profile show better improvement. The latency is reduced by 27.5% compare with EDXY. Under uniform traffic, PARS outperforms DyXY by 56.5% and EDXY by 20.8%. The latency reductions compare with EDXY under 5% hotspots and 10% hotspot traffic are respectively 10.9% and 14.5%.

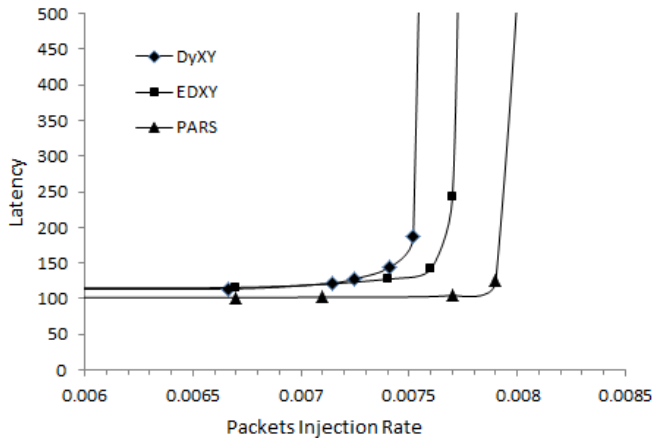


Figure 6. Latency vs. pir. under transpose traffic

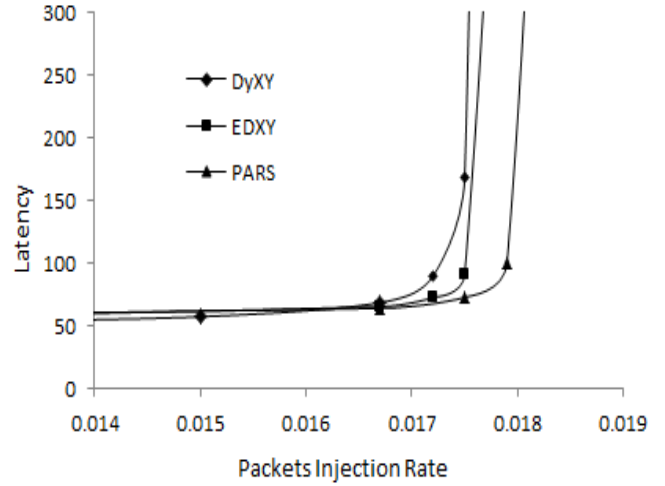


Figure 7. Latency vs. pir. under uniform traffic

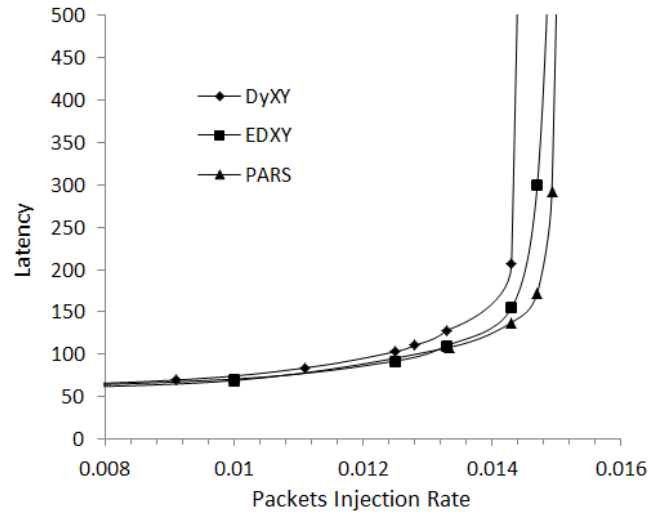


Figure 8. Latency vs. pir. under 5% Hotspot traffic

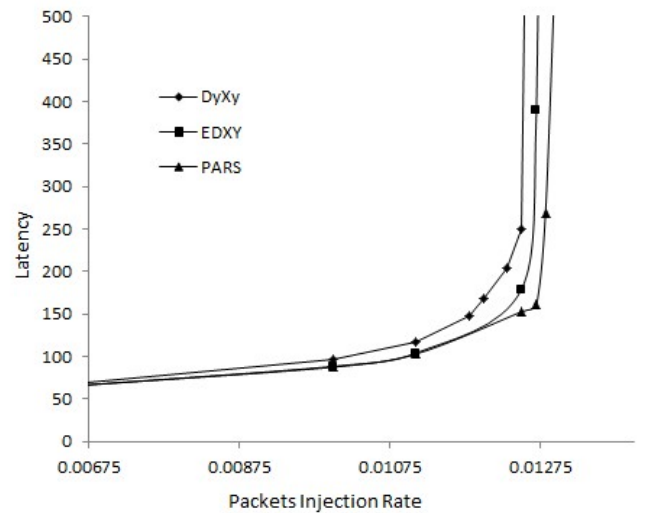


Figure 9. Latency vs. pir. under 10% Hotspot traffic

D. Area Overhead And Power Dissipation

The area overhead and power dissipation of the proposed algorithm are evaluated with Cadence Encounter RTL compiler using UMC L90 Standard Performance Low-K library. We compared the area and power dissipations of three different schemes, DyXY, EDXY, and PARS implemented in NoC. The area overhead of PARS compared with DyXY and EDXY are 5.2% and 2.5%, respectively. The area overhead of logical components in the congestion aware subnetwork (CAS) is negligible. The area overhead of CAS is mostly caused by wires. In total it only occupies 2.4% area in the whole design.

The power dissipation of DyXY, EDXY and PARS are about the same. Under the same circumstance, PARS consumes 787.4mW while DyXY and EDXY consume 796.4mW and 792.8mW respectively. Comparing with DyXY and EDXY techniques, PARS consumes 1.132% and 0.061% less power.

V. CONCLUSIONS AND FUTURE WORK

We introduced a novel congestion-aware routing algorithm PARS for NoC based systems. PARS uses congestion-aware subnetworks to provide more accurate information on current network traffic status. It balances the workload more evenly and thus decreases the latency of the algorithm under all the tested traffic profiles. The area overhead of PARS is affordable and power consumption is reasonable. Furthermore, owing to the high coverage of possible paths, PARS also has great potential to improve fault-tolerance by helping to discover the faulty links. This remains to be investigated as future work.

REFERENCES

- [1] D. Wu, B.M. Al Hashimi, M.T. Schmitz, "Improving routing efficiency for network-on-chip through contention-aware input selection", in proceedings of ASPDAC, pp. 36-41, 2006.
- [2] M. Daneshalab, M. Ebrahimi, P. Liljeberg, J. Plosila, and H. Tenhunen, "Memory-Efficient On-Chip Network with Adaptive Interfaces," IEEE Transaction on Computer-Aided Design of Integrated Circuits and Systems (IEEE-TCAD), Vol. 31, No. 1, pp. 146-159, Jan 2012.
- [3] A. Agarwal, B. H. Lim, D. Kramz, and J. Kubiatowicz, "APRIL: A processor architecture for multiprocessing multiprocessor," in Proc. of the 17th International Symposium on Computer Architecture, pp.148-159, May 1990.
- [4] M. Daneshalab, et. al., "Adaptive Input-output Selection Based On-Chip Router Architecture," Journal of Low Power Electronics (JOLPE), Vol. 8, No. 1, pp. 11-29, 2012.
- [5] P. Gratz, B. Grot, S. W. Keckler, "Regional Congestion Awareness for Load Balance in Networks-on-Chip," in Proc. of the 14th International Symposium on High-Performance Computer Architecture, pp.203-214, Feb. 2008.
- [6] A. Mejia, M. Palesi, J. Flich, S. Kumar, P. López, R. Holmark and José Duato, "Region-Based Routing: A mechanism to Support Efficient Routing Algorithm in NoCs", Very Large Scale Integration (VLSI) Systems, VOL. 17, NO. 3, March 2009.
- [7] M. Ebrahimi, M. Daneshalab, F. Farahnakian, P. Liljeberg, J. Plosila, M. Palesi, and H. Tenhunen, "HARAQ: Congestion-Aware Learning Model for Highly Adaptive Routing Algorithm in On-Chip Networks," in Proceedings of 6th ACM/IEEE International Symposium on Networks-on-Chip (NOCS), May. 2012, Denmark.
- [8] A. Kohler, and M.Radetski, "Fault-tolerant architecture and deflection routing for degradable NoC switch", in Proc of NoC'09, pp. 22-31, 2009.
- [9] M. Ebrahimi, et. al., "LEAR – A Low-weight and Highly Adaptive Routing Method for Distributing Congestions in On-Chip Networks," in Proceedings of 20th IEEE Euromicro Conference on Parallel, Distributed and Network-Based Computing (PDP), pp. 520-525, Feb. 2012.
- [10] M.Li, Q.-A. Zeng, and W. –B. Jone, "DyXY – a proximity congestion-aware deadlock-free dynamic routing method for networks on chip," in Proc. of Design Automation Conference, pp. 849-852, 2006.
- [11] P. Lotfi-Kamran, A.M. Rahmani, M. Daneshalab, A. Afzali-Kusha, Z. Navabi, "EDXY – A low cost congestion-aware routing algorithm for networks-on-chips," in Journal of Systems Architecture 56, pp.256- 264, 2011.
- [12] C.J. Glass, L.M. Ni, "The turn model for adaptive routing," Journal of the ACM 41, pp 894-902, September 1994.
- [13] G.M. Chiu, "The odd-even turn model for adaptive routing," IEEE transactions on Parallel and Distributed Systems 11, pp.729-738, July 2000.
- [14] S. Ma, N.E. Jerger, Z. Wang, "DBAR: An efficient Routing Algorithm to Support Multiple Concurrent Applications in Networks-on-Chip," in ISCA 11' June 2011.
- [15] M. Ebrahimi, et. al, "CATRA-Congestion Aware Trapezoid-based Routing Algorithm for On-Chip Networks," in Proceedings of 15th ACM/IEEE Design, Automation, and Test in Europe (DATE), pp. 320-325, Mar. 2012.
- [16] C. A. Zeferino, A. A. Susin, "SoCIN: A parametric and Scalable Network-on-Chip," in Proc. of the 16th Symposium on Integrated Circuits and Systems Design, 2009
- [17] C. A. Zeferino, M. E. Kreutz, A. A. Susin, "RASoC: A Soft- Core for Networks-on-Chip," in Proc. of the Design, Automation and Test in Europe Conference and Exhibition Designers' Forum ,2004
- [18] L.Mo. Ni, P.K. Mckinley, "A survey of wormhole routing techniques in direct networks," IEEE Computer pp.62-76, 1993
- [19] G. Ascia, V. Catania, M. Palesi, D. Patti, "Implementation and Analysis of a New Selection Strategy for Adaptive Routing in Networks-on-Chip," IEEE transactions on computers, VOL. 57, No. 6, June 2008.
- [20] M. Ebrahimi, et. al., "HAMUM – A Novel Routing Protocol for Unicast and Multicast Traffic in MPSoCs," in Proceedings of 18th IEEE Euromicro Conference on Parallel, Distributed and Network-Based Computing (PDP), pp. 525-532, February 2010, Italy.
- [21] M. Ebrahimi, M. Daneshalab, S. Mohammadi, A. Afzali-Kusha, H. Tenhunen, "An Efficient Dynamic Multicast Routing Protocol for Distributing Traffic in NOCs," in Proceedings of 12th ACM/IEEE Design, Automation, and Test in Europe Conference (DATE), pp. 1064 - 1069 , April 2009, France.
- [22] E. Waingold, M. Taylor, D. Srikrishna, V. Sarkar, W. Lee, V. Lee, J. Kim, M. Frank, P. Finch, R. Barua, J. Babb, S. Amarasinghe, and A. Agarwal. Baring It All to Software: RAW Machines. IEEE Computer, 30(9):86-93, September 1997.
- [23] S. Vangal, J. Howard, G. Ruhl, S. Dige, H. Wilson, J. Tschanz, D. Finan, P. Iyer, A. Singh, T. Jacob, S. Jain, S. Venkataraman, Y. Hoskote, and N. Borkar. An 80-Tile 1.28 TFLOPS Network-on-Chip in 65nm CMOS. In IEEE International Solid-State Circuits Conference, pages 98-99, February 2007.
- [24] M. Dehyadegari, M. Daneshalab, M. Ebrahimi, J. Plosila, and S. Mohammadi, "An Adaptive Fuzzy Logic-based Routing Algorithm for Networks-on-Chip," in Proceedings of 13th IEEE/NASA-ESA International Conference on Adaptive Hardware and Systems (AHS), pp. 208-214, June 2011, USA.
- [25] K. Sankaralingam, et. al., "The Distributed Microarchitecture of the TRIPS Prototype Processor", In International Symposium on Microarchitecture, pages 480-491, 2006.
- [26] J. Duato, S. Yalamanchili and L. Ni. Interconnection Networks – An Engineering Approach. Elsevier Science (USA), pages 179-184, 2006.