

ArR-DTM: A routing-based DTM for 3D NoCs by adaptive degree regulation

Gaizhen Yan^{1,2a)}, Ning Wu¹, Fen Ge¹, Hao Xiao¹, and Fang Zhou¹

¹ College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

² School of Electrical and Electronic Engineering, Anhui Science and Technology University, Chuzhou, 233100, China

a) xucs.yan@126.com

Abstract: Three dimensional (3D) networks-on-chip (NoCs) is beneficial to performance improvement, but suffers severe thermal issues. Because of the fast packet switching activity and small area, an NoC router might even show higher power density than a process element, quite easily forming a thermal hotspot. Throttling based dynamic thermal managements (DTMs) help for fast cooling, but easily results in network congestion and introduces large cost. This paper proposed an efficient routing-based DTM scheme (ArR-DTM), which tries to balance the thermal distribution by providing complementary adaptive degree for lateral planes and finish thermal migration by regulating the adaptive degree in vertical direction. Experiments show that, under the same thermal limit, there is 13.1% to 23.8% performance improvement compared to the fully throttling based DTMs.

Keywords: 3D NoC, DTM, Adaptive routing

Classification: Integrated circuits

References

- [1] B. S. Feero, *et al.*: “Networks-on-chip in a three-dimensional environment: A performance evaluation,” *IEEE Trans. Comput.* **58** (2009) 32 (DOI: [10.1109/TC.2008.142](https://doi.org/10.1109/TC.2008.142)).
- [2] K.-Y. Jheng, *et al.*: “Traffic-thermal mutual-coupling co-simulation platform for three-dimensional network-on-chip,” *IEEE Int. Symp. VLSI Des., Autom. and Test* (2010) 135 (DOI: [10.1109/VDAT.2010.5496709](https://doi.org/10.1109/VDAT.2010.5496709)).
- [3] N. Dahir, *et al.*: “Design and implementation of dynamic thermal-adaptive routing strategy for networks-on-chip,” *Int. Conf. PDP* (2014) 384 (DOI: [10.1109/PDP.2014.70](https://doi.org/10.1109/PDP.2014.70)).
- [4] C.-H. Chao, *et al.*: “Transport layer assisted routing for run-time thermal management of 3D NoC systems,” *ACM Trans. Embedded Comput. Syst.* **13** (2013) 11 (DOI: [10.1145/2501626.2512468](https://doi.org/10.1145/2501626.2512468)).
- [5] Mobile intel pentium 4 processor-M datasheet, <http://www.intel.com> (2003).
- [6] L. Shang, *et al.*: “Thermal modeling, characterization and management of on-chip networks,” *IEEE/ACM Int. Symp. on Microarchitecture* (2004) 67 (DOI: [10.1109/MICRO.2004.35](https://doi.org/10.1109/MICRO.2004.35)).
- [7] K. C. Chen, *et al.*: “Topology-aware adaptive routing for nonstationary irregular mesh in throttled 3D NoC systems,” *IEEE Trans. Parallel Distrib.*

- Syst. **24** (2013) 2109 (DOI: [10.1109/TPDS.2012.291](https://doi.org/10.1109/TPDS.2012.291)).
- [8] A. M. Rahmani, *et al.*: “Design and management of high-performance, reliable and thermal-aware 3D networks-on-chip,” IET Circuits Devices Syst. **6** (2012) 308 (DOI: [10.1049/iet-cds.2011.0349](https://doi.org/10.1049/iet-cds.2011.0349)).
 - [9] Z. Wu, *et al.*: “The adaptive thermal and traffic-balanced routing algorithm based on temperature analysis and traffic statistics,” IEICE Electron. Express **12** (2015) 20150101 (DOI: [10.1587/elex.12.20150101](https://doi.org/10.1587/elex.12.20150101)).
 - [10] W. Huang, *et al.*: “HotSpot: A compact thermal modeling methodology for early-stage VLSI design,” IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **14** (2006) 501 (DOI: [10.1109/TVLSI.2006.876103](https://doi.org/10.1109/TVLSI.2006.876103)).
 - [11] G. M. Chiu: “The odd-even turn model for adaptive routing,” IEEE Trans. Parallel Distrib. Syst. **11** (2000) 729 (DOI: [10.1109/71.877831](https://doi.org/10.1109/71.877831)).
 - [12] S. R. Vangal, *et al.*: “An 80-tile sub-100-w teraflops processor in 65-nm cmos,” IEEE J. Solid-State Circuits **43** (2008) 29 (DOI: [10.1109/JSSC.2007.910957](https://doi.org/10.1109/JSSC.2007.910957)).

1 Introduction

Most recently, 3D NoC is becoming an appealing interconnection solution for 3D integrated circuits (IC), because of the better system scalability [1]. However, in 3D IC, vertically piling up more active devices greatly increases the power density, resulting severe thermal issues [2]. NoC communication contributes a significant part to the total chip power consumption, especially for the highly parallelizable and communication centric applications [3]. Because of the relatively small area, thermal hotspots are easily formed at the NoC routers [4].

Throttling based DTMs have been proposed to enhance the thermal safety of 3D NoC. The state-of-art related schemes mainly include global throttling (GT) [5], distributive throttling (DT) [6] and vertical throttling (VT) [4, 7]. When a node is in thermal emergent state, they would globally, distributively or vertically throttle the router nodes, respectively. In 3D NoC, for fast cooling, most of the schemes are fully throttling based. But the topology becomes time-variant. Maintenance of topology information and topology-aware adaptive routing introduces large cost [7]. Fully throttling a node might also break the traffic dependency chain and then the whole application would stall.

By contrast, routing based DTMs [8, 9] alleviate the thermal issues by balancing the thermal distribution. Rahmani et al. [8] proposed a thermal-aware routing strategy for bus-mesh hybrid 3D NoC. Source or destination layer will be adaptively selected to route in order to migrate traffic to the layer closer to the heatsink. The adaptiveness provided is seemingly limited and only two routing states are available. Our previous work [9] has proposed a thermal and traffic balanced algorithm based on the global sharing information, but introduces additional bandwidth consumption, and thus additional power consumption.

Apart from the adaptiveness and cost, key problem of routing-based thermal management might be the design dilemma in traffic balance and thermal balance. Traffic balance is necessary for high performance, but thermal balance tends to migrate traffic to one layer, leading to an extremely unbalanced traffic distribution. To avoid the additional cost caused by thermal information delivery, increase path

diversity for performance improvement and dynamically make a tradeoff between traffic balance and thermal balance, ArR-DTM proposed in this paper tries to implement thermal management by properly designing and dynamically regulating adaptive degree of flows towards different directions.

2 Problem description

We describe the thermal balance and traffic balance based on the thermal model of 3D NoC, which is adapted from the architecture-level thermal model applied in Hotspot [10], as shown in Fig. 1.

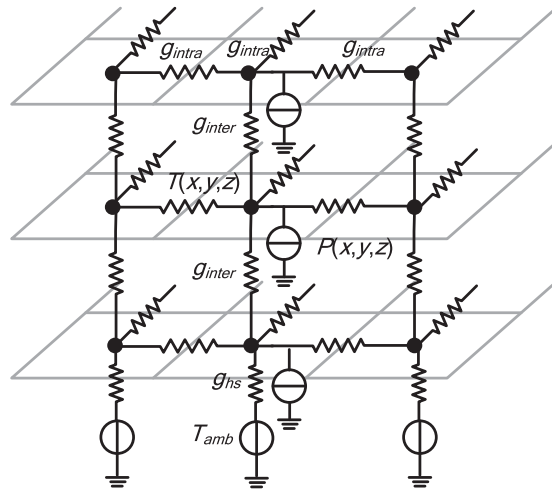


Fig. 1. Thermal model of the 3D NoC

In 3D integrated circuits, the inter-wafer distance is about 50 μm , but the typical processor usually spans more than one millimeter. Therefore, thermal coupling in lateral directions is ignored in this paper. Giving the thermal conductance between two vertical neighbored nodes g_{inter} , thermal conductance of the heatsink g_{hs} and the ambient temperature T_{amb} , the temperature of node at (x, y, z) can be expressed as Eq. (1) and Eq. (2).

$$T_{(x,y)}^1 = T_{amb} + \frac{1}{g_{hs}} \sum_{k=1}^N P_{x,y}^k \quad (1)$$

$$T_{(x,y)}^z = T_{(x,y)}^{z-1} + \frac{1}{g_{inter}} \sum_{k=z}^N P_{x,y}^k \quad (2 \leq z \leq N) \quad (2)$$

where $T_{(x,y)}^z$ and $P_{(x,y)}^z$ are the temperature and power consumption of the node at location (x, y, z) , respectively. N is the farthest layer from the heatsink.

For thermal balanced routing, to meet $T_{(x,y)}^1 = T_{(x,y)}^2 = \dots = T_{(x,y)}^N$, Eq. (3) should be satisfied.

$$P_{(x,y)}^2 = P_{(x,y)}^3 = \dots = P_{(x,y)}^N = 0 \quad (3)$$

That is, all of the traffic in the upper layers should be migrated to the layer closest to the heatsink. Then, 3D NoC would degrade to 2D NoC, and performance will be severely degraded.

For traffic balanced routing, it is expected that traffic in each layer distributed equally, that is, $P_{(x,y)}^1 = P_{(x,y)}^2 = \dots = P_{(x,y)}^N$. Then, thermal gradient between two vertical neighbored nodes can be expressed as Eq. (4).

$$T_{(x,y)}^z - T_{(x,y)}^{z-1} = \frac{N - z + 1}{g_{inter}} P_{(x,y)}^1 \quad (4)$$

The design dilemma in traffic balance and thermal balance strongly motivated us to develop an adaptive routing, which dynamically migrates traffic to the lower layers whenever thermal dissipation is required.

3 Principle of the proposed ArR-DTM

This paper investigates the DTM for mesh-based 3D NoC. The heatsink is placed closer to the bottom layer, whose z coordinate is 0. We propose adaptive degree regulated routing algorithm to realize dynamic thermal management, as shown in Fig. 2, which includes two main ideas. The first idea is to provide complementary adaptiveness for different layers, aiming at reducing the thermal unbalance within a lateral plane. The second is dynamically regulating the adaptive degree in vertical directions based on the thermal emergent state.

3.1 Adaptiveness balance in lateral plane

Odd-even (OE) turn model [11] based adaptive routing can provide relatively even adaptive degree. Based on the OE turn rules, north and south direction are forbidden at even columns for eastward flows while forbidden at odd columns for westward flows, as shown in Fig. 3(a). But for eastward flows, north and east directions are allowed at source columns, even at the even columns. This unbalanced adaptiveness would make the traffic in even columns be slightly higher than the one in odd columns. The resulting thermal difference would be enlarged when multiple layers pill up.

We are motivated to apply different adaptive routing to break the regularity of the forbidden turns in different layers. In this paper, we draw four variants of the OE turn model by rotating the original one 0° , 90° , 180° , and 270° , and then apply them in a lateral plane for every four layers, as described by line 4–11 in Fig. 2. Then, forbidden turns in different layers can complement with each other. Turns forbidden in one layer might find their way in other layers, as shown in Fig. 3(b). This is not only beneficial to balance the thermal distribution of the topmost layer whose temperature is the hottest among all layers, but also beneficial to extend the path diversity which helps for performance improvement.

3.2 Adaptive degree regulation in vertical direction

There is a dilemma in vertical adaptiveness design. Even traffic distribution of each layer, which is better for the network capacity utilization, requires equal adaptiveness in up and down directions. But even thermal distribution should impart higher adaptiveness to downward flows than to upward ones, which tends to result in traffic congestion in lower layers. Therefore, we develop rules for both traffic balanced (Tr_B) and thermal balanced (Th_B) routing, and adaptively regulating the rules based on the thermal state the system is in. They are described by line 13–29 in Fig. 2.

Algorithm: Adaptive degree regulated routing algorithm

Input: $Src(x_s, y_s, z_s); Dest(x_d, y_d, z_d); Cur(x_c, y_c, z_c)$

TherEmerg // Flags to indicate Thermal Emergent State

L_{th} // Layer to split hot region and cool region

Output: *Avail_dirs*

```

1  Avail_dirs  $\leftarrow \emptyset$ ;
2   $e_x = d_x - c_x; e_y = d_y - c_y; e_z = d_z - c_z$ ;
3  IF  $((e_x \neq 0) \parallel (e_y \neq 0))$ 
4  |   IF  $(z_c \% 4 = 0)$ 
5  |   |   THEN Add lateral directions to Avail_dirs based on OE_270°
6  |   ELSE IF  $(z_c \% 4 = 1)$ 
7  |   |   THEN Add lateral directions to Avail_dirs based on OE_180°
8  |   ELSE IF  $(z_c \% 4 = 2)$ 
9  |   |   THEN Add lateral directions to Avail_dirs based on OE_90°
10 |   ELSE IF  $(z_c \% 4 = 3)$ 
11 |   |   THEN Add lateral directions to Avail_dirs based on OE_0°
12 |   END IF
13 |   IF ( $\neg TherEmerg$ )
14 |   |   IF  $(e_z > 0)$ 
15 |   |   |   IF  $((z_d \% 2 = 1) \parallel (e_z \neq 1))$ 
16 |   |   |   THEN Add UP to Avail_dirs
17 |   |   |   IF  $((z_c \neq z_s) \&\& (z_c \% 2 \neq 1))$ 
18 |   |   |   THEN Delete lateral directions in Avail_dirs
19 |   |   ELSE IF  $(e_z < 0)$ 
20 |   |   |   Add DOWN to Avail_dirs
21 |   |   |   IF  $(z_c \% 2 = 1)$ 
22 |   |   |   THEN Delete lateral directions in Avail_dirs
23 |   |   END IF
24 |   ELSE
25 |   |   IF  $((z_s \geq L_{th}) \&\& (z_d \geq L_{th}))$ 
26 |   |   |   IF  $(z_c \neq 0)$  THEN Add DOWN to Avail_dirs
27 |   |   |   ELSE
28 |   |   |   IF  $(e_z < 0)$  THEN Add DOWN to Avail_dirs
29 |   |   END IF
30 |   ELSE
31 |   |   IF  $(e_z > 0)$  THEN Add UP to Avail_dirs
32 |   |   ELSE IF  $(e_z < 0)$  THEN Add DOWN to Avail_dirs
33 |   |   ELSE THEN Add LOCAL to Avail_dirs
34 |   END IF

```

Fig. 2. Proposed adaptive degree regulated routing

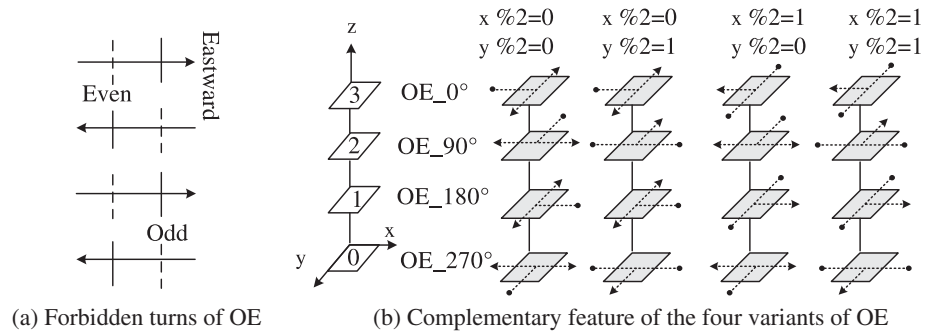


Fig. 3. Complementary adaptiveness in lateral plane

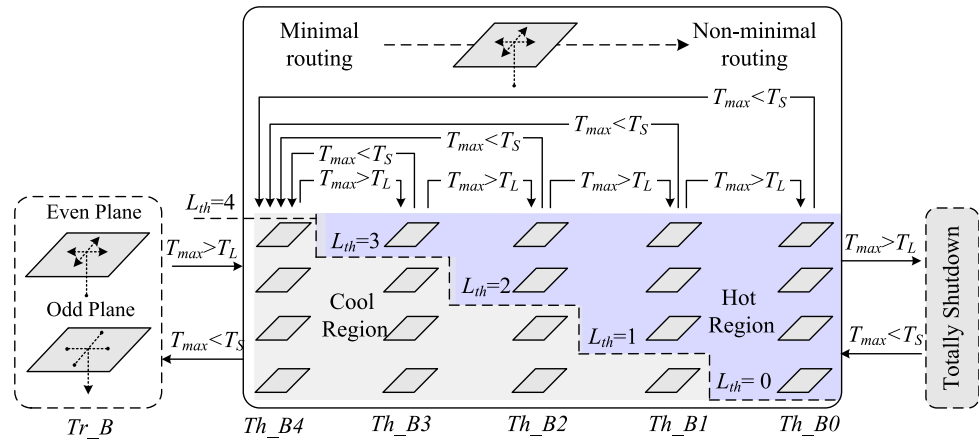


Fig. 4. Routing states and the regulation process

The total dynamic regulation process is shown in Fig. 4. The routing is initiated as traffic balance state (Tr_B). Up-lateral turns are forbidden in even layers and lateral-down turns are forbidden in odd layers, in order to give relatively even adaptiveness to up and down flows. Once the maximum temperature soars up above the thermal limit T_L , the routing algorithm enters thermal balanced state (Th_B). Up-lateral turns are forbidden in all layers and lateral directions are allowed for downward flows to impart higher adaptive degree for downward flows.

Under the thermal balanced state, we further divide the layers into hot region and cool region by threshold layer L_{th} . Traffic between the nodes in hot region are allowed to use non-minimal routing in vertical direction, in order to force more adaptiveness to the traffic in the layers far from the heatsink. If the temperature steadily keeps above T_L , L_{th} would gradually drop down to the bottommost layer. More and more traffic gets the chance to be routed to the lower layer for thermal dissipation. If the temperature drops below the thermal safe threshold T_S , L_{th} would immediately pops above the topmost layer for fast performance recovery, and finally the routing falls back to traffic balanced state. By this way, the proposed algorithm can quickly make a trade-off between thermal balance and traffic balance for the specific traffic load.

Note that, when the traffic load is out of the balance ability of ArR-DTM, that is, there is still thermal emergent node under Th_B0 state, the NoC communication would be totally disabled to enhance the thermal safety.

3.3 Deadlock avoidance strategy of ArR-DTM

The proposed ArR-DTM has two routing modes: traffic balanced routing (Tr_B) and thermal balanced routing (Th_B). Deadlock freeness can be guaranteed under the both routing modes. Anyway, because of different routing rules applied in vertical direction, deadlock and livelock quite possibly be met when the routing mode switches between the two. To solve the problem, ArR-DTM will not allow packet injection and routing mode switching until flits in the network have all been drained, whenever the routing mode should be changed between (Tr_B) mode and (Th_B) mode.

We use global barrier network based on the dimension-wise aggregation scheme to detect the traffic drained state, as shown in Fig. 5. When the routing

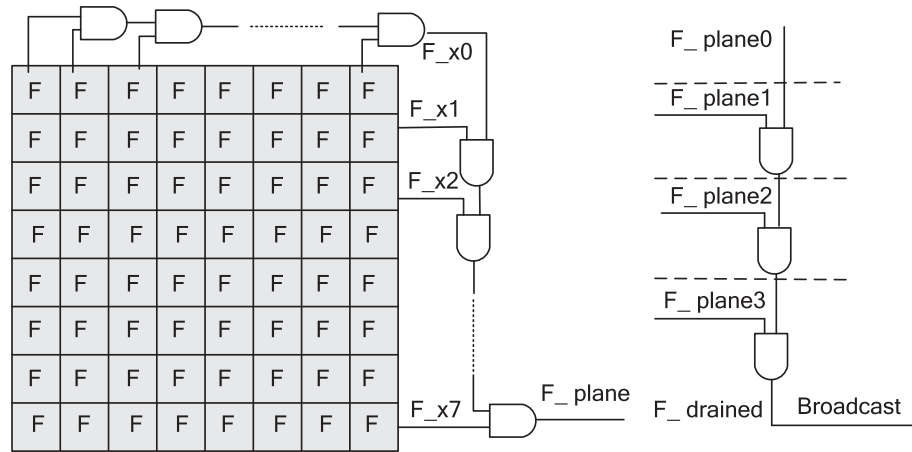


Fig. 5. Global barrier network to detect drained state

mode should be switched between (Tr_B) and (Th_B), the flag F indicating empty buffers at each router would be enabled and fed to the logic AND chain along x direction. The results would be then propagated along the y and z direction. The flag to indicate traffic drained state $F_drained$ will be broadcast to all the nodes along the reverse path. Therefore, for a topology with $m \times n \times k$ nodes, it takes at most $2(m + n + k)L$ time for flag aggregation and broadcast, where L is the propagation latency per hop. Based on our experiment, for a $8 \times 8 \times 4$ mesh network, the total traffic drained process lasts not more than 1000 cycles. Compared to the thermal control period (usually not less than 10 ms, 10^7 cycles for 1 GHz clock frequency), the performance cost caused by traffic draining process is negligible.

4 Experiments

We evaluate the ArR-DTM in the thermal-traffic co-simulation environment AccessNoxim [2]. Network topology is set as $8 \times 8 \times 4$ 3D mesh. Router buffer depth is set as 4 flits, and the flit size is randomly distributed between 2 to 8 flits. The traffic pattern is selected as uniform random. The packet injection rate is set to saturation point to meet the maximum system workload. Tile geometry and power model is the same as Intel 80-core processor [12].

4.1 Effectiveness of the adaptive degree regulation approach

With local buffer level selection strategy, we first validate the effectiveness of the proposed dynamic adaptive degree regulation policy. Layer traffic distribution and throughput under each routing state of ArR-DTM are shown in Fig. 6. Tr_B routing has the most balanced layer traffic distribution and thus highest throughput. But under Th_B routing, with the L_{th} moving towards the heatsink, more traffic are evicted to the bottom layer, and thus the throughput degrades more.

4.2 Cooling effect and performance

We have evaluated five DTMs: DLAR+GT, DLAR +DT, DLAR +VT, TAAR+VT, and the proposed ArR-DTM. Both DLAR [4] and TAAR [7] are topology aware

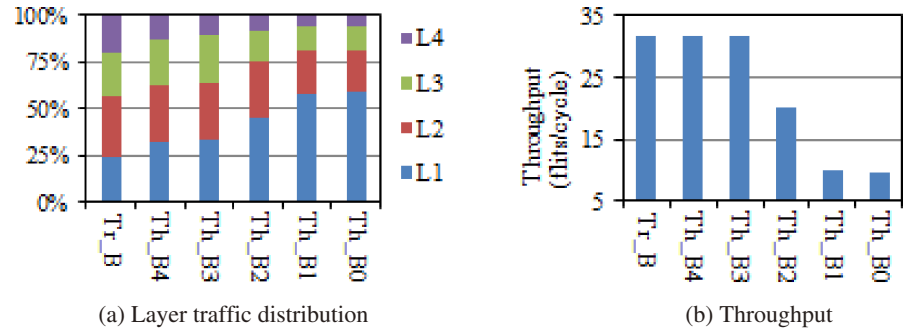


Fig. 6. Layer distribution and throughput under each routing state

adaptive routing, which is necessary for the throttling based DTMs. They check the routability before packets sending to the network based on the maintained topology information.

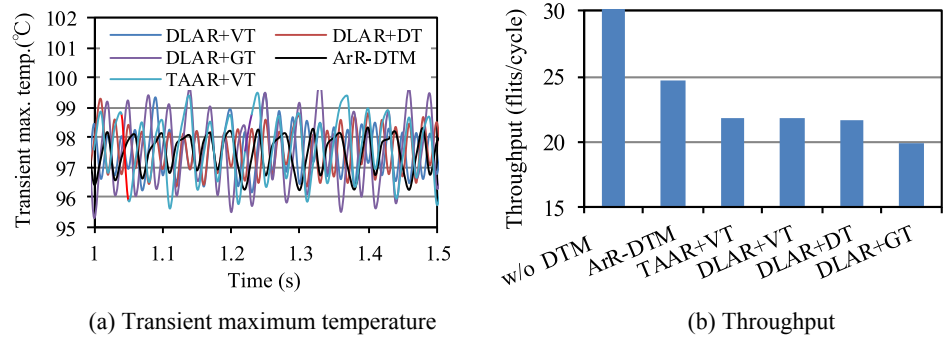


Fig. 7. Comparison of the five involved DTMs

Transient maximum temperature and throughput under the five DTMs are shown in Fig. 7(a) and Fig. 7(b). It can be seen that all the DTMs can limit the temperature below the hard thermal limit, 100°C. But the proposed ArR-DTM seems showing least thermal control oscillation. On the meanwhile, compared to the routing scheme without DTM, ArR-DTM degrades performance least among the involved DTMs. And compared to the other four DTMs, ArR-DTM can improve throughput by 13.1% to 23.8%.

4.3 Area cost

ArR-DTM introduces two additional modules: adaptive routing shown in Fig. 2 and the routing state control shown in Fig. 4. They are implemented in TSMC 90 nm CMOS technology, and the total area cost is shown in Table I. Compared to the additional cost of DLAR and TAAR under UMC 90-nm technology [7], the cost of ArR-DTM is quite negligible.

Table I. Area cost comparison of the involved schemes

Scheme	ArR-DTM	DLAR	TAAR
Area Cost (μm^2)	2851	12976	48546

5 Conclusion

In this paper, we have proposed a routing based thermal management scheme. Unlike the throttling based thermal management schemes, it keeps the original topology unchanged. The topology information need not be maintained anymore, thus saving the area cost. Experiments show that our proposed ArR-DTM scheme is effective for thermal management in 3D NoC. And compared to the fully throttling based DTMs, at most 23.8% performance improvement has been observed under the same thermal limit.

Acknowledgments

The authors acknowledge the financial support by National Natural Science Foundation of China (61376025, 61504059), the Jiangsu Innovation Program for Graduate Education (KYLX15_0283), and the Natural Science Foundation of Jiangsu Province (BK20160806).