

# Traffic- and Thermal-aware Adaptive Beltway Routing for Three Dimensional Network-on-Chip Systems

Kun-Chih Chen, Che-Chuan Kuo, Hui-Shun Hung, and An-Yeu (Andy) Wu

Graduate Institute of Electronics Engineering, National Taiwan University, Taipei, Taiwan (R.O.C.)

kunchih@access.ee.ntu.edu.tw

**Abstract**—The distribution of traffic and temperature in a high-performance three dimensional Network-on-Chip (3D NoC) system become more unbalanced because of chip stacking and applied minimal routing algorithms. To regulate the temperature under a certain thermal limit, the overheated nodes are usually throttled by run-time thermal management (RTM). Therefore, the network topology becomes a Non-Stationary Irregular Mesh (NSI-Mesh) and leads to heavy traffic congestion around the throttled nodes. Because of the traffic imbalance in the network, the system performance degrades sharply as temperature rises. In this paper, a Traffic- and Thermal-aware Adaptive Beltway Routing (TTABR) is proposed to balance both the distribution of the traffic and temperature in the network. The proposed TTABR can be applied to NSI-Mesh and regular mesh. The experimental results show that the proposed TTABR can achieve more balanced both traffic and temperature distribution, and the network throughput is improved by around 3.4~113% with less than 18% area overhead.

## I. INTRODUCTION

As the complexity of System-on-Chip (SoC) grows with Moore's law, the three dimensional Network-on-Chip (3D NoC) is promising to reduce wire complexity problems and provide larger bandwidth with lower power consumption [1]. For some high-performance NoC-based microprocessors such as Cray T3E [2], the minimal routing algorithm is a popular way to achieve high-speed requirement. However, the minimal routing algorithm leads to traffic congestion in the network and traffic load becomes unbalanced. The heavy traffic imbalance easily causes thermal hotspot in the 3D NoC because of die stacking. Therefore, the traffic and thermal imbalance problems become more severe in a high-performance 3D NoC system [3][4].

To solve the thermal problem in 3D NoC, a Thermal-Aware Vertical Throttling (TAVT) scheme was proposed in [4]. Based on the level of thermal emergency, TAVT determines different throttling level to perform temperature control. However, the TAVT makes the network topology become time-varying Non-Stationary Irregular Mesh (NSI-Mesh), as shown in Fig. 1(a). Because the routers cannot transmit any packet until the temperature becomes non-overheated, many packets are blocked in the network for a long time, and the throughput degrades rapidly. Therefore, the traffic load becomes more unbalanced in NSI-Mesh.

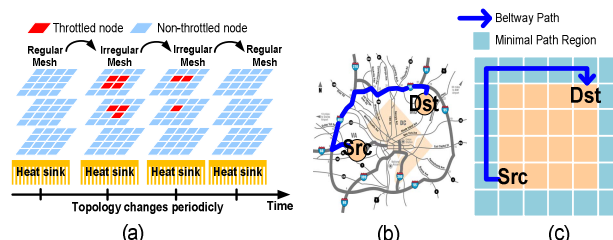


Fig. 1 (a) TAVT results in NSI-Mesh, (b) Beltway in Washington DC, and (c) Packet detours the minimal path region through beltway path.

In the modern public transportation system, the beltway system is a popular way to mitigate the traffic load in the downtown. The highway in Washington DC is an example, as shown in Fig. 1(b). To solve the problem of traffic- and thermal-imbalance in 3D NoC system, we follow the design concept of the Capital Beltway in Washington DC and propose a Traffic- and Thermal-aware Adaptive Beltway Routing (TTABR). TTABR provides an extra beltway path, which is a path beyond the minimal path region (*i.e.*, the region contains all minimal paths), as shown in Fig. 1 (c). The extra beltway path can increase the lateral path diversity, which make the packet have a chance to detour the congested minimal path region or the throttled node. Based on the traffic information of the network, the proposed TTABR adaptively selects path between the minimal path and beltway path. Hence, the network can achieve more balanced traffic and temperature distribution. The contributions of this paper are summarized as:

- 1) Propose a beltway path, which is a non-minimal path, to scatter the traffic load in the minimal path region.
- 2) Adaptive selection between minimal path and beltway one to reduce the latency of packet delivery.

Through the traffic-thermal mutual coupling co-simulation platform [6], the experimental results show that the proposed TTABR can achieve more balanced traffic and temperature distribution. Besides, the network throughput is improved by around 3.4%~113%.

The rest of this paper is organized as follows. In Section II, we introduce some related routing schemes for 3D NSI-Mesh. In Section III, the proposed TTABR is described. In Section IV, the experiments and implementation are shown and discussed. Finally, we conclude this paper in Section V.

This work was supported by the National Science Council under NSC 100-2220-E-002-013 and NSC 100-2220-E-002-016

## II. RELATED WORKS

### A. Transport Layer Assisted Routing (TLAR) [3]

For successful packet delivery, Chao *et al.* proposed to use the information of both transport layer and network layer to ensure there is at least one routable path from the source node to the destination node [3]. If the source layer (*i.e.*, the location silicon layer of the source node) is routable, *TLAR* will decode to use lateral routing first. Otherwise, the downward routing will be used instead. To realize this scheme, the authors proposed Downward Lateral Adaptive Deterministic Routing (*TLAR-DLADR*), which combines the adaptive and XY routing as the lateral routing algorithm and downward routing as the vertical routing. However, the *TLAR* still suffers from traffic congestion due to the insufficient lateral path diversity.

### B. Traffic-balanced Topology-aware Multiple Routing Adjustment (TTMRA) [5]

In [5], Chen *et al.* proposed a Traffic-balanced Topology-aware Multiple Routing Adjustment (*TTMRA*) to dynamically adjust the routing mode of each routing packet based on the throttling situations in the *NSI-Mesh*. *TTMRA* proposed a novel cascaded routing to make the routing packet detour the throttled node for increasing the lateral path diversity and routing flexibility. *TTMRA* selects a cascade node to extend the lateral routable area of each silicon layer. As the experimental results shown, *TTMRA* enhances the successful packet delivery rate by lateral routing path instead of downward routing path, which may cause the traffic congestion in other silicon layer. Nevertheless, the minimal routing still causes the traffic congestion in the minimal path region, which may make the thermal stress become more serious.

### C. Dynamic Thermal-Balance Routing (DTBR) [7]

In [7], Liu *et al.* proposed a Dynamic Thermal-Balance Routing algorithm (*DTBR*) on 2D NoC. The *DTBR* uses the thermal information to represent the thermal condition of neighboring nodes. With the thermal condition of the neighboring nodes, the *DTBR* adaptively selects the routing direction to balance the temperature distribution in the network. Although the *DTBR* can balance the temperature distribution in the network, it still suffers from the traffic congestion in the minimal path region due to the applied minimal routing.

## III. TRAFFIC- AND THERMAL-AWARE ADAPTIVE BELTWAY ROUTING (TTABR)

As mentioned before, the traffic imbalance easily results in thermal imbalances and thermal hotspot. To solve the performance issue caused by the traffic- and thermal-imbalance, we propose a Traffic- and Thermal-aware Adaptive Beltway Routing (*TTABR*). *TTABR* provides the both minimal path and the beltway path. Based on the congestion information of the network, the proposed *TTABR* can adaptively select the minimal routing path or beltway one. Because of more lateral path diversities, the *TTABR* can reduce the traffic congestion and balance the traffic. The detail will be described later.

### A. Baseline Adaptive Beltway Routing

The beltway path is the one without going into the area of minimal path region, as shown in Fig. 1(c). Although the beltway path is a non-minimal path, the packet may be delivered to the destination with shorter latency due to the smaller queueing waiting time. It is a trade-off to select the minimal routing path and non-minimal one. Fig. 2 shows an example of *TTABR* in the source layer. Because of the less lateral path diversity while using minimal adaptive routing, it is easily results in heavy traffic congestion in the minimal path region, as shown in Fig. 2 (a). However, the traffic congestion can be reduced while using beltway routing adaptively, as shown in Fig. 2 (b). To make the packets be routed through the non-minimal beltway path, the previous proposed cascaded routing in [5] can be adopted to restrict the routing path as a beltway path through two-phase XY routing, as shown in Fig. 2 (b).

To detour the traffic hotspot, the Regional Congestion Awareness (*RCA*) selection scheme, which was proposed in [10], is applied in this paper. In addition to the congestion information, the throttling information should be considered, because it may cause heavy traffic congestion around the throttled node. If the router is throttled, the packets will be blocked in the input buffers, and the input buffers of the throttled routers have high probability to be full. To consider the worst case design, if the node is throttled, the propagated *RCA* information will be the same as the one when the input buffers are all full, which can be described as:

$$\text{Local Information} = \begin{cases} \text{Free buffer length, if the router is not throttled} \\ \text{Full, if the router is throttled} \end{cases} \quad (1)$$

By following the *RCA* selection scheme, we choose the feasible output channel based on the information of free buffer length of the effective direction (*i.e.*, local information in (1)), which is shown as:

$$\text{Output Channel} = \max\{\text{Buffer length of each effective direction}\} \quad (2)$$

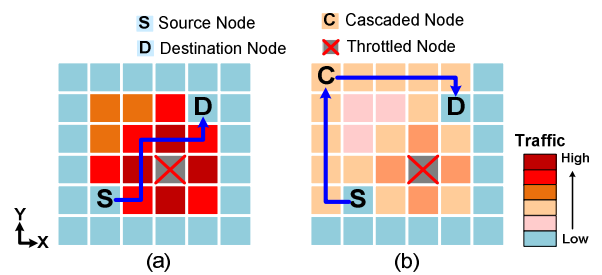


Fig. 2 (a) Minimal adaptive routing results in heavy traffic congestion, and (b) *TTABR* can balance the traffic load significantly.

### B. Advanced Adaptive Multiple Beltway Routing

Because the *TTABR* adopts two-phase XY routing to construct a beltway routing path, the cascaded node, which is used for the beltway routing, has to locate on the diagonal line of a mesh-based NoC platform. By this way, the *TTABR* can provide multiple beltway path levels, which can reduce the latency of the packet deliveries through non-minimal beltway path, as shown in Fig. 3. Based on the *RCA* information, the

*TTABR* can also dynamically adjust the employed beltway path level to balance the traffic load.

The operation flow of the proposed adaptive beltway routing is shown in Fig. 4. Fig. 3 illustrates an example. If the packet in the source node (*i.e.*,  $S_a$  and  $S_b$  in Fig. 3) is decided to be route through the beltway path based on the *RCA* information. The beltway level one is set as the initialized beltway path for simplifying the computation complexity (*i.e.*, the cascaded node is set as  $C_1$ ). When the packet is routed to the beltway level two, the *TTABR* may adaptively adjust its cascaded node to  $C_2$ , if the beltway level two is more feasible based on the *RCA* information. For the cascaded node of the minimal routing, the routing behaviors are the same as the one of *TTMRA* [5]. As the proof in [5] and [9], this kind of hybrid routing algorithm is deadlock free.

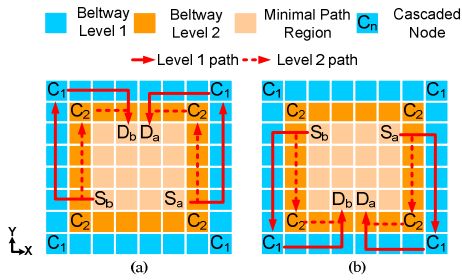


Fig. 3 Beltway path in XYZ routing (a) is a case of using north cascaded node and (b) is a case of using south cascaded node.

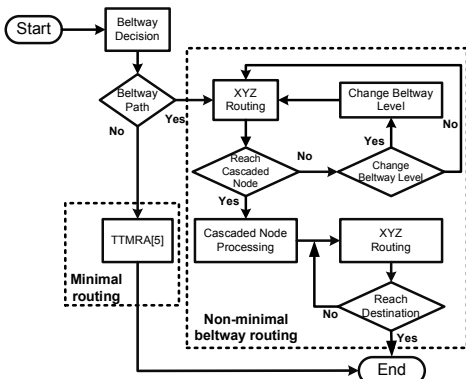


Fig. 4 Flowchart of the proposed *TTABR*.

#### IV. PERFORMANCE EVALUATION AND IMPLEMENTATION

##### A. Performance Evaluation of Proposed *TTABR*

Through the traffic-thermal co-simulation platform [6], the performance and the statistical traffic load distribution (*STLD*) of the proposed *TTABR* are evaluated. For each router, the depth of each input buffer is 8 flits without virtual channel, and the packet size is randomly from 2 to 10 flits.

Fig. 5(a) and (b) show the *STLD* and temperature distribution of the irregular mesh under random traffic pattern. Because of the larger channel utility, the proposed *TTABR* can mitigate the traffic hotspot in *TLAR* and *TTMRA*, which leads to smaller standard deviation ( $\sigma$ ) of the traffic load and

temperature distribution, as shown in Fig. 6. Therefore, the traffic load and temperature distribution become more balanced in *TTABR*. Compared with the previous works, the proposed *TTABR* can improve the performance by around 6.3% to 113%, as shown in Fig. 5(c). For the regular mesh, the proposed *TTABR* still validates the effectiveness of balanced traffic load and the temperature distribution (as shown in Fig. 7(a) and (b)), which leads to smaller  $\sigma$ , as shown in Fig. 8. Compared with the previous works, the *TTABR* can improve the performance by 3.4% in regular mesh, as shown in Fig. 7(c). The performance behaviors of *TLAR* and *TTMRA* are identical, because the two routing algorithms, which are designed for the *NSI-Mesh*, both become West-First adaptive routing in regular mesh.

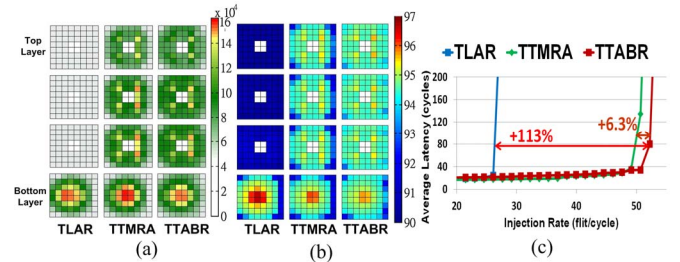


Fig. 5 (a) *STLD*, (b) Temperature distribution, and (c) Performance evaluation of irregular mesh Case 1.

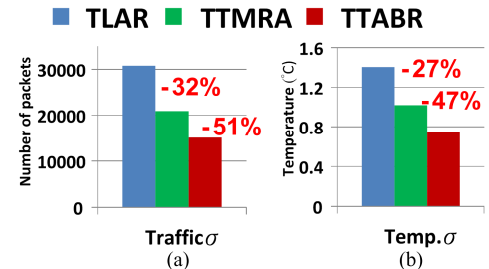


Fig. 6  $\sigma$  of (a) Traffic load distribution and (b) Temperature distribution of irregular mesh.

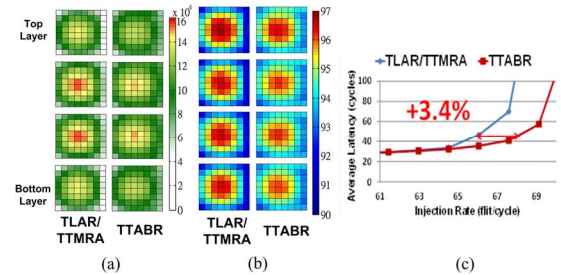


Fig. 7 (a) *STLD*, (b) Temperature distribution, and (c) Performance evaluation of regular mesh.

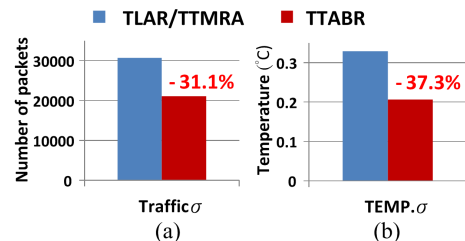


Fig. 8  $\sigma$  of (a) Traffic load distribution and (b) Temperature distribution of regular mesh.



### B. Implementation of Proposed TTABR

In this section, we show the implementation results of the proposed TTABR, including a NI and a router, which will be designed for an 8×8×4 3D NoC. For each flit, it contains 22-bit header and 42-bit payload [12]. By following the design concept in [9], the proposed architecture of TTABR is shown in Fig. 9, which is composed of six parts:

- **Baseline Datapath and Tx/Rx Queues (Tx/Rx):** Tx deals with the message from application layer and packetize the message into packets to network layer. Rx receives packet from network layer and combines to message to application layer.
- **Topology Table (TT):** This table stores the throttling information of each destination, and the TT is updated as the topology changes.
- **Routing Mode Memory (RMM):** RMM stores the routing mode (i.e., the minimal routing or non-minimal beltway routing) toward each destination. The RMM is updated while the topology changes.
- **Cascaded Node Memory (CNM):** CNM stores the cascaded node toward each destination. The CNM will be updated while the topology changes.
- **Control Logic (CL):** CL controls the functionality of Tx/Rx and checks the routing mode. A Finite-State Machine (FSM) is to implement in CL for the timing and signaling controls.
- **Router:** The 3D router requires extra two physical channels for vertical connections. Besides, there are two dedicated channels (i.e., L1 and L2) to connect the NI and the router for supporting cascaded routing.

Based on the synthesis results using the UMC 90nm technology, the comparison of area overhead between the previous works and the proposed TTABR is shown in Table 1. To support the cascaded routing and reduce the latency overhead, there are two dedicated I/O channels in NI (i.e., L1 and L2). Therefore, the hardware cost of Tx/Rx of TTABR is bigger than the previous works. The hardware cost of router of TTABR is bigger than the previous works because of the more complex crossbar. Table 2 shows the comparison of area efficiency after normalization. Obviously, the proposed TTABR can significantly improve the throughput with acceptable hardware cost overhead.

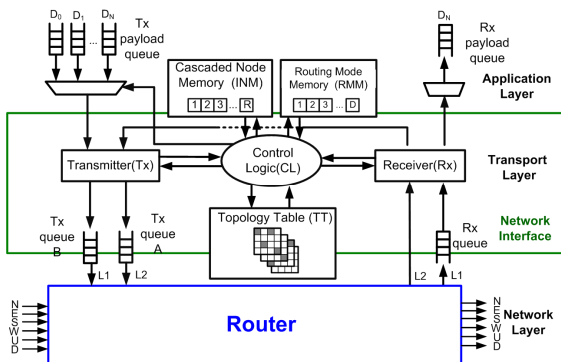


Fig. 9 Proposed architecture of TTABR

Table 1 Area comparison of the TTABR and the related works.

	TLAR [3]	TAMRA [5]	TTABR
NI: Tx/Rx/CL	41,961	51,872	56,018
NI: RMM	4,476	4,476	4,476
NI: TT	2,448	2,448	2,448
NI: CNM	0	2,978	2,978
Router	155,901	171,835	182,599
Total	209,344	233,609	248,519

Unit:  $\mu\text{m}^2$

Table 2 Throughput/Area comparison after normalization.

	TLAR [3]	TAMRA [5]	TTABR
Throughput improvement (T)	1.000	1.625	1.809
Hardware cost overhead (H)	1.000	1.115	1.181
Area efficiency (T/H)	1.000	1.457	1.534

### V. CONCLUSIONS

In this paper, to balance the traffic and temperature distribution in 3D NoC, we propose a Traffic- and Thermal-balanced Adaptive Beltway Routing (TTABR) scheme. In TTABR, the beltway path is provided to scatter the traffic load in the minimal routing region. Because the traffic and temperature distribution become more balanced, the proposed TTABR scheme can improve 3.4~113% network throughput with more efficient area utilization.

### REFERENCES

- [1] R. Marculescu, U.Y. Ogras, L.S. Peh, N. E. Jerger, and Y. Hoskote, "Outstanding research problems in NoC design: System, microarchitecture, and circuit perspectives," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 28, no. 1, pp. 3–21, Jan. 2009.
- [2] S. Scott and G. Thorson, "The Cray T3E Network: Adaptive Routing in a High Performance 3D Torus," *Proceedings of HOT Interconnects IV*, pp. 147–156, Aug. 1996.
- [3] C.H. Chao, Tzu-Chu Yin, Shu-Yen Lin, An-Yeu Wu, "Transport Layer Assisted Routing for Non-Stationary Irregular Mesh of Thermal-Aware 3D Network-on-Chip Systems," *IEEE Intl. SOC Conference (SOCC)*, Sept. 2011.
- [4] C.H. Chao, K.Y. Jheng, H.Y. Wang, J.C. Wu, and A.Y. Wu "Traffic- and Thermal-Aware Run-Time Thermal Management Scheme for 3D NoC System" *IEEE Intl. Symp. Network-on-Chip (NOCS)*, pp.223-230, May 2010.
- [5] K.C. Chen, S.Y. Lin and A.Y. Wu, "Traffic-Balanced Topology-Aware Multiple Routing Adjustment For Throttled 3D NoC Systems," *IEEE Workshop on Signal Processing Systems (SiPS)*, pp. 120-124, Oct. 2012.
- [6] K.Y. Jheng, C.H. Chao, H.Y. Wang, and A.Y. Wu, "Traffic-thermal mutual-coupling co-simulation platform for three-dimensional Network-on-Chip," *IEEE Intl. Symp. on VLSI Design, Automation, and Test*, Apr. 2010.
- [7] F. Liu, H. Gu, and Y. Yang, "DTBR: A dynamic thermal-balance routing algorithm for Network-on-Chip," *Computers & Electrical Engineering*, vol. 38, pp. 270-281, 2012.
- [8] G. M. Link and N. Vijaykrishnan, "Hotspot Prevention Through Runtime Reconfiguration in Network-On-Chip," presented at the *Proceedings of the conference on Design, Automation and Test in Europe - Volume 1*, 2005.
- [9] C.H. Chao, K.C. Chen, T.C. Yin, S.Y. Lin and A.Y. Wu, "Transport Layer Assisted Routing for Run-Time Thermal Management of 3D NoC Systems," to be appeared in *ACM Transactions on Embedded Computing Systems*.
- [10] Gratz, B. Grot, and S. W. Keckler, "Regional congestion awareness for load balance in networks-on-chip," in *Proc. IEEE Symp. High Performance Computer Architecture (HPCA)*, pp.203-214, Feb. 2008
- [11] L. Shang, L. Peh, A. Kumar, and N. K. Jha, "Thermal modeling, characterization and management of on-chip networks," in *Proc. IEEE/ACM Int. Symp. Microarchitecture (Micro'04)*, pp. 67–78, Dec. 2004.
- [12] B.S. Feero and P.P. Pande, "Networks-On-Chip in a Three Dimensional Environment: A Performance Evaluation," *IEEE Trans. Comput.*, vol.58, no.1, pp.32-45, Jan. 2009.