Cool Elevator: A Thermal-Aware Routing Algorithm for Partially Connected 3D NoCs

Ebadollah Taheri¹, Ahmad Patooghy² & Karim Mohammadi¹
e_taheri@elec.iust.ac.ir, {patooghy, mohammadi}@iust.ac.ir
¹Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran.
²Dependable Systems and Architectures Laboratory, School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran.

Abstract—Despite of high throughput and low fabrication cost of vertically partially connected 3D NoCs, thermal difficulties arise from poor heat dissipation and inappropriate traffic distribution of these kinds of 3D NoCs. This paper proposes an adaptive routing algorithm in order to manage thermal challenges in partially connected 3D NoCs. In the proposed routing algorithm, vertical links declare their availability/unavailability status to their neighbor nodes due to their current temperature. In this way, hot vertical links have a chance to reduce their traffic load and to cool down. In a predefined time periods vertical links update process is done to determine current hot and cool vertical links. In an updating time periods, cool vertical links are announced to the routers of each layer for transmitting packets to other layers. Access Noxim simulator is used to evaluates the routing algorithm in different partially 3D networks. Results show that the proposed routing algorithm decreases the number of overheated nodes by at least 74% and improves the thermal variance by at least 13%. These results are achieved within at most 10% overhead in average latency delay.

Keywords—Thermal-aware, Routing algorithm, 3D NoC, Deadlock-free.

I. INTRODUCTION

Network on chip (NoC) is introduced as a communication platform for connection between several processing cores on a single chip [1], [2]. Design and fabrication of NoCs have been leading to Three Dimensional Networks on Chip (3D NoC) as VLSI fabrication technology enables to integrate many processing cores onto a single chip [2], [3]. Using 3D NoCs the network diameter which is the maximum distance between two cores of a chip is significantly reduced. This consequently reduces the power consumption and packets transmission delay in 3D NoCs [3]. On the other hand, fabricating 3D ICs is made a set of serious threats for systems reliability e.g., thermal challenges is among the main reliability threats of 3D NoCs [4]. The majority of faults in chips such as delay fault, electromigration, dielectric breakdown happening in a 3D chip are due to thermal difficulties [5].

3D NoCs use vertical links to connect different cores located in different layers of a 3D chip. Vertically links in a 3D chip is fabricated using Through Silicon Via (TSV) technology that is preferred to other methods because of less space occupation and less packaging overhead [6]. However, the TSV technology increases the possibility of fault occurrence and increases the chip fabrications cost as well because of some extra manufacturing processes.

Recently, 3D NoCs with partial vertical links are proposed as a solution to simultaneously reduce TSV challenges and fetch advantages of 3D NoC [7]. However, routing in partially connected 3D NoCs is more complicated because of the disruption of order in the network by considering deadlock problem. In addition, in 3D NoCs with partially links several routers have to use a common vertical links. Using common links leads to traffic increment in the vertical links and nodes around them which in turn produces more heat in such nodes. In the absence of temperature management in 3D NoC particularly in area near to vertically links, the networks reliability will be threatened.

The main challenge in partially 3D networks on chip is cooling restriction for layers away of the chip heat-sink. If heat production in chip is not managed, low cooling rate for these layers is a threat for chips reliability [2], [8]. Heat in addition to fault production and giving speed to fault occurrence process leads to power consumption increment due to leak current increment [9]. Power increment by heat produces a positive feedback which leads to thermal increment in a short time in the considered point and consequently damage to chip [9], [10]. Thermal management techniques in networks are divided to reactive and proactive categories [11]. The reactive techniques [11], [8] temporarily prevent the routers activity which has reached to the critical temperature which consequently provides time for the router to become cool. In this technique routers safety is increased in terms of not encountering very high temperature, however networks performance will be decreased because of interruption in routers activity for cooling process and considering that cooling speed is low. In the proactive techniques [12], [13], the rise in temperature and reaching to critical temperature is prevented by controlling routers activity.

In this paper a routing algorithm for heat distribution in 3D networks on chip with vertically partially links is presented. The proposed routing algorithm using two virtual channel is deadlock and livelock freedom. In order to desirable thermal management, an adaptive vertical link assignment method for each pair of source and destination is presented. In the specified periods of time, the network becomes updated and routers temperature are measured in these periods. After temperature measurement, vertically link of each router which are hotter than a thermal threshold is not announced to the routers as an available vertical link. In this way, the proposed method give some times to overheated routers to become cool.

The remainder of the paper is organized as follows. In Sec-
tion 2 related works and motivation of the paper are studied. Section 3 presents the proposed routing algorithm. Section 4 discusses about simulation results. Finally, conclusion remarks are shown in Section 5.

II. RELATED WORK AND MOTIVATION

A. Vertically partially connected 3D NoCs

In order to take the advantages of 3D NoCs against 2D NoCs and reduce the manufacturing costs, vertically partially connected 3D NoCs have been recently proposed. The main concern of routing algorithms for such networks are deadlock freedom [7], vertical link assignment to each packet according to its source and destination [14], and reliability [15], [8], [16].

Ying et al. in [17] proposed a deadlock-free routing algorithm with some restriction on the placement of vertical links. The routing algorithm selects minimal paths to route packets to vertical links without any extra virtual channel. However, the restrictions on vertical placement result in 1) incompetent placement because the restrictions of placement limit designers to make a desirable placement and 2) low reliability because of low redundancy in the routing algorithm.

Elevator-first routing algorithm [7], using two virtual channels, have high capability to being used in partially connected 3D NoCs without deadlock occurrence. By assigning a fixed vertical link for each source node, each packet is routed to the vertical link as an intermediate destination in the source layer. After passing vertical dimension, the packet is routed to the destination in the destination layer. In order to achieve a deadlock-free routing algorithm, packets are divided into two sets, 1) the destination is upper than the source (vertically upper) and 2) the destination is lower than the source (vertically lower). The first and second virtual channels are assigned to the first and second sets respectively. Each packet is routed in its virtual channel according to the destination location (vertically upper or lower). Elevator-first routing algorithm improves the network performance with low hardware overhead. However, Low adaptation in vertical link selection results in low path diversity and poor reliability.

Salamat et al. in [15] proposed a fault tolerant routing algorithm in order to improve reliability in vertically partially connected 3D NoCs. Using 2 virtual channels for solving deadlock problem, the proposed routing tolerates the vertical link faults. This routing algorithm by dynamic vertical link assignment after reaching a faulty vertical link, selects another vertical link to tolerate the vertical link fault. However, reassignment of vertical link results in packets detouring and low throughput.

Foroutan et al. [14] proposed an algorithm to optimize vertical links assignment. The strategy of a routing algorithm for selecting vertical links plays an important role in routing algorithm for partially connected 3D NoCs. A routing algorithm in partially connected 3D NoCs, as determiner for the path between source and destination sends packets first, form source node to vertical link in the source layer and second, from the vertical link to the destination node in the destination layer [7], [15]. So, main determinant for the path is vertical link location according to the source and the destination. In other words, vertical link assignment for each packet is a main determination for traffic and throughput of partially connected 3D NoCs. In this paper to distribute traffic in order to improve thermal challenges, the proposed routing algorithm utilizes a proposed technique for adaptive vertical link assignment. Overheated router after sensing temperature in the updating time periods, are eliminated from candid elevators to the assignment strategy.

B. Thermal management in 3D NoCs

In order to improve reliability and reduce cooling costs of the chip, thermal-aware routing algorithms distribute chip temperature by keeping the temperature of routers below the threatening degree [8], [18], [12]. The thermal-aware routing algorithms with reactive or proactive approaches are employed to manage the overheated router in the network. In order to design a thermal-aware 3D NoC, each router is equipped with thermal sensor to monitor thermal state [12], [19], [20], [16]. Therefore, the routing algorithm distributes the traffic of network according to the sensed temperature from these sensors.

Reactive techniques in routing algorithms [8], [18], [19] are commenced after reaching router temperature to a thermal threshold. The overheated router in such techniques is throttled for a predefined time or until it becomes safe[21]. In other words, each throttled router does not send and receive any packet until the router becomes safe. Therefore, the network with throttled routers became a non-stationary irregular mesh (NSI-Mesh). Reactive routing algorithm [8], [11] are employed in the throttled 3D NoC to detour the packets that are blocked with a throttled router. Chen et al. in [8] proposed a reactive routing algorithm for fully connected (not partially connected) 3D NoCs. The routing algorithm with decision in transport layer, detours packets that have a throttled node in theirs default path to an intermediary node. However, the throttling state of the router result in a high delay for transmitting packets and performance degradation.

On the other hand, proactive routing algorithm [16], [13] distributes heat generation to reduce thermal variance and limits temperature of each node. Lee et al. in [13] proposed a routing algorithm with dynamic regulation of buffer depth in order to deliberate increasing congestion in overheated node. Traffic aware routing after increasing congestion in the overheated region results in thermal distribution. However, this method suffers from low buffer utilization and traffic increase which result in performance degradation.

All of the proposed routing algorithms for thermal management are designed for fully connected 3D NoCs. This paper proposes a proactive techniques to distribute heat in partially connected 3D NoCs.

III. THERMAL-AWARE ROUTING ALGORITHM

A. Routing procedure and deadlock freedom

The main challenge in designing routing algorithms for partially connected 3D NoCs, is solving deadlock problem in the vertical dimension. In the proposed routing algorithm that uses two virtual channels, packets that the destination of them are located in other layers, first are routed in the first virtual channel. Packets that the destination of them are located
in upper layers, as shown in Figure1, are switched to the second virtual channel when exiting from the source layer. For example, the packets are routed from S3 and S4 to D3 and D4, first are routed in the first virtual channel then are switched to the second virtual channel after finishing routing in the source layer. Packets that the destination of them are located in lower layers are switched to the second virtual channel when arriving the destination layer. Packets with source and destination in same layer, are randomly routed in the first or second virtual channels, see S1 and S2 to D1 and D2.

**Theorem III.1.** The proposed routing algorithm is deadlock-free in vertical dimension in each virtual channel.

**Proof:** In each virtual channel to formation cyclic dependency in vertical direction both turning to up and down directions should exist together. In the routing algorithm 1) There is no turn to up direction in the first virtual channel 2) There is no turn to down direction in the second virtual channel. All turns in the proposed routing algorithm for the first and second virtual channels are shown in Figure.2.

In each layer the proposed routing algorithm uses XY routing algorithm. So, the routing algorithm is deadlock-free in the first and the second virtual channel. Notice that each vertical link in this algorithm are considered as a pillar, i.e. the vertical link pattern in each layer is similar to the vertical link pattern in the other layers. Therefore, all turns in vertical dimension are in source and destination layer, as shown in Figure.1. No cyclic dependency is occurred between two virtual channels because there is no switching from the second virtual channel to the first virtual channel. So, the proposed routing algorithm is deadlock-free.

The procedure of the proposed routing Algorithm is shown in algorithm.1. As shown in line 2-4, randomly selection between virtual channels is employed in order to distribute utilization of the virtual channels for the state that the source and destination are in the same layer. Each packet gets elevator id from the source router that is saved in the update stage, see Subsection B. Switching to the second virtual channel occurs in the destination layer for downward routing and in the source layer for upward routing, see line 10 and 23.

**B. Elevator assignment**

One of the main steps in designing a routing algorithm for partially connoted 3D NoCs, is identifying the location of vertical links, which are also called elevators, to the source routers. In order to distribute the traffic for thermal management, the proposed routing algorithm eliminate utilization of the elevator of overheated routers. Each router is equipped with thermal sensor in high throughput 3D NoCs [13], [18], [21]. An updating time is considered to online sensing temperature by the sensors. After sensing temperature, the overheated routers are considered like the router that do not have elevator. In updating stage, the routers that have elevator and the temperature of them is below the predefined thermal threshold

![Figure 2](image-url)
send themselves id number to other routers in each layer. As shown in Figure 3 an id number is given to each router. The routers after sensing temperature in the updating stage sends itself id number to the neighbor router. In this example, routers with id number 3, 9 and 19 are the routers that have elevator and the temperature of them are below the predefined thermal threshold. In first cycle, as shown in Table I, routers with one hop to routers 3, 9 and 11 receive and save the router id that have the cool elevator, i.e. 2, 5, 7, 8, 10, 13, 15, 18. The router that have saved an elevator id number, sends this id number to neighbor in next cycle. After 4 cycles for this example all routers save the near elevator address. Notice this algorithm in update stage, transmits only the data of elevator location (unlike the normal state).

The proposed routing algorithm employs a technique like proposed method in [22] for updating thermal information. In updating stage, the network is stopped to inject and absorb the packets until the end of this stage, i.e. only thermal information are shared in this stage. If D is the diameter of network, the maximum time for updating in the worst case is D cycles. In the simulation of this paper, see section 4. In the simulation of this paper, see section 4, thermal sensing is updated after 100000 cycles and the diameter of network is 17. So, the time overhead for the proposed algorithm is 17/100000. The time overhead is below 1%, so the overhead is negligible.

As shown in Algorithm 2, if local router has the cool elevator, it sends the local id to the neighbors, for up and down elevator apart. If local router does not have cool elevator but has a router id that has cool elevator, sends the id number to the neighbors, see line 5. In the condition that local router receive an id number before, rejects any request to updating the elevator location from the neighbors. Therefore, 1) no cyclic requests for data transmitting of the elevators occurred and 2) each router is able to receive id number of first elevator, i.e. the nearest elevator. If router receive several request from neighbors at the same time, selects one of them randomly.

**Algorithm 2** update elevator location.

1. while (D cycles) do
2. if (router has cool up (down) elevator) then
3. send itself id to the neighbors
4. else
5. send up(down) elevator id to the neighbors (if exist)
6. end if
7. if (neighbor routers send id to the current router) then
8. if (up(down) elevator variable has no updated) then
9. save received id in up(down) elevator variable
10. end if
11. end if
12. end while

IV. RESULTS AND DISCUSSION

The proposed routing algorithm is evaluated against Elevator-first routing algorithm [7] in terms of thermal condition and average latency. Access Noxim simulator [23]. with capability of online monitoring the temperature of each node, is used for this evaluation. Access Noxim, an extended version of Noxim [24], is a traffic and thermal mutual-coupling simulator that uses Hotspot thermal model [25].

The simulator is equipped with the proposed routing algorithm and the extra virtual channel. An 8x8x4 network with uniform random traffic for 1000000 cycles is simulated (for Elevator-first and the proposed routing algorithm). The depth of buffers are considered to be four flits and each packet contains 8 flits.

The thermal distribution for each layer, with considering the temperature threshold 80°C in the proposed routing algorithm, are shown in Figure 4. As shown in the figure, the proposed routing algorithm distributes the heat generation and decreases the temperature of the overheated regions. The third layer temperature is lower than the other layer because of close distance to the heatsink and high cooling efficiency of this layer. The overheated regions are in the layer that are away from the heatsink. As shown in the figure, the overheated regions are eliminated in the proposed routing algorithm.

Comparison of the thermal conditions for 2 different states of the presence of elevators are shown in Table 2. (20% elevators means 20% of all elevators exist, i.e. this value for fully connected topology is 100% elevators). Thermal variance in the proposed routing algorithm improves around 74%. This means the distribution of the proposed routing algorithm is improved because of heat generation management. The number of

<table>
<thead>
<tr>
<th>Cycle</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>Id number</td>
<td></td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Id number</td>
<td></td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Id number</td>
<td></td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Id number</td>
<td></td>
<td>1</td>
<td>3</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
<td>9</td>
<td>9</td>
<td></td>
<td></td>
<td>9</td>
<td>9</td>
</tr>
</tbody>
</table>

Figure 3. Sharing elevator location (a) Id number of each node (b) elevator assignment after updating for each layer.
of overheated nodes decrease around 13%-30% after 100000 cycles. Therefore, the reliability of system is improved by using the proposed routing algorithm.

Average latency in the network with the proposed routing is shown in Figure. 5. The overhead of the proposed routing algorithm is under 10% for 20%-50% presence of elevators. As shown in the figure, the overhead in low presence of elevators is negligible.

V. Conclusion

Using two virtual channels, this paper proposes a deadlock-free routing algorithm for vertically partially connected 3D NoCs. These kinds of 3D NoCs are susceptible to different faults because poor cooling efficiency and undesirable traffic distribution that deteriorates into the thermal challenges. The paper also proposes a dynamic vertical link assignment in order to distribute heat generation in the layers of the network. The proposed routing algorithm relieves overheated nodes at least 74%, improves thermal variance around 13%-30%. The overhead of the proposed algorithm is lesser than 10% in average latency.

REFERENCES
