6th International Conference On Advances In Computing & Communications, ICACC 2016, 6-8 September 2016, Cochin, India

PT-BAR: Prioritized Thermo-Buffer based Adaptive Routing Protocol for Network-on-Chip

R.Suraj\textsuperscript{a,}\textsuperscript{*}, P.Chitra\textsuperscript{b}

\textsuperscript{a,b}Thiagarajar College of Engineering, Madurai-625015, India

Abstract

The Network-on-Chip (NoC) is an important technology that replaced the traditional bus-based architecture for the future of System-on-Chip (SoC). The NoC system provides better scalability, performance, reliability, etc. to the SoC networks by implementing the principles of interconnection networks and packet switching. One of the major problems in NoC is the increase in temperature of the nodes that leads to unbalanced thermal management within the network. This further leads to performance degradation due to damaged nodes. A novel thermal management scheme is proposed in this paper, which makes use of the thermal state and buffer state of the nodes for routing the packets. The proposed Prioritized Thermo-Buffer based Adaptive Routing (PT-BAR) protocol maintains a thermal region based on the thermal balance of the nodes within the region. Only high priority packets are transferred through the thermal region to preserve thermal balance throughout the network. During packet transfers the thermo-buffer model considers the energy consumption and thermal conductivity of nodes to calculate temperature and the current buffer state of the node. This factor is used to choose a path that will provide better performance in end-to-end delay and throughput while preserving the thermal balance of the network. Proposed PT-BAR reduces the temperature by 1.5°C than the existing protocol by making use of the efficient thermal model and buffer models for the efficient routing strategy.

© 2016 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the Organizing Committee of ICACC 2016

Keywords: Prioritization; Thermal balance; Network-on-Chip; Adaptive routing; Packet switching

1. Introduction

Highly integrated System-on-Chip (SoC) technology has been used recently where large numbers of transistors are being integrated into a single small chip [1]. The major cause for thermal glitches within the System-on-Chip networks are the continuous increase in consumption of power and the system cool down cost that leads to degradation in system performance and reliability. This further increases the delay incurred and the power leakage in
the network [2]. Bottlenecks like these in the SoC network can be avoided by replacing the traditional bus-based architecture [3] of the on chip network with the Network-on-Chip (NoC) that is built using interconnection networks and packet switching. Issues such as end-to-end delay, low throughput, scalability and bandwidth are handled using the NoC network. The NoC architectural design still suffers from thermal glitches and problems even though the performance is high [4].

The major challenges in any NoC network are to provide the following: (1) an effective routing algorithm that obtains high performance and throughput with less delay, (2) enhanced thermal management scheme to provide balanced temperature throughout the network and (3) a reliable deadlock recovery mechanism to improve network reliability. The routing algorithm applied in the network should consider these aspects to provide a model that can obtain efficient packet delivery and can still preserve the thermal balance of the network. The most critical issue of the above is the thermal management of the network that affects the network performance completely.

Two cases of thermal issues have been noted in any given NoC system. The first is the regional temperature differential that occurs due to unbalanced distribution of temperature of the nodes in the network. The thermal unbalance within the system will eventually lead to system failure and performance degradation as it will affect network delay. The second is the occurrence of hotspot nodes due to a high temperature of any node that happens due to processing too much data or by consuming large amounts of energy. This may cause the hotspot node will stop working normally or get damaged with increase in temperature. Unique strategies should be used to obtain good thermal balance.

The routing algorithm implemented in the NoC system should also consider the thermal balance of the network before forwarding packets. The routing algorithm selects a path from source to destination based on a routing function and selection strategy to transfer the packets. The set of available output ports from the current nodes are obtained using the routing function and a suitable path is selected using the selection strategy using a weight factor.

Another dimension in efficient routing for managing thermal balance is by forwarding limited number of packets through high temperature nodes, and no packet at all, through hotspot nodes. The priorities of packets should be maintained in such a way that all high priority packets get delivered without delay. The next aim of routing algorithm is to effectively handle deadlocks [10] during packet transfer in the NoC system. Deadlocks can be prevented and avoided by implementing certain deadlock avoidance and deadlock recovery mechanisms [11]. Unwanted use of buffer space and traffic congestion can be avoided [11] by handling deadlocks. One such mechanism is to forward the deadlocked packets using a special deadlock recovery channel within the network.

A novel and dynamic routing protocol is proposed in this paper that works as an adaptive protocol in a given NoC system and aims to obtain thermal balance throughout the network and still provide good throughput and performance. The proposed PT-BAR routing protocol is designed using a priority based T-B(Thermo-Buffer) state model that considers both the thermal state and the buffer states of the nodes when forwarding packets. The thermal model is built by identifying node temperature and energy consumption based on energy transformation and thermal conductivity. The prioritizations of the packets are considered when forwarding them through high temperature regions. The proposed routing algorithm is implemented in a 3D mesh NoC system.

The key contributions of the proposed routing algorithm are as follows: (1) the T-B state model can achieve good performance in end-to-end delay and throughput while conserving node temperature, (2) by using prioritization of nodes in packet transfer hotspot nodes and unbalanced thermal distribution can be avoided and finally (3) deadlocks can be avoided and especially in case of high priority nodes the delivery ratio obtained will be high.

The remainder of the paper is organized as follows. Section 2 talks about existing methods in thermal management and routing protocols. Section 3 explores the working principle of the proposed PT-BAR protocol using the T-B state model and the prioritization principle. Section 4 shows the simulation of proposed routing strategy using a 3D NoC system and results are compared with existing methods. Section 5 provides the conclusion of the paper.

2. Existing Methods and Motivation of Work

2.1 Network-on-Chip

The Network-on-Chip is a network architecture that provides a detailed structure of various elements of the network. The various processing elements of NoC network can be classified as:
- **IP cores** – The IP cores are the main processing nodes of the network.
- **Network Interface (NI)** – The NI provides the various connection protocols of the nodes.
- **Routers** – Each IP core is connected to the neighbour cores using a router.
- **Physical Links** – The physical links are used for connecting all the routers.

The IP cores of the NoC system are the main functional units and they contain processing elements such as a processor, memory, CPU, input and output units, etc. The NI provides characteristics to each IP cores based on the underlying network protocol used. The efficiency in communication and data transfer is dependent mainly on the routers of the NoC system. In this paper 3D mesh based NoC network is considered for implementation that consists of large number of nodes. The sample design of a basic 3D mesh NoC system is shown in Fig. 1 that consists of 48 nodes in 3 layers with a 4x4 2D mesh in each layer.

### 2.2 Existing Routing Protocols

Though there are any routing protocols available in the literature, the adaptive based routing protocol is used in this paper since they are suitable for thermal balance [12]. Nilsson et al. [13] proposed a unique selection strategy in adaptive routing namely as Proximity Congestion Awareness (PCA) that exploits the load information of neighbour nodes to choose the next hop path. Li et al. [6] proposed the dynamic XY protocol based on PCA that selects the next hop node with free buffer space. Ascia et al. [14] proposed Neighbours-on-Path (NoP) to obtain more information about the network using the adjacent nodes. Gratz et al. [15] proposed the Regional Congestion Awareness (RCA) to get both local and non-local states but it suffers from inter and intra-region interferences. Ma et al. [16] proposed the Destination-based Adaptive Routing (DBAR) to avoid such interferences. Apart from these other adaptive routing protocols such as DyAD routing [17], odd-even routing [18], turn based routing [19] and table-based minimal routing algorithms [20] have been used.

In recent times many other heuristic and optimization algorithms have also been proposed for routing in the NoC systems. Giuseppe et al. [21] proposed a selection policy that aims to minimize the average delay and power consumption within the NoC system. They make use of a cost function generated using Fuzzy Rule Base System (FRBS) with downstream router buffer and link power dissipation as input and the cost function as output. Later Sani et al. [22] proposed a NoC architecture using the Globally Self-Adaptive and Scalable Routing Algorithm (GSASRA) that makes use of the Particle Swarm Optimization (PSO) algorithm. Here global routing decisions are made using PSO and it is better than Fully-adaptive and XY-routing in terms of latency, energy consumption and throughput. Hsien-Kai et al. [23] proposed an Ant Colony Optimization based adaptive routing protocol for NoC by considering the Network Information Region (NIR). The NIR indicates the combinations of various network and the routing information. Here both the spatial and temporal network information are integrated using the ACO-based pheromone diffusion (ACO-PhD) adaptive routing strategy. This approach improves the overall network
performance.

2.3 Existing Thermal Management Techniques

An effective thermal model should be designed for the NoC system that is suitable with the routing protocol to provide efficient thermal management scheme. By using thermal sensors a direct and accurate thermal distribution scheme can be provided as proposed by Lionel et al. [24] and Rao et al. [25]. But this requires large number of control links to transmit thermal signals and they need to be arranged precisely. To reduce hardware costs, a thermal profile or scheme is designed for the NoC. Current thermal management schemes are categorized as Design-Time Optimization (DTO) and Dynamic Thermal Management (DTM) [26].

In DTO, thermal balance is achieved during offline design of the NoC system by considering the worst-case situations. The DTO approach is effective only if the network is stable, since optimization of NoC thermal distribution is obtained with the loss of data transmission performance. Chunsheng et al. [27] proposed hardware based thermal management scheme where each node is allocated with different voltages and clock frequencies based on on thermal information of the nodes. Wei et al. [28] proposed a thermal balance approach by optimizing positions of IP cores.

The DTM strategy maintains thermal distribution of NoC system dynamically and it can be further categorized as proactive and reactive DTM [29]. Emilio et al. [30] proposed a DTM scheme by considering both hardware-based and software-based methods such as regulating voltages and clock frequencies across nodes and providing task migration. Yang et al. [31] proposed a multi-clock frequency based thermal management scheme where different clock frequencies are allocated to nodes dynamically to regulate thermal balance to each node separately. Greg et al. [32] proposed a workload migration strategy where formation of hotspot is avoided by migrating hotspot traffic to other nodes, hotspot traffic managed accordingly as proposed by Sharifi et al. [33] proposed a method for implementing a thermal management scheme using workload migration and DVFS simultaneously.

Recently, a new class of Runtime Thermal Management (RTM) scheme has been proposed by Nizar et al. [34]. They introduced the Dynamic Programming Network (DPN) for control logic in adaptive routing and the temperature is sensed using the Ring Oscillators (ROs). The issues associated with the DPN convergence and sensing are handled. The overall thermal reliability and the thermal distribution of the NoC system is reduced here compared to the DTO and DTM based schemes.

3. Methodology

This paper presents a novel adaptive routing protocol that makes use of a T-B state model and prioritization of packets to obtain thermal balance throughout the NoC system. The proposed method also aims to attain good performance in terms of delay and throughput. The T-B state model defines the temperature and energy consumption of the nodes along with buffer space of the nodes. The prioritization algorithm implements the routing algorithm based on the thermal regions of the 3D NoC network. The T-B state model is built using two aspects, one is the Temperature model (T) and the next is the buffer model (B).

3.1 Temperature Model of T-B

The T-B temperature model combines both the energy transformation and thermal conductivity of the network to measure a node temperature of the NoC system. Let $N$ be the number of nodes in the NoC system and $T_i^0$ is the initial temperature for node $i$ where $i = 1, \ldots, N$. Current temperature $T_i$ of a node is given as in Eq. (1):

$$T_i = T_i^0 + T_i^E + T_i^T$$

(1)

$T_i^E$ is the energy transformation temperature associated with node $i$ that shows increase in temperature due to energy consumption. The energy transformation for node $i$ is given in Eq. (2) where $E_i^{total}$ is the total energy consumption of node $i$.

$$T_i^E = \epsilon_i * E_i^{total}$$

(2)

$T_i^T$ is the thermal conductivity temperature, that is generated due to heat conduction. It depends on parameters such as, material of the chips and the distance between routers. A thermal coefficient $R$ is defined based on these two factors. If temperature difference between any two nodes $i$ and $j$ is $T_{ij}^{diff}$ and heat conductivity between them is $R_{ij}$, then thermal conductivity is given as in Eq. (3):
Substituting Eq. (2) and Eq. (3) in Eq. (1) the temperature of node $i$ can be expressed as below in Eq. (4):

$$ T_i = T_i^0 + (\varepsilon_i * E_i^{total}) + \sum_{j=1,j\neq i}^{N} T_{ij}^{diff} * R_{ij} $$

The difference in energy consumption $E_{ij}^{diff}$ between nodes $i$ and $j$ obtained using Eq. (2) is substituted in Eq. (4) to express node temperatures based on energy consumption of any node $i$ as in Eq. (5):

$$ T_i = T_i^0 + (\varepsilon_i * E_i^{total}) + \sum_{j=1,j\neq i}^{N} E_{ij}^{diff} * R_{ij} $$

The energy consumed by forwarding flit packets is used to calculate total energy consumed by node $i$. In the proposed NoC system, a virtual channel router is used that transfers flit packets using various stages such as Receive Flit (RF), Routing Computation (RC), Virtual-channel Allocation (VA), Switch Allocation (SA) and Forward Flit (FF). Data Flits cross all these process and Head Flits contains only the RF and FF operations. Let all the packets transmitted in the NoC network have a fixed number of flits $n$. The average energy consumed to transmit one flit is given in Eq. (6):

$$ E_{flit} = E_R + E_{SA} + E_F + \frac{1}{P}(E_{RC} + E_{VA}) $$

The energy consumed in node $i$ is given in Eq. (7) for transmitting $P_i$ flits through the node.

$$ E_i^{total} = E_{flit} * P_i $$

Substituting Eq. (7) in Eq. (5) and replacing the energy difference $E_{ij}^{diff}$ with traffic difference $\beta_{ij}^{diff}$ that shows the difference in number of packets transferred in nodes $i$ and $j$, the temperature can be represented as in Eq. (8):

$$ T_i = T_i^0 + (\varepsilon_i * E_{flit} * P_i) + \sum_{j=1,j\neq i}^{N} \beta_{ij}^{diff} * R_{ij} $$

By assuming that traffic distribution is same across all nodes, the total temperature of the NoC system can be expressed as in Eq. (9):

$$ T = T_0 + \varepsilon_i * E_{flit} * \beta $$

Here $T_0$ is the initial total temperature of the network and $\beta$ is the total packets transmitted in the network.

### 3.2 State Model of T-B

The T-B buffer state of a given node can be calculated using the available buffer space of the node and the total energy consumption. The buffer state value of the available output ports or next node is given as in Eq. (10):

$$ S_o = (w_1 * B_0 - w_2 * E_o^{total}) * w_3 $$

Here $B_o$ and $E_o^{total}$ are the available buffer in the output port and the total energy consumption of the next hop node $o$. $B_0$ represents normalized value of buffer space calculated using mean and standard deviation of all buffer spaces of nodes. $w_1$ and $w_2$ are the weights associated with buffer state and energy state of the nodes respectively. $w_3$ indicates if the node is available or not. The weight values represent the importance of the respective parameter used. The weight values are chosen in such a way that the sum of $w_1$ and $w_2$ equals to 1 and the value of $w_3$ is set as per the state of nodes (0-Slept, 0.5-Half Slept and 1-Active).

### 3.3 Thermal Region and Prioritization

The thermal region indicates a specific region within the NoC system with certain characteristics. The packets transferred through the network are classified as low priority for read, write and block transfer messages and high priority for real time messages and signal messages. A node with temperature higher than a specific threshold is called as hotspot node. Based on the results and transistor parameters studied by Chao et al. [35] a threshold probability of $0.8 \%$ was used by Junhui et al. [36]. $80^\circ C$ is the temperature at which the transistors and the materials used for making the nodes starts to get damaged. Whenever a hotspot occurs it is denoted as a half-slept node that can only forward high priority packets. The region around this node is the half-slept region. In case a node fails or gets damaged due to high temperature it is denoted as a slept node and the area around this node is a slept
region and no packets can be forwarded here. This can be described clearly as shown below in Fig. 2.

![Fig.2. PT-BAR protocol scenario with different nodes](image)

3.4 Prioritized Thermo-Buffer based Adaptive Routing

The proposed Prioritized Thermo-Buffer based Adaptive Routing (PT-BAR) protocol makes use of the thermal information and buffer information to forward the packets to next nodes. The PT-BAR protocol can achieve high performance with respect to delay and throughput by using the prioritization property and thermal regions. Also the thermal balance of network is kept under control by using the state information of T-B, this reduces delay in the network.

When the node temperature cross the threshold value the router was made to half slept, So only high prioritized packet can enter into the node. The flit head carrying RT data are classified into high priority packets and which contain Read/Write or transfer messages are called low priority packets. For the half-slept node, low priority packets are deflected as the node is closed. On the other hand, high priority packets can go through the half-slept node even if there is high temperature address by Thermo Buffer.

Consider NoC network with 3 layers as in Fig. 2. Let the node (2, 1, 1) be either a faulty node (damaged node) or a hotspot node. Then a half-slept region or a slept-region is formed surrounding this node that is represented in dark grey in Fig. 2. The number of packet transfer within this region is minimized since the hotspot node will conduct its heat to nearby nodes. The region is removed only after the temperature of the hotspot node goes down or if the faulty node is repaired.

Let the source node be (0, 2, 2) and destination node be (3, 3, 2). To transfer packets from source to destination the proposed PT-BAR algorithm should choose a path excluding the slept region and half-slept region to avoid
thermal imbalance. From the source node the packets can be transferred to three different axes X, Y and Z and the list of all possible output nodes are taken as a set. For each of these nodes in this set, the current temperature and current state is calculated as in Eq. (8) and Eq. (10). The node with least temperature and better state is chosen as the next hop node. The process is repeated till the packets are delivered to the destination node. The overall procedure of PT-BAR protocol is explained below.

Prioritized Thermo-Buffer based Adaptive Routing Algorithm

Initialization:

$O$: The set of output nodes from current node, $|O| = M$

$o \in O$: Output node from the set $O$ where $o = 1, 2, \ldots, M$

$S_o$: State information of output node $o \in O$

$T_o$: Temperature of output node $o \in O$

$\Delta T$: Threshold temperature

$\Delta T_{min}$: Minimum temperature

$S$: Source node

$C(c_x, c_y, c_z)$: Current node

$D(d_x, d_y, d_z)$: Destination node

Algorithm:

1. Router in C receives a packet from router in S
2. Find $\Delta X = d_x - c_x, \Delta Y = d_y - c_y, \Delta Z = d_z - c_z$
3. If $\Delta X > 0$ add $(c_x + 1, c_y, c_z)$ to set $O$, else if $\Delta X < 0$ add $(c_x - 1, c_y, c_z)$ to set $O$, else do nothing
4. If $\Delta Y > 0$ add $(c_x, c_y + 1, c_z)$ to set $O$, else if $\Delta Y < 0$ add $(c_x, c_y - 1, c_z)$ to set $O$, else do nothing
5. If $\Delta Z > 0$ add $(c_x, c_y, c_z + 1)$ to set $O$, else if $\Delta Z < 0$ add $(c_x, c_y, c_z - 1)$ to set $O$, else do nothing
6. For each output node $o \in O$, calculate $T_o$ and $S_o$ [Eq. (8),(10)]
7. $\forall o, o \in O$, If $S_o$ not satisfied remove $o$ from $O$
8. $\forall o, o \in O$, if $T_o > \Delta T$ mark $o$ as hotspot node and generate half-slept region
9. $\forall o, o \in O$, If $o$ in slept region remove $o$ from $O$
10. $\forall o, o \in O$, If $o$ in half-slept region and packet is low prioritized remove $o$ from $O$, else do nothing
11. Set $\Delta T_{min} = 80$
12. $\forall o, o \in O$, if $T_o < \Delta T_{min}$ set $\Delta T_{min} = T_o$
13. $\forall o, o \in O$, if $T_o == \Delta T_{min}$ set $C = o$
14. Repeat steps 2 to 13 till $C = D$

To prevent the occurrence of deadlocks a virtual channel based deadlock recovery mechanism is implemented along with the PT-BAR protocol. An additional virtual channel is used as an escaping path for the packets during a deadlock. To remove the deadlock, the deadlocked packet should release all allocated resources and must be forwarded through the virtual channel to escape into the network. By removing the allocated resources from the deadlocked packet, the network can now recover from a deadlock condition and function normally.

4. Simulation and Result

4.1 Initial Setup

A 3D mesh NoC system of size of 4x4x4 is simulated using a cycle accurate simulator noxim. This system contains 64 nodes and the wormhole switching technique [31] is used to make accurate execution of the routing algorithms. During simulations energy consumed and raise in temperature are monitored continuously for each node. The factor that are monitored includes the number of packets transferred by all the nodes are of same length during the execution. Temperature is measured by using of the Hotspot [33] simulator tool. To make a detailed study on the routing algorithms a large total of 50000 packets were transmitted during the time of simulation throughout the 3D NoC network.

Varying traffic conditions were taken to compare the efficiency of the protocols under these conditions. The traffic patterns such as (1) uniform traffic, (2) transpose traffic and (3) hotspot traffic were considered here. In case of uniform traffic the probability of transferring packets from a given source node to a given destination node is always the same. In hotspot traffic some of the nodes in the NoC system are taken as hotspot nodes initially during
the start of simulation. In these hotspot nodes only 5% of the packets are sent to other nodes of them, 80% of the packets sent to the rest nodes with the same probability and in case of proposed PT-BAR protocol only the high priority packets were transmitted. The remaining nodes have the same probabilities to transfer packets. Finally in case of transpose traffic a packet that is generated by a source node can be sent only to a specific fixed destination node. Here in the simulation both synthetic traffic and real traffic conditions were used. Throughput and end-to-end delay for performance evaluation is provided under synthetic traffic, and the performance related to thermal distribution is provided under real traffic.

4.2 Comparison of Results

The proposed PT-BAR routing algorithm is compared with the other existing 3D NoC such as MAR, XYZ, Cube NoC Routing and DyXYZ routing algorithm. The initial temperature of all the nodes in the network is taken as 25°C during the simulation of the algorithm existing and the proposed. Fig. 3 through Fig. 8 illustrates the distribution of throughput and end-to-end delay recorded under various traffic patterns for different algorithm for the 4x4x4 NoC system. Fig. 9 through Fig. 13 shows the overall distribution of the temperature of all the nodes in the NoC network after the simulation has been completed.

The proposed PT-BAR protocol obtains a higher throughput and lesser end-to-end delay in all the traffic conditions compared to the other 3D NoC routing protocols. The proposed PT-BAR also makes use of the slept and half-slept regions to reduce the high temperature of nodes and at the same time used the buffer model of the T-B model to obtain better throughput and lesser end-to-end delay.

Fig.3. Throughput under uniform traffic

Fig.4. Throughput under transpose traffic
Fig. 5. Throughput under hotspot traffic

Fig. 6. End-to-End delay under uniform traffic

Fig. 7. End-to-End delay under transpose traffic

Fig. 8. End-to-End delay under hotspot traffic
Fig. 9. XYZ Thermal Distribution

Fig. 10. MAR Thermal Distribution

Fig. 11. DyXYZ Thermal Distribution

Fig. 12. Cube NoC Thermal Distribution

Fig. 13. PT-BAR Thermal Distribution
The worst thermal distribution is noted in the XYZ protocol and this is because of the fact that it ignores the states of the nodes when forwarding packets. The thermal distributions in MAR and DyXYZ routing protocols are better than the XYZ protocol because they are more dynamic in packet forwarding and can adapt themselves based on the situation of the network. The thermal distribution of Cube NoC routing is still better than XYZ, MAR and DyXYZ as it considers the current network state and node state when forwarding the packets and the energy consumption here is less.

In PT-BAR routing, thermal distribution is reduced by an amount of about 1.5°C compared to the existing routing by making use of the efficient buffer and thermal models for efficient routing strategy.

5. Conclusion

In this paper, a novel thermal balanced routing algorithm PT-BAR, is proposed that can attain balanced temperature throughout the network and can maintain the balance even during occurrence of faulty nodes and hotspot nodes. The T-B state model and the priority schema introduced in this algorithm achieve thermal balance and also improve the throughput. The end-to-end delay is also reduced and hence guarantees the performance of high priority packet delivery. Thermal regions are generated for faulty nodes and hotspot nodes to forward packets based on their priorities. The proposed method provides better performance in cases of thermal distribution, delay and throughput compared to existing methods. In future, this protocol can be enhanced further by providing a better model for prioritization of the packets by considering deflection based routing model.

References