

Noc Evolution and Performance Optimization by Addition of Long Range Links: A Survey

Vaishali Maheshwari¹

¹ NMIMS, Mumbai

Received: 7 February 2015 Accepted: 28 February 2015 Published: 15 March 2015

Abstract

In the last decade, Networks-on-Chips became the leading edge technology due to the growing requirements of electronic systems. Basically, NoC is an advancement of bus interconnect technology. The challenge is to interconnect existing components such as processors, controllers, and memory arrays in such a way that there is an optimal utilization of communication resources necessitating optimization of the various dominant factors like energy/power consumption, interconnection delay, latency, throughput, etc. In this paper, we focused on the evolution of NoC. Then we studied and have shown through an example that when application specific long-range links are inserted among the tiles whose communication frequencies are high, there is a reduction in the average packet latency and an energy efficient architecture is build up with high throughput. We also discussed the turn model which is deadlock free and the energy model for NoC.

Index terms— NoC, long range link, 2D, turn model

1 Introduction

resent day VLSI technology permits to build systems with millions of transistors on a single chip to meet the growing computational applications and to achieve high performance with lowpower/energy requirements. A system on a chip or system on chip (SoC) is an integrated circuit (IC) that integrates all components of a computer or other electronic system into a single chip but the interconnection between each other is a challenging issue [2]. Traditionally, the interconnection networks used for communication among the components in a system on chip were bus-based and point to point links. In a shared bus interconnection network, many masters and slaves share the bus with each other but only one master at a time can use the bus, and the other masters have to wait for their turn so a bus arbiter makes a decision among multiple bus access requests. The shared bus architecture was used because of its lowcost and simple installation characteristics. However, it has limitation in its scalability because only one master at a time can use the bus at a time. It becomes a communication bottleneck when the number of bus requesters grows tremendously and the bandwidth is limited. Dedicated point-to-point links improves latency, power usage and bandwidth availability. But as the number of IPs grows, the number of dedicated point-to-point interconnections grows exponentially resulting in a larger realization area [3]. A solution that satisfies scalable bandwidth requirement is "Network on Chip (NoC) architecture". Dally and Towlesproposed that NoC eliminate ad-hoc global wiring and introduced on-chip interconnection network with modular design which gives better performance, higher bandwidth, and scalability and reusable components. NoCs use packets to route data from the source Processing Element (PE) or node to the destination PE or node via a network fabric that consists of network interfaces, routers and interconnection links [1].

A research on network-on-chip (NoC) depicts a change of state from computation-centric to communication-centric design paradigm by development of scalable communication structures and thus achieving global communication in SoC [4].

A 2D mesh topology is a regular grid-like NoC architecture which has a simpler design layout. They have well-controlled electrical parameters and reduced power consumption on the global wires. Such architectures

have long packet latencies although the shortest path algorithms are used because here the packet have to travels many hops to reach to the destination node which leads to an energy inefficient architecture [6]. In this survey paper, we summarize the evolution and necessity of NoC. We have shown the NoC architectures for a standard mesh network and also modified the same by addition of long range links. For the convenience in Section II, computations are shown to calculate average packet latency and Manhattan distance for a standard mesh network and for network using long range links. In Section III, turn model is discussed. In Section IV, energy model is shown. Lastly, conclusion followed by the future scope is discussed in section V.

II.

3 Computations

For random traffic patterns categorized by the communication frequencies f_{ij} , τ_{ij} is computed by the author [5] written as follows:

Eq. 1 where $d(i,j)$ is the distance from routers i to router j , and t_r , t_s , t_w are the is architectural parameters representing time to make the routing decision, traverse the switch and the link, respectively. Finally, L is the length of the packet, while W is the width of the network channel.

For the standard mesh network, the Manhattan distance (d_M) is used to compute $d(i, j)$ where x and y denote the x - y coordinates. [5] Eq. 2

For the routers with long-range links, [5] Eq. 3

In the above equation, $l(i,k)$ means that node i is connected to node k via a long-range link [5].

An example shown below demonstrates that the use of long range link decreases the number of hops traversed by a packet for reaching to the destination node.

4 Turn Model

A packet traverses through various nodes as it moves from source to destination. At each hop a decision is to be made whether to move straight or take a turn along a routing path. This decision is significant because a wrong turn may lead to cyclic dependencies which may further cause a deadlock due to which packet is unable to reach destination. Some combinations of turns are proposed which are deadlock free. This is called turn model [10]. A packet has to follow turn model till it reaches to destination.

5 Dynamic Communication Energy Model

The dynamic communication energy model for the network on chip can be defined as: $E_{bit}(t_i, t_j) = n \text{ hops} \times E_{rbit} + (n \text{ hops} - 1) \times E_{Lbit}$ Eq. 4

Where $E_{bit}(t_i, t_j)$ is the average energy consumption for sending one bit of data from t_i to tile t_j , n hops is the number of routers the bit traverses from tile t_i to tile t_j , E_{rbit} is the energy consumed by the router for transporting one bit of data and E_{Lbit} is the energy consumed by unit link/channel for transporting one bit of data.

To reduce energy consumption, it is important to identify the energy efficient architectures. Therefore the energy-performance trade-offs need to be considered. Depending on the parameter selected an efficient methodology is proposed that is to be designed based on the selected performance metrics with help of long range interconnects for standard grid based network [9].

V.

6 Conclusion

We have summarized the evolution and necessity of NoC by showing the disadvantages of buses over NoC architectures. We have shown the NoC architectures for a standard mesh network and also modified the same by addition of long range links. By taking an example we have shown that by adding long range links to a mesh network number of hops traversed by a packet. Hence, we found that there is a significant reduction in the average packet latency and increase in network throughput.

VI.

8 Future Scope

For exploring more network structures and to reduce energy consumption and wire delay, we can intend to explore the effect of long rank interconnects/links based on the application specific traffic load for constant bit rate traffic with practical constraints on link length and the port availability per tile. The feasibility of the proposed design is planned to be explored for 2D as well as 3D NoCs. A 3D NoC is the stack of 2D NoC in such a way that each stack is again connected to its front and rear stack. It was proposed that in 3D NoC, the numbers of hops traversed by the flit towards the destination are reduced, thereby, have the following advantages such as high throughput, reduced message latency and energy dissipation as compared to the traditional 2D NoCs [10]. So we



Figure 1: Figure 1 :

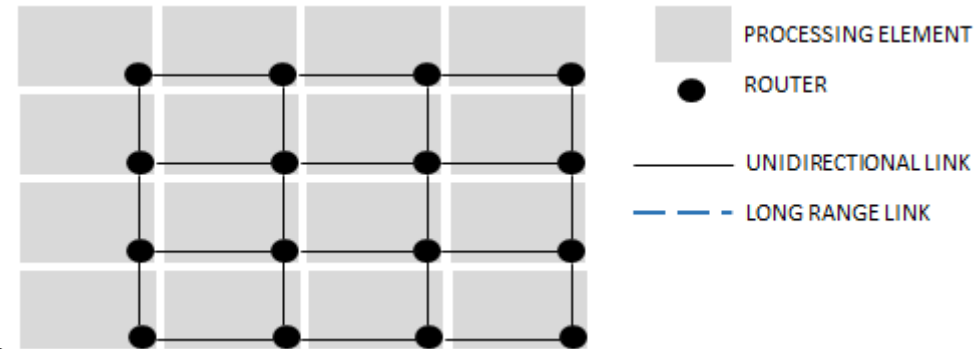


Figure 2: Figure 2 :

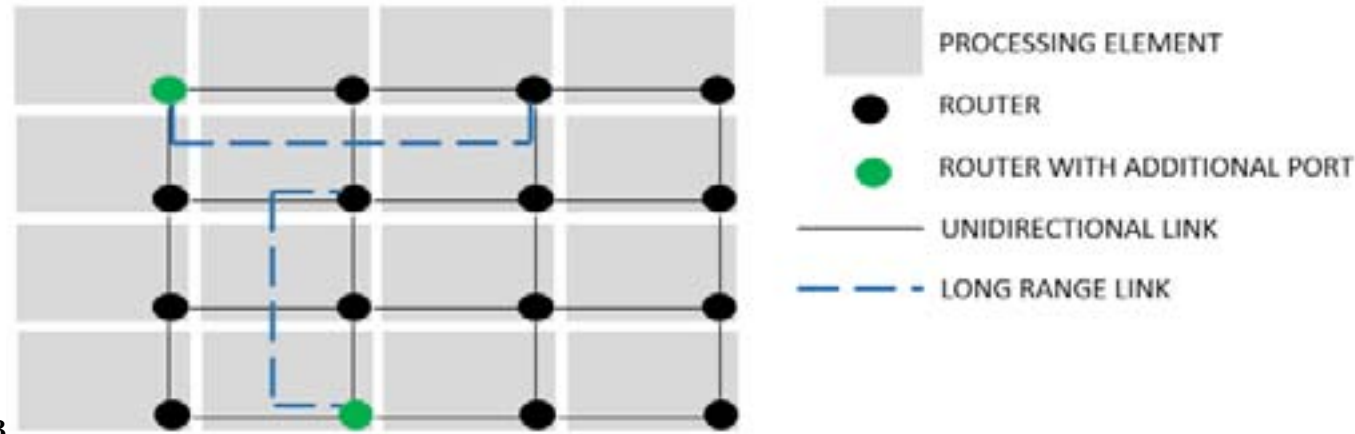


Figure 3: Figure 3 :

$$\tau_0 = \sum_i \sum_{j \neq i} f_{ij} \left[d(i, j)(t_r + t_s + t_w) + \max(t_s + t_w) \left\lceil \frac{L}{W} \right\rceil \right]$$

Figure 4: Figure 4 :

$$d_M(i, j) = |i_x - j_x| + |i_y - j_y|$$

Figure 5: Figure 5 :

98 propose that the long range links when inserted in a 3D NoC topology will give drastically good results in terms
99 of high network throughput and reduced average packet latency. ^{1 2}

¹© 2015 Global Journals Inc. (US) 1

²© 2015 Global Journals Inc. (US)

$$d(i, j) = \begin{cases} d_M(i, j) & \text{if no long-range link is attached to } i \\ \min(d_M(i, j), 1 + d_M(k, j)) & \text{if } l(i, k) \text{ exists} \end{cases}$$

Figure 6: Figure 6 :

[New SoC Communication Infrastructure Paradigm. International Journal of Soft Computing and Engineering (IJSCE)]
, *New SoC Communication Infrastructure Paradigm. International Journal of Soft Computing and Engineering (IJSCE)* 2231-2307. 1 p. .

[Bjerregaard and Mahadevan ()] ‘A Survey of Research and Practices of Network-on-Chip’. T Bjerregaard , S Mahadevan . *ACM Computing Surveys* 2006. 38.

[Umit et al. ()] *Application-Specific Network-on-Chip Architecture Customization via Long-Range Link Insertion*, Y Umit , R Ogras , Marculescu . 2005.

[Zhang et al. ()] *Comparison Research between XY and OddEven Routing Algorithm of a 2-Dimension 3X3 Mesh Topology Network*, W Zhang , L Hou , J Wang , S Geng , W Wu . 2006.

[Umit et al. (2006)] ‘It’s a Small World After All: NoC Performance Optimization Via Long-Range Link Insertion’. Y Umit , R Ogras , Marculescu . *IEEE* July 2006. 14 (7) .

[Benini and Micheli ()] ‘Networks on chips: A new SoC paradigm’. L Benini , G De Micheli . *IEEE Computer* 2002. 35 (1) .

[Hyung Gyu Lee et al. ()] ‘On-chip communication architecture exploration: A quantitative evaluation of point-topoint, bus, and network-on-chip approaches’. N Hyung Gyu Lee , Chang , Y Umit , R Ogras , Marculescu . *ACM Transactions on Design Automation of Electronic Systems* 2007. 12 (3) p. 23.

[Puthal et al. ()] *Performance Comparison of 2D and 3D Mesh NoC*, M K Puthal , M S Gaur , V Laxmi . 2011.

[Dally and Towles (2001)] ‘Route packets, not wires: Onchip interconnection networks’. W Dally , B Towles . *Proc. DAC*, (DAC) June 2001.

[Glass and Ni ()] ‘The turn model for adaptive routing’. C J Glass , L M Ni . *Proc. ISCA*, (ISCA) 1992. p. .