

Review of XY Routing Algorithm for Network-on-Chip Architecture

Shubhangi D Chawade
M-Tech (Electronics)
B. D College of Engineering
Sewagram(Wardha)

Mahendra A Gaikwad
Professor
B. D. College of Engineering
Sewagram(Wardha)

Rajendra M Patrikar
Professor
VNIT
Nagpur

ABSTRACT

The Network-on-Chip (NoC) is Network-version of System-on-Chip (SoC) means that on-chip communication is done through packet based networks. In NOC topology, routing algorithm and switching are main terminology. The routing algorithm is one of the key factor in NOC architecture. The routing algorithm, which defines as the path taken by a packet between the source and the destination. As XY routing algorithm mainly used in NOC because of its simplicity. This paper basically review of XY routing algorithm in which we study a different type of XY routing algorithm. The classification of XY routing algorithm is totally depend upon the environment and requirement. Such that IX/Y routing algorithm is for less collision in network, for deadlock-free and livelock-free DyXY is used, for fault-tolerant XYX routing algorithm is proposed and Adaptive XY routing algorithm is used for fully utilization of network resource.

Keywords

Network-on-chip; XY routing algorithm; XYX routing algorithm; IX/Y Routing algorithm; DyXY Routing algorithm.

1. INTRODUCTION

Traditionally, System-on-Chips (SoCs) utilize topologies based on shared buses. Dally and Towles proposed replacing dedicated, design specific wires with general purpose, (packet-switched) network [1], hence marking the beginning of Network-On-Chip (NOC) era. According to NOC design approach, designers use network design technology to analyze and design SOC. In other words, designers view a SOC as a micro-network of components. SOC interconnection design can be done using the micro-network stack paradigm, which is an adaptation of the protocol stack. NOC can be defined as “*network-on-chip is a communication network targeted for on chip*”. The NOC approach was proposed as a promising solution to these complex on-chip communication problems [2][3][4][5]. For the NOC architecture, the chip is divided into a set of interconnected blocks (or nodes) where each node can be a general-purpose processor like a Digital Signal Processing this processor commonly known as processing element (PE). A router is embedded within each node with the objective of connecting it to its neighbouring nodes. The router has four ports (West, South, East, North) to connect with other routers and a local port to connect with PE means have five input and five output. as shown in figure 1. Router basically consist of Routing algorithm and switching techniques, where routing algorithm proposed selection

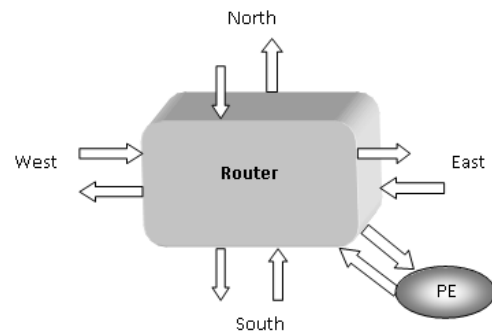


Figure 1: Generic Router structure

strategy which based on the concept of neighbours-on-path. The routing algorithm can be used in order to avoid faulty or congesting ports. Switching techniques determine when and how internal switches connect their inputs to outputs and the time at which message components may be transferred along these path.

The overall performance of a NOC depends on several network characteristics, such as topology, routing algorithm, flow control, and switching technique[6].

2. OVERVIEW OF NOC DESIGN APPROACH

Switching method, topology and Routing algorithm are three important techniques in the design of an NOC.

2.1 Switching Technique

There are two major switching techniques: circuit switching and packet switching. In Circuit switching virtual or physical link is establish between source and destination before a message is being transferred. Form [7] and [8] circuit switching is usually employed when Quality of Service (QoS) is considered because there is no contention delay as the link is held until all the data are transmitted so its behaviour is more predictable. For the simultaneous data communication between source-destination pair packet switching is used. Messages are transfers per-hop basis. In this type of switching first messages are divided into number of packets at source node and then send into the a network, this packets then get next router address using routing algorithm and traverse accordingly in a series of network nodes and finally reached at destination node. As it providing simultaneous data communication it utilized most of NOC platforms. Packet switching further classified into three classes: store and

forward (SAF), virtual cut through (VCT), and wormhole switching.

2.2 Topology

Topology means how nodes are placed and connected to each other. For better result of latency and bandwidth topology of the network is consider. There are different types of topology such as mesh, torus, mixed and custom topology proposed in [9]. Some researchers have proposed the application-specific topology that can offer superior performance while minimizing area and energy consumption [10]. The most common topologies are 2D mesh and torus due to their grid-type shapes and regular structure which are the most appropriate for the two dimensional layout on a chip.

Mesh topology is easy to implement as all nodes are in equally distance as shown in Figure 2 and also makes addressing of the cores quite simple during routing. A mesh topology has four inputs and four outputs from/to other routers, and another input and output from/to the PE.

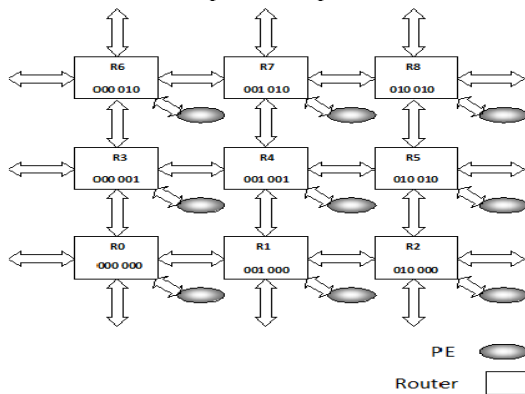


Figure 2: 3X3 Mesh Topology

2.3 Routing Algorithm in NOC

Routing algorithm is a key factor which affects the efficiency of the communication of NOC. The routing algorithm, which defines the path taken by a packet between the source and the destination, is a main task in network layer design of NOC. According to where routing decisions are taken, it is possible to classify the routing in source and distributed routing [11]. In source routing, the whole path is decided at the source router, while in distributed routing each router receives a packet and decides the direction to send it. According to how a path is defined to transmit packets, routing can be classified as deterministic or adaptive. The routing algorithms also be classified based on their adaptability, fault-tolerant capability, a centralized controller controls the data flow in a system and depends on the number of destination. But widely classified as Deterministic, Oblivious and Adaptive. Deterministic routing algorithm always choose the same path between a pair of nodes, Balance of load is very poor in this case, but they are commonly used due to easy implementation. In Oblivious routing algorithm packets route without considering the network's state. Adaptive routing algorithm use information about the network's state (e.g. channel load information, length of queues for resources, etc.) to make routing decisions. A well-designed router should provide adaptive routing to route a packet through the less congested channels. In adaptive routing, each router has congestion information of its surrounding neighbourhood. The channel congestion metric can be based on the number of free virtual channels, the number of free buffers, the demand for switch output or a combination of these parameters [12]. Based on the

congestion information, the router routes the packet to its destinations through the less congested channels. Gratz *et al.* propose regional congestion aware routing [13] which abstracts the regional congestion level into four single numbers to determine which direction is more likely to have low congestion. Ascia *et al.* [2] introduce *Neighbors-an-Path* adaptive routing algorithm for a router with no virtual channels. It uses immediate neighbour's congestion level for adaptive routing. Some more classification are possible like Fault-Tolerance-routing Algorithm - in which dynamically detects the faulty components while routing the packets. And Routing Through Reconfiguration- a cycle-free contour around a faulty link to use new unique paths instead of the broken paths.

3. TYPE OF XY ROUTING ALGORITHMS

3.1 XY

The most popular routing algorithm that is XY routing is proposed by Wang Zhang and Ligang Hou[15] in which packet have header flit, tail flit and data flit where the header flit having the destination node address, For implementation 2-Dimension mesh topology and wormhole switching mechanism are used. Each router having co-ordinates as (x,y), for the routing current address (Cx,Cy) is compare with the destination router address (Dx,Dy) of the packet, depends up on the comparison output of routing algorithm router routes the packets. if (Dx > Cx) head flit moves to East else it take West turn upto (Dx = Cx) become equal this passion is called as horizontal alignment. Now (Dy,Cy) undergoes compression, if it is found that (Dy < Cy) then packet's header flit moves towered South else North upto (Dy=Cy). For simulation an environment maintained as the packet size is 8 bytes with a random destination mode, the percentage load is 50% which mean that 50% of maximum bandwidth is used, the interval between successive flits is 2 clock cycles, the simulation runs 1000 clock cycles and the clock frequency is 1 GHz, Synthetic traffic generators generate traffic in the first 300 clock cycles with warm-up period of 5 clock cycles. After simulation authors define an average performance parameter P to evaluate the average performance of the algorithm that is P where P= Average Throughput of Network/ Average Latency per packet of Network which is 0.86 for the XY algorithm. So they come to conclusion that the implement of XY routing algorithm is simple as well as the X-direction channel latency is averagely larger than Y-direction channel latency and X-direction channel throughput is averagely all square with Y direction throughput.

3.2 IX/Y

For less collision Ahmad M. Shafiee, Mehrdad Montazeri, and Mahdi Nikdast developed new routing algorithm called as Intermittent Routing Algorithm [16] in which both XY and YX routing algorithm are used and name it as IX/Y routing algorithm. For study the performance of this routing algorithm, Ahmad Shafiee et al. used 4 X 4 mesh topology with NOXIM Simulator. In IX/Y Boolean variable 0 or 1 is set circular manner that is for first packet 0 is set and XY routing algorithm takes place, for the second packet Boolean variable 1 is set and YX routing algorithm is used. For the next packet again 0 is set and XY routing take place and same processer follow for next packets. this will give the help to have fewer collisions. While comparing the result, delay with respect to injection rate with other routing algorithm like XY, Dyad T,

Negativefirst, Northlast, westfirst and oddeven algorithm author [16] found that intermittent algorithm is roughly remained stable below 30 cycles, and also notice that this statistics is approximately the same in a 10*10 mesh-based and is an enduring algorithm in higher injection rates.

3.3 DyXY

A novel routing algorithm, namely dynamic XY (DyXY) routing algorithm which provides adaptive routing based on congestion condition in the proximity, and ensures deadlock-free and livelock-free routing at same time was proposed by Ming Li, Qing-An Zeng, Wen-Ben Jone [17]. Using DyXY routing algorithm packet only route a shortest path between source and destination. If multiple shortest paths are available then router choose the path for the packet depends upon the congestion condition of the network. The detailed routing algorithm is summarized as follows:

- 1 Read the destination of an incoming packet.
- 2 Compare addresses of the destination and the current router.
 - 2.1 If
 - the destination is the local core of the current router, send the packet to the local core;
 - 2.2 Else
 - 2.2.1 If
 - the destination has the same x (or y-axis) address as the current router, send the packet to the neighbouring router on the y-axis (or x-axis) towards the destination;
 - 2.2.2 Else
 - check the stress values of current router's neighbours towards the destination, and send the packet to the neighbour with smallest stress value.

Where the stress value is a parameter representing the congestion condition of a router, that is the number of occupied cells in all input buffers and each stress value is updated based on an event-driven mechanism. As With static XY routing, the length of a path traveled by packets for a give pair of source and destination is a constant, which equals to shortest path length. For DyXY routing, although the routing path is not static, it is always a shortest path and hence the length is still the shortest path length. Therefore, the average packet path length is only affected by the communication pattern. For the performance calculation author [17] used an event-driven simulator using C++. By injecting more than 140,000 packets into the network in each simulation, and the NOC was warmed up for 20,000 packets before measuring latencies. After that they compared results of XY and DyXY and they found that

3.3.1 By size varied from 3X3 to 9X9 with average packet injection rate increasing from 0.1 to 0.3

3.3.2 The DyXY routing algorithm achieves better balance in load distribution

3.3.3 The average mean response time for all routers can be effectively estimated using the analytical lower bound and upper bound.

3.4 XYX

Ahmad Patooghy and Seyed Ghassem Miremadi [18] proposed deadlock free and Fault-Tolerant routing algorithm for the higher traffic load using XY and YX routing algorithm. The basic idea behind the XYX routing algorithm is that the source node adds some detection code to each packet and the destination node checks the contains of received packets, If error is found then the source node request to resend corrupted packet to destination using NACK signal. Actually At the source node original packet is converted into two types of packet as original and redundant packet this is possible using indicator bit set as "0" and "1" respectively in the header flit of

packet. Using the indicator bit router comes to know that current packet is either the original or the redundant. In XYX routing algorithm the original packets follows XY routing algorithm and redundant packets YX routing algorithm which is inversed version of XY routing algorithm. At the destination node when a new original packet is received the parity bit will be regenerated and checked with those detection code embedded in the header of the packet. If two parity sets are equal, the packet will be accepted at the destination node otherwise it will be dropped detection code waits to receive the redundant copy of the packet. For experimental analysis Author [18] create simulation platform as, 6x6 mesh topology, a minimum of 20000 packets have been delivered, a fixed length of 32 flits per packet, mean rates of spanned from 0.001 to 0.02 packet per cycle per node, and in the reliability evaluation experiments, a wide range of error rates have been also injected to the network to precisely investigate the reliability improvement of the proposed routing algorithm. In this experiment the network has been simulated with error injection rate approximately 0.5, 1, 2, 5, 10, 20 and 50 percent of all flits experience a bit flip error. In result it found that at lower traffic rate XYX and other algorithms like S2S have correctly delivered more than 90% of the packets and their behavior are almost the same, In high error rate condition, the XYX routing could tolerate more than 60% of errors while it imposes lower performance and power consumption overheads and the XYX routing has the best performance almost in all working conditions.

3.5 Adaptive XY Routing

for fully utilizing network resources Mohsen Nickray, Masood Dehyadgari and Ali Afzali-kusha [19] proposed Adaptive XY The Adaptive XY routing algorithm is a derivative of classic XY routing algorithm. In proposed routing algorithm author used a context aware agent which route packets based on the entire network status and other members of society. This algorithm could operate as a deterministic or adaptive routing depends on neighbor agents' (switches) load condition. When congestion becomes high, agent tries to route packets through less congested path. The agents find less congested paths using 2-bits quantized load value so each agent has four quantized value that are sent to four North, South, East, and West neighbours. Configuration of load values for agents is operated using configuration packets. Configuration packets are generated according to information received from all neighbours then all neighbours collaborate each other to route packets. As showed in Table 1

Table 1

Value	DESCRIPTION
00	Port is free and ready to receive
01	Threshold0 (Initial value= 50% of buffer is full)
10	Threshold1 (Initial value= 75% of buffer is full)
11	Port is quite busy and can't receive

Emergence of a new route containing two parts: Firstly packets are routed by XY routing, when the destination port determined by XY routing is congested (a threshold is defined for each port of switch which is a criterion in order to realize a port is congested or not. This threshold involves the position

of a switch for example switches are located in the center of mesh have threshold0 and the others have threshold1)author do not route packet to the destination where declared by XY routing rather author[19] compare quantized load value of neighbors and select the path that its quantized load value is minimum. This algorithm causes user don't have to wait until the crowded route release. Author also perform study the other routs which do not have any heavy load. Simulation perform on 3X3 mesh topology using SystemC language in which environment keeping base class as SystemC . For results use two mode of load generation random test and test using multimedia applications. Tables 2 and 3 compare simulation results which show better results for context-aware agent based algorithm and the proposed algorithm shows more improvement when number of packets increases that is in higher loads. When sent packets in a period are changed from 500 to 25000, the improvement of latency is changed from %11 to %25 and the improvement of max latency is changed from 0% to 33%.

Table 2 : Conventional XY routing

	500	2000	10000
Send Packet	450	3200	25000
Latency(Avg.)	17	30	43
Max Latency(Avg.)	34	70	93

Table 3: Context-aware agent based routing

	500	2000	10000
Send Packet	450	3200	25000
Latency(Avg.)	15	25	33
Max Latency(Avg.)	34	49	62

4. CONCLUSIONS

A different type of XY routing algorithm is used in different network condition. The selection of XY routing algorithm it totally depends upon the application and the traffic of packet in the network. As simplicity in implementation is important in all architecture so an XY routing algorithm widely used . For fewer collisions Intermittent Routing Algorithm is preferred. The DyXY routing algorithm achieves better balance in load distribution as well as provide deadlock-free and livelock-free facility. Were we can't compromised with accuracy of received data , we go for the fault-tolerant routing like YXX .If application is focused on network resources utilization the Adaptive XY routing algorithm is best choice. Finally we can say that choice of XY routing algorithm is totally depends upon environmental condition of NOC architecture.

5. REFERENCES

[1] W. J. Dally and B. Towles. 2001 Route packets, not wires: On-chip interconnection networks.

- [2] L. Benini and G. DeMicheli. 2002 Networks on chips: Anew SoC paradigm.
- [3] Wang Zhang, Ligang Hou, Jinhui Wang, Shuqin Geng, Wuchen Wu. Comparison Research between XY and Odd-Even Routing Algorithm of a 2-Dimension 3X3 Mesh Topology Network-on-Chip
- [4] A. Hemani, A. Jantsch, S. Kumar, A. Postula, J. Oberg, M. Millberg, and D. Lindqvist. 2000 Network on a chip: An architecture for billion transistor era.
- [5] S. Kumar, A. Jantsch, M. Millberg, Oberg, Soininen, M. Forsell, K. Tiensyrja, and A. Hemani. 2002 A network on chip architecture and design methodology.
- [6] G. Acsia, V. Catania, M. Palesi, D. Patti. 2008 A New Selection Strategy for on-Chip Networks
- [7] K. Chang, J. Shen and T. Chen. 2006 Evaluation and design trade-offs between circuit-switched and packet-switched NOCs for applicationspecific SOCs.
- [8] P. Marchal et al. 2005 Spatial division multiplexing: a novel approach for guaranteed throughput on NOCs,
- [9] E. Salminen, A. Kulmala, and Timo D. Hamalainen, 2008 Survey of Network-on-Chip proposals
- [10] Yuan-Long Jeang, Tzuu-shaang Wey, Hung-Yu Wang, Chung-Wei Hung, and Ji-Hong Liu. An Adaptive Routing Algorithm for Mesh-Tree Architecture in Network-on-Chip Designs
- [11] J. Duato, S. Yalamanchili, L. Ni, Interconnection networks: an engineering approach, Morgan Kaufmann, Revised Edition, 2002.
- [12] P. Gratz, B. Grot, and S.W. Keckler. Regional congestion awareness for load balance in networks-on-chip. *High Performance Computer Architecture*, 2008. *HPCA 2008. IEEE 14th International Symposium on*, pages 203-214, Feb.2008.
- [13] Mohsen Nickray, Masood Dehyadgari, Ali Afzali-kusha "Adaptive Routing Using Context-Aware Agents for Networks on Chips"
- [14] W. Dally and B. Towles, "Route packets, not wires: on-chip interconnection networks," in *DAC*, 2001, pp. 684–689.
- [15] Wang Zhang, Ligang Hou, Jinhui Wang, Shuqin Geng, Wuchen Wu "Comparison Research between XY and Odd-Even Routing Algorithm of a 2-Dimension 3X3 Mesh Topology Network-on-Chip"
- [16] Ahmad M. Shafiee, Mehrdad Montazeri, and Mahdi Nikdast , "An Innovational Intermittent Algorithm in Networks-On-Chip (NOC)"
- [17] Ming Li, Qing-An Zeng, Wen-Ben Jone, "DyXY - A Proximity Congestion-Aware Deadlock-Free Dynamic Routing Method for Network on Chip"
- [18] Ahmad Patooghy and Seyed Ghassem Miremadi "YXX: A Power & Performance Efficient Fault-Tolerant Routing Algorithm for Network on Chip"
- [19] Mohsen Nickray, Masood Dehyadgari, Ali Afzali-kusha, "Adaptive Routing Using Context-Aware Agents for Networks on Chips"