

# Performance Evaluation of Different Routing Algorithms in Network on Chip

Jayant Kumar Singh  
Department of Electronics  
and Communication Engineering,  
National Institute of Technology,  
Rourkela, Odisha  
Email: jayantsngh101@gmail.com

Ayas Kanta Swain  
Department of Electronics  
and Communication Engineering  
National Institute of Technology  
Rourkela, Odisha  
Email: swain.ayas@gmail.com

Tetala Neel Kamal Reddy  
Department of Electronics  
and Communication Engineering  
National Institute of Technology  
Rourkela, Odisha  
Email: tneelkamalreddy@gmail.com

Kamala Kanta Mahapatra  
Department of Electronics  
and Communication Engineering  
National Institute of Technology  
Rourkela, Odisha  
Email: kkm@nitrrkl.ac.in

**Abstract**— Network on Chip (NoC) is a new paradigm to make the interconnections inside a System on Chip (SoC). By the developments achieved in integrated circuits (IC) manufacturing there have been attempts to design vast amounts of network on the chips in order to achieve more efficient and optimized chips. A better routing algorithm can enhance the performance of NoC. XY routing algorithm is a distributed deterministic algorithm. Odd-Even (OE) routing algorithm is distributed adaptive routing algorithm with deadlock-free ability. Every NoC should satisfy some performance requirements like low latency, high throughput and low network power. Here we demonstrated the impact of traffic load variations on average latency, average throughput and total network power for two routing algorithms XY and OE on a 3x3 2- dimensional mesh topology. The simulations have been performed on NIRGAM NoC simulator version 2.1 for constant bit rate (CBR) traffic condition. The simulation results contains overall average latency (clock cycles per packet), average throughput (in Gbps) and total network power (in mW). Performance metrics (P) is calculated for both routing algorithms and compared.

**Keywords**—*Network on Chip, Router, XY routing algorithm, OE routing algorithm, Performance metrics*

## I. INTRODUCTION

A rapid progress in Very Large Scale Integration (VLSI) in the past recent years has resulted in the fabrication of millions of transistors on a single silicon chip. This advancement in the micro-electronics leads to the integration of various components of a computing system or any other electronic system on a single Integrated Circuit (IC) to implement a complete system on a chip. Thus a paradigm called System on Chip (SoC) came into existence that refers to the system made up of interconnected cores or Intellectual Property (IP) blocks on a single chip. So an effective communication system is required between the IP blocks of SoC [1]. Nowadays direct interconnections and mostly shared busses are used for onchip communication. The problem with direct interconnections is that they are not scalable and become inefficient with an increase in the number of cores [2]. Shared buses do not give satisfactory results when scaled beyond 8 to 10 cores.

Contention for the bus and arbitration also slows down the data movement. They are only good for the systems with less number of connections. So Network on Chip (NoC) is being considered as the most suitable candidate for implementing interconnections in core based system on chip (SoC) design. In NoC paradigm, cores are connected to each other through a network of routers and they communicate among themselves through packet-switched communication [2]. A SoC with NoC infrastructure is shown in Fig 1.

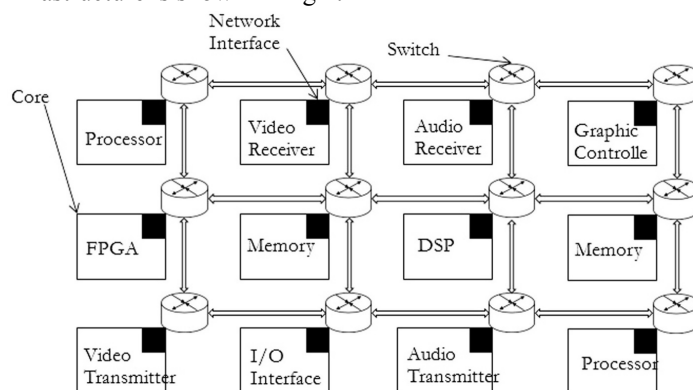


Fig.1. SoC with NoC

## II. NOC ARCHITECTURE AND FUNCTION LAYER

### A. Overview Of NoC Architecture

A typical NoC architecture consists of multiple segments of wires, routers, IP cores (Resources), RNI (Resource Network Interface) as shown in Figure .1. In a tiled, city-block style of NoC layout, the wires and routers are placed similar to street grids of a city, while the clients (e.g., IP cores or Resources) are placed on city blocks separated by wires. The IP cores or resources can be General Purpose Processors, DSP, memory, application specific hardware component, I/O controller, Graphic controller, Mixed signal Module, RF unit etc. Resource should be implemented by using same technology as that of switch network. RNI enables the resource to send data to router. RNI consists of 2 parts i.e. Resource dependent part and Resource independent part (Fig 2).Resource independent part is designed in such a way that

RNI appears as another router to the connected router. If homogenous resources are used then resource dependent part can be reused. Resource dependent part is responsible for flitization, deflitization and implementing encoding scheme where flitization is formation of flits by dividing the packets and deflitization is formation of packet by combining the flits. Router in NoC is required to forward an incoming packet to destination resource if it is directly connected to it or to forward packet to another router connected to it. For an incoming packet a route can be selected either by looking up a routing table already stored in the router memory or route can be computed by executing a routing algorithm. Router implements layers of communication protocols below and including the network layer (OSI model).

### B. NoC Function layers

The NoC function can be classified into several layers: application, transport, network, data link, and physical layers. A NoC router must contain both software and hardware implementations to support functions of the layers. Physical layer deals with the actual transfer of data. It is responsible for clock signals for every connection, number of wires, control signals, electrical level and medium of transfer etc. Data link layer is concerned with packet formation, node to node communication, error detection and correction, flow control, encoding scheme etc. Network layer performs packetization, routing of packets from source to destination, resource addressing, packet buffering and congestion control. The transport layer addresses the congestion and flow control issues to prevent buffer overflow and to avoid traffic congestion. At the application layer, target applications will be broken down into a set of computation and communication tasks such that the performance factors like energy and speed can be optimized.

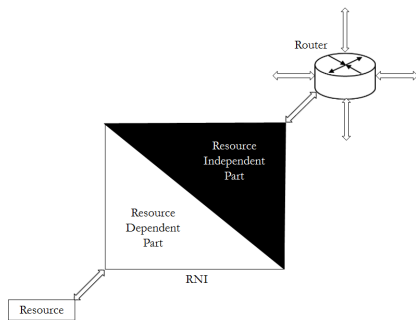


Fig.2. Resource Network Interface

### III. BASIC NOC ROUTER ARCHITECTURE

A generic router consists of five ports i.e. east, west, north, south and local port and a central crosspoint matrix [10]. Each port has an input channel and an output channel. Data packets move into the input channel of one port of router by which it is forwarded to the output channel of other port. The input and output channels have their own decoding logic which enhances the performance of the router. Buffers work as temporary storage of data. Here the buffering method used is store and forward. Control logic is required to make

arbitration decisions. Thus, a communication is set up between the input and output ports. This connection or configuration between these ports is formed by the central crosspoint matrix.

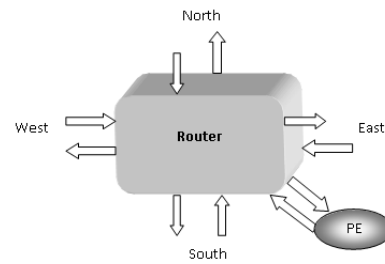


Fig.3. Generic NoC Router

### IV. OVERVIEW OF NOC DESIGN CONCEPTS

There are three concepts for NoC design i.e. topology, switching techniques, routing algorithm.

#### A. Topology

Interconnection networks are made up of a set of shared router nodes and channels, and *topology* of the network implies the arrangement of these nodes and channels. The topology of an interconnection network is analogous to a roadmap. A good topology exploits the characteristics of the available packaging technology to meet the bandwidth and latency requirements of the application at minimum cost. There are various topologies for NoC e.g. Mesh Torus, Tree, Polygon, Butterfly, Spidergon, Star etc. But Mesh and Torus topology is very simple in design point of view. A mesh-shaped network consists of  $m$  columns and  $n$  rows. The routers are situated in the intersections of two wires and the computational resources are present near routers. Addresses of these routers and resources can be easily identified by x-y coordinates in mesh. One of the advantages of Mesh Topology is the easy detection and isolation of faults in the network. It is also more protected since messages go through a dedicated line and the message will only reach its intended addressee. One of the disadvantages of mesh topology is it will require a huge investment if the network is large. A Torus network is an improved version of basic mesh network. A simple torus network is a mesh in which the heads of the columns are connected to the tails of the columns and the left sides of the rows are connected to the right sides of the rows. Torus network has comparatively better path diversity than mesh network, and it also has more minimal routes.

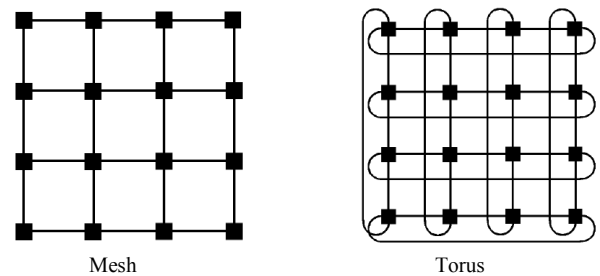


Fig.4. Two simple topologies for NoC

## B. Switching Techniques

Switching techniques are used to move data from the input channel to the output channel of a router. Latency in NoC mostly depends upon the switching technique implemented. Examples of switching techniques are packet switching (Store and forward), circuit switching, virtual cut-through and wormhole switching. In store and forward, also called packet switching, a complete packet moves from one router to the next. It has a greater communication delay because a packet can only be forwarded when it is entirely received. Moreover a large buffer of at least equal to one packet size is required. In circuit switching, an electrical connection is established between source and destination before transmitting the data. The connection is free after the data has been sent out. Although, this switching guarantees reliable transfer of data but network resources are underutilized. In virtual cut-through switching, the router starts forwarding the packet as soon as header arrives and the destined output channel is idle. There is no need to linger for the entire packet arrival. But if the intended output channel is busy then the complete packet has to be stored by the switch. Therefore the buffer constraint in switches is of at least one packet. In wormhole switching, a packet is divided into flits. Flit is a flow control digit. There are generally three kinds of flits i.e. head, body and end. The head flit carries the control and routing information of a packet; body flits carry the payload and end flit contains payload as well as end of packet information. Thus buffer required in routers is very small and can be as small as one flit. The flits appear to move through the network like a worm and hence the technique gets the name wormhole switching [4].

## C. Routing algorithm

It determines the path that each packet follows between source and destination pair. There are many ways to classify routing in NoC. Most commonly used classes are discussed below.

- Deterministic and Adaptive routing
- Minimal and Non minimal routing
- Static and dynamic routing

In deterministic routing the path from the source to the destination is completely determined in advance by the source and the target addresses. In adaptive routing, several paths from the source to the destination are possible. When a packet enters a router, destination address is read from the header and accordingly, the routing function computes all possible output ports where this packet can be forwarded to. Then a routing function selects one of the admissible output ports to forward the packet. The selectivity of output port depends upon the dynamic network conditions such as congestion and link faults. There also exist partially adaptive routing algorithms which restrict certain paths for communication. They are simple and easy to implement compared to adaptive routing algorithms. A routing which uses shortest possible paths for communication is known as minimal routing. It is also possible to use longer paths for data transfer from source to destination. This possibility results from the adaptivity offered

by a routing algorithm. The type of routing which uses longer paths for communication although shortest paths do exist is known as non-minimal routing. However, non-minimal routing has some advantages over minimal routing including possibility of balancing network load and fault tolerance. In static routing, the communication path cannot be altered after a packet departs the source. In dynamic routing, a path can be changed any time depending upon the network conditions. Source routing is static while distributed routing can be static or dynamic depending upon the routing algorithm used. It should be noted that even when adaptive routing algorithms are used to compute paths for source routing, it remains static unless some sophisticated selection technique is introduced in the network.

## V. PERFORMANCE REQUIREMENTS FOR NOC

Performance requirements that every NoC must satisfy

- Small latency
- Guaranteed throughput
- Path diversity
- Sufficient transfer capacity
- Low power consumption
- Fault and distraction tolerance
- Architectural requirements of scalability and programmability

But there are three important parameters for NoC which we are considering in this paper i.e. Network Latency, Network Throughput and Total Network Power.

Network Latency is measured from the time its head flit is generated by the source to the time its tail flit is consumed by the destination. Let  $L_{ij}$  be the packet  $j$  and  $N_i$  be the number of packet received by processor  $i$  (After warm-up time).  $N$  is the number of processors in the platform.

$$\text{Average Network Latency} = L_{\text{avg}} = (1/N) \sum (1/N_i \sum L_{ij}) \quad (1)$$

for  $i = 1, 2 \dots N$  and for all  $j$

Network Throughput is defined as the rate at which the network can successfully accept and deliver the injected packets. Let  $T_{\text{sim}}$  and  $T_{\text{warm}}$  be the simulation time and warm-up time respectively.

Average Network Throughput (in packets per unit time per node) is given by

$$T_{\text{avg}} = 1 / (N (T_{\text{sim}} - T_{\text{warm}})) \sum N_i \quad (2)$$

where  $i = 1, 2 \dots N$

Total Network power estimation is based on activities of components while running a certain traffic pattern. It should be less for all NoC.

## VI. ROUTING ALGORITHM DESCRIPTION

### A. XY Routing Algorithm

The XY routing algorithm is one kind of distributed deterministic routing algorithms. For a 2-Dimesion mesh topology NoC, each router can be recognized by its coordinates  $(x, y)$ . Here the packet is transmitted first in X direction then in Y direction (Figure 5). The XY routing algorithm compares the current router address  $(C_x, C_y)$  to the

destination router address (Dx,Dy) of the packet, stored in the header flit. Flits must be routed to the core port of the router when the (Cx,Cy) address of the current router is equal to the (Dx,Dy) address. If this is not the case, the Dx address is initially compared to the Cx (horizontal) address. When  $Cx < Dx$ , the flits will be routed to the East port, when  $Cx > Dx$ , to West and if  $Cx = Dx$  the header flit is already horizontally aligned. If this last condition is true, the Dy (vertical) address is compared to the Cy address. Flits will be routed to South when  $Cy < Dy$ , to North when  $Cy > Dy$ . If the chosen port is busy, the header flit and all succeeding flits of this packet will be blocked [5].

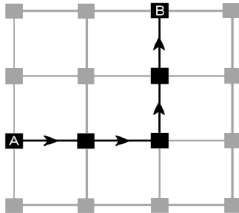


Fig.5. XY routing from router A to router B

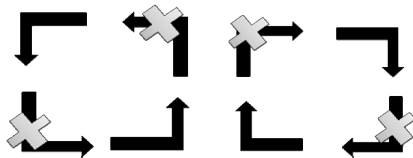


Fig.6. Allowed Turns in XY routing

### B. OE Routing Algorithm

OE routing algorithm is a distributed adaptive routing algorithm which is based on odd-even turn model. It applies various restrictions, for avoiding and averting deadlock occurrence [11]. Odd-even turn model assists deadlock-free routing in two-dimensional (2D) meshes with no virtual channels. Elucidation of some definitions is essential in order to represent this algorithm. In a two-dimension mesh with dimensions  $X*Y$  each node is identified by its coordinate (x, y). In this model, a column is called even if its x dimension element is even numerical column. Also, a column is identified as odd if its x dimension element is an odd number [6]. A turn engages a 90- degree change of traveling direction. A turn is a 90-degree turn in the following description. There are eight kinds of turns, as per the traveling directions of the related channels. A turn is called an ES turn if it engages a variation in direction from East to South. Similarly, we can classify the other seven kinds of turns, namely EN, WS, WN, SE, SW, NE, and NW turns, where E, W, S, and N specify East, West, South, and North, respectively. All together, there are two main theorems in odd-even algorithm:

- Theorem1: No packet is permitted to do EN turn in each node which is located on an even column (Fig.7). Also, No packet is permitted to do NW turn in each node that is located on an odd column (Fig.8).
- Theorem 2: No packet is permitted to do ES turn in each node that is in an even column (Fig.7). Also, no

packet is permitted to do SW turn in each node which is in an odd column (Fig.8).

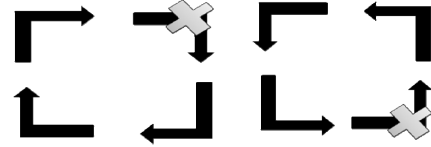


Fig.7. Allowed turns in OE routing (even column)

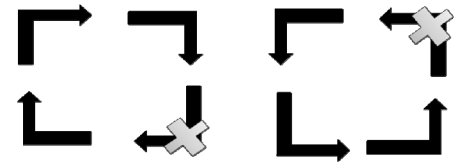


Fig.8. Allowed turns in OE routing (Odd column)

## VII. EXPERIMENTAL SETUP

NIRGAM (NoC Interconnect Routing and Applications' Modeling) is a systemC based simulator [8]. It is used in this experiment to calculate performance parameters like average latency, average throughput, total network power in Mesh topology with 3x3 dimension. Here each node is connected with CBR (Constant Bit Rate) traffic generator. CBR value is 12 Gbps. Packet size is 20 bytes and flit interval is 2 clock cycles where clock frequency is 1GHz. This topology will be simulated with warm-up time 800 clock cycles and total simulation time 50000 clock cycles. Here % of Load (% of maximum bandwidth utilized) is going to be varied and according to that the effect on performance parameters will be observed.

## VIII. RESULTS AND GRAPHS

% of Load	Overall Average Latency per channel (clock cycle/packet)	
	XY	OE
10	11.251	10.5371
30	11.4877	10.7535
50	13.5859	12.3141
70	13.1227	13.3141
90	14.7451	12.4611
100	29.4295	26.1582

Table .1. % of Load vs Overall Average Latency

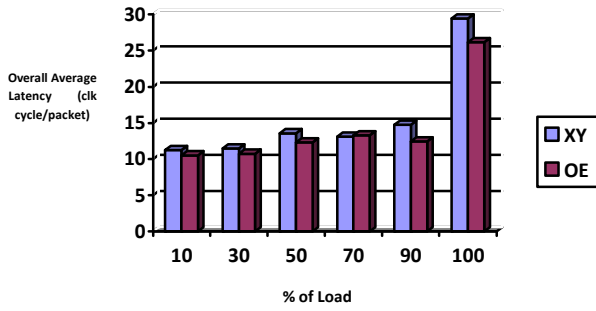


Fig. 9 % of Load vs Overall Average Latency

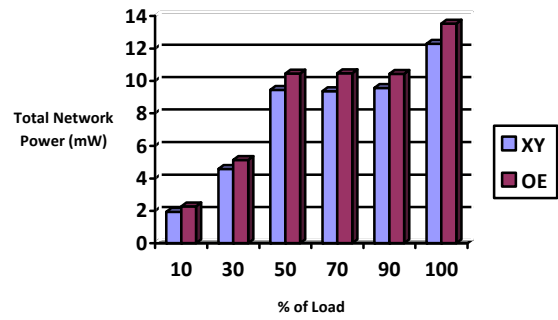


Fig.11 % Load Vs Total Network power

% of Load	Average Throughput (Gbps)	
	XY	OE
10	1.96537	2.28772
30	3.58726	5.12377
50	8.198	9.2098
70	9.33423	10.4799
90	9.96787	10.9844
100	10.2839	11.9231

Table.2. % Load Vs Average Throughput

### IX. PERFORMANCE METRICS CALCULATION

Performance metrics is the ratio between average throughput and average latency. More the “P” better the Routing Algorithm.

$$P = \text{Performance Metrics (Per channel basis)} = \frac{\text{Average Throughput}}{\text{Average Latency (clock cycle per packet)}}$$

For XY Routing (50% Load)  $P = (8.198/13.5859) = 0.6034$   
 For OE Routing (50% Load)  $P = (9.2098/12.3141) = 0.7479$

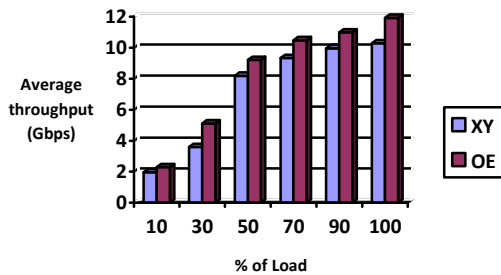
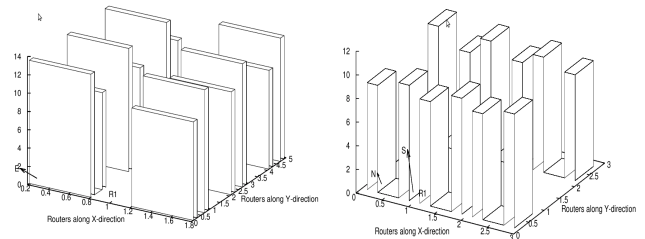


Fig. 10 % Load Vs Average Throughput

% of Load	Total Network Power (in mW)	
	XY	OE
10	1.95374	2.28772
30	4.58726	5.13877
50	9.45833	10.4572
70	9.38296	10.4799
90	9.56787	10.4404
100	12.2839	13.5231

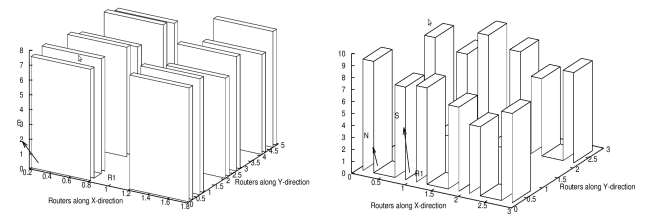
Table.3. % Load Vs Total Network power



X-LATENCY

Y-LATENCY

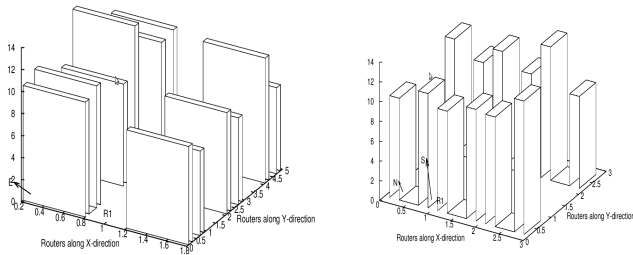
Fig.12. Average latency per packet for XY routing



X-THROUGHPUT

Y-THROUGHPUT

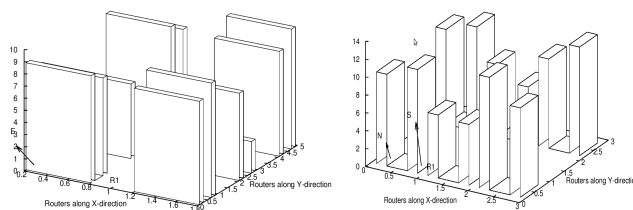
Fig.13. Average throughput for XY routing



X-LATENCY

Y-LATENCY

Fig.14. Average latency per packet for OE routing



X-THROUGHPUT

Y-THROUGHPUT

Fig.15. Average throughput for OE routing

In Fig.12 and Fig.14 Average latency per packet graph is plotted for 9 nodes for both XY and OE routing algorithm where X-latency indicates the latency value in X-direction(East and West) and Y-latency indicates the latency value in Y-direction(North and South) in NoC. In Fig.13 and Fig.15 Average throughput graph is plotted for 9 nodes for both XY and OE routing algorithm where X-throughput indicates the throughput value in X-direction(East and West) and Y-throughput indicates the throughput value in Y-direction (North and South) in NoC.

## X. CONCLUSION

The routing algorithm is one of the research area of network layer for NoC design whose approach of design can be adapted from a protocol stack including functional layers of OSI model. Based on a 2-Dimension 3x3 mesh topology NoC with CBR traffic, two different routing algorithms i.e. XY routing algorithm and OE routing algorithm are simulated on NIRGAM simulator platform and impact of Percentage Load variation is compared with three different parameters namely overall average latency (in clock cycles/packet) , average throughput (in Gbps) and total network power (in mW).

OE routing algorithm has less overall average latency than XY routing algorithm. In average throughput case we can see the complete dominance of OE over XY routing algorithm. But in total network power case OE has more power consumption than XY for all load conditions. So XY routing algorithm is more power efficient than OE routing algorithm. The Performance metrics shows that OE routing algorithm is better routing algorithm than XY routing algorithm in Network on Chip design.

## REFERENCES

- [1] W. Dally and B. Towles, "Route packets, not wires: onchip interconnection networks," Proc. Design Automation Conference, Jun.2001, pp. 684-689.
- [2] S.Kumar , A. Jantsch, J. Soininen, M. Forsell, M. Millberg, J. Oberg. et al, "A network on chip architecture and design methodology," Proc. IEEE Computer SoC Society Annual Symposium on VLSI
- [3] J. Duato, S. Yalamanchili, L. Ni, Interconnection networks: an engineering approach, Morgan Kaufmann, Revised Edition, 2002.
- [4] L. M. Ni, P. K. Mckinley, "A Survey of Wormhole Routing Techniques in Direct Networks", Computer, vol. 26(2),Feb 1993, pp. 62 -76.
- [5] 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", WRI Global Congress on Intelligent Systems, GISIS'09, May 2009, pp. 329-333.
- [6] Ge-Ming Chiu, "The odd-even turn model for adaptive routing", IEEE Transactions on parallel and distributed systems, Jul 2000, pp. 729-738.
- [7] Palesi M., Holsmark R., Kumar S. and Catania V, "Application Specific Routing Algorithms for Networks on Chip", IEEE Transactions on Parallel and distributed systems. Vol.20,No. 3, March 2009.
- [8] L. Jain., "NIRGAM: A Simulator for NoC Interconnect Routing and Applications Modeling", DATE conference, Sep 2007, pp. 1-2.
- [9] R. Marculescu, H. Jingcao, U. Y. Ogras, "Key research problems in NoC design:a holistic perspective," Proc. Third IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis(CODES+ISSS'05),Sept.2005.
- [10] N. Kavaldjiev, G. J. M. Smit, and P. G. Jansen, "A virtual channel router for on-chip networks", IEEE International SOC Conference, pp. 289-93, USA,2004
- [11] A. Patooghy, H. Sarbazi-Azad, "Performance comparison of partially adaptive routing algorithms", 20<sup>th</sup> International conference on Advanced Information Networking and Applications,2006
- [12] P. Guerrier and A. Greiner, "A generic architecture for on-chip packet-switched interconnections", Proc. Design, Automation and Test in Europe Conference and Exhibition (DATE 2000), Mar. 2000, pp. 250-256.