

Performance Explorations of Multi-Core Network on Chip Router

U.Saravanakumar

Department of Electronics and Communication Engineering
PSG College of Technology
Coimbatore, India
saran.usk@gmail.com

R. Rangarajan

Department of Electronics and Communication Engineering
Indus College of Engineering
Coimbatore, India
profrr@gmail.com

Abstract — Due to minimization of communication latency, timing constraints and energy consumption, Network on Chip (NoC) dominated Multi-Core System on Chip (SoC). In order to keep up the balance between power, area, performance and robustness to traffic changes in NoC, many research works conduct by designers. Here, we proposed new router architectures for Multi-Core NoC which gives less slack time than the conventional Wormhole router architecture. Different stages of pipelining method and hierarchical concept in scheduler implemented at the input side of proposed router. The working function of new architecture had verified by simulation and area, power and delay calculated by Synopsys tool in UMC 0.13 μm . FPGAs are identified as an incarnation of NoC, then the proposed architecture implemented in Virtex II Pro. And finally proposed router architecture compared with previous router architectures in terms of power, area and slack time.

Keywords - Multi-Core, System on Chip, on – chip communication, Network on Chip, Pipeline, FPGA, Power, Flit size, Slack time

I. INTRODUCTION

System on Chips (SoCs) offer better solutions to modern electronic devices, telecommunication applications, mobile and multimedia. According to International Technology Roadmap for Semiconductor devices (ITRS), at the end of decade, SoC designs have faced many design challenges like functionality, reliability, synchronization, and electrical noise due to crosstalk and (Electromagnetic Interference) EMI [1]. Another design challenge in SoCs is determinism, which causes by transmitting a digital data via wires, design components with higher level of design abstraction, granularity and control system for distributed communication [2], [3].

In order to reduce the above challenges, designers still searching and developing new paradigms in various ways. One of such methods is adapting network technology in to SoCs because it has viewed as micro – network of components. But this micro – network architectures are different from wide area network due to their unique characteristics such as energy constraints and design time specialization. Micro – network architecture for SoCs has developing with various architectures like shared medium networks, direct network, indirect network and hybrid network. Network architectures identified as a major part to affect the area, power consumption and communication latency of Network on Chips (NoCs). At the same time, designers must aware of switch design, link design and Network Interface design to offer better NoCs [4]. To offer efficient NoC architecture, many researchers have studied about routing algorithms, topologies and switch fabric. While comparing, router and switch fabric consuming more power and area. With various techniques designers tried and evaluated to reduce area, power and latency [5].

Remainder of this paper is organized as follows, second chapter will discuss about NoC architecture and third chapter will deal with basic elements of NoC router architecture. Switching mechanisms, related works, result analysis and conclusion parts will be discussed in chapter fourth, fifth, sixth and seventh respectively.

II. MULTI-CORE NOC ARCHITECTURE

The Network architecture for single chip multiple processor describes the topology and physical connections. Perhaps signal propagation time (delay) is greater or lesser than processors cycle time, physical property of network have a strong reason on performance of Network. A network architecture works based on topology with the help of switching, routing and flow control. At the same time designers have verity of choices to select topology to meet their specific requirements. The elements of network architecture are processing and storage elements also refer as nodes, switches and physical links where nodes incorporate with switches.

Based on physical link between nodes, network architecture have classified as shared medium network (single physical link can be shared by all nodes in a network but only one node can send data through physical link), direct network (every node has a switch, switches have point to point connection in a network), indirect network (every node has connected to the switch), hybrid network (combination of different network architectures) [6], [7], [8]. Network architectures play an important role to realize on chip effectiveness such as static and dynamic costs. And finally NoC designers must consider properties of network architecture for single ship multiple processors. The main properties of network architectures are deadlock, livelock and starvation. Figure 1 describes the 2 dimensional mesh

based network architecture for multiple processors System on Chip.

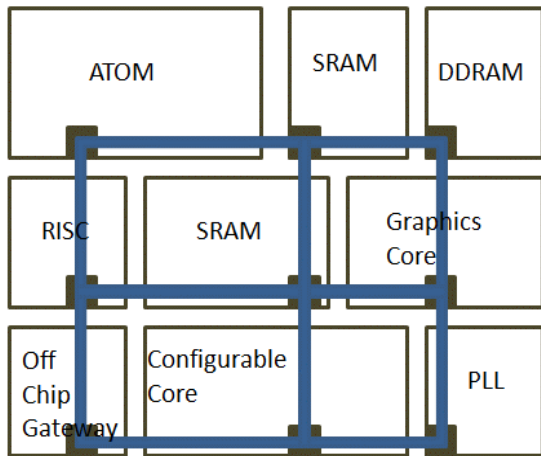


Figure. 1 Typical 2D Mesh NoC Architecture for Multi-Core Single Chip

III. BASIC ELEMENTS OF NOC

From the above discussions, NoC is a structure with collection of finite number of routers; each router has five input/output ports. Each router in a network structure is identified using its coordinate numbers or positions. Among five input/output ports, four input/output ports are dedicated to communicate with neighbour router, and remaining one is dedicated to communicate with local or core elements. As shown in Figure 2, the main elements of router are Switch Fabric (SF), Core Interface (CI), Network Interface (NI), Physical link (Bi-directional). The detailed explanations of each element are discussed below [9], [10].

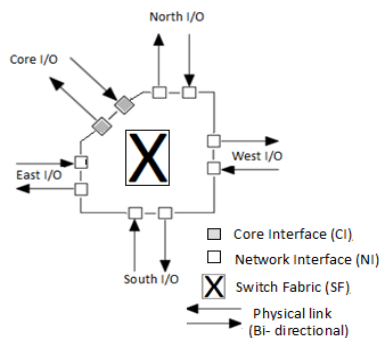


Figure 2 2-D Mesh Router Architecture

A. Physical Link

Physical link of router consists of bidirectional ports. The following sub-sections describe the explanations of input

and output ports. Each input port has its own control and routing logic, buffer. In most of the router design, First in First out (FIFO) used as a buffer and FIFO is designing with shift register or dual line Static RAM. The size and depth of FIFO is fixed based on designers' requirement. The position of FIFO fixes communication process, which is described in Figure 3.

```

if (FIFO = empty)
    control signal => high
    data write => start
    communication => start
Else if (FIFO = full)
    control signal => low
    data write => stop
    data read => start
    forward read data to destination => start
    
```

Figure 3 Status of FIFO buffer

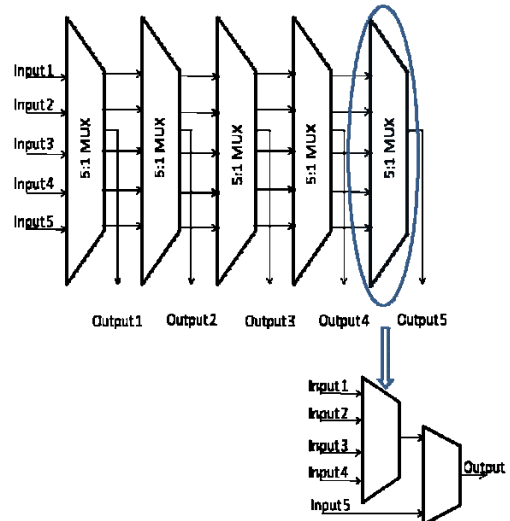


Figure 4 Switch Fabric using MUX

Each output port has its own functional logics as similar with input port except arbiter. The main function of arbiter is to avoid contention at output port. If there are more than one request from input ports to one output port, arbiter selects anyone of the requests based on highest priority, because arbiter uses rotating priority method. Among the five input/output ports, input port from east direction has highest priority. The main need of arbiter is, to improve the communication service via contention reduction [11], [12], [13].

B. Switch Fabric

In router design, this switch fabric part have considered as important and central part. Switch fabric has two different kind of design method, one is Pass transistor based

and another one is Multiplexer-based. According to input/output ports count of router, switch fabric will have cross point connections (Eg. For our design, switch fabric have 25 cross point connections because we have designed 5 input/output ports router). Figure 4 describes the architecture of switch fabric [14].

IV. SWITCHING MECHANISMS

In network, each router communicates with neighbouring router, and it maintains forwarding an incoming packet until it reaches destination. Switching mechanisms have direct impact on network performance and buffer size through controlling packet forward process. Following sub-sections deal about different switching mechanisms.

A. Store and Forward

Data packet is split in to multiple flits as Header flit, Body flits and Tail flit. Each and every router receives header flit first, and then body flits. In order to avoid buffer overflow, keep the buffer size larger than total packet size. After receiving tail flit, router decides to send header flits of incoming data packet to neighbouring routers. The main advantage of this mechanism is simple control process and drawback is buffer size. Figure 5(a) describes the store and forward mechanism.

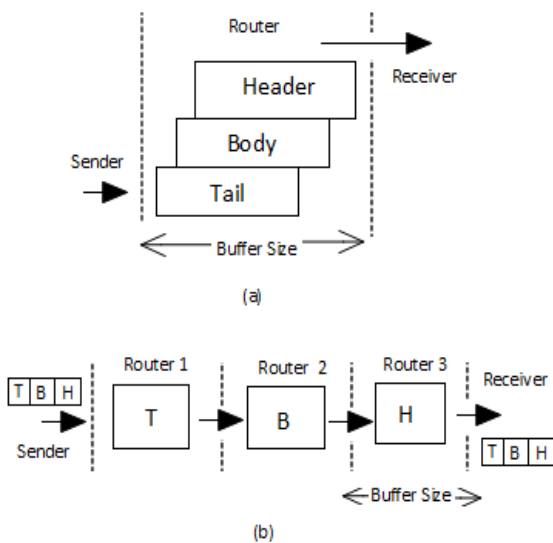


Figure 5 (a) Store and Forward (b) Wormhole

B. Wormhole

Wormhole mechanism is different from earlier method. In this mechanism, routers store only one flit and send the stored flit before arriving next flit to neighbouring routers as shown in Figure 5(b). Therefore, parts of a packet can be stored in multiple routers. The main advantage of this

mechanism is, less buffer size but larger than header flit. But the main drawback of wormhole mechanism is blocking of continuous process due to contention of header flit.

C. Virtual Cut Through

Virtual channel buffers employ to store the whole packet and reduce the blocking of continuous process. This technique is called as Virtual Cut Through mechanism and it has the advantage of low latency [15].

V. PREPARE YOUR PAPER BEFORE STYLING

For our study analysis, we chose Wormhole (WH) router due to its merits. The detailed working functionality of this router and proposed pipeline techniques will be given in the following sub sections.

A. Conventiaonl WH Router

Both the Circuit Switched (CS) and Virtual Cut-Through (VCT) routers use the concept of Cut-Through, which offer low latency network. The latency of these routers is independent of path length. VCT router requires extra buffer for blocked packets and CS router faces the problem of channel sharing. Therefore, Wormhole (WH) router introduced to overcome the above challenges without any compromise in latency.

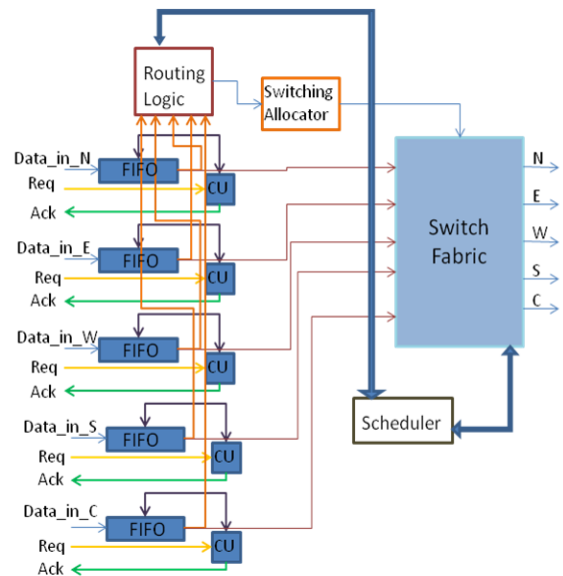


Figure 6 (a) Conventional Wormhole router

1 bit cntrl	32-bit data packet	3-bit Source	3-bit destination
----------------	--------------------	-----------------	----------------------

Figure 6 (b) Data Packet format

Figure 6(a) shows the block diagram of WH router. The main blocks of WH router are Buffer, Buffer control unit (CU), Switch Allocator, Switch Fabric and Scheduler. The primary functions of this router include flow control, routing, switching, multiplexing physical connections and clock recovery. Here, switch fabric is designed with the help of Multiplexer and Stop – Go mechanism used to manage buffers. The depth of buffer is 8 bit. CU is used to check the status of the buffer whether it is full or empty. And finally scheduler is designed with i-SLIP algorithm and Round Robin Arbiter to provide congestion free data transmission from input to output port if multiple requests to access single output port [16]. Figure 6 (b) shows the packet format, it consists of 32 – bit data, 3 – bit source and destination addresses and 1 – bit control.

B. Case 1: Proposed Work

In order to improve the performance of WH router, here we introduced 5 – stage pipeline technique in this router and pipeline stages are Input Fetching, Path Allocation, Crossbar Stage, Switch Allocator Stage and last Destination Stage. The reason to introduce 5 stage pipeline rather than 2 or 3 or 4, length of the longest stage provides maximum frequency of router. Figure 7 describes the each stages of pipeline. Horizontal position shows the number of clock cycles and vertical position shows the stages in the router.

PS/CC	1	2	3	4	5	6	7	8	9
1	IF	BS	SA	SWS	DS				
2		IF	BS	SA	SWS	DS			
3			IF	BS	SA	SWS	DS		
4				IF	BS	SA	SWS	DS	
5					IF	BS	SA	SWS	DS

IF - Input Fetching ; BS - Buffer Stage ; SA - Switch Allocator Stage
 SWS - Switch Fabric Stage ; DS - Destination Stage
 PS - Pipeline Stages ; CC - Clock Cycles

Figure 7 Five stages of pipeline technique

The registers were added to store the data in the five stages such as IF, BS, SA, SWS and DS (i.e. Registers are added in the input channels, Path allocation stage, Switch Allocator, Cross bar and output stage). The proposed router takes 3 clock cycles to transfer the 1st data, but after 3 cycles the data will be received to the destination for each cycle. The speed of operation has been increased and reduces the critical path. But the area, power and latency has been increased slightly due to the additional registers.

C. Case 2: Proposed Work

Current trends in Multi – core technology involved to balance trade – off between area, power and speed. In case

1, we proposed 5 – stage pipeline architecture to reduce slack time, so it increased the processing speed of router for NoC but considerable amount of area and power are increased. Here we implemented a WH router based on hierarchical approach in scheduler which is developed based on iSLIP algorithm with Round Robin Arbiter (RRA) and 5 – stage pipeline. Generic RRA consists of two blocks, one is input selector and another one is pointer updater. Input selector selects one of the multiple requests from input ports and pointer updater generates a grant signal for next possible request. In Hierarchical RRA, first process is to divide number of inputs in to k subsets and which are called as Sub RRA (do the functions same as generic RRA), and they have local request. Local requests and grants are done in multiple stages of sub RRAs based on priority settings. A simple pass signal is used for smooth transitions between sub RRAs. These sub RRAs are arranged in two stages as hierarchically. First stage consists of sub_RRA_0s (select one request among multiple request signals to Hierarchical RRA based on priority, if no request, No request signal will be generated), second stage consists of sub_RRA_1s (generate grant signal for selected request signal by sub_RRA_0s, if no request, No request signal will be generated). Figure 8 shows scheduler designed with Hierarchical RRA [17].

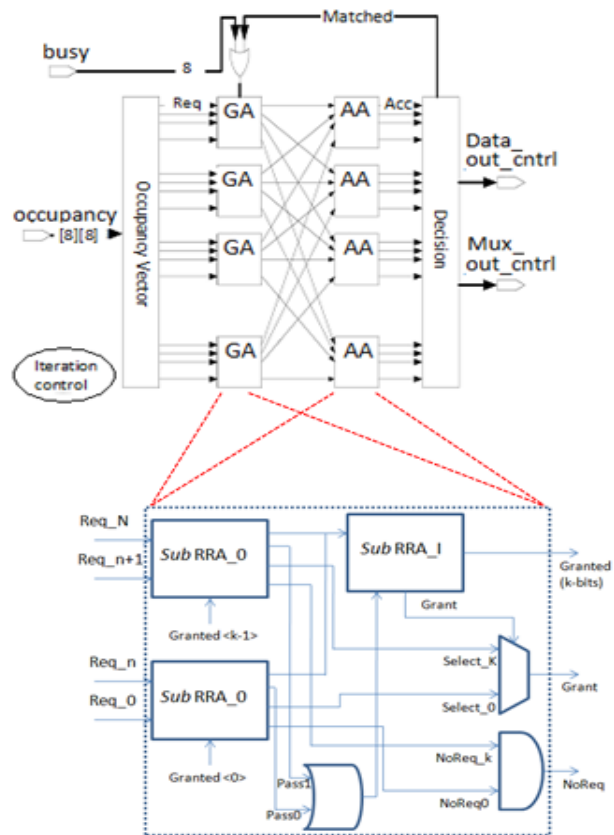


Figure 8 Scheduler with Hierarchical RRA

At last, we designed a novel router with 5 stage pipeline and hierarchical round robin arbiter to provide less area, less power and less slack time.

VI. RESULTS AND DISCUSSIONS

Both routers WH and Wormhole with Pipeline (WHWP) were described using Verilog HDL. Modelsim XE is used for the functional and logic simulation. The Xilinx integrated tool environment ISE 09.2i is used for the automated logic synthesis, mapping, placing and routing of circuits. Tools included in this environment generate reports describing the area, power and delay of implementation, a netlist used for timing simulation, and a bitfile used to configure on targeted FPGA device.

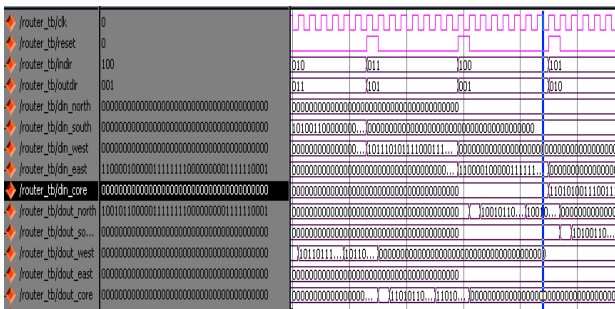


Figure 9 (a) Simulation result of conventional WH router

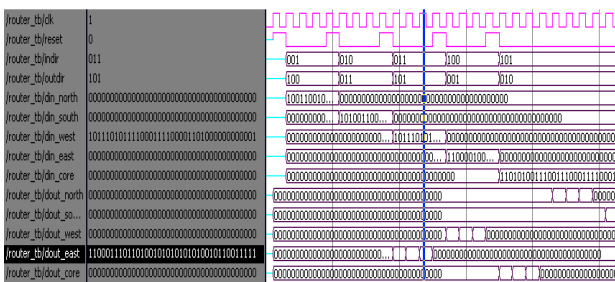


Figure 9 (b) Simulation result of Proposed (case 2) router

Simulated output of WH and proposed (with combination of case 1 and 2) routers are described in Figure 9 (a) and (b). Verilog HDL code of proposed routers was synthesized and implemented in XILINX Virtex II Pro (selected device: 2v40cs144-6) FPGA Board. RTL schematic view of proposed router was achieved by synthesizing Veilog Code using Xilinx. Some hardware utilization feature parameters were measured and noted in Table 1. Other parameters like area, power and timing were calculated in 130 nm CMOS technology using Synopsys tool with nominal operating voltage of 1V and packet size is 39 bits. Library file used to calculate the area, power and slack time in Synopsys is fsc0h_d_sc_bc. By applying pipeline technique, slack time is reduced from 3.95 ns (data required time = 4.81, data arrival time = -0.86) to 2.04 ns (data required time = 3.83, data arrival time = -1.79). In order to

verify the performance of this proposed architecture is then compared with Circuit Switch (CS) router and CS with Pipeline (CSWP) router. Figure 10 compares the cell internal and switching power of all the above four routers with packet size of 39 bits. Table 2 shows the comparison of area, power and slack time of CS, CSWP, WH and proposed routers. Similar with WH and proposed routers, CS and CSWP routers are described with Verilog HDL. Power, Slack and Area of CS and CSWP routers also calculated using Synopsys Design Compiler Tool with UMC 130nm Technology. The cell internal power is increased by 9% while net switching power is reduced by 11% with pipeline technique. Table 2 compares all the parameters of proposed router with conventional WH router. Based on the data in the table, total area of proposed router (case 1) is increased from 38516 μm^2 to 44623 μm^2 by introducing pipeline scheme, total dynamic power consumption is increased from 8.2233 mW to 11.254 mW and leakage power is increased from 6.2120 μW to 9.5163 μW .

TABLE I. SYNTHESIZED PARAMETERS

Parameters /Routers	CS	CSWP	WH	Proposed (case 1)
CB Area (μm^2)	4190	4040	8821	7450
NCB Area(μm^2)	7989	15549	29695	37173
TC Area (μm^2)	12179	19589	38516	44623
DP (mW)	4.112	6.8256	8.2233	11.254
LP (μW)	2.7402	4.2768	6.2120	9.5163
ST (ns)	4.12	3.29	3.95	2.04

CB = Combinational Block, NCB = Non Combinational Block, DP = Dynamic Power, LP = Leakage Power, ST = Slack Time.

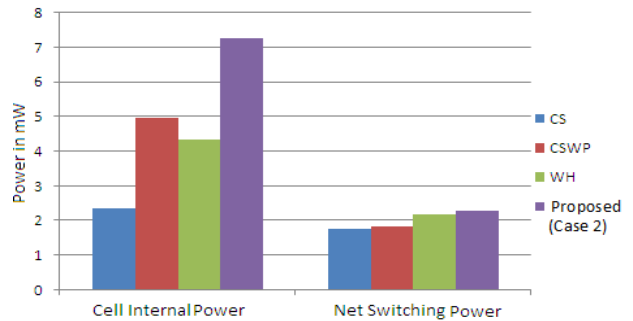


Figure 10 Power consumption of CS, CSWP, WH and Proposed (case 2) routers

TABLE II

COMPARISON OF AREA, POWER AND SLACK TIME OF CS ROUTER AND WH ROUTER AND PROPOSED ROUTER (CASE 1)

Details	WH	Proposed (case 1)
Number of Ports	392	392
Number of Nets	435	633
Number of Cells	197	211
Number of References	5	5

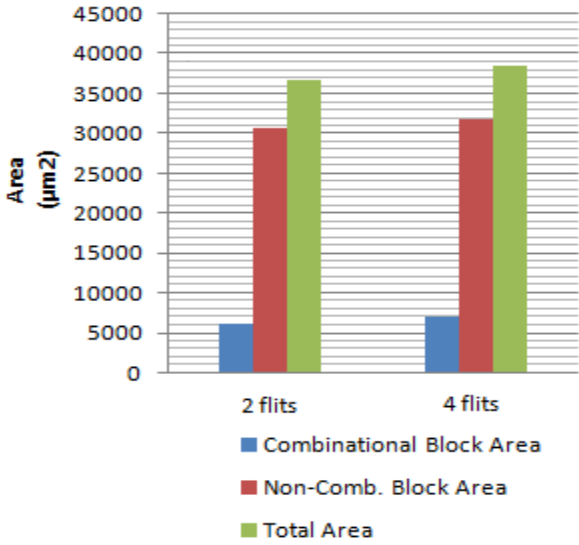


Figure 11 Area of Proposed (case 1) Router with different flit sizes.

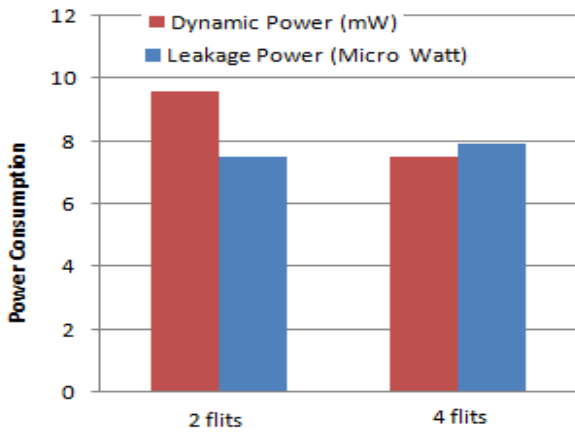


Figure 12 Power consumption of Proposed (case 1) with different flit sizes

TABLE III
COMPARISON OF AREA AND POWER OF CONVENTIONAL WH ROUTER AND PROPOSED ROUTERS (CASE 1 AND 2)

Parameters	Conventional WH Router	Proposed router	
		Case 1	Case 2
Dynamic Power (mW)	8.2233	11.254	7.640
Leakage Power(Micro Watt)	6.212	9.5163	6.4603
Total Cell Area (Micro m2)	38516.19	44623.26	35252.38

By combining pipeline technique and hierarchical approach to generic RRA, conventional WH router achieved less slack time with less power and area consumption. Proposed router (case 2) reduced 21.01% of total cell area of

conventional WH router and 32.11% of total power consumption. The detailed comparison of conventional WH router, Proposed Case 1 and 2 routers shows in Table 3.

A. Impact of Different Flit Sizes

System designers have another degree of freedom to analyze the performance of NoC is represented by different flit Size with fixed payload size for each and every transmission. This parameter is related to NoC bandwidth and it determines network buffering resources with respect to overall NoC power and area. Flit size can be customized and mainly affects the packetization logic at Network Interface block. Figure 11 and Figure 12 compare the impact of flit size on area and power consumption of Proposed (case 1) router. When moving number of flits from 2 to 4, area penalty of combinational and non combinational blocks can be achieved but it reduces the dynamic power consumption. But leakage power slightly increased with decreases in flit size. Therefore we conclude this section, flit size and area of router are the trade-off.

VII. CONCLUSIONS

In this paper, we introduced pipeline scheme in WH router for NoC to meet better efficiency. By applying pipeline technique, we achieved less slack time of Proposed (case 1) WH than conventional WH router. Then, hierarchical method proposed (case 2) in scheduler unit to achieve less power and area consumption. These two routers were described by Verilog HDL and synthesized using Xilinx respectively. Synopsys tool was used to compute power, area and slack time of two routers in UMC 130 nm technology with nominal voltage of 1V and packet size of 39 bits. In addition, we analyzed that the flit size have a strong effect on power and area of router. Finally we conclude, proposed router (case 2) provide better working performance than CS, CSWP, and conventional WH router.

ACKNOWLEDGEMENT

We would like to thank all the anonymous reviewers for their valuable suggestions. We also extend our sincere thanks to VLSI Design centre, Department of Electronics and Communication Engineering, PSG College of Technology, Coimbatore.

REFERENCES

- [1] International Technology Roadmap for Semiconductors [Online]. Available: <http://public.itrs.net>
- [2] D. Sylvester and K. Keutzer, A Global Wiring Paradigm for Deep Submicron Design, IEEE Trans. CAD/ICAS, pp. 242- 252, 2000.
- [3] R. Ho, K. Mai, and M. Horowitz, The Future of Wires, Proc. the IEEE, April 2001, pp. 490-504.
- [4] L. Benini et al., Networks on chips: A new SoC paradigm, IEEE Computer, vol. 36, no. 1, Jan. 2002, pp. 70-78.
- [5] Ming Liu, Improving the Performance of a Wormhole Router and Wormhole Flow Control, MS Thesis, Sweden, 2005.
- [6] A. Kumar, D. Manjunath, and J. Kuri, Communication Networking: An Analytical Approach, Morgan Kaufmann, 2004.

- [7] P. T. Wolkotte, G. J. M. Smit, G. K. Rauwerda, and L. T. Smit, An energy-efficient reconfigurable circuit-switched network-on-chip, Proc. 19th IEEE International Conference on Parallel and Distributed Processing Symposium, 2005, pp. 155-163.
- [8] Felperin, S.A et. al, Routing techniques for massively parallel communication, Proc. Of IEEE, Vol. 79, 1991, pp. 488 – 503.
- [9] D.C. Grunwald, D.A. Reed, Networks for parallel processors: Measurements and prognostications, Proceedings of the 3rd Conf. on Hypercube Concurrent Computer Applications, New York, NY, 1998, pp. 610–619.
- [10] L.M. Ni, P.K. McKinley, A survey of wormhole routing techniques in direct networks, IEEE Computer., 26 (2), 1993, pp. 62–76.
- [11] K. M. Al-Tawil, M. Abd-El-Barr, and F. Ashraf, A survey and comparison of wormhole routing techniques in mesh networks, IEEE Network, vol. 11, 1997, pp. 38–45.
- [12] Sorin Manolache, Petru Eles, Zebo Peng, Buffer space optimisation with communication synthesis and traffic shaping for NoCs, Proceedings of the conference on Design, automation and test in Europe: Proceedings, March 06-10, Munich, Germany, 2006.
- [13] Michihiro Koibuchi, K.Anjo, Y. Yamada, A. Jouraku, H. Amano, A Simple Data Transfer Technique Using Local Address for Networks-on-Chips, Parallel and Distributed Systems, IEEE Transactions on, vol.17, no.12, Dec. 2006, pp.1425-1437.
- [14] Schroeder. M. D, Birrell. A. D, Burrows. M, Murray. H, Needham. R.M, Rodeheffer. T.L, Satterthwaite. E.H, Thacker, C.P, Autonet: a high-speed, self-configuring local area network using point-to-point links, Selected Areas in Communications, IEEE Journal on, vol.9, no.8, Oct 1991, pp.1318-1335.
- [15] Federico Angiolini , Paolo Meloni , Salvatore Carta , Luca Benini , Luigi Raffo, Contrasting a NoC and a traditional interconnect fabric with layout awareness, Proceedings of the conference on Design, automation and test in Europe: Proceedings, Munich, Germany, March 06-10, 2006.
- [16] Federico Angiolini, Paolo Meloni, Luca Benini, Salvatore Carta, Luigi Raffo, Networks on Chips: A Synthesis Perspective, ParCo 2005, 2005, pp. 745-752.
- [17] U. Saravanakumar, K.Rajasekar, R. Rangarajan, Implementation of scheduling algorithms for on chip communications, International Journal of Computer Applications, 2013.