Research Journal of Applied Sciences, Engineering and Technology 4(13): 1997-2002, 2012
ISSN: 2040-7467

Submitted: March 14, 2012 Accepted: April 07, 2012 Published: July 01, 2012

A Scalable and Minimized Butterfly Fat Tree (SMBFT) Switching Network for On-Chip Communication

[1]Sheraz Anjum, [2]Imran Ali Khan, [2]Waqas Anwar, [1]Ehsan Ullah Munir and [2]Babar Nazir
[1]COMSATS Institute of Information Technology, Wah Cantt, Pakistan
[2]COMSATS Institute of Information Technology, Abbottabad, Pakistan

Abstract: This study proposes a scalable and cost effective Network on Chip (NoC) based architecture that is a modified version of Butterfly Fat Tree (BFT) network and is known as Scalable and Minimized Butterfly Fat Tree (SMBFT) switching network. The corresponding floor plan and scalable routing algorithm for the proposed network is also presented. Component Based Interconnection Network Simulator (CINSIM) was used to evaluate the steady state as well as transient behaviors of SMBFT, BFT and Binary Tree switching networks for average delay at targets. Results show that the proposed on-chip network outperforms the other two in terms of average delay, area and cost. SMFBT also comprises of less number of routers, links and levels. Hence the proposed network of switches is superior to BFT and Binary Tree and can efficiently be used for on-chip communication networks.

Key words: Component based Interconnection Network Simulator (CINSIM), Networks on Chip (NoC), Scalable and Minimized Butterfly Fat Tree (SMBFT), switching techniques

INTRODUCTION

According to ITRS (2004) and Grecu *et al*. (2004) the realization of complex Systems on a Chip (SoCs) consisting of billions of transistors fabricated in low end technologies will soon be feasible. Such SoCs imply the seamless integration of heterogeneous resources performing different functions, operating at different clock frequencies and hence give rise to new challenges (Pierre and Alain, 2000).

Existing bus based SoCs have various constraints. Dedicated wires cannot accommodate more and more IP cores (William and Brian, 2001) as well as global synchronization (Hemani *et al*., 1999). To overcome the problems of bus based SoC designs, Network on Chip (NoC), paradigm (Hemani *et al*., 1999) was introduced in late 90s. To solve the global wire delay problem, all the related studies of Wingard (2001), William and Seitz (1986), Ilkka *et al*. (2002), Martti (2002), Faraydon *et al*. (2001), Partha *et al*. (2003a, b) and Sheraz *et al*. (2009, 2011) have proposed some kind of architecture for interconnection of heterogeneous resources.

This study proposes a scalable and minimized butterfly fat tree switching network. The floor plan and the routing algorithm for the proposed network of switches is also presented. CINSIM (Walter *et al*., 2004) is used to simulate and compare the steady state as well as transient behavior of SMBFT, BFT and binary tree networks for average delay at targets. The rest of this study consist of four sections named rchitecture and Floor Plan of SMBFT outing for SMBFT imulation Environment and Conclusions respectively.

SMBFT ARCHITECTURE

The detailed architecture of SMBFT switching network is shown in Fig. 1. The proposed network is a minimized version of BFT which in turn is a derivative of fat tree network (Pierre and Alain, 2000). In this network, switches are placed at vertices and IPs at the leaves.

Coordinates pair (l, p) identifies each node. where p represents position and l represents a node level. Lowest level contains N IPs ranging from 0 to (N-1). IPs are connected to switches through their Resource Network Interface (RNI). Each switch has four child, three neighbouring and one parent port and is denoted by $S(l, p)$. Levels depend on number of IPs and is given in equation-1:

$$\text{Number of Leves} = \log_4 N - 1 \quad (1)$$

The l-th level of the SMBFT has $N/4^l$ switches. If the switches are less than 4, the level is omitted. Therefore, for IP at $(0,i)$ the parent would be at $(1, p)$ where,

$$P = \lfloor i / 4 \rfloor \quad (2)$$

Similarly, for each switch at (l, i) the parent would be at $(l+1,P)$. As each switch has three neighbours (left, right and next or cross), therefore the left neighbour would be at coordinates (l, L_i), right neighbour at (l, R_i) and the next

Corresponding Author: Sheraz Anjum, COMSATS Institute of Information Technology, Wah Cantt, Pakistan

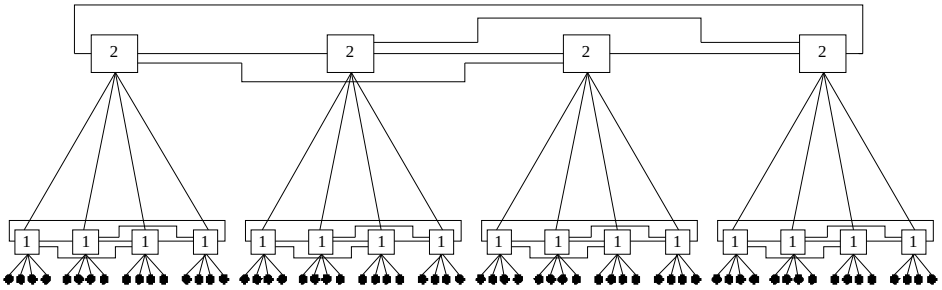

Fig. 1: SMBFT Switching Network

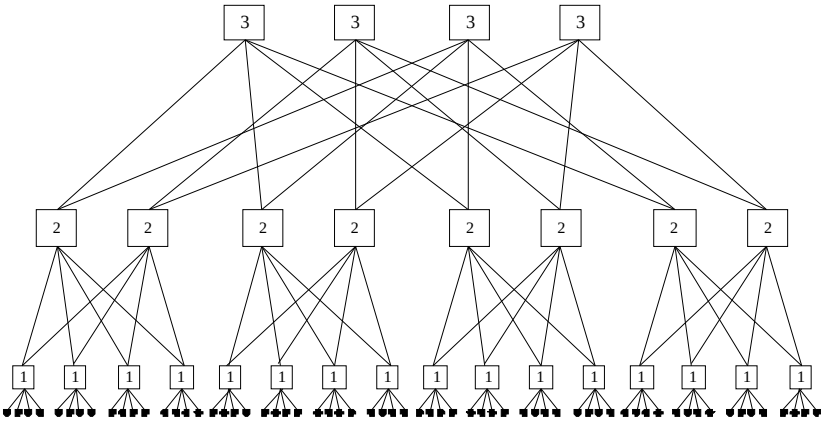

Fig. 2: BFT Switching Network

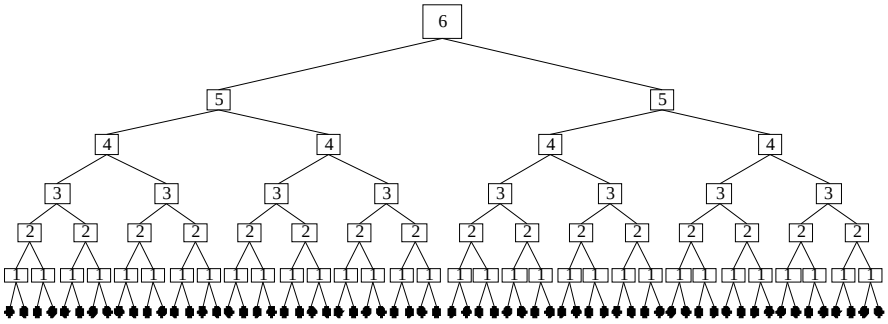

Fig. 3: Binary Tree Switching Network

neighbour would be at (l, N_i) where L_i, R_i and N_i are given by equations 3 to 5:

$$L_i = \lfloor i/4 \rfloor * 4 + (i+3) \bmod 4 \quad (3)$$

$$R_i = \lfloor i/4 \rfloor * 4 + (i+1) \bmod 4 \quad (4)$$

$$N_i = \lfloor i/4 \rfloor * 4 + (i+2) \bmod 4 \quad (5)$$

The index *'i'* for any IP or switch at each level *'l'* starts at 0 and increments from left to right in sequence. Fig. 2 and 3 show the BFT and binary tree interconnection networks respectively. All the three networks of Fig. 1 to 3 are scalable and shown only for N = 64 resource.

Table 1: Comparison of the 3 Networks containing 64 Nodes

	SMBF	TBFT	Binary Tree
No of switches	20	28	63
No of links	46	48	62
No of levels	2	3	6

Table 1 shows the details of comparison between SMBFT, BFT and Binary Tree switching networks. It is clear from the table that our proposed architecture (SMBFT) uses less number of switches, links and levels for the same number of resource nodes and effectively reduces the cost of the switching network.

SMBFT floor plan: Figure 4 shows the regular and scalable floor plan for SMBFT. The floor plan is shown

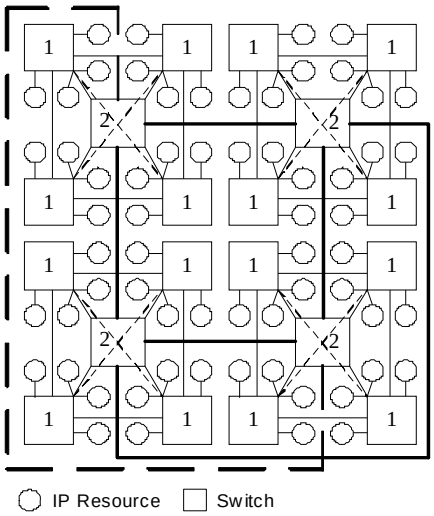

Fig. 4: SMBFT floor plan

for 64 IP nodes where it can be implemented in 2D with 3 metal layers on a chip. Figure 4 shows four types of wires/links. Thin solid wires, thin dotted wires, thick solid wires and thick dotted wires. These four types of wires can be implemented in three different metal layers as follows. Thin and thick dotted lines should be implemented on one layer and the solid lines must be implemented on two separate layers.

Routing for SMBFT: The routing algorithm determines the destination for each header flit in an intermidiate switch. The data flits simply follows. Let:

$$\text{Own Rang Start} = \text{ORS} = i * 4^j \quad (6)$$

$$\text{Left Neighbour Rang Start} = \text{LNRS} = i * 4^j \quad (7)$$

$$\text{Right Neighbour Rang Start} = \text{RNRS} = i * 4^j \quad (8)$$

$$\text{Next Neighbour Rang Start} = \text{NNRS} = I * 4^j \quad (9)$$

Also each switch has four children links ranging from 0 to 3, then the complete routing algorithm used in SMBFT is given in Table 2.

Table 2: Routing Algorithm for SMBFT

```
If (dest ³ ORS && dest < ORS + 4^l )
    If(l = = 1) ## For Level 1 Routers only
        If(dest mod 4 = = 0)
            Deliver to child link 0
        Else if(dest mod 4 = = 1)
            Deliver to child link 1
        Else if(dest mod 4 = = 2)
            Deliver to child link 2
        Else if(dest mod 4 = = 3)
            Deliver to child link 3
    Else ## For all except Level 1 Routers
        If( ⌊(dest mod 4^l) / 4^(l-1)⌋ = 0)
            Deliver to child link 0
        Else if(⌊(dest mod 4^l) / 4^(l-1)⌋ == = 1)
            Deliver to child link 1
        Else if(⌊(dest mod 4^l) / 4^(l-1)⌋ =-= = 2)
            Deliver to child link 2
        Else if(⌊(dest mod 4^l) / 4^(l-1)⌋ == = 3)
            Deliver to child link 3
Else if(dest ³ LNRS && dest < LNRS + 4^l)
    Route to Left Neighbour
Else if(dest ³ RNRS && dest < RNRS + 4^l )
    Route to Right Neighbour
Else if(dest ³ NNRS && dest < NNRS + 4^l )
    Route to Next Neighbour
Else
    Route to the Parent
```

SIMULATION ENVIRONMENT

As in this study it is important to evaluate the steady state and transient behaviors of proposed switching network against BFT and binary tree. Also router in SMBFT needs to implement wormhole switching technique for on-chip packet communication. Therefore, depending on the requirements as mentioned above and the facilities available, the component based interconnection network simulator (Walter *et al*., 2004) was selected for the purpose of simulation and evaluation of all the three switching networks. The networks as well as the important simulation components are readily available in CINSIM library for building and evaluation of any kind of network on chip simulations.

- The Network components that are used in this study to describe all of the three networks are listed below.
- Source Buffers Used to specify traffic sources
- Non-Shared Buffers Used as intermediate data (Flit) holders
- Routers Used to route packets
- Target Buffers Used to analyze received packets

In addition to the network components, wormhole switching strategy was used to analyze the average delay for steady state (10 simulations) and transient behaviors (for 50 clock cycles) of all the three networks. The buffer and packet sizes are incremented for each simulation runs.

Figure 5 shows the implementation of SMBFT with 64 nodes using cinsim-gui. It mainly contains two types of Meta elements. The small one with 4 I/Os represents a router at level 1 that is why they are 16 and the larger one represents a router at level 2, hence they are 4 in number. The connections between these Meta nodes are dictated by the proposed switching network of Fig. 1. Figure 6 shows the detailed implementation of any level 1 router along with four IP nodes in a SMBFT network. It has four I/Os used to connect to its three neighbors and one parent router. Four source (S0 to S3) and four target buffers (T0 to T3) used to simulate four IP nodes sending and receiving packets in discrete components known as flits. Router R0 is used to connect all the four I/Os and the

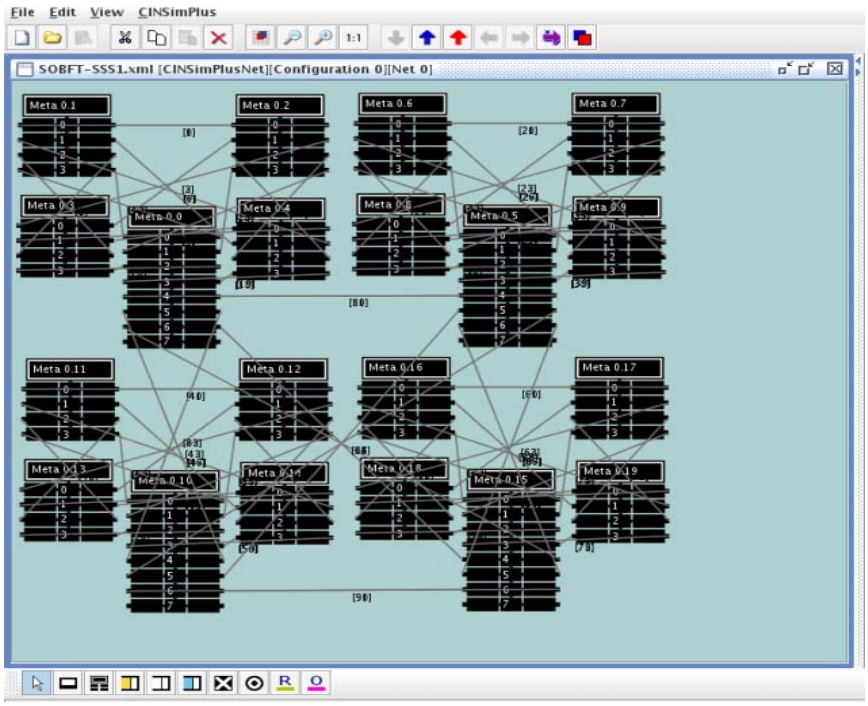

Fig. 5: SMBFT implementation using CINSIM

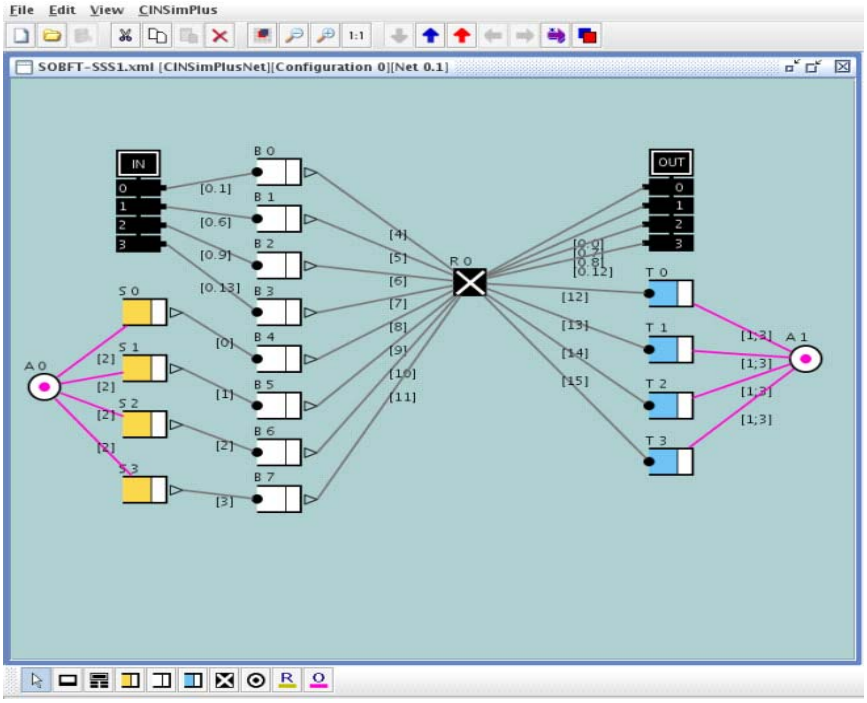

Fig. 6: Level one router in SMBFT

sources/target buffers through simple buffers (B0 to B7). Analyzer A1 is used to record the average delay at targets. A traffic source is attached to each source buffer for the generation of packets according to geometric distribution.

SIMULATION RESULTS

The Scalable and Minimized Butterfly Fat Tree (SMBFT), BFT and Binary Tree switching networks were

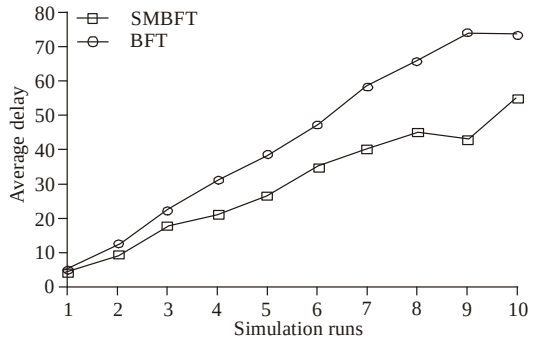

Fig. 7a: Delay vs simulation runs of SMBFT against BFT

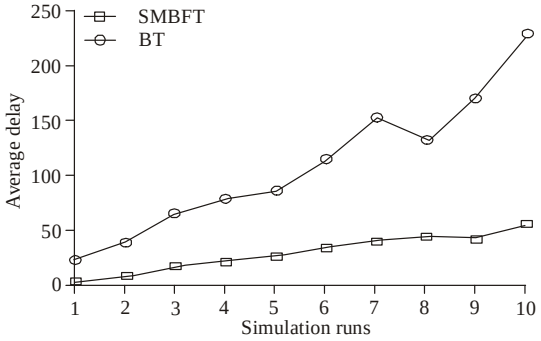

Fig. 7b: Delay vs simulation runs of SMBFT against BT

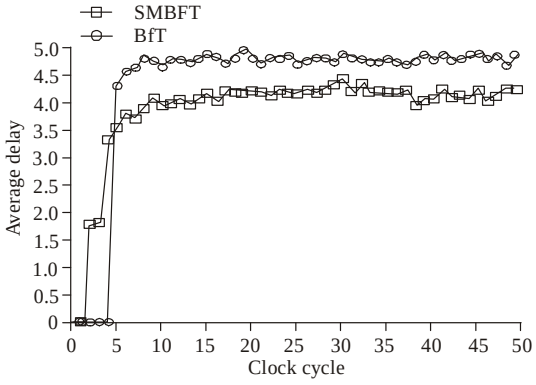

Fig. 8a: Delay vs clock cycle of SMBFT against BFT

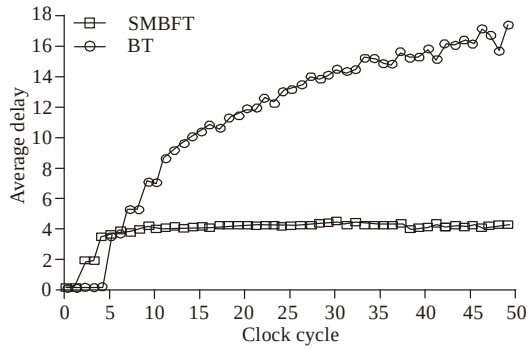

Fig. 8b: Delay vs clock cycle of SMBFT against BT

compared for average delay at targets using both the steady state and the terminating simulations. Steady state simulation is used to determine the performance characteristics for the steady state of an interconnection network and the terminating simulation is used to analyze the transient behavior of any interconnection network with-in a given range of clock cycles, say 1-50.

The steady state simulation is computed ten times and each time approx. 10% traffic load is added to the previous one. Figure 7a and b shows the average delay verses simulation runs of SMBFT, BFT and BT. It is evident from the Figure that SMBFT comprises of less average delay than the other two switching networks. In addition SMBFT performs even better in heavy traffic loads.

Figure 8a and b shows the average delay at target Vs clock cycles (transient behaviors) of SMBFT, BFT and BT for the first 50 clock cycles. Again for terminating simulation the estimated average delay of SMBFT is less than BFT and even lesser than BT whose delay is larger due to increased number of switch levels.

CONCLUSION

This study proposed a scalable and minimized BFT interconnection network suitable for on-chip packet switched communication and compares its performance against BFT and BT using Component based Interconnection Network Simulator (CINSIM). The performance was evaluated in terms of average delay at targets for steady state and transient behaviors of the networks. Results show that SMBFT has less estimated average delay than BFT and even lesser than BT for both types of simulations. The floor plan and the associated routing algorithm for the proposed architecture are also presented. In addition SMBFT comprise of less number of routers, links and levels. Therefore, it is concluded that the proposed switching network is superior to BFT and BT and can be efficiently used for on-chip communication networks. Future study is intended to implement the proposed switching network and the associated router on an FPGA for computation of total silicon area and maximum attainable speed.

ACKNOWLEDGMENT

This study is sponsored by COMSATS Institute of Information Technology, Pakistan.

REFERENCES

Faraydon, K., N. Anh, D. Sujit and R. Ramesh, 2001. On-chip communication architecture for OC-768 network processors. Proceedings of DAC, pp: 678-683.

Grecu, C., P.P. Partha, I. Andr and R. Saleh, 2004. A scalable communication-centric soc interconnect architecture. Proceedings of 5th International QED Symposium, pp: 343-348.

Hemani, A., T. Meincke, S. Kumar, A. Postula, T. Olsson, P. Nilsson, J. Oberg, P. Ellervee and D. Lundqvist, 1999. Lowering power consumption in clock by using globally asynchronous locally synchronous design style. Proceedings DAC, pp: 873-878.

Ilkka, S., S.T. David and N. Jari, 2002. Interconnect IP node for future system on chip designs. In Proceedings of the First IEEE International Workshop on EDTA, pp: 116-120.

Martti, F., 2002. A scalable high-performance computing solution for networks on chips. IEEE Micro, 22(5): 46-55.

Partha, P.P., C. Grecu, Ivanor and R. Saleh, 2003a. Design of a switch for network on chip applications. Proceedings of ISCAS, pp: 217-220.

Partha, P.P., C. Grecu, A. Ivanor and R. Saleh, 2003b. High-throughput switch-based interconnect for future SoCs. Proceedings of 3rd IEEE International Workshop on SoC for Real-Time Applications, pp: 304-310.

Pierre, G. and G. Alain, 2000. A generic architecture for on chip packet-switched interconnections. Proceedings of DATE, pp: 250-256.

International Technology Roadmap for Semiconductors (ITRS), 2004. Retrieved from: Public.itrs.net/.

Sheraz, A., J. Chen, P.P. Yue and L. Jian, 2009. A delay optimized architecture for on-chip communication. J. Electr. Sci. Techn. China, 7(2): 104-109.

Sheraz, A., U.M. Ehsan and W.N. Muhammad, 2011. Simulation and performance evaluation of network on chip architectures and algorithms using CINSIM. J. Basic Appl. Scient. Res., 1(10): 1594-1602.

Walter, A., M. Kuhm, D. Tutsch, D. Ludtke and C. Zimmermann, 2004. CINSim Handbook: Installation and User's Guide [M], Technische Universit Berlin, Real-Time Systems and Robotics Copyright 2004-2007.

William, J.D. and C. Seitz, 1986. The Torus Routing Chip, Technical Report 5208:TR: 86. Department of Computer Science, pp: 1-19.

William, J.D. and T. Brian, 2001. Route packets not wires: On chip interconnection networks. In Proceedings of DAC, pp: 684-689.

Wingard, D., 2001. Micro-network based integration for SoCs. In Proceedings of DAC, pp: 673-677.