High Throughput Architecture for High Performance NoC

Mohamed A. Abd El ghany, Magdy A. El-Moursy* and Mohammed Ismail**
Electronics Engineering Dept., German University in Cairo, Cairo, Egypt
Electronics Research Institute, Cairo, Egypt.
*Mentor Graphics Corporation, Cairo, Egypt.
**Electrical Engineering Dept., The Ohio State University, Columbus, USA. The RaMSiS Group, KTH, Sweden.
E-mails: mohamed.abdel-ghany@guc.edu.eg, magdyaelmoursy@gmail.com, ismail@ece.osu.edu

Abstract— High Throughput Butterfly Fat Tree (HTBFT) architecture to achieve high performance Networks on Chip (NoC) is proposed. The architecture increases the throughput of the network by 38% while preserving the average latency. The area of HTBFT switch is decreased by 18% as compared to Butterfly Fat Tree switch. The total metal resources required to implement HTBFT design is increased by 5% as compared to the total metal resources required to implement BFT design. The extra power consumption required to achieve the proposed architecture is 3% of the total power consumption of the BFT architecture.

Index Terms – BFT, NoC, Throughput, Latency.

I. INTRODUCTION

As the number and functionality of intellectual property blocks (IPs) in System on Chips (SoCs) increase, complexity of interconnection architectures of the SoCs have also been increased. Different researches have been published in high performance SoCs, however, the system scalability and bandwidth are limited. As described in [1]-[5], NoCs are emerging as the best replacement for the existing interconnection architectures. NoC provide different set of constraints in the design paradigm.

High throughput and low latency are the desirable characteristics of a multi processing system. Many NoC architectures have been proposed in the past, e.g., SPIN [3], CLICHE [2], Folded Torus [1], Octagon [4] and Butterfly fat-tree [5]. Among those, the butterfly fat tree (BFT) has found extensive use in different parallel machines and shown to be hardware efficient. The main advantage of the butterfly fat tree is that the number of switches in the network converges to a constant irrespective of the number of levels in the tree network.

The main focus of this paper is to present a high throughput interconnect architecture of butterfly fat tree. High Throughput Butterfly Fat-Tree (HTBFT) architecture with 16 IPs is shown in Figure 1. The effect of the architecture on the overall efficiency and performance is discussed. The throughput is improved while the latency is preserved. The used metal resources required to achieve this throughput improvement is negligible with respect to the available metal resources. The relation between the number of virtual channels and the frequency of switch is characterized.

This paper is organized as follows: In Section II, The new HTBFT architecture is explained. In section III, the architecture of proposed port circuit is presented. In Section IV, the performance analysis is demonstrated. In Section V, the overhead of the proposed architecture is discussed. Finally, conclusions are provided in section VI.

II. HTBFT ARCHITECTURE

The interconnect template of butterfly fat-tree architecture was proposed in [5]. This structure assumes a 4-ary tree with switches connected 4 down links and 2 up links. Each group of 4 leaf nodes needs one switch. At the next level, half as many switches are needed (every 4 switches on the lower level need 2 switches at the next level). This relation continues with each succeeding level. According to the results presented in [6] [7], the throughput varies with the number of virtual channels. When the number of virtual channels is increased beyond four, the throughput is saturated. On the other hand, the average message latency increases with the number of virtual
channels. To keep the latency low while preserving the throughput, the number of virtual channels is constrained to four. The analysis in [6], [7] lack circuit optimization. Only four virtual channels are allowed to maximize throughput.

A novel interconnect template to integrate IP blocks using HTBFT architecture is proposed as shown in Figure 1. In the proposed architecture, rather than using a single interconnect bus between each two switches, two buses are employed. Each group of 4 IPs (no. 0, no. 1, no.2 and no.3) needs one switch (no.4). Each switch in the first level (no. 4) connects to each switch in the second level (no. 5) by 2 buses as shown in Figure 1. The number of virtual channels can be doubled to get higher throughput. Each bus will support half number of virtual channels. Average latency could be maintained by using half number of virtual channels per interconnect bus. The effect of increasing the number of virtual channels on the frequency of the switch is discussed in Section IV.

Increasing the number of buses between two switches could improve the throughput by optimizing the design on the circuit level as shown in the following section. However, using two buses to connect two switches implies a consumption of the metal resources and may be silicon area for the repeaters within long interconnect bus. The overhead of the proposed architecture is discussed in Section V. The port architecture is described in Section III.

III. PORT ARCHITECTURE

The switch of BFT has six ports, four children ports and two parent ports. Each port of the switch includes input virtual channels, output virtual channels, a header decoder, controller, input arbiter and output arbiter as shown in [5]. When the number of virtual channel is increased, the throughput increases. The input arbiter is used to allow only one virtual channel to access a physical port. The input arbiter consists of a priority matrix and grant circuits [6].

The priority matrix stores the priorities of the requests. The grant circuits generate the granted signals to allow only one virtual channel to access a physical port. The messages are divided into fixed length flow control units (flits). When the granted virtual channel stores one whole flit, it sends a full signal to controller. If it is a header flit, the header decoder determines the destination. The controller checks the status of destination port. If it is available, the path between input and output is established. All subsequent flits of the corresponding packet are sent from input to output using the established path. The flits from more than one input port may simultaneously try to access a particular output port. The output arbiter is used to allow only one input port to access an output port.

Virtual channels consist of several buffers controlled by a multiplexer and an arbiter which grants access for only one virtual channel at a time according to the request priority. Once the request succeeds, its priority is set to be the lowest among all other requests. In the proposed architecture, rather than using one multiplexer and one arbiter to control the virtual channels, two multiplexer and two arbiters are employed as shown in Figure 2. The virtual channels are divided into two groups, each group controlled by one multiplexer and one arbiter.

Let us consider an example with the number of virtual channels of 8 channels. In BFT architecture, 8x8 input arbiter and 8x1 multiplexer are needed to control the input virtual channels as shown in Figure 2 (a). The 8x8 input arbiter consists of 8x8 grant circuit and 8x8 priority matrix. In the proposed architecture, two 4x4 input arbiters, two 4x1 multiplexers, 2x1 multiplexers and 2x2 grant circuit are integrated to allow only one virtual channel to access a physical port as shown in Figure 2 (b). The 4x4 input arbiter consists of 4x4 grant circuit and 4x4 priority matrix. The area of 8x8 input arbiter is larger than the area of two 4x4 input arbiters. Also, the area of 8x1 multiplexer is larger than the area of two 4x1 multiplexers. Consequently, the required area to implement the BFT switch is higher than the required area to implement the proposed switch with the proposed architecture. The performance analysis is presented in section V.

![Figure 2](image.png)

IV. PERFORMANCE ANALYSIS

Neglecting circuit optimization techniques in BFT, the change in the maximum frequency of the switch with the number of virtual channels is shown in Figure 3. When the number of virtual channels is increased beyond four, the maximum frequency of the switch is decreased. The throughput is saturated when the number of virtual channels is increased beyond four [6] for different number of ports.

The proposed architecture is implemented using the synthesis tools from Xilinx ISE, used for 90nm technology node. Physical design characteristics of the switch are obtained using ModelSim SE simulation. The proposed switch is less complex than the BFT switch. Therefore, the maximum frequency of the switch is improved. The change in the maximum frequency of the proposed switch with the number of virtual channels is
shown in Figure 4 for 6 ports. When the number of virtual channels is higher than eight, the maximum frequency of proposed switch is degraded.

Given the throughput definition [6]:

\[
TP = \frac{\text{Number of messages completed} \times \text{Message length}}{(\text{Number of IP blocks}) \times \text{total time}}
\]  

(1)

The throughput is proportional to the number of completed messages. The number of completed messages increases with the number of virtual channels. Total transfer time of messages decreases with the frequency of switch. Therefore the throughput can be improved by increasing the number of virtual channels or by increasing the frequency of switch. In the proposed architecture, the number of virtual channels is doubled while preserving the average latency. Therefore, the throughput of using eight virtual channels in the HTBFT is double the throughput of four virtual channels in BFT. The average latency of HTBFT with 8 virtual channels equals to the average latency of BFT with 4 virtual channels. Considering the Maximum frequency and the number of completed messages for HTBFT, the throughput of HTBFT is determined. The variation of throughput with the number of virtual channels for HTBFT and BFT is shown in Figure 5. In our architecture, when the number of virtual channels is increased beyond eight, the throughput saturates. The architecture increases the throughput of the network by 38%. In the following section, the overhead of the proposed architecture is considered.

The area required to implement the BFT switch and HTBFT switch is shown with different number of virtual channels in Figure 6. In HTBFT, the virtual channels are divided into two groups. Each group is controlled by one control unit. The area of two HTBFT control units is less than the area of BFT control unit. The HTBFT architecture decreases the area of switch by 18%. Consequently, a system with eight virtual channels achieves high throughput, high frequency and low latency while the area of design is optimized.

V. SYSTEM OVERHEAD

It is possible to organize the butterfly fat tree so that it can be laid out in \(O(N)\) active area (IPs and switches) and \(O(\log(N))\) wiring layers. The basic strategy for wiring is to distribute tree layers in pair of wire layers – one for horizontal wiring \(H_{a+1,a}\) and one for vertical wiring \(V_{a+1,a}\). The length of horizontal part \(H_{a+1,a}\) equals to the length of vertical part \(V_{a+1,a}\) given that the chip is squared. More than one tree layer can share the same wiring trace. High throughput architecture has the same number of active area, but the number of wires will be doubled. The length of inter-switch wire depends on the number of levels in BFT, which depends on the system size. The interswitch wire length is given by the following expression [5]:

\[
w_{a+1,a} = \frac{\sqrt{\text{Area}}}{2^{\text{levels}-a}}
\]  

(2)

Where \(w_{a+1,a}\) is the length of the wire spanning the distance between level \(a\) and \(a+1\) switches, where \(a\) can take integer values between 0 and \((\text{levels}-1)\). In the circuit...
implementation of HTBFT, a bus between each two switches has 12 wires, 8 for data and 4 for control signals. The number of metal levels increases reaching twelve according to the latest ITRS. Metal resource on chip is increased. Considering a chip size of 20 mm x 20 mm, technology node of 90 nm, and a system of 1000 IP blocks, the number of interswitch links are obtained. The number of switches equals to 500 switches. Given the optimal global interconnect width $W_{opt}$ of 935 nm, optimal global interconnect spacing $S_{opt}$ of 477 nm [8], the length of $H_{a+1,a}$ and $V_{a+1,a}$ are calculated. Using the critical interconnect length of 2.54 mm, optimal repeater size of 174 [8], the number of repeaters equals to 384 repeaters. The area of repeaters required to implement the HTBFT interswitch links equals to 8352 $\mu$m² (it equals to the double area of repeaters required for BFT interswitch links). The power consumption of repeaters and switches required to implement the BFT and HTBFT is presented in Table I.

Table I. Power consumption of repeaters and switches for BFT and HTBFT

| Arch. | No. of repeaters | Power dissipation of repeaters and interswitch links (mW) | Power dissipation of switches (mW) | Total power dissipation | Percentage of power dissipation of repeaters and interswitch links (%) |
|------|-----------------|----------------------------------------------------------|-----------------------------------|------------------------|---------------------------------------------------------------|
| BFT  | 192             | 943.99                                                   | 41210                             | 42153.99               | 2.2                                                           |
| HTBFT| 384             | 1887.98                                                  | 41255                             | 43142.98               | 4.4                                                           |

The number of BFT levels is seven. The global interconnect pitch is $W_{opt} + S_{opt}$. Assuming all of global interconnects have the same line width and line spacing, then the number of global interconnects $N$ per layer is given by

$$N = \frac{\sqrt{\text{Area}}}{W_{opt} + S_{opt}} \quad (3)$$

The horizontal wiring is distributed in the metal layer no. 11 and the vertical wiring is distributed in the metal layer no. 12. The total length of horizontal wires needed equals to 3796 mm (it is 5 % of the total metal resources available in the metal 11). The same for total length of vertical wires, it requires 5 % of the total metal resources available in the metal 12. For the proposed design, Rather than using a single bus between each two switches, two buses are employed. Therefore, the total metal resources required to implement the proposed architecture will be 10%. The metal resources of HTBFT architecture equals to the double metal resources of BFT architecture. The extra metal resources required to achieve the proposed architecture is negligible as compared to the metal resources.

The power consumption of interswitch links and repeaters for BFT and HTBFT is shown in Table II. On the average, the extra power consumption required to achieve the proposed architecture is 3% of the total power consumption of the BFT architecture. Therefore, the extra power consumption required to the proposed architecture is negligible as compared to the total power consumption of the network.

Table II. Power consumption of interswitch links and repeaters for HTBFT and BFT

| Technology node | No. of IPs | No. of levels | Power consumption of interswitch links and repeaters for BFT (mW) | Power consumption of interswitch links and repeaters for HTBFT (mW) |
|-----------------|-----------|--------------|---------------------------------------------------------------|---------------------------------------------------------------|
| 130 nm          | 500       | 6            | 1024.22                                                   | 2048.44                                                   |
| 90 nm           | 1000      | 7            | 943.99                                                   | 1887.98                                                   |
| 65 nm           | 2500      | 9            | 712.33                                                   | 1424.66                                                   |
| 45 nm           | 7500      | 10           | 948.36                                                   | 1896.72                                                   |

VI. CONCLUSIONS

In this paper, the HTBFT architecture is proposed to increase the throughput of the switch in NOC. The proposed architecture can also improve the latency of the network. The throughput is increased by 38% while preserving the latency. The area of HTBFT switch is decreased by 18% as compared to the area of BFT switch. The total metal resources required to implement the proposed design equals to 10%; It is increased by 5% as compared to the total metal resources required to implement BFT design. The extra power consumption required to achieve the proposed architecture is 3% of the total power consumption of the BFT architecture. The relation between throughput, number of virtual channels and switch frequency is analyzed. The simulation results demonstrate the performance enhancements in terms of throughput, number of virtual channels and switch frequency. It is shown that optimizing the circuit can increase the number of virtual channels without degrading the frequency. The throughput of BFT is also improved with the proposed architecture.

REFERENCES

[1] W. J. Dally and B. Towles. “Route packets, not wires: on-chip interconnection networks”, In Proceedings of Design Automation Conference, pp 684–689, June 2001.
[2] S. Kumar et al., “A Network on Chip Architecture and Design Methodology,” In Proceedings of the IEEE Computer Society Annual Symposium on VLSI, pp. 117-124, 2002.
[3] P. Guerrier and A. Greiner. “A generic architecture for on-chip packet-switched interconnections”, In Proceedings of Design Automation and Test in Europe Conference and Exhibition, pp. 250–256, Mar. 2000.
[4] F. Karim, A. Nguyen, and Sujit Dey, “An Interconnect Architecture For Networking Systems on Chips,” IEEE Micro, vol. 22, no. 5, pp. 36–45, Sept./Oct. 2002.
[5] P.P. Pande, C. Grecu, A. Ivanov, and R. Saleh, “Design of a Switch for Network on Chip Applications,” In Proceedings of The 2003 International Symposium on Circuits and Systems (ISCAS), vol. 5, pp. 217-220, May 2003.
[6] P. P. Pande, C. Grecu, M. Jones, A. Ivanov, and R. Saleh, “Performance Evaluation and Design Trade-Offs for Network-on-Chip Interconnect Architectures”, IEEE Transaction on Computers, vol. 54, no. 8, Aug. 2005.
[7] P. Pande, C. Grecu, A. Ivanov, R. Saleh, "High-Throughput Switch-Based Interconnect for Future SoCs", the 3rd IEEE International workshop on SoC for real-time Applications, pp 304-310, July 2003.
[8] X.-C. Li, J.-F. Mao, H.-F. Huang, and Y. Liu, “Global interconnect width and spacing optimization for latency, bandwidth and power dissipation,” IEEE Transactions on Electron Devices, vol. 52, no. 10, pp. 2272–2279, Oct. 2005.