A Centralized Controller as an Approach in Designing NoC

Abdolhossein Fathi and Keihaneh Kia
Department of computer Engineering and Information Technology, Razi University, Kermanshah, Iran
Email: a.fathi@razi.ac.ir, kiya.keyhaneh@stu.razi.ac.ir
*Corresponding Author Tel:+989166615160

Abstract—This paper presents a new NoC architecture to improve flexibility and area consumption using a centralized controller. The idea behind this paper is improving SDN concept in NoC. The NoC routers are replaced with small switches and a centralized controller doing the routing algorithm and making control decisions. As one of the main desirable property of NoC is flexibility, in this work with the help of centralized controller, having different topologies and also having two separate networks in a single platform is possible. The other effects of this new scheme are power and area consumption which are investigated. Performance of the NoC is also studied with an analytical model and compared with the traditional NoC. The proposed NoC is implemented in VHDL, simulated and tested with ISE Xilinx.

Index Terms—Network on chip (NoC), software define network (SDN), centralized controller, flexibility, reconfigurability.

I. INTRODUCTION

With improvement of technology more processing cores, and also a system, can integrate on a single chip which is called system on chip (SoC). The SoC elements can be computational and graphical processors, IO units, memories and etc. One of the main problems with SoC is communication between cores. Type of these communications are generally partitioned into three groups as bus based structure, point to point structure and network on chip (NoC). The first two types are not appropriate in high communication rate and especially when the number of cores is many. NoC presents a new solution to eliminate the shortcomings of traditional types of communication [1]. Communications in NoC is abstracted from network concept in such a way that there is some IPs which produce traffic and are connected to the other nodes through network interfaces. All IPs are connected to the others with routers and the routers are linked according to a network topology and a routing algorithm [2-4]. There are some problems with this method of communication like traffic congestion and lack of flexibility and so on. Researchers try to overcome the problems by employing different topologies, none blocking routing algorithms, 3D NoC and also some improvement in hardware implementation and so on[5-8].

Growing need for simplification, results in revolution of software define network (SDN). This idea is based on plane separation. In SDN the control plane which is consisted of controlling parts of the routers is disported from the data plane. In data plane there is flow tables that determines the output port of each incoming flow. All the control plans of the router are collected and centralized in the controller. Controller is responsible for the managing and routing functions in the network. It leads to decrease cost of development and has a simplified device and other benefits.

Recently the idea of software define network (SDN) [9, 10] is proposed in NoC. This is a good solution and can overcome some of the NoC’s defects. Mapping SDN concept in NoC needs considering changes in implementation, which is not considered yet. This is the innovation of this paper, a new NoC architecture with a centralized controller. The main property of the proposed NoC is high flexibility and low area consumption. The power consumption and the latency can decrease.

This paper is organized into the following parts as follow: After the introduction there is related works and background. In section III there is the proposed architecture and in the next section the advantages and disadvantages are mentioned. In section V the implementation and results are presented. At the end there is the conclusion.

II. RELATED WORKS

An NoC consists of four main parts, including: IP core, router, network interface and communication links [4]. The IP cores are traffic generator like memory, IO, CPU, … and not considered part of the NoC design. They produce traffic to the networks. The rule of the router is as the routers in conventional networks with a less degree of complexity. Each router has a set of ports which are used to connect router with its neighboring routers and with the IP cores of the system [11, 12]. The network interface is responsible for the conversion between the high level protocol (HLP) that the IP uses and the packet-based communication protocol of the NoC and other higher level network issues. The NoC’s main parts are shown in Fig. 1.

Routers are responsible for receiving and forwarding packets inside the network based on NoC parameters. The router is the main component in an NoC and its area and
Communication mechanisms in NoC are flow control and switching mode which deals with the allocation of channels and buffers to data including [17]: packet switching (PS) and circuit switching (CS).

Switching Mode only exists in PS networks and defines how packets move through the network. They are including store-and-forward (SAF), virtual cut-through (VCT) and wormhole (WH).

An NoC can be evaluated by parameters like performance, area, energy/power consumption, quality of service (QoS) and flexibility. In this paper flexibility means scalability that is ability to add more cores and design re-uses that allows more parameters to be changed. In other words ability to use the same NoC architecture for multiple applications and with extendable number of cores.

According to our investigations, using the concept of SDN in designing NoC is recently proposed in few works implicitly or explicitly [18].

In [19] the main objective is based on openflow in SDN, so the router architecture is divided into two separated layers: data plane and control plane layers. In data plane layer there are switches for transferring data and in the control plane layers, there are routing algorithm and management modules. Therefore, it is a reconfigurable NoC to meet the requirement of different applications. The main capability of it, is coordination with changes in protocols and redesign of routing algorithms. There is not a central controller and the controlling parts of routers can be configured individually.

In [20] there is a photonic network for data plane and a centralized controller for improvement in utilizing photonic resources. In controller there is some queues for request of getting and releasing a route in the photonic switches. With the difference in the speed of a photonic resources and the conventional and none-photonic controller, it seems to be not applicable regardless of controller as a bottleneck of NoC. It suffers from the lack of implementation and details of design.

An NoC based on SDN is presented in [21] for many cores systems. In this work the abstraction layers of SDN are mapped in NoC and modifying according to the requirement of many cores systems and the concept is presented without details of implementation. In [22], a performance evaluation of SDNoC is presented.

In this paper the following specifications is considered for the proposed design. Flow control is packet switching in which a packet is divided into some flits. The first flit is header including destination address. After that there is one or more data flits and finally is the tailor flit. Switching mode which is only a parameter of PS networks in this design is wormhole and the routing algorithm is selected XY routing for its simplicity.

As the main part of an NoC is router, we mainly focus on this part. The router was designed with 4 ports for communication with neighboring routers, North, East, South, West and a local port for communication with IP cores and a controller port. Also we use a controller for having different topologies in one NoC platform. The proposed NoC is flexible in such a way that can be divided into two completely separate NoCs. Also the controller is reconfigurable so different routing algorithms can be used.
A single controller is replaced the controlling parts of the routers, so the area is saved too. For some applications that need half of the nodes, the unused parts can be switched off so the power consumption will decrease in this case.

Another advantage of the scheme is the simplification based on less processing in packets. The controller process the header flit for the first time then the other flits of the packet will go through the network to the destination without any processing delay. Area and performance of the NoC is investigated.

### III. PROPOSED ARCHITECTURE

For decreasing the complexity and area of the NoC routers as one of the most effective elements, and increasing flexibility of the whole NoC, we separate the controlling parts of router from the hardware of it. Accordingly the routers changes to switches that conduct a flit from input port to one of the output ports according to commands of the controller. So we can refer to the router as switch after that in this paper. Figure 2 shows the proposed architecture for NoC with central controller.

---

The main part of this NoC architecture is the controller. It can decide on a flit to rout through the switches according to routing algorithm that in this paper is XY for its simplicity. The details of this controller will be discussed later.

The NoC is composed of some nodes and a controller. In this paper we consider 16 nodes as an example and it is extendable. Each node consisted of an IP core, network interface and a switch. The IP core connected to the local port of each switch through the network interface. Each switch is connected to its neighboring switches with the four primary ports and to the controller through the control port.

When a core produces a packet, send it to the network interface. The network interface changing the packet into some flits. The first flit is header with the destination address and the next flits are data flits and at the end is the tailor flit. The network interface will send the flits to the switch.

When a switch receives a header flit, send the destination address to the controller and wait for the controller response which is the ID of all the ports to the final destination. The switch send it to the destination port instead of the header flit and all the other related flits to that port. When the next switch receives the header, can recognize the related ID port from it and will send the header and the other flits of the packet to that port without the interference of the controller. If a switch notices itself as the last switch from the header, will send the header and other data flits to the local port to go to the IP core.

The header is just like an array of destination ports of each switch in the path. As it is obvious the controller engages each packet once. This seems to decrease the...
load of the controller to a great extent. In a conventional NoC in all the routers in the path of a flit, the routing algorithm should run. Whatever the amount of computation of the NoC decreases the power consumption will decrease too. As the routers are replaced with small switches the area will save. The controller consumption area is less than the summation of all the controlling parts of the routers. The amount of saving area determines analytically in section V.

The central controller monitors the input ports from all the switches in a priority order. In an improved controller the routing algorithm can avoid busy switches and determines a free path for the packet. In behalf of being centralized, it is possible since the controller is aware of the status of all the switches. The main problem with this controller is that, it is the bottleneck of the chip. In order to overcome this defect, we can have two controllers that share the traffic of the NoC.

The proposed NoC is capable of supporting more than one topology. For this purpose we select the torus topology. This topology can be considered a mesh with neglecting the outer edges of the chip by the controller in its decisions. It can also be one or two ring networks. In this paper we consider the first half nodes as a separate ring network from the second half nodes. It is suitable for some applications that need less than half of the IPs for communication. In this case the other part will be turned off which results in decrease in static power consumption. These structures for Noc with 16 nodes are shown in the figures 3, 4 and 5. Mode is an input of NoC, determines which topology should be considered. This NoC has the flexibility of running in different topologies as each application is more consistent with one of them. Having two separated ring is a strength. In this case the floor planning of nodes should be in such a way that related nodes placed in the same ring. The controller will choose the related routing algorithm according the selected topology. The flits structures are depicted in Fig. 6.

IV. ADVANTAGES AND DISADVANTAGES

The advantages and disadvantages of the proposed architecture from different aspects such as flexibility, area consumption, power consumption and lattency are considered in below.

Fig. 3. Torus topology.

Fig. 4. Two ring shown with different colors

Fig. 5. Mesh topology with red links.
The main advantage of the proposed design is the high flexibility. The desired topology for each application can be selected. In addition, with a good floor planning of cores we can have two separated NoCs. It can be improved by having two controllers. In this paper we just consider a unique controller.

From the area consumption point of view, the proposed NoC is better than conventional NoCs. Because all the controlling parts of the routers are replaced with a controller. With considering additional control links, as the number of switches are considerable, the amount of saving area dominants the additional links and the controller.

One of the important objects of NoC is power consumption. As mentioned before in a NoC, the routing algorithm will run on a flit in every node in the path to the destination. But in our NoC this operation will run just ones in the beginning of the path. This means less processing operations that leads to decrease dynamic power consumption. For decreasing static power, if for running applications half or less than half of the IPs are required, the unused switches are turned off by the controller based on mentioned ring topology.

The performance of the NoC depend on the work load. It is investigated in the next section.

B. Disadvantages

The controller is the bottleneck of the NoC. When the load is high the controller responses the nodes with a delay that may not be acceptable. By using a hierarchical or multiple controllers the load of network will share among them.

One of the main issues that should be considered is that if the controller fails, the NoC will fail too. For solving this issue we should design a hardened controller or designing two controllers, one of them as a primary unit and the other as a spare. If a primary fails or is very busy, the spare controller will start running.

Solving the above problems will left for future works.

V. IMPLEMENTATION AND RESULTS

The proposed NoC is implement in VHDL, simulated and tested with ISE Xilinx 14.7, shown in Figure 7. One the aims of this work, is improving the NoC architecture for being flexible in such a way that can easily work with different topologies or changes to two or more separate NoCs. We choose FPGA for implementation and testing the accuracy of it. We use a uniform traffic for testing the functionality and flexibility of the NoC [23].

In addition to the above property, the NoC will decrease the area. As we know, the routers replace with small switches, instead of all the controlling parts of the switches, there is a controller.

The area of an NoC is the summation of the area of 16 simple switches and a controller and 16 communicative and 16 control links in the example of the 16 node NoC. The network interfaces and the IPs are merely the same in all the models and ignored. We have:

\[ A_{\text{total}-\text{new}} = 16A_{\text{sw}} + 16A_{\text{com-link}} + 16A_{\text{ctrl-link}} + A_{\text{controller}} \]  

\[ A_{\text{total}-\text{old}} = 16A_{\text{sw}} + 24A_{\text{com-link}} + A_{\text{controller}} \]

We should prove that \( A_{\text{total}-\text{new}} < A_{\text{total}-\text{old}} \):

\[ 16A_{\text{sw}} + 24A_{\text{com-link}} + A_{\text{controller}} < 16A_{\text{sw}} + 16A_{\text{com-link}} \]

\[ 16A_{\text{sw}} + 8A_{\text{com-link}} + A_{\text{controller}} < 16A_{\text{sw}} \]

As \( A_{\text{controller}} \) is equivalent to \( A_{\text{sw}} + A_{\text{controller}} \) and \( A_{\text{controller}} \) is equivalent to \( A_{\text{controller}} + \epsilon \), we can rewrite (5) as follow:

\[ 16A_{\text{sw}} + 8A_{\text{com-link}} + A_{\text{controller}} + \epsilon < 16A_{\text{sw}} + 16A_{\text{controller}} \]

\[ \Rightarrow \quad 8A_{\text{com-link}} + \epsilon < 15A_{\text{controller}} \]

Since the \( \epsilon \) is small and is the difference between the area of the main controller and the controlling part of a router, it is clear that \( 8A_{\text{com-link}} + \epsilon < 15A_{\text{controller}} \). Even if the controlling part of a router is so small that is comparable to a link, because of the coefficient of \( A_{\text{controller}} \) which is twice the coefficient of \( A_{\text{com-link}} \), the Inequality (7) is correct, so we can conclude that \( A_{\text{total}-\text{new}} < A_{\text{total}-\text{old}} \).

For having a comparison on the performance of the new scheme relation to the conventional NoC, we use a mathematical analysis which is proposed in [24, 25]. In this paper the authors use a mathematical model for router based on the average number of packets in the buffers of the router. After that the model generalized to NoC and the final performance parameters are average buffer utilization of each buffer, the average latency per flow and the maximum network throughput. As the basic assumptions of [25] are similar to the proposed NoC, we use this mathematical model to analyze our model with the average packet latency.
It is notable to say that Ref. [25] is compared with the simulation results with the error less than 10%. The model proposed to avoid the time consuming simulation process and we use this model for sake of simplicity and high speed. According to Ref. [25], the relationship between the packet injection rate and the average number of packets waiting in the input port of each router is as follow:

$$\lambda_j = \frac{N_j}{t_j}$$  \hspace{1cm} (8)

$\lambda_j$ is the injection rate to the port $j$ of the router and $N_j$ is the average number of packets in the buffer and $t_j$ is the average time that packets are waiting in the buffer that includes service time for the header flit and transferring time of the other data flits of the packet and the service time of other packets coming before the packet. Here we don’t mention the details of the model and just mention the final equation:

$$N = (1 - TC)^{-1} R$$  \hspace{1cm} (9)

$N$ is a matrix for the average number of the packets in buffers of the router and $T$ is the service time and $C$ is the packet contention probabilities matrix and $R$ is the residual time and $\alpha$ is the traffic arrival rate matrix.

The same assumption are considered in this paper as XY routing algorithm, the wormhole flow control and a mesh topology.

Like the method used in [25] to prove the better performance of the NoC, we should prove that:

$$N_{\text{new-NoC}} < N_{\text{old-NoC}}$$  \hspace{1cm} (10)

$$(1 - T_{\text{new}}C_{\text{new}})^{-1}R_{\text{new}} < (1 - T_{\text{old}}C_{\text{old}})^{-1}R_{\text{old}}$$  \hspace{1cm} (11)

The injection rate and $R$ and $C$ are equal. $C$ is the probability that two ports compete for the same output port which is the same for both models. We should prove that:

$$(1 - T_{\text{new}}C)^{-1} < (1 - T_{\text{old}}C)^{-1}$$  \hspace{1cm} (12)

As $C$ is a positive coefficient, we can write the above relation as below:

$$\frac{1}{1 - \alpha T_{\text{new}}} < \frac{1}{1 - \alpha T_{\text{old}}} \Rightarrow (1 - \alpha T_{\text{new}}) > (1 - \alpha T_{\text{old}})$$  \hspace{1cm} (13)

$$\Rightarrow T_{\text{new}} < T_{\text{old}}$$  \hspace{1cm} (14)

That is true that the average service time of the incoming packets to the router in the proposed design is less than the conventional NoCs, since in the new NoC, just the packets coming from the local port are routed by the routing algorithm and packets coming from other ports are just send to the destination port. But in traditional design for each packet coming from any port, a process should be done to decide the destination port.

In the referenced article the latency is modeled as follow:

$$L_{\text{ad}} = W_s + \sum_{\{i,j\}} (w_j + H_s) + \left[ \frac{s}{W} \right]$$  \hspace{1cm} (15)

$L_{\text{ad}}$ is the average latency for each pair of source/destination. $W_s$ and $w_j$ are the input queuing delays for the source and other ports respectively, $H_s$ is the header flit’s service time and $S$ is the size of each packet and $W$ is the bandwidth of the NoC. The total packet latency of the NoC is as:
\[ L = \frac{1}{\sum_{s,d} x_{sd}} \sum_{s,d} x_{sd} \times L_{sd} \]  

(16)

\( x_{sd} \) is the injection rate from source \( s \) to destination \( d \). The result of comparison shows in figure 8. From the figure it can be concluded that when the traffic load is low the proposed NoC has less latency in comparison to the conventional NoC with the same specifications and common parameters.

![Fig.8. The average latency for different traffic loads.](image)

In high work load the controller has a negative effect on the latency and the conventional NoC works with less delay. It can be concluded that the controller is the bottleneck of the NoC in high traffic load and the proposed solution for it is using hierarchical or multiple controller to overcome the overflowing. It may investigated in future works.

As we know the power consumption of circuit depends on the technology of implementation, so by implementing a design in FPGA we can’t have the real power consumption of a circuit. Judging about the power consumption of the design needs technology dependent analysis, which is left for future work.

VI. CONCLUSION

For getting more benefits of using SDN concept in NoC, some improvements according to the requirements of the NoC are presented in this paper. Flexibility and scalability are some of the main approaches of the design. The topology of this NoC is capable of being two separate rings, mesh and torus and also a configurable controller for having different routing algorithms. The area will decrease too. Latency of the packets in new scheme in low terrific rate is better than the old NoC. Because the controller is the bottleneck, it can be improved with multiple controllers.

ACKNOWLEDGMENT

The authors wish to thank DR Amir Rajabzadeh and Dr. M. Ahmady, faculty of Department of computer Engineering, Rzi University, for their kindness guidance.

REFERENCES


Authors’ Profiles

Abdolhossein Fathi was born in Delfan, Iran, in 1978. He received his B.Sc. in Computer Engineering from Iran university of Science and Technology in 2001 and his M.Sc. in Computer Architecture from Sharif University of Technology in 2003. He also received his Ph.D. Degree from the University of Isfahan in 2012. He is an assistant professor at Razi University, Kermanshah, Iran. His research interests include Computer Architecture, HDL-based design, Signal and Image Processing, Biometric, Data Compression, Computer Vision, Pattern Recognition and Medical Image Analysis.

Keihaneh Kia received the B.Sc. degree (with highest honors) from Shahid Bahonar University of Kerman, Kerman, Iran, in 2007, and the M.Sc. degree from Islamic Azad University (IAU), Science and Research Branch, Tehran, Iran, in 2010, both in computer engineering. She is currently working toward the Ph.D. degree in computer architecture with the Department of Computer Engineering, Razi University. Her research interests include network on chip, reliability and energy efficiency in computer systems.