

Signaling Approached Based NoC Architecture for Improving Quality of Service

¹Dr. Swati S. Godbole, ²Jaya R. Surywanshi

¹Department of Electronics Engineering, Shri Ramdeobaba COE and Management, Nagpur, India

²Department of Information Technology, NDMVPS's KBT COE, Nashik, India

Email: godboless@rknc.edu, suryavanshi.jaya@kbtcoe.org

Received: 20th September 2018, Accepted: 11th October 2018, Published: 31st October 2018

Abstract

System on chip (SoC) is generally referred to as the integration of software and silicon hardware which are designed to perform certain functions for various applications like smart phones, digital television etc. Network-On-Chip (NoC) can be seen as the novel communication mechanism that can be adopted for utilization with SoC. Router plays a very important role for NoC. The foremost dispute in these NoC routers is related to the quality of service (QoS) functionalities like better bandwidth, inadequate transaction times, rate of losses, and low congestion. Throughput, Latency or Delay, Errors, Bandwidth, Dropped Packets, Jitter or variation in delay, Packet loss, Out-of-order packet delivery are certain parameters that are to be considered while designing the NoC for better QoS. To improve Quality of Service (QoS) parameters such as throughput and latency, Virtual Output Queuing (VOQ) signaling technique based router architecture is proposed. The number of pipeline stages of a packet transfer is reduced to less number of stages with switch allocation and traversal based on VOQ application in buffering. The study is undertaken to improve the Quality of Service in NoC architecture by proposing a novel modeling technique.

Keywords

Dynamic XY (DyXY) Routing Algorithm, Modified Cuckoo Search Optimization, Network-on-Chip (NoC), Quality of Service (QoS), System-on-Chips (SoCs), Virtual Output Queuing (VOQ).

Introduction

The field of electronics is widely developing in the modern times. The electronics these days have become much more compact accommodating a ton of services in a tiny area. This is facilitated by the integrated architecture. All the different components are developed as a single chip according to the requirement of the application. According to predictions as per the Moore's Law, the density of an integrated circuit roughly doubles with every passing year [1].

The term System on chip (SOC) is highly popular for compact systems and fundamentally used to refer the chip which has the integration of software with the on-chip silicon hardware and is specially designed so as to perform specific functions. The integrated chip which is system on chip may have digital signals, analog signals, mixed signal or even radio frequency functions - all capable from a single chip developed on a single substrate thereby making it compact as well as functional. They are widely popular in a variety of applications the most common of it being smart phones. They are also found in digital television sets, wearable computing devices and many others [2].

The popularity of SoC could be attributed to the variety of applications they offer on a single integrated module. However, there are certain drawbacks for the SoC like unpredictable performance, scaling problems, and reduced efficiency, to name a few. There have been numerous researches undertaken globally to tackle these issues and improve the performance [3]. The advancements in the field of VLSI integration densities have reached towards the Nano scale era; they have also witnessed a paradigm shift in the designing of the chip from computation centric designs to the newly accepted communication centric designs which have the capability of incorporating very large number of simple cores. Numerous schemes of the traditional design advocate the point to point scheme using buses and crossbars. The point to point communications schemes help achieve fast and efficient communication; however, the wire density is considered as a barrier for adapting them into the architectures of multi core SoC. The buses which are considered extremely simple in design have intrinsic drawbacks like arbitration, scalability as well as the bandwidth bottleneck with the increase in the number of cores in a SoC. In the case of crossbars, its applicability is limited due the high requirements of power and area posed by it. Hence, when considering multi-core architectures like Chip Multiprocessors (CMP) or the Multi-processor System-on-Chip (MPSoCs), the traditional solutions are rendered useless for the purpose of handling communication and thus to overcome these communication challenges some efficient communication infrastructure is required [4].

The answer to all the above issues of limited density, scalability, bandwidth and power can be the adoption of Network On chip (NoC) to facilitate the interconnection and the communication in the multicore SoC. The technology of NoC finds implementation in the SoC interconnects consistently because it is significantly smaller, has improved speed and is deemed to be much more power-efficient than the traditional crossbars as well as decoupled busses [5].

Research Methodology and Implementation

The principal objective of the research is to improve the Quality of Service (QoS) in NoC architecture. In proposed work QoS parameter like throughput and latency has considered to improve QoS. A novel NoC architecture will be designed which is the aid in improving QoS. Virtual Output Queuing method [6] with signaling approached has introduced. Various network topologies will be employed, and QoS parameters like throughput and latency are analyzed. The Figure 1 shows the proposed scheme-flow diagram. It considers the VOQ signaling scheme that depends on router architecture for NoC. Figure 2 shows the arrangement of sixteen cores and routers in a 4×4 mesh topology [7]. The 2D mesh topology of Network-On-Chip (NoC) consists of IP core, routers, Network interfaces and links. In NoC each core is linked to a switch by a network interface. The purpose of a network interface is to define how data packets are formatted for transmission and routing. Routers direct data over several links and the links connect switches with network interfaces.

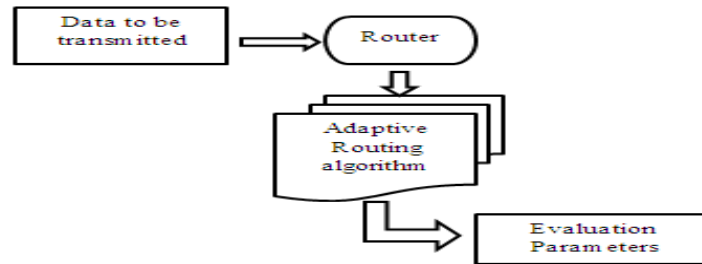
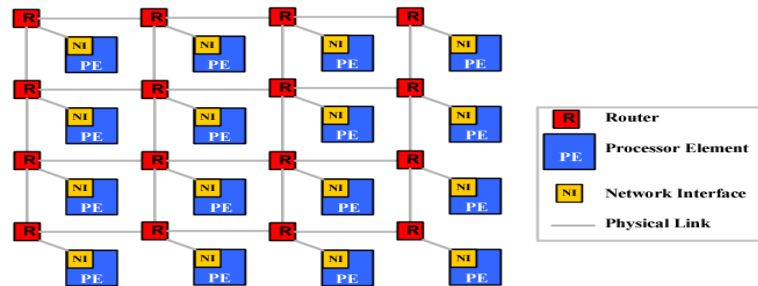


Figure 1: Proposed Scheme-Flow Diagram of NoC.



incoming signaling flit. These stored states are required at the header flit management during the path establishment processes. Finally, this block is cooperating with the routing block during the data flit routing. The routing block is the main block of the router. It ensures the routing processes of the different incoming flits. This block uses the wormhole technique and a Dynamic X-Y [10] routing algorithm basing on signaling information in order to find the best path for the data processes.

| S.N | Components | Specifications |
|-----|---------------------------------|---------------------------|
| 1. | Topology | 4×4 2D topology |
| 2. | Routing Algorithm | Dy XY routing |
| 3. | Packet switching | Wormhole packet switching |
| 4. | Buffer size | 32 bits |
| 5. | Flit Size | 32 bits |
| 6. | Packet size | 16 flits |
| 8. | No of virtual channel | 4/port |
| 9. | Depth of virtual channel | 4 flits/VC |
| 10. | Width of the communication port | 32 bits |
| 11. | Channel width of the network | 32 bits |

Table 1. Specifications Details of Proposed Router

▪ The Virtual Channel Output Queue

Virtual output queue buffer system has maintained a separated queue for each output port at each input side. That's why it is called as virtual output queue buffer. In VOQ architecture input side has a separate VC has reserved for each output port. In proposed work has combinely used this buffering system with signaling block to remove the deadlock avoiding conditions.

▪ Flow Control

Flow control mechanism shows how efficiently the network resources have allocated. Two flow control methods are available. One is buffered flow control and another is buffer less flow control methods. Buffer less mechanism has a simpler and it more power and space efficient as compare to buffered. But the limitation are packets loss and misrouting the packets inside the network in case of congestion and conflict of resources. In that case buffered is useful. Different buffered flow control methods are available. For proposed router Credit Based flow control has preferred.

▪ Credit Based Flow Control

Whenever any input buffer has available a free buffer space then this information has generated in the form of the credit. This credit information then converted into the stress value of the node. If free buffer space is more than stress value is less and if free buffer space is less stress value is more. This stress value then forwarded by the singling block to neighboring router to manage the congestion condition. The crediting mechanism always works hand in hand with arbitration, providing additional information to the arbiter to come up with a more correct decision.

▪ Router Arbiter Mechanism

The Router Arbiter work has defined where it is responsible for forwarding the incoming packet to the desirable output port. Router arbiter takes the decision who will assign the output port form incoming requests. When there are multiple requests has generated simultaneously for a single device or channel then there is conflict. This conflict has resolved by arbitration mechanism. Round robin descending order clockwise direction mechanism has implemented here for arbitration work. The arbiter can collect the requests of packets simultaneously from the five nodes. In presented architecture Arbitration mechanism, Signaling block and Routing block has work together. Arbiter block has fully connected with all input and output port. When any packets arrived inside buffer then this information forwarded to signaling block. Signaling block also propagate this information to all neighbors routers. Physical paths are presents there for forwarding this information to other routers. This signaling path is bidirectional. Same path has used for collecting the neighboring buffer status information. Routing block always checks the status of signaling block while preparing the routing path. Then it will provide the control to arbitration mechanism. Before transferring the packets to the next hope it check the status of READY signal of output port. If it is ready then and then only packets has switched to other hope. This mechanism has removed the congestion creation probability form network. When arbiter are actually sending the packets from one node to another node it generate the CONNECT signal high. So there is no other channel try to access for same node. Arbiter completed their work when it found TAIL flit of packets. When arbiter received the last TAIL FILT it has release the CONNECT signal. So other input port can generate the request for same port. Arbitration mechanism works on 5X5 mux based architecture.

▪ Round Robin Arbitration Mechanism

Round robin arbitration method has a round robin mechanism where each port has given a priority on circular fashion in clockwise or anti-clockwise direction. Required input has generated its request to arbiter. When more than one request is there then arbiter apply the algorithm. The ascending order clockwise direction round robin algorithm used here. The highest priority request granted first then it goes at last priority level. The round robin mechanism works on time slot allotment based. In a predefine time slots it will check which ports need the service. This works on token passing mechanism. This token has continuously moved ahead. Suppose there is one port that has a highest priority own the token for request but its output port do not have a ready state then token has automatically forwarded to other higher priority node whose node is ready to share the resources. In this way it works for assigning of sharing resources.

Arbiter keeps the input output resource allotment information in a table .This table is called as history table. Arbiter refers this table while applying routing algorithm and simultaneously update that entry as per assignment the resources.

| Port | Virtual channel 1 (Local port) | Virtual channel 2 (Local port) | Virtual channel 3 (Local port) | Virtual channel 4 (Local port) |
|------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| N | 1 | 1 | 1 | 1 |
| E | 1 | 1 | 1 | 1 |
| W | 1 | 1 | 1 | 1 |
| S | 1 | 1 | 1 | 1 |

Table 2. Arbiter History Table at Initial State

History table shows the port request in row wise where request generation for virtual channel in column wise. Suppose at the initial condition all channels are free and available to all four ports. So values are 1 which represents the free VC buffer status. Now the four ports simultaneously generate the request for VC1 then what happened. All four channels have needed same VC. At that time arbitration apply the round robin algorithm. As per the algorithm suppose port 'W' port has highest priority it will own the token. Arbiter gives VC to 'W' port and updates the history table contents as follows. Assign the VC to 'W' represents by high signal. It means not accessible to others.

| Port | Virtual channel 1 (Local port) | Virtual channel 2 (Local port) | Virtual channel 3 (Local port) | Virtual channel 4 (Local port) |
|------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|
| N | 0 | 1 | 1 | 1 |
| E | 0 | 1 | 1 | 1 |
| W | 1 | 1 | 1 | 1 |
| S | 0 | 1 | 1 | 1 |

Table 3. Arbiter History Table State After Channel Allotment

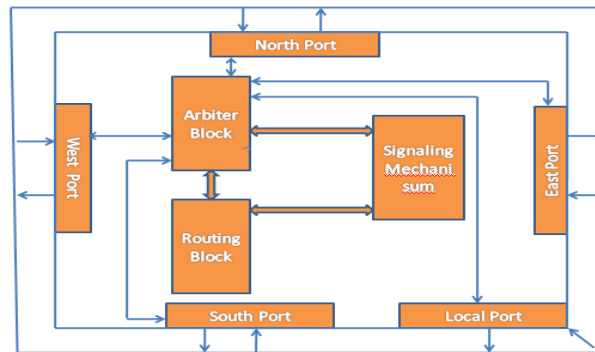


Figure 3: Router Architecture with Signaling Block

Signaling Mechanism

Congestion in the context of NoC defines different states of network such as a node or link which carries large data, traffic gets mixed with lots of other traffic, incoming packets arrive simultaneously for the same outgoing link etc. It may degrade the network service quality because of packets are delayed and no free space available in input buffer. Lots of reasons are there for congestion occurrence. Suppose the bandwidth is not sufficient and network data traffic exceeds capacity then congestion will be created inside the network .In presented work a Signaling block introduce inside router to fight congestion problem.

The signaling block performs some major task for congestion control. It generates and forward signaling information from asynchronous flits in order to update the neighboring routers data with new traffic load states. The important work of this block is to analyze newly received signaling flits in order to update the internal routing table. The information stored in signaling table for making the routing decision.

The block will be collected congestion information from neighboring port by checking status of virtual output queues. If the buffer has free space for storing packets it means no congestion is there otherwise congestion is presented. All neighboring blocks send its current status information to the signaling block through physical signaling control signal. So this signaling block has the information of all neighboring block. This is required to manage a head-flit during the path establishment process.

In the signaling block, the signaling process is done by the generation of signaling flit .Singling information is nothing but the information of congestion present in neighboring router by checking the free buffer space availability of routers. Singling flits provides the congestion information of neighboring routers. Signaling block holds the neighboring routers information for routing purpose. This information is required to the header flit management for the path establishment processes. Signaling block directly connected to arbiter block and routing block. This information then forward to routing block. Routing block then decided the new route to forward the packets. It always prefers the free route where the no conjunction is presented.

| Port | Virtual channel 0 | Virtual channel 1 | Virtual channel 2 | Virtual channel 3 |
|------|-------------------|-------------------|-------------------|-------------------|
| N | 1 | 0 | 0 | 1 |
| E | 1 | 1 | 1 | 0 |
| W | 0 | 0 | 0 | 0 |
| S | 1 | 0 | 1 | 0 |
| L | 0 | 1 | 0 | 0 |

Table 4. Stress Value Table of Signaling Block

Table shows the stress values entry of all neighboring port. From table it is clear that '1' value presents the high stress value where '0' value represents less stress value of each node. Stress value signal has update by the stress value control signal. These are bi-directional physical path signal.

Routing Algorithm

Routing algorithm is a major factor which concerns the efficiency of the communication of NoC. DyXY provides adaptive routing based on congestion condition.

Results and Discussion

The proposed system will be implemented for simulation in the Xilinx 14.1 Simulator software package to observe the performance and analyze it. The parameters like throughput and Latency would be checked at the output to come to the agreement of the efficiency of the proposed system. The router architecture is implemented by VHDL software package to examine the performance. The QoS parameters such as throughput and latency are estimated at the output which verifies the efficiency of the proposed system. For the experimentation, injection rate of packets are considering 1 flit per node per cycle that is maximum value where working frequency is 500 MHz. The verification of results of this research work is been finalized based on the benchmarks parameters. These benchmark parameters are finalized on the basis of the existing literature survey where the working frequency is 500 MHz[11][12][13]. The QoS parameters are evaluated for various topologies such as 2D, Tree, Mesh, Octal, Ring and Star topology. Table 5 shows the throughput and latency results for different topologies which verify that 2D topology achieves high throughput and low latency among them.

| S. N | Topology | Average Latency in Clocks | Average Throughput in % |
|------|----------|---------------------------|-------------------------|
| 1 | 2D | 23 | 57.7320 |
| 2 | Tree | 41 | 42.9760 |
| 3 | Mesh | 45 | 36.7290 |
| 4 | Octal | 62 | 53.2700 |
| 5 | Ring | 43 | 42.9760 |
| 6 | Star | 50 | 33.8100 |

Table 5. Throughput and Latency Results for Different Topologies.

We have estimated the average throughput and latency for different routing algorithms with 2D mesh topology. Table 6 shows the comparison of throughput and latency results for different routing algorithms. The results verify the DyXY routing algorithm with 2D mesh network achieves the better QoS parameters. The performance factors like Throughput and Latency are estimated to analyze the effectiveness of our proposed research and will be compared with existing techniques.

| S. N. | Topology | Routing Algorithm | Average Latency in Clocks | Average Throughput in % |
|-------|----------|-------------------|---------------------------|-------------------------|
| 1. | 2 D Mesh | XY | 32 | 26.8 |
| 2. | 2 D Mesh | Y priority | 30 | 26.8 |
| 3. | 2 D Mesh | XYX | 28 | 53.6 |
| 4. | 2 D Mesh | DyXY | 27 | 59.4 |

Table 6. Throughput and Latency Results for Different Routing Algorithms.

Conclusion

In this paper, we have presented Virtual Output Queuing (VOQ) signaling approach based router architecture for improving the throughput and latency. The proposed router is implemented for various routing algorithms and topologies using VHDL simulation software and evaluated the performance of throughput and latency. The experimental results show that the proposed Dynamic XY (DyXY) routing algorithm with 4×4 2D network topology reduces the communication latency by 27 in clocks and throughput by 59.4% when compared to other routing algorithm such as XY, Y priority and XYX routing algorithm.

References

- [1] Patti, R. S. (2006), "Three-dimensional integrated circuits and the future of system-on-chip designs", Proceedings of the IEEE, 94(6), 1214-1224.
- [2] Wang, L. T., Stroud, C. E., & Toubia, N. A. (2010), "System-on-chip test architectures: nanometer design for testability. Morgan Kaufmann"
- [3] Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012), "Internet of things: Vision, applications and research challenges", Ad Hoc Networks, 10(7), 1497-1516.
- [4] Jerraya, A., & Wolf, W. (2004), "Multiprocessor systems-on-chips", Elsevier.
- [5] Kumar, S., Jantsch, A., Soininen, J. P., Forsell, M., Millberg, M., Oberg, J. & Hemani, A. (2002), "A network on chip architecture and design methodology", In VLSI, 2002. Proceedings. IEEE Computer Society Annual Symposium on (pp. 117-124). IEEE.
- [6] Fadoua Hassen, Lotfi Mhamdi, "A Scalable Packet-Switch Based on Output-Queued NoCs for Data Centre Networks. 2016 IEEE International Conference on Communications (ICC).
- [7] Mohammed Amine Meghabber, Abdelkader Aroui, Toufik Djeradi, "A flexible network on-chip router for data-flow monitoring", 2017 5th International Conference on Electrical Engineering - Boumerdes (ICEE-B)
- [8] M.Mareli, B.Twala, "An adaptive Cuckoo search algorithm for optimization", Applied Computing and Informatics Volume 14, Issue 2, July 2018, Pages 107-115
- [9] A.S.Joshi, Omkar Kulkarni, G.M.Kakandikar, "Cuckoo Search Optimization- A Review", ScienceDirect Volume 4, Issue 8, 2017, Pages 7262-7269
- [10] Ming Li, Qing-An Zeng, Wen-Ben Jone, "DyXY - a proximity congestion-aware deadlock-free dynamic routing method for network on chip", 2006 43rd ACM/IEEE Design Automation Conference
- [11] "Local Congestion Avoidance in Network-on-Chip", Minghua Tang, Xiaola Lin, and Maurizio Palesi, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 27, NO. 7, JULY 2016
- [12] "Shield: A Reliable Network-on-Chip Router Architecture for Chip Multiprocessors", Pavan Poluri, and Ahmed Louri, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 27, NO. 10, OCTOBER 2016
- [13] "A Modular Router Architecture Design For Network on Chip", Brahim Attia, Wissem Chouchene, Abdelkrim Zitouni, Nouredine Abid, and Rached Tourki, Faculty of Sciences of Monastir, 2011 8th International Multi-Conference on Systems, Signals & Devices.