IOSR Journal of VLSI and Signal Processing (IOSR-JVSP)
Volume 7, Issue 4, Ver. I (Jul. - Aug. 2017), PP 29-35
e-ISSN: 2319 – 4200, p-ISSN No. : 2319 – 4197
www.iosrjournals.org

Network-on-Chip: A State-of-the-art Review

Seema [*], Pawan Kumar Dahiya[**]
[*]*Scholar of M. Tech. in ECE (VLSI Design), Deenbandhu Chhotu Ram University of Science & Technology, Sonepat, India*
[**]*A.P., ECED, Deenbandhu Chhotu Ram University of Science & Technology, Sonepat, India.*
Corresponding Author: Seema

Abstract: *Large scale System-on-Chip (SoC) has been enabled by the scaling of microchip technologies. As data intensive applications have emerged and processing power has increased, the threat of the communication components on single-chip systems introduced network on chip (NoC). NoC provides the concept of interachip communication. In this paper a study treats an outstanding concept for system-on-chip communication introduced as communication network on-chip (NoC). This paper includes the NoC basics, network topology, relevant research issues and different abstraction levels.*

Keywords**:** *NoC, SoC, FPGA, VLSI***.**

Date of Submission: 29-06-2017 Date of acceptance: 15-07-2017

I. Introduction

As complexity of the embedded system increasing and modern silicon technology scope spreading, there is a movement towards separate systems having many processing elements with specifications, implemented on one chip only, defined as a On-Chip-System or System-On-Chip (SoC). IC design has been advanced into SoC era by increasing integration levels offered by the current semiconductor technologies [1]. Better performance and functionality of the instant and future SoCs is achieved by integration large amount of processing elements on a single-chip [2]. Chip design includes four essential and distinct aspects: memory, communication, I/O and computation. Many challenges of communication in SoC have attracted increasing attention towards NoC. On Chip network (NoC) provides a concept for the intrachip communication [3].

Communication within single chip: According to Moore's law every second year, processing elements get double on a single chip. This increasing rate slowdowns to doubling every three year for fixed area of a chip and now forcing this rate exponentially. With the maturity of technology a paradigm shift occurs such as room-to rack-level systems (LSI-1970s) where a chip was system module's component, rack-to board-level systems (VLSI 1980s) where a chip was system-module and board-to chip-level systems (ULSI-1990s) where a chip makes the whole system therefore System-on-chip comes into picture. Fig.1 shows the paradigm shift as technology matures. SoC

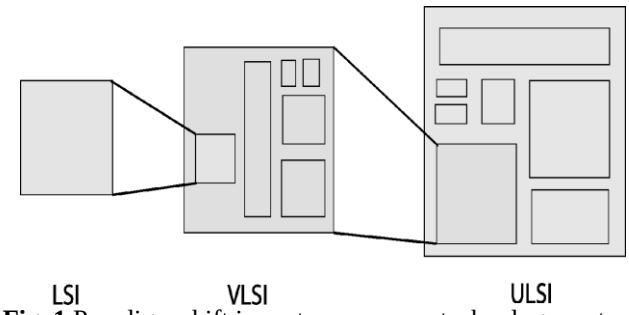

Fig. 1 Paradigm shift in system scope as technology mature

uses complex pipelined processing and firmly interdependent processes for the communication structure with wide ranges. A shared segmented SoC communication structure use common approach that is based on network and known as network-on-chip. Many ingredients related to communication address by NoC. NoC research area provides two widely taken concepts for NoC that are: (1) SoC's subset NoC (2) SoC's extension NoC. This survey consists the basics of NoC including topologies with the all the design abstraction levels and description of all recent trends involved in intra chip communication. We have summarized many relevant key issues such as architectural issues. Section 3 & 4 is related to research being done related to all

these issues. The rest of the survey is organized as: Section 2 includes review the NoC basics, simple NoC example addressing some important issues related to system-level architectures and research areas. Section 3 includes more research related to NoC coordinated. Section 4 consist of modeling of NoC will be discussed.

II. Network On Chip Basics

This section incorporate NoC basics starting with an attention on component-based view that will introduce the basic constructive blocks of a typical NoC then a glance on issues related to system-level architectural in NoC-based SoC and in the last abstraction-based layered NoC presented here.

2.1 Network on Chip example:

Fig.2 presents the various types of NoC topologies. Instead of these topologies and devoted node-to-node links, a more relevant design is modified, implementing grid routing nodes, expanding across the chip and connecting them by communication links as shown in Fig.3. This figure shows a 4×4 grid NoC structured sample which gives global-level-communication.

NoC have the following elementary components: --*Network adapters*: These provide the interface through which IP cores make connection to NoC. Network adapters provide service of separation between computation (cores) and communication (the network).

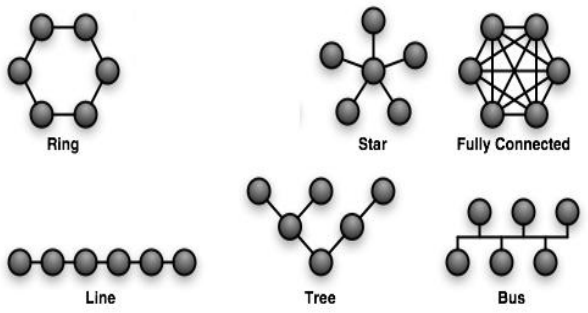

Fig.2. Network on Chip Topologies

--*Routing nodes* provide services of routing to the data corresponding to selected protocols.

--*Links*: These connect the nodes, providing raw bandwidth.

Routing is defined as a process of selecting a path for the massage from its source to destination corresponding to network topology. Routing algorithms are formed with two major objectives as (1) Balanced Network Traffic (2) Short Network Path Length. Adaptive routing and deterministic routing algorithm are used to provide route in NoC.

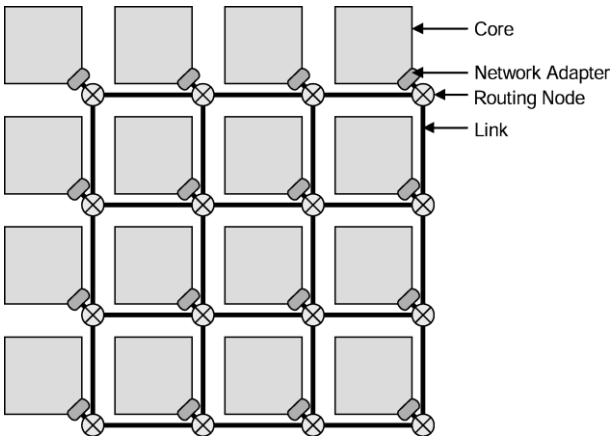

Fig.3. A 4×4 grid structured NoC topology specifying the elementay components.

2.2 Architectural Issues: System composition and clustering are architectural issues that affect the diversity of communication. Fig.4 shows how categorization of system composition done along the axes of homogeneity and granularity of system cores.

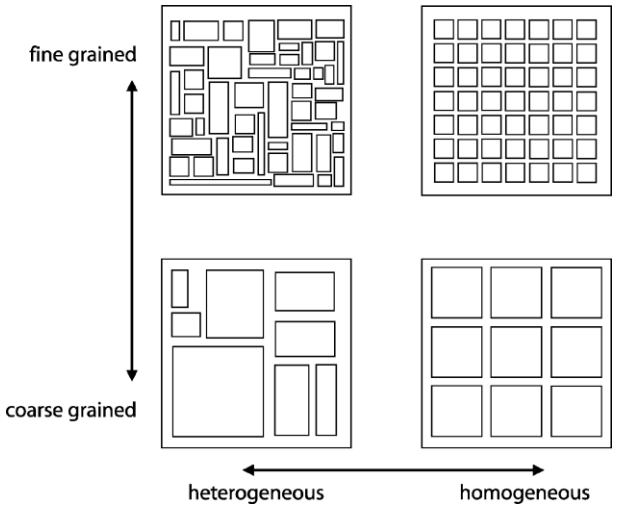

Fig. 4 System composition's categorization along the axes of homogeneity and granularity.

Clustering is related with the system's localization. It can be of two type (1)logical (2) physical. Implementing the hardware primitives in the network and programming tools are related to logical clustering. While physical clustering related to the co-existing knowledge of traffic pattern in the system and can reduce total communication cost, power, performance and global communication. Re-configurability can be related in NoC-based system as how the NoC a very flexible communication structure can be used to make the system reusable from the application pint of view. Routing nodes and links are the reconfigurable resource rather than wires in NoC. The main architectural issue is balancing of flexibility, hardware cost and performance of the system as a whole.

2.3 Network-level-Absraction:

Today NoC researches include the range from physical implementation based on gate-level to design methodologies and tools along with system layout aspects and application. For a better understanding of NoC, NoC research spectrum is separated into four levels as following: (a) system-level (b) network adapter-level (c) network-level (d) link-level. Fig.5. presents the OSI layers, the fundamental aspects of NoC and the data flow through the network connecting these research fields.

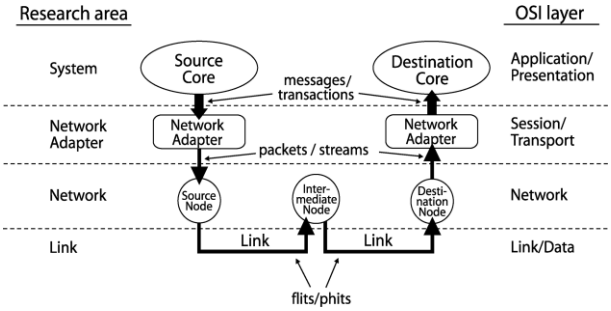

Fig. 5 OSI layer, different NoC research levels and flow of data connecting these levels.

The system enclosed architectures (cores and network) and applications (processes). Almost network implementation details are masked in this level. The NA-level separate the cores from network. At this level, massages or transactions are encapsulated and packets are made which enclose the information regarding their destination. Fig.6. shows a dummy of packetization.

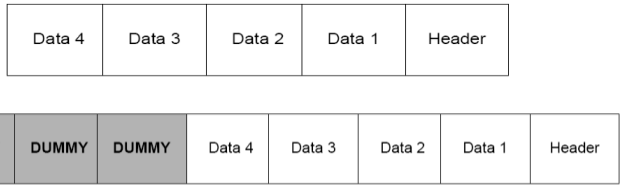

Fig.6 Packet with fixed size having dummy data (top) and packet with variable size having same data (bottom)

The network includes all routing nodes, links etc, implementing the protocol, defining the topology and flow control from one node to other node. Lowest level of NoC research fields is link level. At link level, basic data grams are flits and phits. Flits are flow control units and phits are physical units. Streams and packets are made by these node level atomic units. Most commonly phits and flits are equivalent, while a sequence of phits could be used to make a flit to employ highly serialized links. Encoding and synchronization issues are considered in link-level research.

III. Netwrok-On-Chip Research

This section includes review of various concepts of research groups. Fig. 7 shows a simplified classification of this research.

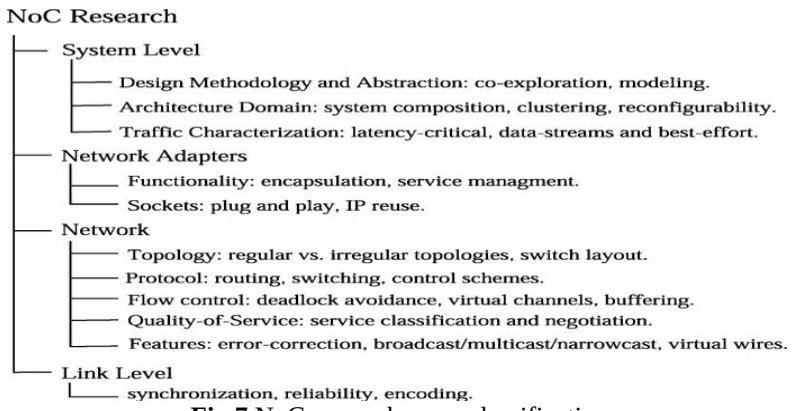

Fig.7 NoC research areas classification.

3.1 Network Adapter level:

Main objective of this level is to provide interface between the cores and the network and to provide very clear communication services. At this level, the barrier between communication and computation is stated. As shown in Fig.8, the NA component implements interfaces on the both network and core side. Interface implemented at the network side named as network interface (NI) while interface at core side named as core interface (CI). The objective of NA is to provide high-level and transparent communication services to the core by utilization of the basic services provided by network hardware.

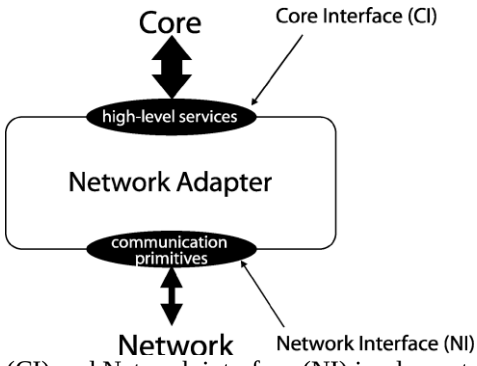

Fig. 8. Core interface (CI) and Network interface (NI) implemented by Network Adapter

Here, we address the use of standard socket firstly, then services of NA and finally NA implementations.

3.1.1. Sockets: At the core side of NA the core interface (CI) may be implemented to comply with a System-on-Chip socket standard. Main focus of socket is to orthogonalize communication and computation. Reusability is facilitated by socket because the core interface adheres to the socket standard alone separately from network hardware. Generally used socket standards are Open Core Protocol (OCP)and Virtual Component Interface (VCI).

3.1.2. Services of NA: Mostly, encapsulation of the traffic is provided by NA for the basic communication media and these services are managed by the network. Encapsulation concerns flow control in network, routing tasks and global addressing, data acknowledgement and reorder buffering and management of buffer from network congestion.

3.1.3. NA Implementations. For a successful NoC design a transparent understanding of NA contribution is crucial for a successful NoC design. [1] State synchronization issues and proposed a design of an asynchronous envelope for the use in practical GALS design. Here globally asynchronous wrappers adapt interfaces of synchronous modules. A synchronous/asynchronous mixed NA architecture proposed in [2]. Here, a synchronous domain where packets formed OCP interface connected to asynchronous domain where flits sequencing done. A complete NA design proposed in [3] for the ÆTHEREAL NoC which provides a shared-memory abstraction to the cores and give compatibility all on-chip existing protocols as AXI, DTL and OCP with the cores. AS indicated in the HERMES NoC [4], the transactions are made 50% slower than the native core interface with a trade-off issue that is partitioning hardware and software. In [5] a comparison established between software and hardware implementations for packetization task and provides different cycles used by software and hardware. Software took 47 cycles while hardware took only 2 cycles. A hardware implementation of the whole NA introduced in [6] which gives pipelined architecture to get maximum throughput with latency of between 4 and 10 cycles.

3.2 Network level

The objective of this level is to transfer data from source to designated destination. The topology and the protocol implementation by a network mainly define an on-chip network. Topology consists the connectivity and layout of the nodes and links between the nodes on the chip. These links and nodes are defined by protocol.
3.2.1. Topology: Topologies for on-chip network are in two for (1) regular topology (2) irregular topology. Regular forms of topologies are widely used for NoC. Fig. 9 shows the regular topologies termed as k-ary n-cube(grid type) where k is degree of every dimension and n is the total number of dimensions..

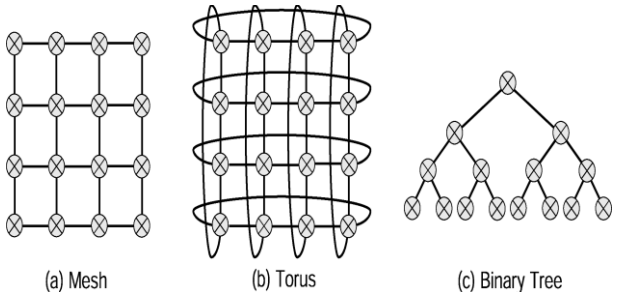

Fig. 9 Regular topologies (a) 4-ary 2-cube mesh (b) 4-ary 2-cube torus (c) 2-ary binary tree

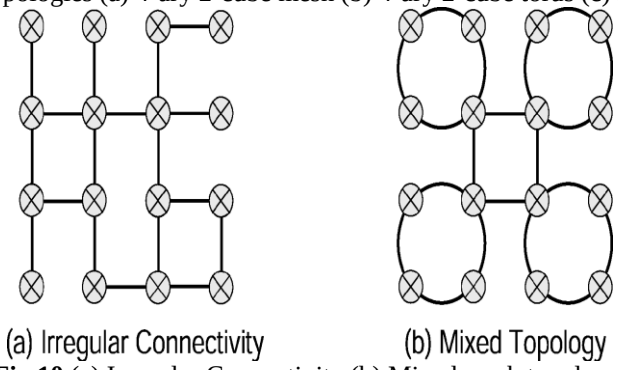

Fig.10 (a) Irregular Connectivity (b) Mixed mesh topology

For the increasing size of the regular topologies area and power consumption of the network are scaled.Reguler topologies are mesh, torus, binary tree ,k-ary tree and k-ary n-dimensional fat tree topologies. Irregular form of topologies includes the topologies which are managed by mixing different form in asymmetric pattern, hybrid or hierarchy as shown in Fig. 12[8].

3.2.2 Protocol: Protocols defined as set of rules used to moving the data through the NoC. Protocols addresses few aspects such as circuit or packet switching, connection oriented or connectionless mechanism, adaptive or determine routing, minimal or non-minimal routing and central or distributed control. The protocol defines the use of available resources. The fundamental components of routing nodes are switch, arbitration unit, buffers, routing and link controller [10].
3.2.3 Flow Control: The flow or movement of packet through the network path is determined by the mechanism known as flow control. Therefore, flow control includes local and global issues. Issues related to correct operation of the network, optimal utilization of network resources and predictable performance of communication services are address by flow control. Main motive of flow control of the network is to avoid

Deadlock and livelock firstly. Livelock comes into picture when resources constantly change their state waiting for other to finish while deadlock comes into picture when resources are in waiting state for each other to be released. Local and global methods can be used to avoid livelock and deadlock. Dimension-ordered routing, the turn-model and odd-even turn-model are popular choice for NoC to avoid deadlocks.

3.2.4 QoS: QoS is quality of services and defined as the service quantification given by the network to the demanding core. Two main aspects are involved by Qos are defining the services which are represented by quantification and balancing the services demands. These services can be of low latency, low power and high throughput. Many special services such as broadcast, multicast, narrowcast, virual wires and complex operations are also covered by QoS.

3.3. Link level

Node to node links is address by link level. These links includes one or more physical or virtual channels. Synchronization, implementation, encoding and reliability are the main research area at this level. In multi clock SoC, for link-level synchronization a FIFO design presented in [11]. A ring used by this FIFO for the storage purpose of elements and full and empty state of the FIFO indicated by tokens. Detection of the state of FIFO (full or empty) makes the synchronization robust. Hence globally asynchronous and locally synchronous become advance over it due to less leakage and asynchronous design style of NoC. The power consumption and link delays increase due to the effects of wires as the chip technologies scaling down. A few physical level issues related to implementation of NoC links are wire segmentations, pipelining and low swing drivers. Communication reliability issues become relevant during designing global interconnect in DSM technologies. Main sources of power supply noise, crosstalk such as intersymbol interference and electromagnetic interference (EMI) effect the reliability. Main motive of encoding is reduction in power usage per communicating bit with maintaining the noise margin and high speed for communication of NoC. Encoding also includes error management of on-chip communication that contains error detection and correction occurred in the network. Encoding is applicable to packets, flits and phits.

IV. Noc Modeling

Network-on-chip is defined as a subset of SoC and integral part of SoC design methodology and architecture. NoC design includes broad design space and implementation decisions. For the NoC design flow, integration and verification of NoC design simulation and modeling is essential. Here, in this section NoC modeling related issues uncovered then a glance on design methodology which is used for system-level impact study on NoC and in last traffic characterization.

4.1 Modeling

Modeling of NoC provides abstract software models where first consideration is to understand the required NoC architecture and traffic within it. Basically the main objectives of NoC modeling are (1) to search broad design and feature space (2) to figure out the trade-offs among design-time, power, and area etc. (3) to adhere application requirement on one-side with technology constraints on other side. NoC modeling includes the three aspects as modeling environment, abstraction level and result analysis.

4.1.1 Modeling Environment: NoC models are modeled on the basis of analytical and simulation and used to model the communication across all the abstraction levels. Allocator, scheduler and synchronizer are used to model NoC. The allocator provides the route traversal requirements of data in terms of resources required such as buffers, bandwidth etc. The allocator also concerns about minimization of resources conflicts. The scheduler objective is to transfer the data according to the requirements of network services and it pay attention about minimize resource occupancy. The synchronizer provides dependencies among concurrent communicating data. Therefore NoC architectures can be broadly described by these three components and real-time simulation can be performed. An example of NoC simulator is OPNET which gives a convenient tool for modeling of a hierarchical network. For the evaluation of performance and power of NoC a VHDL-based RTL model is used. Power and delay are calculated for fine-grained aspects of NoC routers and links. Spice simulator is used for evaluation of power and delay.

4.1.2 Different abstraction levels NoC modelling: A modular NoC-based system-on-chip design needs a full range of abstraction levels and this broad range is provided by simulation at different level of abstraction using hardware description languages for example a library of C++, SystemC and System Verilog. Many researchers working at different level of abstraction define NoC architecture in many ways. For example according to some researchers network topology basically acquire the use of multihop communication structure while some working at higher level of abstraction define NoC as a multiport blackbox communication structure having multiple ports for communication. Working at the highest level of abstraction provides a great degree of freedom from issues related to low levels.

4.2. Design methodology:

Maximum use of characteristics according to the requirements of application is considered in a SoC design methodology where communication infrastructure needed. NoC design is very flexible such as parameterizable singular IP cores, flexible building blocks, tailored at network layer for design and reusability into application specification. At different level of abstraction, different NoC researchers give unique customized NoC architecture. Parametrizabilityat system-level and granularity of NoC can be used to compare different NoCs. .

4.3. Traffic Characterization

For NoC design firstly, design decision are made which are based on the traffic. Therefore characterizing the expected traffic for respective system is the first step of NoC designing. For an relevant communication of traffic pattern include different combinations of topology, protocol, packet sizes and flow control mechanism. Different-different communication fashion is acquired by NoC. Traffic characterization of NoC categories into three parts as: (1) Latency Critical: Traffic latency critical have lowest payload and includes the latency demands for example memory access, critical interrupts etc. (2) Data Streams: This traffic category has the high payload and it demands for Quality-of-Service in form of bandwidth. (3) Miscellaneous: This category of traffic has not any specific requirements from the network.

V. Conclusion

The advancement in SoC introduced a platform known as network-on-chip (NoC). The NoC provides on-chip communication that was getting more and more complex in SoC. Different network topologies such as ring, star, fully connected, line, tree, bus, mesh and torus etc. are used to provide an arrangement of the NoC routers. Routers are designed according to the OSI layer model of NoC. There are four layers similar to OSI model as system, network adapter, network and link which gives different level of abstraction in NoC. The scope of NoC is very widely because of many research areas relating to each layer. Many research have been done in the modeling style of NoC and still a wide area for the research.

References

[1] TOBIAS BJERREGAARD AND SHANKAR MAHADEVAN, "A Survey of Research and Practices of Network-on-Chip," ACM Computing Surveys, Vol. 38, March 2006, Article 1.
[2] MUTTERSBACH, J., VILLIGER, T., AND FICHTNER, W. 2000. Practical design of globally-asynchronous locally synchronous systems. In Proceedings of the 6th International Symposium on Advanced Research in Asynchronous Circuits and Systems (ASYNC). IEEE Computer Society, 52–59.
[3] BJERREGAARD, T. 2005. TheMANGOclockless network-on-chip: Concepts and implementation. Ph.D. thesis, Informatics and Mathematical Modeling, Technical University of Denmark, Lyngby, Denmark.
[4] RADULESCU, A., DIELISSEN, J., GOOSSENS, K., RIJPKEMA, E., ANDWIELAGE, P. 2004. An efficient on-chip network interface offering guaranteed services, shared-memory abstraction, and flexible network configuration. In Proceedings of Design, Automation and Testing in Europe Conference (DATE). IEEE, 878–883.
[5] OST, L., MELLO, A., PALMA, J., MORAES, F., AND CALAZANS, N. 2005. MAIA—a framework for networks on chip generation and verification. In Proceedings of the Asia and South Pacific Design Automation Conference (ASP-DAC). IEEE.
[6] BHOJWANI, P. AND MAHAPATRA, R. 2003.Interfacing cores with on-chip packet-switched networks. In Proceedings of the 16th International Conference on VLSI Design. 382–387.
[7] DALLY, W. J. 1990. Performance analysis of k-ary n-cube interconnection networks. IEEE Trans. Comput. 39, 6 (June) 775–785.
[8] MAI, K., PAASKE, T., JAYASENA, N., HO, R., DALLY, W. J., AND HOROWITZ, M. 2000. Smart memories: A modular reconfigurable architecture. In Proceedings of 27th International Symposium on Computer Architecture,161–171.
[9] KUMAR, S., JANTSCH, A., SOININEN, J.-P., FORSELL, M., MILLBERG, M., OBERG, J., TIENSYRJ¨A, K., AND HEMANI, A. 2002. A network-on-chip architecture and design methodology.In Proceedings of the Computer Society Annual Symposium on VLSI (ISVLSI). IEEE Computer Society, 117–124.
[10] DUATO, J., YALAMANCHILI, S., AND NI, L. 2003. Interconnection Networks: An Engineering Approach. Morgan Kaufmann.
[11] CHELCEA, T. AND NOWICK, S. M. 2001. Robust interfaces for mixed-timing systems with application to latency-insensitive protocols. In Proceedings of the 38th Design Automation Conference (DAC). IEEE, 21–26.
[12] Qiang Liu, IP Protection of Mesh NoCs using Square Spiral Routing, IEEE Transactions on Very Large scale Integration, Vol. 24, No. 4, April 2016.

IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) is UGC approved Journal with Sl. No. 5081, Journal no. 49363.

Seema. "Network-on-Chip: A State-of-the-art Review." IOSR Journal of VLSI and Signal Processing (IOSR-JVSP) 7.4 (2017): 29-35.