An Integrated Perspective of Crucial Research Issues in NoC Router Design

Trupti Patil, Anuradha S

Abstract: ***A new reliable high throughput NOC router design is proposed with FSM based smart arbiter module for 4X4 Mesh architecture. This design is based the XY routing algorithm with prioritized round robin arbitration and synthesis of the proposed design is done on Spartan III FPGA. An enhanced work is also done in this paper to explore the drawbacks of the exceptional techniques of the existing generation and to research the scope for overall performance improvisation of the NoC designing.***

Key words: Arbiter, FPGA, Network on chip (NOC), XY routing, prioritized round robin arbitration.

I. INTRODUCTION

Advancing IC era allows the SoC designs to incorporate different processor cores ranging from simple Embedded Memories to Application specific component on a single chip, the communication between these processor core can be efficiently achieved by using Network on Chip (NoC) approach. NoC provides the physical interconnection to transfer data between processing cores, it acts as communication background, the increasing quantity of transistors in emerging chip technologies, which follows the Moore's regulation [1], lets in to benefit from the concept of putting multiple cores (but less difficult ones) in a single chip. Many slower cores show a higher overall performance and power liaison than a monolithic single high-performance core. Although placing more cores in a sole die is feasible, however it needs to be finished through thinking of a new demanding situation like achieving consistency among all cores, reliable interconnection among the cores and to gain full benefits from all cores, in an effort to design an efficient architecture.

The Layout of Network on Chip involves selection of several main components consisting Routing technique, selecting suitable Topology and mapping community nodes as consistent with application. Performance of traditional single core systems is constrained by the clock frequency. Thermal troubles like heating and energy dissipation (the electricity wall hassels) prevent engineers from escalating the clock frequency. Though NoC barrows the ideas from matured area of computer networks, it cannot just use as it is for the IC technology because the layers in NoC are more closely bound than micro network [3], it incurs some layout issues as the processor cores are divers in nature, proper communication between cores requires careful synchronization and selection of optimal network topology between the cores. The design issues of NoC can be categorized into different level

i. System level :Design methodology, Architecture selection, traffic characterization
ii. Network adapter :Service management, Encapsulation
iii. Network level: Topology selection, Routing technique, Flow control and Error detection
iv. Link level: Synchronization, Encoding and Reliability.

Router plays an vital role in coordinating the data flow in network. Hence designing an efficient Router is main requirement of an on chip network based SoCs. Till now very few issues are addressed in NoC Router designing, a proper formulation of the main research themes is still missing. In this paper we discuss some major design issues and also propose an efficient Asynchronous router design. The paper is organized as follows: In Section II general description NoC router architecture. Section III enumerates design problem of NOC router. In section IV a efficient Asynchronous router design is proposed. Finally, in section IV conclusion remarks are presented.

II. ROUTER DESIGN

The Architecture of network on chip with detailed router diagram is shown in figure 1, it includes three main components Router(R), Network interface (NI) and Links, which are responsible for the communication between different Processing Elements (PE). The NI acts as a mediator linking the router and the nodes, it separates the computation part from the communication part. Router is in charge for transmitting the information from sender node to receiver node.

A classic router consist five input port and five output port, input port has buffers to store packets, routing controller and switch allocator, which select the path for packets through links. Link connects different nodes in the network according to the chosen topology.

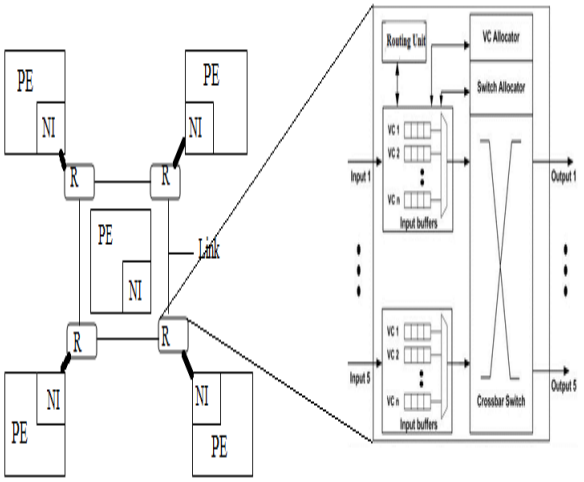

Fig 1: NoC Architecture with Router diagram

Revised Manuscript Received on April 16, 2020.
***** Correspondence Author
Trupti Patil , M.Tech in Embedded system from JNTU, Hyderabad
Dr Anuradha M Sandi, Associate Professor, GNDEC, Bidar,.

Router is very important part of the Network on chip which facilitates the interaction between the processing elements. The design of a router involves determining the number of virtual channels, buffer organization, routing technology used, switching techniques, topology selection ande. Routing controller chooses the proper route according to stored Routing table, switch allocator creates path between input and output ports by set/resetting crossbar switch. A appropriate router needs to be carefully designed while sticking to the energy budget and clock frequency. The layout problems of NoC router wishes a cautious modeling as they have got extreme impact in phrase of vicinity, energy and performance

III. DESIGN PROBLEMS

A.Network Architecture.

The conventional method of designing an entirely synchronous chip with a single global clock is not profitable anymore because of smaller process geometries, larger wire delays, better degree of integration of multiple cores on large dies. The massive attempts are required for skew control. The noteworthy power consumption of the global clock calls for specific clocking techniques. Indeed, in addition to various frequencies, different cores could have their own most favorable supply voltage to allow better energy consumption and performance management. Strategies which includes asynchronous or mesochronous clocking are the better options that hold the promise of simplifying timing closure and global clock distribution.

The Network on Chip architecture could be either Synchronous or Asynchronous. In Synchronous NoC architecture the Router is driven by global clock hence it consumes more power and it makes circuit more prone to clock skew problem. Synchronous designs are area efficient and fast , but implementation of high frequencies will be challenging [2] and they suffers from EMI (Electromagnetic interference) [3]. In an Asynchronous NoC design, the Routers are self timed and they are liberal to process variations, the whole chip in Asynchronous design can be divided into many isolated clock domains, it reduces the overall design time but the design will be comparatively slow. Asynchronous designs are most preferable choice in real time application where the transmission of small data packets within limited power requirement. Synchronous designs are good choices where large data packets and continuous transmission are involved for example multimedia applications.[4].

To utilize the advantage of both Synchronous and Asynchronous design technique the researchers have came up with technique called GALS (Globally Asynchronous and Locally Synchronous), in this technique the whole system is divided into smaller synchronous regions which removes the need for global clock.

So far now many Asynchronous NoC designs are presented by researchers such as Bundled data logic which provides high throughput with simple hardware requirements yet they will be very sensitive timing variations, QDI (Quasi delay insensitive) technique where the application performs correctly, independent of delays. The QDI techniques implemented many different encoding techniques such as LETS (Level Encoded Transition signaling) and LEDR (Level Encoded Dual Rail encoding)[5], the first one uses four phase encoding, for this throughput is low as the communication step are more, whereas in LEDR technique two phase encoding is used which reduces the communication steps hence increases the Throughput but they can applied only small applications (Eg: Links) [5] as handling the two phase information is difficult task.

B. The Topology Synthesis Problem

In on chip networks, physical arrangement of the nodes and connection between the nodes are defined by the network Topology. The path followed by the data packet is resolute by the underlying topology, the number of routers through which the packet passes and length of the link depends on the chosen topology hence the network topology influences the latency of the network and as packet traverse through different routers, each router consumes some amount of energy hence the topology directly affects the network energy consumption. Therefore selection of proper topology is a matter of importance.

Typically used topologies in on chip networks are Mesh, Tree, Ring, Polygon, Torus and double chained topology. Some researchers have also proposed Application precise topologies for the better performance [27]. Study shows that custom topology generation techniques do better than usual techniques in terms of both area and delay. Network partitioning generation techniques brings about topologies with lower area. Long-range insertion, which provides long-range links to standard mesh topology, is determined to be the most delay-efficient approach. Researchers have made efforts to improve energy and power with special topology called Z mesh topology [26].As topology plays a vital function in network overall performance, its miles the obligation of designer to pick out topology that is compatible for their applications.

C.The Buffer Sizing Problem

In NoC the buffers at input channel of each router have significant effect on the overall area. The buffer size is directly proportional to area of implementation of Router and it is inversely proportional to the network latency. Thus, the overall use of buffering asset must be minimized to decrease the realization overhead. But in real time programs with heterogeneous traffic styles which needs for more greater buffer area, one viable solution can be reconfiguration methods, to dynamically trade the buffer size to each channel, primarily based at the need of the application and thus energy efficiency of the system is increased increasing the power efficiency of the system to assign more buffer space to heavily loaded channels [28]. Hence it is the responsibility of the designer to select proper buffer size.

D.Routing Protocol Selection

The selection of routing technique depends on two main things the complexity of the implementation and the network performance. Latency, Throughput, Power dissipation, QoS and Reliability of the on chip network depends on the selected topology. The routing protocol dictates the real path for the data packet and it affects all network parameters, therefore selecting routing protocol suitable for the application is very important task for the designer.

Many different routing techniques are designed by the researchers to improve the performance of the network such as Deterministic routing where the path through which data traverse determined before sending data whereas in Adaptive routing multiple paths are available for data hence it provides better throughput but this technique need more resources. The routing techniques such as Turn- around routing and Node Labeling Technique , provides solutions to Live lock or dead lock problems[12]. Some advanced routing techniques are also available for example congestion Look ahead routing technique[13], in which the For instance, in deflective routing packets are sent to one of the free output channels belonging to a available shortest path; if this isn't possible, then packets are misrouted. Many variation are proposed to the conventional routing to get the information regarding congestion or block at any router that helps in avoiding such routes, one such algorithm is "XYX" deadlock free algorithm, it insures fault tolerance by sending a replica of every packet via all the available paths and then adopting error detection approach at destinations. "NARCO", that is based on OE turn model combined with packet duplication to make sure fault tolerance, and uses two VCs to avoid deadlock and with a view to identify faulty routes. "PR-WF" is modified version of turn model ensure fault tolerance and deadlock freedom, in this method 3 level buffer structure is adopted inside each node: cache in IP core, pseudo receiving buffer in NIs and a buffer in router port, which helps to avoid prohibited turns caused by west-first algorithm and hence increases the number of links in the network.

For 3D topology, many routing algorithm such as "LA-XYZ", "TTABR" and "DyXYZ" are proposed as a low latency, high throughput routing algorithms which not only improves the performance but also solves the deadlock, livelock and congestion issues.

In the case of irregular and/or reconfigurable topologies, many routing strategies are proposed one such is a "Reconfigurable routing algorithm (RRA)", based on a 4-bit register to solve any faulty router CDG (Channel Dependency Graph) to make design deadlock free.

For the topologies which are not necessarily rectangular Algorithm like "MPA" is based on turn model which makes design deadlock free that is able to support faulty regions.

Another more simple protocol is shown in called "TRAIN" where mesh topology is transformed into a tree to take benefit of resulting shortcuts. The algorithm guarantees fault tolerance by isolating faulty routers and links when detected, forming a sub network of non-faulty components. It also uses virtual cut-through switching technique to avoid deadlock.

Literature is rife with many efficient routing algorithm, designer has to choose proper routing technique according to applications demand.

E. The Switching Problem

Once the routing decisions are made, how the messages are passed along the switches and how the switches inside the routers are set/reset are determined by switching techniques. The switching technique imposes some satisfying constraints by the application. Selection of suitable switching technique trades-off between area and performance

Many switching techniques are available for on chip networks, one such design is Circuit switching [15, 16] in which source and destination are connected directly, it is static in nature and hence design is complex but it provides guaranteed services. Wormhole switching [18,19], is a most suitable technique for on chip network which demands higher Latency with small buffer demand. Some real time application demands both guaranteed services and higher network performance in such cases the designer has to choose a hybrid switching technique [20]

F. Power and Thermal Management

In recent year network on chip designed power efficiently, as applications demands longer battery life, hence power modeling of on chip network becomes as important as the performance of network. Thermal evils like heating and energy dissipation (the power wall problem) checks engineers from raising the clock frequency. In order to reduce the on chip power consumption the router and links has to be designed efficiently by selecting efficient signaling techniques and best utilization of available channel width. Low power SoC demands the research in power modeling of on chip network.

IV. PROPOSED ROUTER DESIGN

In this paper we have proposed Asynchronous Router design for Embedded applications. Asynchronous routers are more suitable for real time embedded applications [4] which provides high throughput but they have design complexity, in order to fully utilize advantage of asynchronous architecture we have to work on simplification of design. In Asynchronous design using four phase encoding throughput will be reduced due to communication complexity and if we go for two phase encoding which includes two types of phase information i.e., Even and Odd. In LEDR the encoding technique throughput will be high but handling two phase information is difficult which makes the design of router complex. Thus, the router can be implemented without considering the phase information, which reduces the number register and communication steps are reduced to half.

General configuration of the planned router is shown in fig 2, which consists of five input units and five output units. The input unit includes encoder units and routing controller. The output unit includes arbiter and decoder . Each input unit is connected to other four output units except its corresponding output unit. This router has five input and output ports. Flits are transmitted from one port to one of other ports.

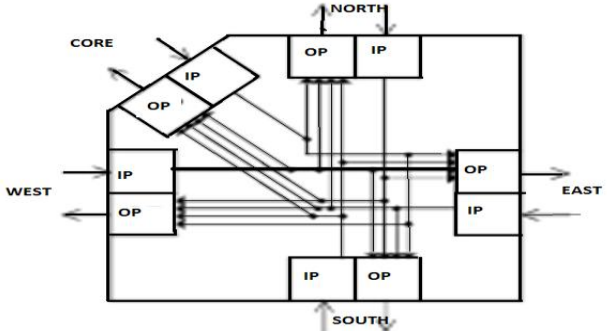

Fig 2: LEDR based Asynchronous NoC Router

Figure 3 shows the architecture of the single node. This node is designed for two dimensional networks and would be connected to 4x4 node on SoC. The node has inbuilt priority and arbiter module that is used for internal port allocation. Local port input and local port output is used to connect the node with the local processor or controller.

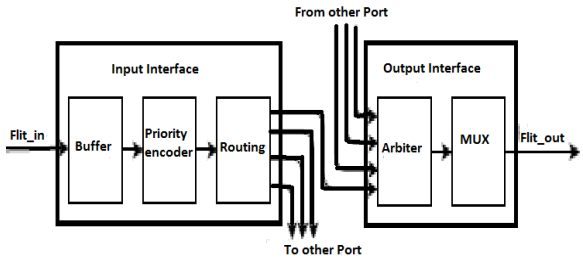

Fig 3: Block diagram for Single node

Whereas directional (north,east west and south) ports are used to connect to other nodes in the network. In this design we employed XY routing algorithm comes under distributed deterministic routing algorithm. It never runs into deadlock or live lock condition and follows the shortest path. This algorithm is suitable for regular and irregular network topologies.

At its most basic, an arbiter is a device that takes as input N requests, and outputs a single grant, in the form of a one-hot. A one-hot is a group of bits of arbitrary size consisting of all zeros except for one; i.e., one bit is logic high, or "hot." In this way, the arbiter looks at its set of inputs and allows a single device access to the resource. Arbiter in this design works on the rule that the lowest priority is assigned to the request which was just served during the next round of arbitration. Arbiter controls the arbitration of the ports and resolves conflict difficulty. It maintains the status reports of all the ports and is aware of which port is communicating and which one is free. Round robin scheduling will be used in case of Packets with the same priority and for the same output port are arrived.

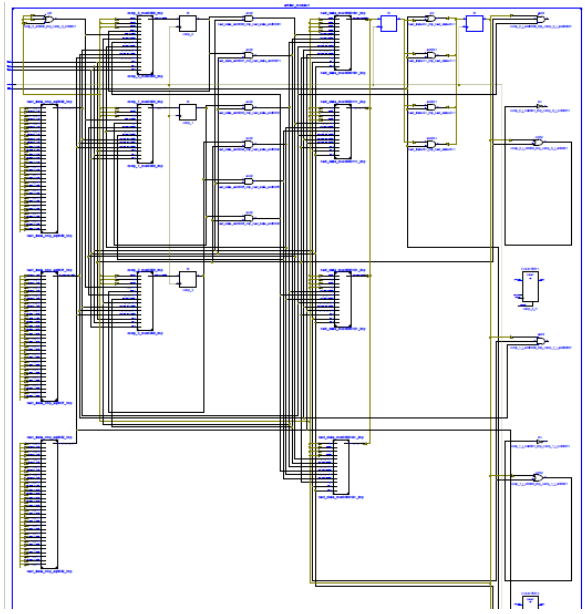

Fig 4: RTL schematic for Arbiter

Table I: Device utilization summery for arbiter module

Device Utilization Summary (estimated values)			
Logic Utilization	Used	Available	Utilization
Number of Slices	180	14752	1%
Number of Slice Flip Flops	110	29504	0%
Number of 4 input LUTs	254	29504	0%
Number of bonded IOBs	8	250	3%
Number of GCLKs	1	24	4%

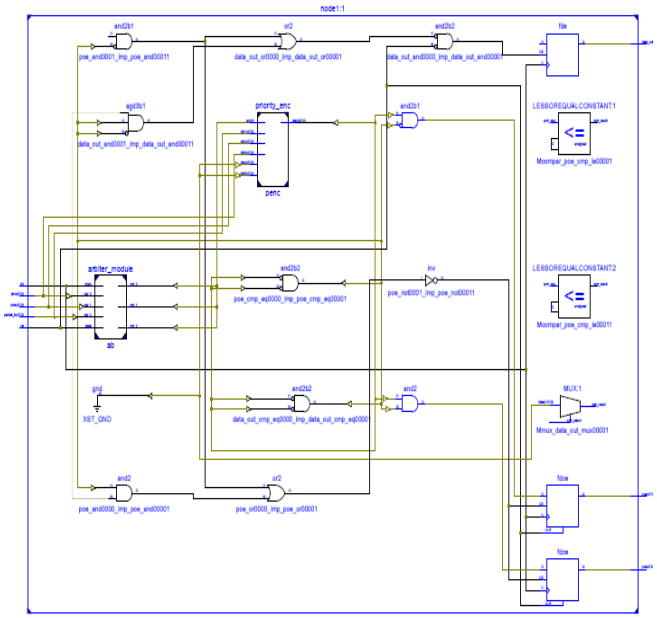

Fig 5:Single node RTL schematic

Figure 6 shows the simulation result for single node working. Input is feed on packet_in. as the destination port is selected as 00 which means that the data coming on any port belongs to this destination. Hence output is appearing on data_out. sel is used to determine which port input should be passed on the output port hence when we change the sel corresponding input port data from in4 is passed on to output port.

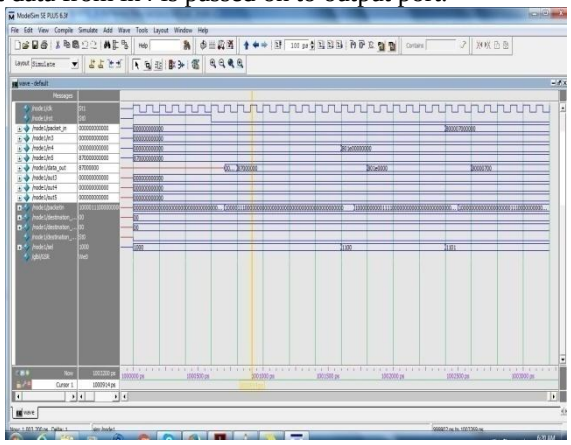

Fig 6: Simulation results

Table II: Design Summary (Estimated values)

Device Utilization Summary			
Slice Logic Utilization	**Used**	**Available**	**Utilization**
Number of Slice Registers	18	207,360	1%
Number used as Flip Flops	18		
Number of Slice LUTs	29	207,360	1%
Number used as logic	29	207,360	1%
Number using O6 output only	29		
Number of occupied Slices	28	51,840	1%
Number of LUT Flip Flop pairs used	47		
Number with an unused Flip Flop	29	47	61%
Number with an unused LUT	18	47	38%
Number of fully used LUT-FF pairs	0	47	0%
Number of unique control sets	2		
Number of slice register sites lost to control set restrictions	6	207,360	1%
Number of bonded IOBs	57	960	5%
IOB Flip Flops	13		
Number of BUFG/BUFGCTRLs	1	32	3%
Number used as BUFGs	1		
Average Fanout of Non-Clock Nets	2.25		

V. CONCLUSION

Investigation is done on various methods shows the shortfalls of the exclusive strategies and the present design .It also urges us to develop something new to overcome the discussed issues. The new design in contrast to the above will be more helpful as the new architecture for network-on-chip (NoC) router using asynchronous delay-insensitive data-transmission method which is more efficient and has lot of advantages when implemented on FPGA.

REFERENCES

1. Gordon E. Moore, "Cramming more components onto integrated circuits", Electronics,1965
2. W. 1. Dally and B. Towles, “Route Packets, Not Wires: On-Chip Interconnection Networks”, Proceedings of the 38th Design Automation Conference, p.684 (2001).
3. A. Jantsch and H. Tenhunen, “Network on Chips”, Kluwer Academic Publishers, Boston, 2003.
4. S. Kumar, A. Jantsch, J-P. Soininen, M. Forsell, M. Millberg, J. Oberg, K. Tiensyrja, and A. Hemani, “A network on chip architecture and design methodology”, IEEE Computer, pp. 117-124, 2002.
5. Masashi Ima, Thiem Van Chu , Kenji Kise, Tomohiro Yoneda I “The Synchronous vs. Asynchronous NoC Routers: An Apple-to-Apple Comparison between Synchronous and Transition Signaling Asynchronous Designs”
6. Naoya Onizawa, Atsushi Matsumoto, Tomoyoshi Funazaki, and Takahiro Hanyu, “High-Throughput Compact Delay-Insensitive Asynchronous NoC Router” IEEE Transactions On Computers, Vol. 63, No. 3, March 2014
7. A. Lines, “Asynchronous Interconnect for Synchronous SoC Design,” IEEE Micro, vol. 24, no. 1, pp. 32-41, Jan./Feb. 2004.
8. J. Sparsø and S. Furber, “Principles of Asynchronous Circuit Design: A Systems Perspective”. http://books.google.com/, Jan. 2001.
9. M. Dean, T. Williams, and D. Dill, “Efficient Self-Timed Level-Encoded 2-Phase Dual-Rail (LEDR),” Proc. Univ. of California/Santa Cruz Conf. Advanced Research in VLSI , pp. 55-70, 1991.
10. P. McGee, M. Agyekum, M. Mohamed, and S. Nowick, “A Level-Encoded Transition Signaling Protocol for High-Throughput Asynchronous Global Communication,” Proc. IEEE 14th Int’lSymp. Asynchronous Circuits and Systems, pp. 116-127, Apr. 2008.
11. J. Hu, R. Marculescu. “Application-specific buffer space allocation for Networks-on-Chip router design.” In Proc. ICCAD, 2004
12. G. Varatkar, R. Marculescu. “On-chip traffic modeling and synthesis for MPEG-2 video applications”, IEEE Trans. on VLSI, 12(1), 2004.
13. J. Duato, et. al. Interconnection Networks: An Engineering Approach. Morgan Kaufmann, 2002.
14. C. J. Glass, L. M. Ni. “The turn model for adaptive routing”, Proc. ISCA, May 1992.
15. J. Hu, R. Marculescu. “DyAD-Smart routing for Networks-on-Chip”. Proc. DAC, June 2004.
16. L. Shang et. al. “PowerHerd: Dynamically satisfying peak power constraints in interconnection networks”, Proc. Intl. Symp. On Supercomputing, June 2003.
17. P. Pratim Pande, C. Grecu, M. Jones, A. Ivanov, and R. Saleh, “Performance evaluation and design trade-offs for network-on-chip interconnect architectures”, IEEE Transactions on Computers, vol. 54, no. 8, pp. 1025-10 40, 2005
18. D. Kim, Manho Kim, and G.E. Sobelman, “CDMA-based NoC architecture”, Proc. IEEE Conference on Circuits and Systems, vol. 1, pp. 137-140, 2004.
19. E. Bolotin, et. al. “QoS architecture and design process for Networks-on- Chip” Journal of Systems Arch., vol. 50, 2004.
20. T. Bjerregaard, J. Sparso. “A Router architecture for connection oriented service guarantees in the MANGO clockless Network-on-Chip” Proc. DATE, 2005.
21. J. Dielissen, “Concepts and implementation of the Philips Network-on-Chip. IP-based SoC Design”, Nov. 2003.
22. M. Millberg, “Guaranteed bandwidth using looped containers in temporally disjoint networks within the Nostrum network on chip”, Proc. DATE, 2004
23. K. Lee. “ A 51mW 1.6GHz on-chip network for low-power heterogeneous SoC platform”. Intl. Solid-State Circuits Conf., 2004.
24. J. Liang, “An architecture and compiler for scalable on-chip communication”, IEEE Trans. on VLSI Systems, 12(7), July 2004.
25. N.Prasad Priyajit Mukherjee Santanu Chattopadhyay IndrajitChakrabarti “Design and evaluation of ZMesh topology for on-chip interconnection networks”, Journal of Parallel and Distributed Computing Volume 113, March 2018, Pages 17-36

AUTHORS PROFILE

Trupti Patil completed her BE from VTU,Belagavi in the year 2011 and M.Tech in Embedded system from JNTU, Hyderabad in 2013. Currently doing research design from VTU (Belagavi) India on Efficient NoC Router designs. Published 3 Papers in International and National Journals, 2 papers in International and National Conferences and published a book titled “Microcontroller and Microprocessor”. Her Main research interests and activities include design of asynchronous networks-on-chip.

Dr Anuradha M Sandi completed her BE, Mtech and PhD degrees from Gulbarga university, India in the the year 1996,2005 and 2014, respectively. She is currently working as a Associate professor in GNDEC, Bidar,. She has more than 21 year s of experience in teaching field. Published more than 9 Papers in International and National Journals **and** 10 papers in International and National Conferences from Research. Her main research interests and activities include CMOS VLSI Technology, Embedded System Design and network security . She is a member of the **LMIETE (** Life Member of Institute of Electronics and Telecommunication Engineering) **LMISTE** (Life Member of Indian Society for Technical Education).