JOURNAL OF CRITICAL REVIEWS

ISSN- 2394-5125 VOL 07, ISSUE 09, 2020

Multi constraint shortest path optimization using metaheuristic algorithms in asynchronous network on chip

Nidhi Mehra,
Asst. Professor, SOC (School of computing),
GEHU-Dehradun Campus
DOI: 10.48047/jcr.07.09.581

Abstract: Very-Large-Scale Integration (VLSI) design makes use of Network-On-Chip (NOC) because it provides an appropriate connectivity topology and a practical alternative to system-on-chip. The lack of a clock in an asynchronous Network-On-Chip (ANOC) layout results in reduced power consumption. However, the computational complexity of attaining optimum path routeing in an ANOC is high. The total cost and the performance of an interconnection network are two important design concerns. The network's performance is determined by the topology configuration chosen and used by the routeing algorithm. The topology, switching process, and routeing algorithm are the primary determinants in ANOC design. In this study, we use an asynchronous mesh topology based on Hopfield Neural Networks (HNNs). The shortest path issue has been combined with a number of algorithms in an effort to improve routeing efficiency.

Keywords: Hopfield Neural Network (HNN), asynchronous Network-On-Chip (ANOC), Very-Large-Scale Integration (VLSI)

Introduction

In a system-on-chip (SOC), many intellectual property cores are combined and orchestrated on a single chip substrate. It's capable of doing a wide variety of tasks, from analogue to digital to mixed circuit to radio frequency. A system-on-chip (SOC) combines hardware and software. Although SOCs have numerous useful uses, their restricted performance is a result of their dependence on global interconnectivity. Both platform-based SOC design and fixed-function SOC design, also known as application-specific SOC design, fall under the category of embedded and energy-constrained SOCs. Reduced energy consumption while meeting performance expectations is a goal that may be achieved by optimising both fixed-function and platform-based SOC. Information regarding the traffic, such as the bandwidth between individual cores, provides an opportunity for optimisation. One cutting-edge design that may exist inside a system-on-a-chip is the network-on-a-chip (NOC). In order to make use of and benefit from the inefficient shared bus design of the SOC chip, a NOC is now used. The SOC is superseded by the NOC due to the former's superior design and operation. The NOC idea is a revolutionary improvement over traditional on-chip layouts; its efficient connectivity architecture solves the worldwide cable latency issue. Several approaches have been proposed to improve NOC's efficiency and performance. NOC technology not only makes use of networking theory, but also helps with the growing issue of SOC scalability. NOC has more power efficiency than SOC. Because of its improved throughput, superior scalability, and effective reusability in multi-core systems, NOC has emerged as a viable option for

effective and efficient on-chip communication. Typical on-chip network design is seen in Figure 1.

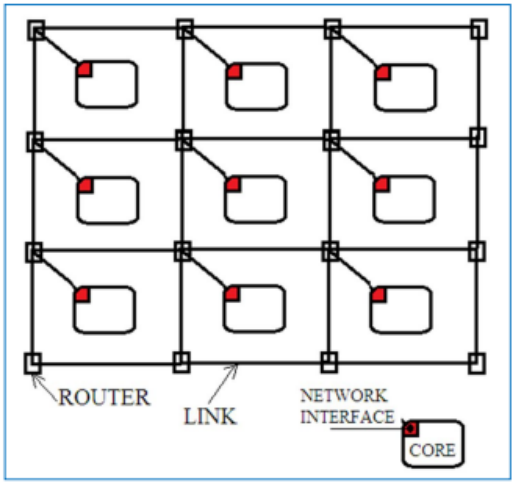

Figure 1. Typical on-chip network architecture

Connected appropriately, nodes, connections, and the network interface make up an on-chip network. Each node has both a processor and a router. The connections provide the actual means of communication between the network's nodes. A router may act as a smart buffer and execute routeing operations. Since each IP core may use a unique interface protocol, a network interface block is required to create the logical link between the IP cores and the network. The topology of the on-chip network is specified by a graph representing the design of the network. It is the topology of a network that determines how its nodes and connections are connected and which pathways a message may take to reach its destination. In order for a message to go from its origin to its final resting place, a routeing algorithm must decide which way to take based on a predetermined set of criteria.

Mesh, rings, and toroids are the most often investigated topologies. Since it is straightforward and straightforward to implement in real time applications, mesh topology is often prefered. The total system cost and the required performance of an interconnection network are two important design problems. In addition, a network's performance is affected by its topology and the efficiency of its routeing algorithm.

The task of synchronising the multiple digital parts of a synchronous digital chip is made possible by a clock signal. There is no information carried by a clock signal, and it does not do any meaningful work, yet it constantly requires a large amount of power, ideally between 40 and 70 percent. Power dissipation is an issue because of excessive energy use. Complex clock tree designs are required for synchronous design, which in turn requires a lot of space and a lot of power. The complexity of the clock distribution issue grows as the network grows in size. It's possible for clock power consumption to equal or exceed logic power use. As a result, the additional complexity to the design and the resulting increased power consumption is an inevitable consequence of such clock distributions. Figure 2 depicts a common clock tree arrangement.

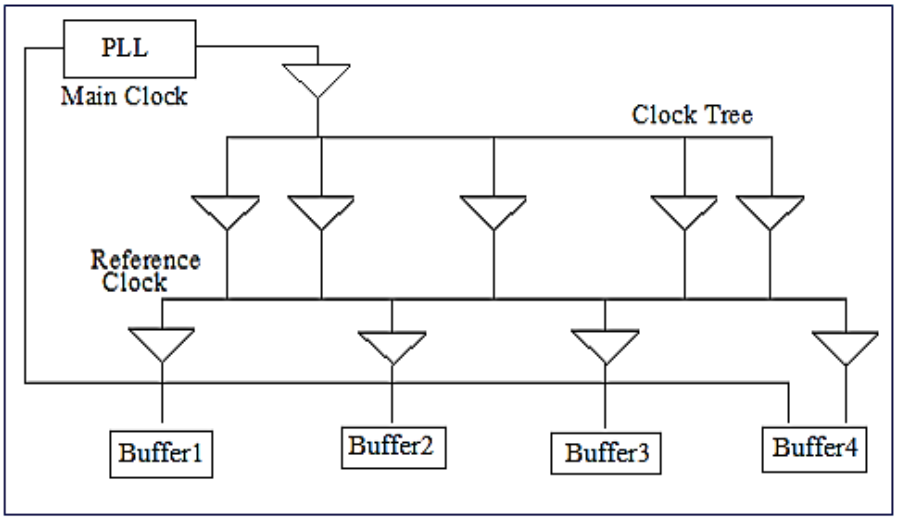

Figure 2 Clock tree structure

Asynchronous Network-On-Chip (ANOC)

The lack of the clock in an asynchronous architecture reduces power usage significantly in comparison to a synchronous system. In addition, there are no interruptions in the flow of data while using an asynchronous architecture because of the use of handshake protocols. It provides features like shutting down unused components and adjusting to new settings automatically.

The asynchronous connections and the resistance to voltage and temperature swings are two advantages of an asynchronous architecture. It's resilient to problems caused by shifting manufacturing methods. Asynchronous interconnects are either purpose-built for a particular application or created with power-constrained SOCs and sparsely provided resources in mind.

Commonly utilised in 4G telephony evolution, a Flexible Architecture of Unified System for Telecom (FAUST) chip is essentially a platform-based SOC that employs an asynchronous mesh-based NOC. Guaranteed Services (GS) across OCP interfaces are made available by the ANOC that relays messages. FAUST's successor, the MAGALI chip, provides an open platform for the construction of several modes of the LTE standards and makes use of a system-on-chip (SOC) to facilitate mode reconfiguration and the data link between the heterogeneous blocks. In a traditional synchronous NOC with a mesh topology, moving the data storage to the links greatly decreases the network's power consumption at the expense of some performance. In order to increase the NOC's reliability while keeping the throughput constant, error-detection link pipeline circuits may be implemented.

Using the connections as distributed FIFO buffers, we may implement elastic buffers that are functionally equivalent to asynchronous buffers and so reduce the complexity of the routers. These buffers analyse the effects of bulky communications and many energy-efficient suggestions, such as the pipelined connections used by several sources. The buffers use SOC and mesh topology for evaluation but do not improve performance or take precedence over

the amount of connections. The two main models for classifying asynchronous circuits are the delay-insensitive model and the limited delay model. The bounded delay model takes into account the cumulative effect of gate and wiring delays. To achieve proper circuit performance, this type requires substantial timing. The functioning of a delay in sensitive (DI) circuit is unaffected by delays in the gates and wires since it is an asynchronous circuit. The 2-mode protocol only requires one communication round trip per financial transaction.

Toggle REQ to start the transaction, then toggle ACK to confirm data reception and finish the exchange. It is not required that the signals reset to zero after each transaction. Distributed routers in each subsystem are linked to their neighbours in all four cardinal directions through global asynchronous links. The switch used here has input and output ports for each of the four routers. The router's local port is earmarked for use with converters to link to the IP cores of other devices. Asynchronous Synchronous First In First Out (AS-FIFO) and Synchronous Asynchronous First In First Out (SA-FIFO) interfacing technology is used to transmit data between the intellectual property cores and the local input ports.

Literature Review

Yong Chen et.al.,(2020) Under both simulated and real-world application traffic, the suggested solution is compared to both centralised and distributed best practises. The testing findings show a tenfold increase in success rate compared to centralised software solutions and a 5-10% increase compared to centralised hardware solutions, all while showing a two-orders-of-magnitude improvement in allocation performance. In comparison to the most current distributed solution proposal, it is up to 8 times faster at allocating resources and has a 29 percent greater success rate.

M. Thilagavathi et.al.,(2019) Because of its great performance and intelligent structure, network-on-chip has been the subject of much discussion. The Destination Tag is the name given to the type of on-chip communication used for high-throughput applications. Noc has lately included the destination Tag method. This paper's objective is to detail the procedure known as the particle swarm optimisation method, which is used to locate the best possible answer for shortest pathways. PSO is a solution we've come up with. The goal is to build and optimise the multiprocessor NOC traffic regulation unit's algorithms. The efficiency and effectiveness of the network may be improved. The routeing algorithm PSO is optimised to perform as well as possible. Particle swarm optimisation methods integrated into a multi-NOC architecture for use across several networks; used to keep power gating and packet scheduling from becoming unmanageable bottlenecks. We zeroed down on a low-power, scalable design using PSO. This technique improves performance over a wide range of network loads. It stores the surplus energy that the grid generates over time. Using Xilinx, the simulation results for the proposed router were shown experimentally, and the code was written in VHDL.

Yao Hu et.al.,(2019) In this research, we compare the effects of application mapping on performance across random and non-random network architectures. We provide a number of application mapping strategies and evaluate their efficiency in work scheduling under the

assumption that the computer system cannot forecast the communication patterns. The evaluations with an extensive compound application workload demonstrate that the average turnaround time may be reduced by up to 39.3 percent using a random connected mapping approach and up to 72.11 percent using a diameter/ASPL-based mapping method when using random topologies instead of non-random ones.

Methodology

The complexity of the search space means that classical algorithms such as Dijikstra's, Floyd-Warshall's, and Bellman-Ford's cannot guarantee solutions in polynomial time. They need a lot of processing power and can't be scaled to a bigger network. The Hopfield neural network is used to discover the best pathways in an asynchronous Mesh topology by utilising three optimisation techniques: PSO, FF, and ACO.

Asynchronous Mesh Topology

The Mesh topology, together with wormholes and packet switching, has been used to construct several different kinds of NOC prototypes. Therefore, the Mesh topology is chosen for this study. Every node in a Mesh topology has a specific connection point from one node in the network to another, meaning that each link only serves the two nodes it links. HNN with asynchronous Mesh topology is used because it provides the robust approach necessary for the ANOC system to work efficiently**.**

The parameters used for Mesh topology in this study are given in Table 1.

Parameter Description	Assigned Value
Start ID for node in topology	1
Finish ID for node in topology	16
Number of nodes in the topology	16
Maximum allowable links	30
Number of links	16

Table 1 Parameters of Mesh topology

Hopfield Neural Network Algorithm

In this study, we combine the Hopfield network with a Mesh topology to find the shortest route. After proper asynchronous updating, the network goes through a certain number of iterations. The neurones in the network are then retrieved to determine whether or not the pattern is a good match for the network. HNN need a weight matrix to keep track of the patterns. Therefore, these recurring elements must be included in the supplied data. Content addressable memory is a weight matrix that stores patterns generated by the network's dynamics. The shortest path in the intercommunication pathways is then identified with the help of the optimisation method. The Mesh topology's linkages are given numerical values,

called "weights," through HNN. The connection distances in HNN are calculated using a custom algorithm built in MATLAB and then run. In order to address the routeing issue and extract important parameters, the HNN's output values are put in to the Mesh topology, which is then subjected to several optimisation methods.

The HNN structure that is generated using the above algorithm is shown in Figure 3

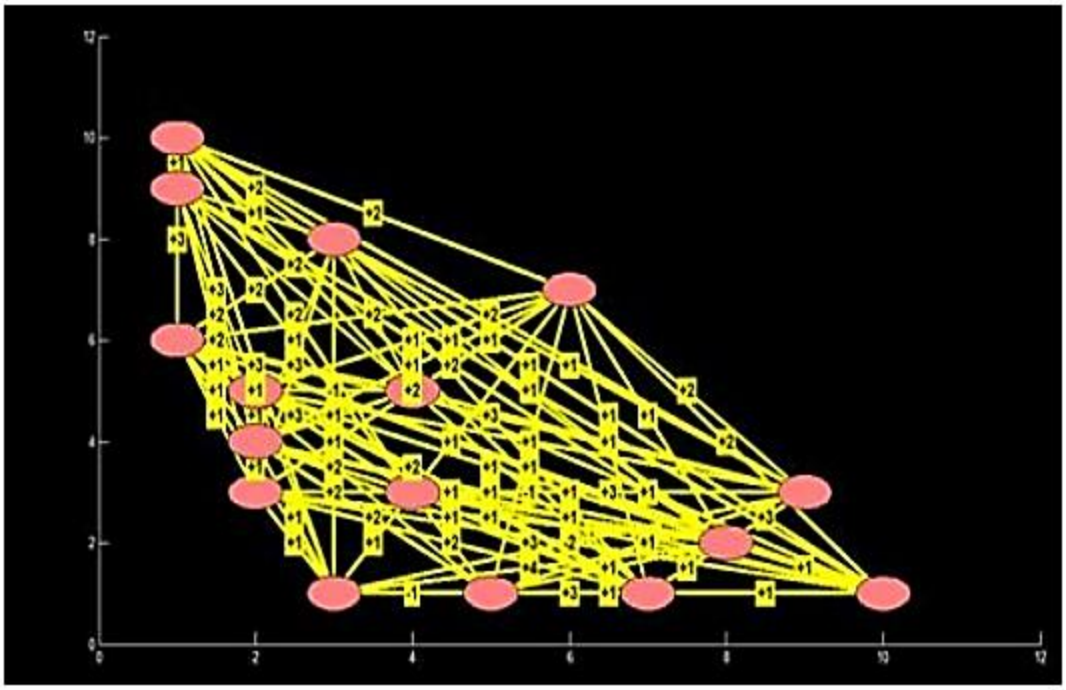

Figure 3. HNN structure employed in this work

HS optimisation, along with CS, ACO, FF, and PSO algorithms, are compared to one another in an attempt to determine the best method for picking the shortest route in the ANOC. Analysis by comparison is done in measures of performance such network topology, throughput, computation time, fitness function, path precision, path success rate, route failure rate, energy dissipation, energy residual, and average residual energy.

In order to establish the multi constraint optimum shortest route, the HNN Mesh topology has been chosen as the underlying data structure for all of the optimisation techniques utilised in this study. To find the shortest path, the HS method performs a global optimisation as a minimal optimisation problem. Next, by inverting the benchmark function, the answer is transformed into a maximisation issue.

Performance Analysis of PSO, FF, ACO, CS AND HS Optimization Algorithms

Twenty runs of one hundred iterations every have been performed for each optimisation strategy used in this study. The average values of the simulation's most crucial parameters are shown in Table 2. The best method for determining the optimum route is determined by comparing the values of the specified metrics acquired for various optimisation algorithms from the outcomes of the simulation.

Table 2. Values of metrics obtained from optimization algorithms

PARAMETERS	NODE	PSO	FF	ACO	CS	HS
Distance (mm)	16	23.00	20.00	17.50	15.36	12.82
Throughput (%)	16	22.00	31.00	48.00	83.12	92.62
Computation Time (Sec)	16	13.29	9.71	6.96	5.34	3.91
Fitness	16	29.02	19.18	14.09	5.22	1.36
Path Accuracy (%)	16	67.50	80.98	88.00	91.48	95.40
Success Rate (%)	16	12.00	13.00	29.00	42.00	77.00
Route Failure Rate (%)	16	42.30	21.50	16.00	6.74	4.35
Dissipated Energy (J)	16	0.17	0.12	0.08	0.06	0.05
Residual Energy (J)	16	4.00	3.33	2.86	2.50	2.22
Average Residual Energy (J)	16	3.19	2.14	1.54	1.22	1.20

Distances between nodes for determining multi-constraint shortest paths using the five techniques are compared in Figure 4. The latency between the nodes in ANOC decreases linearly with the number of iterations, however the FF method performs better as the number of iterations increases compared to PSO. The PSO algorithm has been shown to perform less reliably than competing methods. The ACO algorithm has lower inter-node latency than the FF does. Initial iteration latency is roughly 1.57 times faster than the FF. It has reliable results in CS algorithm tests. This might be because the delay does not decrease after several iterations. The selection of nodes and the latency calculation are done at the nearest neighbor nodes in ANOC.

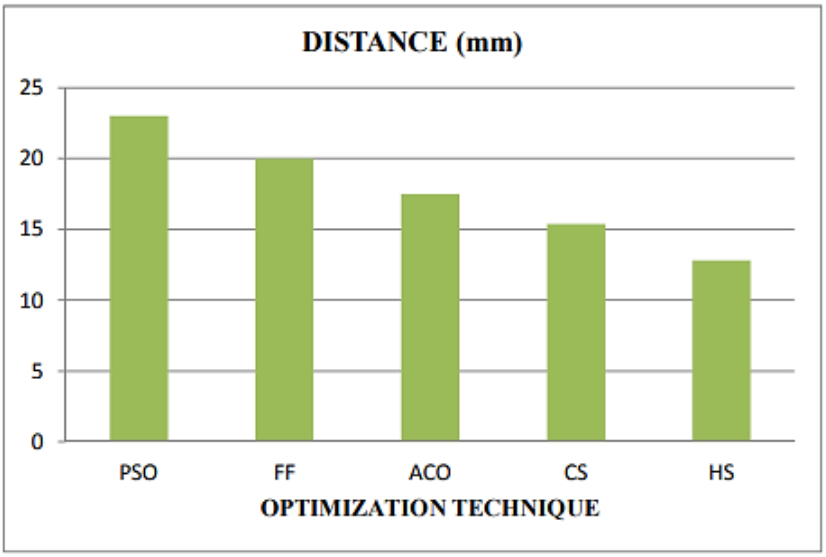

Figure 4. Distance between nodes

When comparing the CS to the PSO and the ACO, it is clear that the CS performs better throughout the iteration for node selection. However, HS algorithm has the advantage over all other approaches since it chooses the nodes from the closest neighbours.Among the five methods, the suggested HS optimisation method has the smallest distance, proving its viability.

NOC's efficiency in data transmission is quantified by its throughput. Throughput comparison between 3x3 and 4x4 mesh topologies for PSO, FF, ACO, CS, and HS is shown in Figure 5.to locate the most direct route. In the ANOC architecture, the suggested HS algorithm has better throughput than the other algorithms. The throughput follows a linear ramp slope as the number of nodes in the network grows, demonstrating the HS's effectiveness. As a result, effectiveness of the HS algorithm is justified in terms of throughput

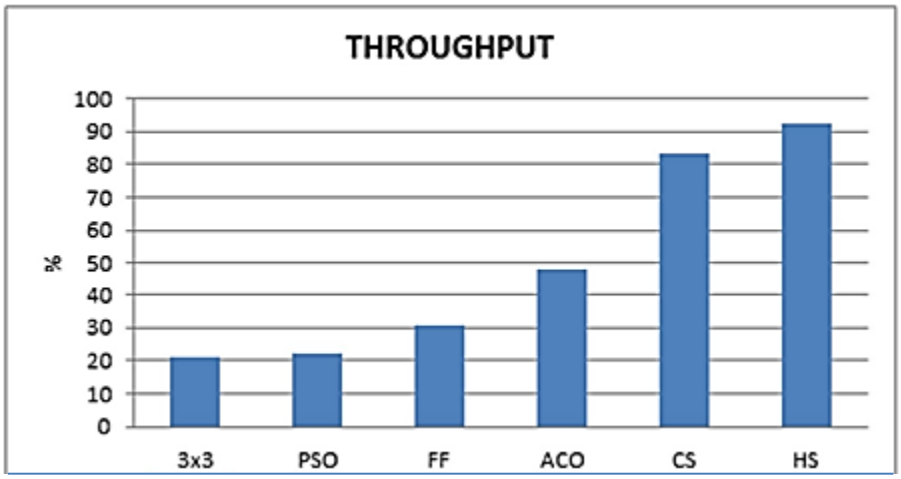

Figure 5. Throughput

Conclusion

In this study, we used Hopfield neural networks with Particle swarm optimisation, Firefly optimisation algorithms, Ant colony optimisation, CS optimisation, and HS optimisation for asynchronous network protocol design.Network-On-Chip uses various optimisation techniques to identify the multi-constraint shortest route. The aforementioned methods were evaluated using a number of criteria, and the optimal method for optimising the Asynchronous Network-On-Chip was selected as a result. Even though five algorithms were examined, HS and CS's performance was scrutinised since the other three algorithms fell short. The multi-constraint shortest route was found by the HS method in only 20 iterations, whereas its nearest competitor, CS, required 100 iterations. The data moved 15 percent farther than CS after 100 repetitions. In addition, the convergence rate in HS was much greater than in CS. The output of HS was higher than that of CS by a factor of 10%. HS's shortest route calculation was 25% faster than CS's. Similar fitness functions were observed between HS and CS.

References

1. Y. Chen, E. Matus, S. Moriam and G. P. Fettweis, "High Performance Dynamic Resource Allocation for Guaranteed Service in Network-on-Chips," in *IEEE Transactions on Emerging Topics in Computing*, vol. 8, no. 2, pp. 503-516, 1 April-June 2020, doi: 10.1109/TETC.2017.2765825.
2. M. Thilagavathi and J. S. Sadiq, "High Performance Energy Efficient Grid Based Routing Algorithm for Multi Network on Chip," *2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)*, Tamilnadu, India, 2019, pp. 1-3, doi: 10.1109/INCOS45849.2019.8951411.
3. Y. Hu and M. Koibuchi, "The Impact of Application Mapping on Non-Random and Random Network Topologies," *2019 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, Xiamen, China, 2019, pp. 16-25, doi: 10.1109/ISPA-BDCloud-SustainCom-SocialCom48970.2019.00015.
4. H. Nakahara, N. A. V. Doan, R. Yasudo and H. Amano, "XYZ-Randomization using TSVs for low-latency energy efficient 3D-NoCs," *2017 Eleventh IEEE/ACM International Symposium on Networks-on-Chip (NOCS)*, Seoul, Korea (South), 2017, pp. 1-8.
5. L. Wang, X. Wang and T. Mak, "Adaptive Routing Algorithms for Lifetime Reliability Optimization in Network-on-Chip," in *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2896-2902, 1 Sept. 2016, doi: 10.1109/TC.2015.2506571
6. N. Mohanapriya and C. Priya, "A fault tolerant router with optimized an colony algorithm," *2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, Ramanathapuram, India, 2016, pp. 167-170, doi: 10.1109/ICACCCT.2016.7831622.
7. M. R. Mallick, S. S. Pradhan, J. Padhi and S. Panigrahi, "A novel approach to maximize network life time by reducing power consumption level using CGNT model," *2016 International Conference on Circuit, Power and Computing Technologies (ICCPCT)*, Nagercoil, India, 2016, pp. 1-6, doi: 10.1109/ICCPCT.2016.7530247.
8. Kendaganna Swamy S, A. Jatti and Uma B V, "Performance enhancement and area optimization of 3×3 NoC using random arbiter," *2015 IEEE International Advance Computing Conference (IACC)*, Banglore, India, 2015, pp. 11-15, doi: 10.1109/IADCC.2015.7154672.
9. Z. Li, A. Qouneh, M. Joshi, W. Zhang, X. Fu and T. Li, "Aurora: A Cross-Layer Solution for Thermally Resilient Photonic Network-on-Chip," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 1, pp. 170-183, Jan. 2015, doi: 10.1109/TVLSI.2014.2300477.
10. Kendaganna Swamy S, A. Jatti and Uma B V, "Performance enhancement and area optimization of 3×3 NoC using random arbiter," *2015 IEEE International Advance Computing Conference (IACC)*, Banglore, India, 2015, pp. 11-15, doi: 10.1109/IADCC.2015.7154672.
11. E. Hansson and C. Kessler, "Global Optimization of Execution Mode Selection for the Reconfigurable PRAM-NUMA Multicore Architecture REPLICA," *2014 Second International Symposium on Computing and Networking*, Shizuoka, Japan, 2014, pp. 322-328, doi: 10.1109/CANDAR.2014.72.
12. E. Hansson and C. Kessler, "Global Optimization of Execution Mode Selection for the Reconfigurable PRAM-NUMA Multicore Architecture REPLICA," *2014 Second*

International Symposium on Computing and Networking, Shizuoka, Japan, 2014, pp. 322-328, doi: 10.1109/CANDAR.2014.72.

13. B. Betkaoui, Y. Wang, D. B. Thomas and W. Luk, "Parallel FPGA-based all pairs shortest paths for sparse networks: A human brain connectome case study," *22nd International Conference on Field Programmable Logic and Applications (FPL)*, Oslo, Norway, 2012, pp. 99-104, doi: 10.1109/FPL.2012.6339247.
14. Hailong Yao, Yici Cai, Xianlong Hong and Qiang Zhou, "Congestion and performance driven full-chip scalable routing framework," *2005 6th International Conference on ASIC*, Shanghai, China, 2005, pp. 856-859, doi: 10.1109/ICASIC.2005.1611440.
15. V. Seth, Min Zhao and Jiang Hu, "Exploiting level sensitive latches in wire pipelining," *IEEE/ACM International Conference on Computer Aided Design, 2004. ICCAD-2004.*, San Jose, CA, USA, 2004, pp. 283-290, doi: 10.1109/ICCAD.2004.1382587.