

Comparative Performance Evaluation of NoC-based Multicore Systems through Traffic Engineering

Savita Gautam *

University Women's Polytechnic, Aligarh Muslim University, Aligarh, 202002, India

E-mail: savvin2003@yahoo.co.in

ORCID iD: <https://orcid.org/0000-0003-3931-8166>

*Corresponding Author

M. Sarosh Umar

Department of Computer Engineering, Aligarh Muslim University, Aligarh, 202002, India

E-mail: saroshumar@zhcet.ac.in

ORCID iD: <https://orcid.org/0000-0002-4186-5938>

Abdus Samad

Department of Computer Engineering, Aligarh Muslim University, Aligarh, 202002, India

E-mail: abdussamad@zhcet.ac.in

ORCID iD: <https://orcid.org/0000-0003-4845-7184>

Abstract: Traffic patterns significantly effects the performance of networks-on-chip (NoC) architectures. The rapid increase and unpredictable behaviour of traffic may cause delay in high-speed packet transmission that ultimately increase the cost of the system. Choice of topology is also a measure concern to cope up the uneven traffic patterns. In this paper, an analytical and experimental evaluation of various NoC architecture is carried out in terms of their performance capabilities. To evaluate the performance of the considered architectures, a number of architectural characteristics such as network diameter, degree and cost are evaluated and simulation results are obtained under different traffic patterns. An organizational model is proposed while considering the problem of delay in traffic engineering using different NoC architectures. The BookSim simulator is chosen for evaluating parameters like network latency, throughput and execution time. This is carried out by implementing different interconnection networks under five routing evaluation traffic models with appropriate selection of NoC architectures. Effect of virtual channels (VC) is also assessed under same traffic pattern. Four regular topologies are used to carryout comparative studies namely standard 4 x 4 Mesh, Folded Torus, DIMB network and recently introduced Linearly Extensible Triangle Network (LECA). Based on the study carried out a high-level modeling of NoCs with appropriate topology is evaluated under certain network parameters used to evaluate the performance of NoC architectures. Research in this direction shows that the number of cores in NoC architectures with appropriate routing techniques effectively reduce the cost and complexity of the system without losing the performance of architecture.

Index Terms: Noc, BookSim Simulator, Traffic Pattern, Folded Torus

1. Introduction

Routing and Packet forwarding are the two fundamental tasks which are essentially carried out by a router. The main issue with NoC architectures is facing the delay while forwarding packets from source node to destination node. Internet and multimedia applications have caused network traffic to grow rapidly. On the other hand, there is a great emphasize how to cope with high-speed routers by overcoming increasing network delay. A typical router uses a routine table to initialize packet forwarding considering the next hop route. The choice of topology in NoC architectures has a significant impact in the design and implementation of best organisational model. As a result, it is equally important to assess the performance of specific routing algorithm on specific NoC architecture. Therefore, having appropriate scheduling mechanism and mapping it effectively on NoC systems is an integral process for a globally-optimized communication along with the design of architecture. An enormous amount of research has been carried out to design and evaluation of the performance of such networks [1], [2], [3].

Designing asynchronous architecture is more modular and do not suffer from issues incurred in synchronous architectural design. However, there is not much support from the Electronic Design Automation industry for asynchronous systems. Thus, a combined idea of synchronous and asynchronous designs is more popular in the NoC architectures [4]. Reuse of computing cores by incorporating modularity certainly enhances design productivity and produces better results in terms of network latency and throughput. Thus, it enables a higher level of abstraction during NoC design through the architectural modelling and simulation. A number of different NoC topologies have been proposed for effective design of NoC based systems, however, mesh topology is considered the most generalised one by designers due to its simplicity and symmetrical layout [5],

[6]. The main issue with the mesh topology is in terms of communication layout as it has longer diameter that increases the communication latency. Torus topology is another better alternative which has reduced latency as compared to mesh. In torus topology wrap-around connections enable the connections of switches on the edges to the switches on the reverse edges which help to provide enhanced communication through wrap-around channels. A 4 x 4 mesh and folded torus are demonstrated in Fig. 1(a) and Fig. 1(b).

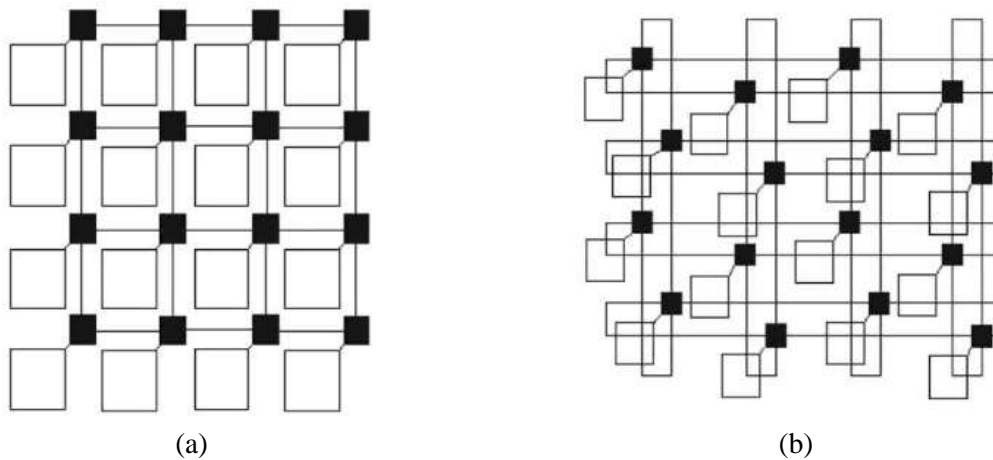


Figure 1. (a) 4 x 4 Mesh (b) 4 x 4 Folded Torus network

Wrap-around connections, however, may increase the routing delay in communication [7]. Network and traffic transmitters (router) need to react against congestions they can sense. This problem could be solved by folding the torus architecture and thus could be a better choice for NoC architectures [8].

The attractiveness of an NoC architecture lies in the effective utilization in the network communication and that how much is it scalable for larger applications. Therefore, testing the network for NoC communication is an important issue which is addressed in this work. We used a BookSim network simulator which is designed to target NoC architecture communication. Many topologies exists in the configuration files and hence Booksim is considered more flexible. Multiple router architectures with diverse routing algorithms are also implemented with synthetic traffic patterns can be injected into the network. There are choices of switching techniques, virtual channels and buffer parameters. It can further be enhanced with new application and routing techniques. We have evaluated performance of four similar NoC architectures with performance metrics latency and throughput for a given set of traffic choices. In particular, simulation results are obtained for 16-cores mesh, folded torus, DIMB and LECA networks with both bit permutation as well as with digit permutation.

The aim of this paper is to propose a performance model for effective design of NoC architectures under different traffic engineering. The paper consists of five sections. Section 2 describes the considered NoC architectures and their characteristics. Section 3 discusses setup used for simulation along with the simulation results. Section 4 discusses the results presented in section 3 and a comparative study is carried out to validate the proposed model. Section 5 concludes the paper.

2. Related Works

Topologies are evaluated for various performance metrics like latency, execution time and throughput. Each solution is generated with the help of a task graph. A number of regular and irregular network topologies has been studied and evaluated in the recent past. An irregular type of topology known as Undefined Topology Network on Chip (UTNoC) is proposed in which each router is connected to any other topology thus forming an interface in the system. However, each router can connect to just one processing element. Routing is carried out in terms of tables which are filled through a broadcast stage [9]. Da Silva et. al. proposed a similar approach to optimize the irregular topologies for real-time applications. However, the author considered the soft real time problem for optimization [10]. To decrease the average latency produced by the network and to enhance the throughput an optimization approach is required particularly for hard real-time applications. For mapping of the routers in to task graph, a heuristic algorithm is used. The algorithm utilizes an irregular topology in an efficient manner and make it capable to increase the number of tasks (packets) in order to meet the deadline without increasing the average latency of the NoC. The proposed work helps in the selection of best model in terms of architecture and routing mechanism that suits for such applications.

The selection of the right topology is another important task along with the appropriate routing algorithm. Both the entities are important for evaluating the performance of a NoC architecture. Some of the famous topologies for on-chip technologies are cube-based topologies, tree networks, fat-tree, ring, torus and 2D-mesh and a

number of variants of mesh and torus [11], [12], [13]. Mesh and torus are considered to be most commonly used architectures for such NoC performance investigations. Apart from the listed conventional architectures, there is a plethora of hybrid NoC architectures that include the characteristics of different class of networks. LECA and DIMB are two such networks reported recently. LECA is a hybrid architecture reported recently which conceives the desirable characteristics of cube-based architecture like small diameter and degree as well it is having simpler and symmetrical architectural layout in analogy to conventional 2D architectures [14]. A 4 x 4 LECA architecture is having a symmetrical structure and scales effectively for greater size of network. The skeleton of LECA network is shown in Figure 2 (a) in which $X_0, X_1, X_2, \dots, X_n$ represents the n-core architectures whereas the 4 x 4 system is derived by using 4 LECA networks and depicted in Fig. 2(b). The LECA network architecture is considered to evaluate and compare the performance with conventional NoC architectures.

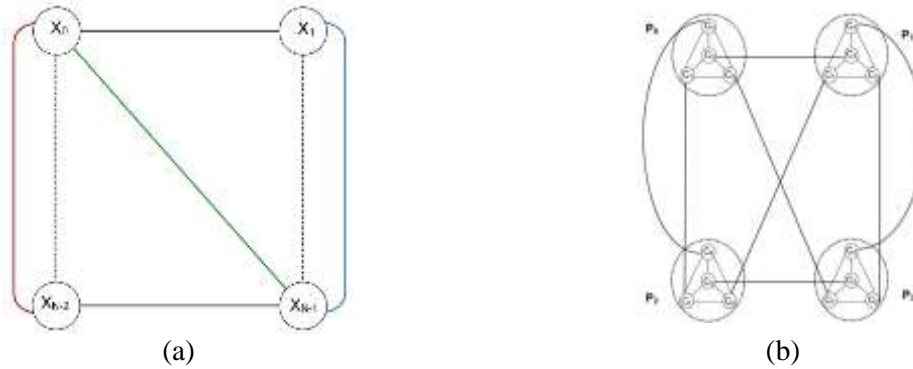


Figure 2. (a) Basic Layout of 4 x 4 LECA

(b) A 4 x 4 LECA on chip architecture

DIMB is another similar architecture which is designed to reduce hop count and hence network latency. The native topology to design DIMB is de Bruijn network which has constant node degree and hence known as de Bruijn inspired Mesh-based (DIMB) topology. The DIMB is similar to LECA network in the sense that it has low diameter and cost. DIMB is formed like an $n \times n$ Mesh topology in which each node has an address of (X, Y) which is inspired from a 2-D Mesh, in which X and Y represent the horizontal and vertical de Bruijns network [15]. To reduce the diameter and the average inter-node distance of the network the nodes at the edges of the DIMB are inter-connected with each others. This will further enhance the performance of an NoC that utilizes the concept of DIMB. We can add wraparound links to make a 2D torus-Bruijn. The layout of $n \times n$ DIMB architecture is demonstrated in Fig. 3.

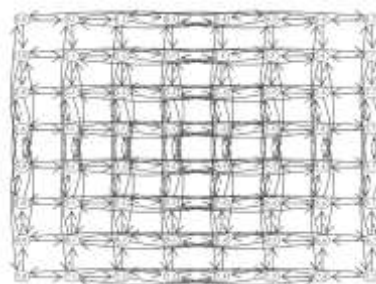


Figure 3. A 8 x8 DIMB network

The topological characteristics of these networks is demonstrated in Table 1. In the present work a simulation study is carried out for conventional mesh, folded torus, DIMB and LECA networks to use them in NoC architectures.

Table 1. Characteristics of NoC Architectures

| Network Property | Level | Number of Nodes | Diameter | Degree | Cost |
|------------------|-------|-----------------|----------|--------|--------------|
| Network | N | | | | |
| Mesh | N | N^2 | $2(N-1)$ | 4 | $4*(2(N-1))$ |

| | | | | | |
|---------------|---|--------|-------------------------|---|----------------------------|
| DIMB | N | N^2 | N | N | N^2 |
| FTorus | N | N^2 | N-1 | 4 | $4(N-1)$ |
| LECA | N | 2^4N | $\frac{[Core/16]}{+ 3}$ | 4 | $4([\frac{Core}{16}] + 3)$ |

3. Simulation Results

For evaluating the accuracy of each considered NoC architecture, every simulation run consists of experiments under certain simulation parameters for different traffic under the same environment. For analysing the performance, we have evaluated the behaviour with uniform, bit permutation and digit permutation type of traffics [16], [17]. In particular a total of five set-up are made each for uniform, bit reversal, bit complement, transpose and tornado traffic patterns. The parameters like virtual channel (VC) plays an important role when evaluating the performance under a variety of traffics [18]. The value of VC can be set to 2, 4, 6 and 8 for different set-up. Experiments are done for several combinations of network sizes, message lengths, VCs and buffer values. However, in this paper simulation results for VC = 6 are obtained and described to get the actual behaviour of the considered networks.

3.1. Effect on Network Latency

Latency is one of the important parameters to evaluate the performance of a NoC architecture. To observe the behavior of latency the simulation run consists of generating packet injection rate for different pattern of traffic and evaluating latency on to various considered NoC architecture with VC equal to 6. The estimation of latency is obtained and the curves are plotted as latency against packet injection rate and are shown in Fig. 4 (a) to Fig. 4 (e).

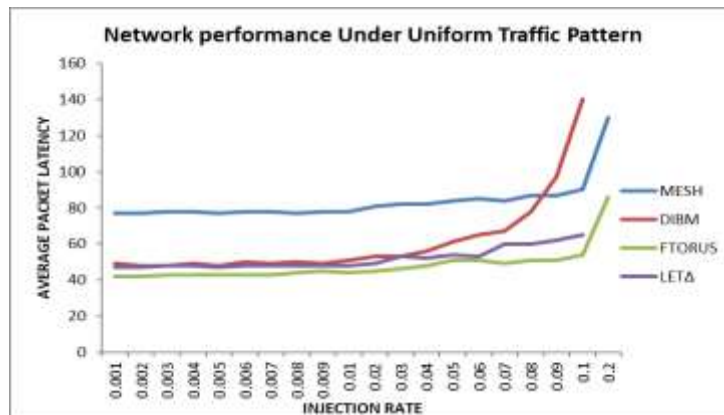


Fig 4 (a) Network Latency with Uniform Traffic

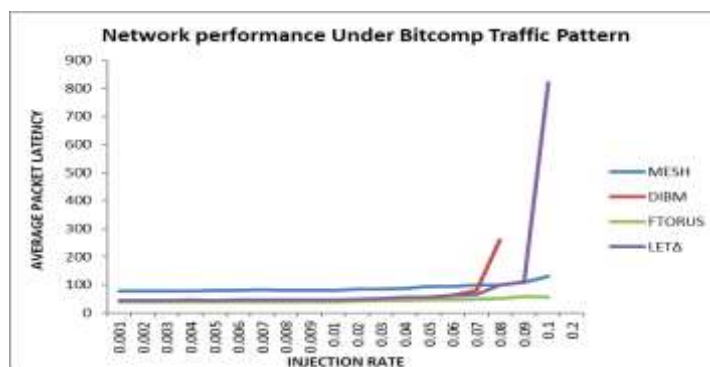


Fig 4 (b) Network Latency with Bitcomp Traffic

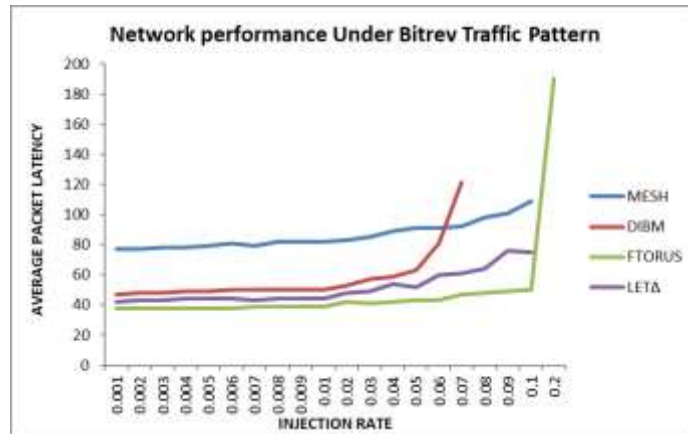


Fig 4 (c) Network Latency with Bitreverse Traffic

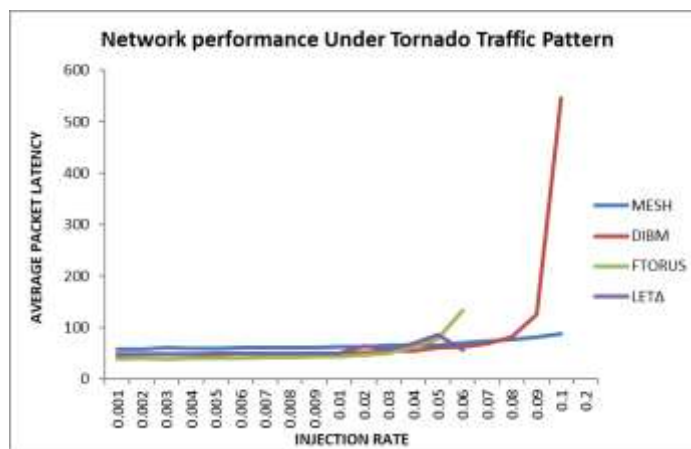


Fig 4 (d) Network Latency with Tornado Traffic

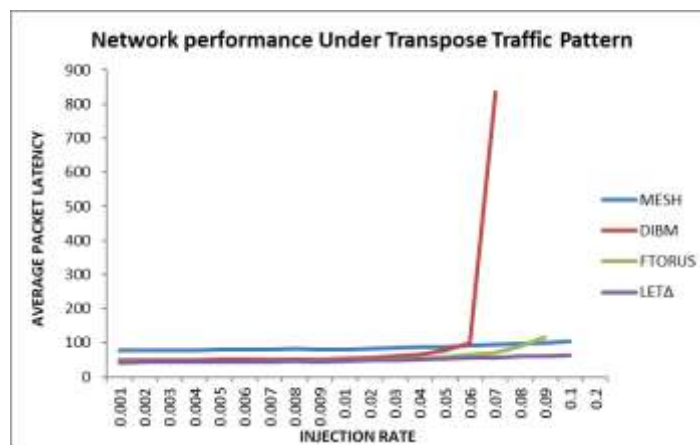


Fig 4 (e) Network Latency with transpose Traffic

Fig. 4. Network performance Under Different Traffic Patterns

The results shown in Figure 4 (a) to Figure 4(e) clearly indicate that the LETA and Folded Torus networks are producing better results under almost all types of traffic engineering except the results obtained with Bit-reversal traffic is little bit. This is due to the fact that LETA network does not have wrap around connections at each level

of the network. From the curves shown it can be argued that the LECA network could be a better choice to design an NoC system.

3.2. Effect on network Throughput

To test the different networks for throughput superiority the average throughput is evaluated for a particular configuration of network with varying injection rate under different scenarios of traffic pattern. The throughput may also vary depending upon the network size. We evaluated throughput for fix 8 x8 size networks and comparative charts are drawn for different networks with different traffic patterns and demonstrated in Fig. 5. Observing the simulation results obtained for different 4 x 4 network in terms of throughput it is analysed that LECA network is producing higher throughput in all the traffic patterns like standard Mesh network. For lesser values of injection rate, the performance of Folded Torus network is very poor. Similarly, in case of uniform traffic even the performance of mesh is also degraded. The DIBM, however, producing comparable results. Therefore, performance in terms of throughput make the new networks notable.

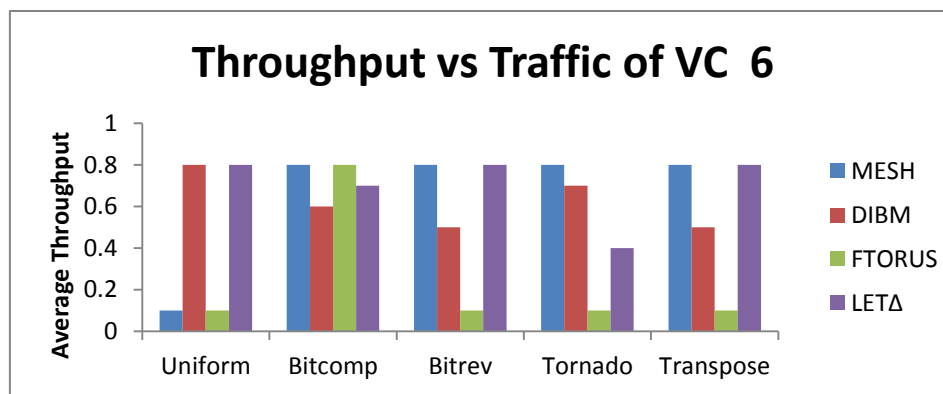


Fig. 5. Network performance in terms of Throughput Under Different Traffic Pattern

3.3. Effect on Execution Time

To evaluate and compare the performance, the total execution is evaluated for attaining the desired throughput with minimum delay in communication under different traffic patterns. The curves are drawn in the similar fashion as that used for evaluating the network latency and shown in Fig. 6 (a) to Fig. 6 (e). Again, as we noticed that the execution time for LECA network is always lesser in all the traffic patterns. In case of Bit-complement traffic, though the LECA network produced greater latency, however, it has lesser impact on execution time. Similarly for Tornado traffic, in which LECA network shows lesser improvement in throughput but it does not have greater impact on execution time. Increasing the injection rate does not have much effect on the execution time. It shows that the LECA network equally performs for higher levels of task structures or in fact producing even better results for higher levels.

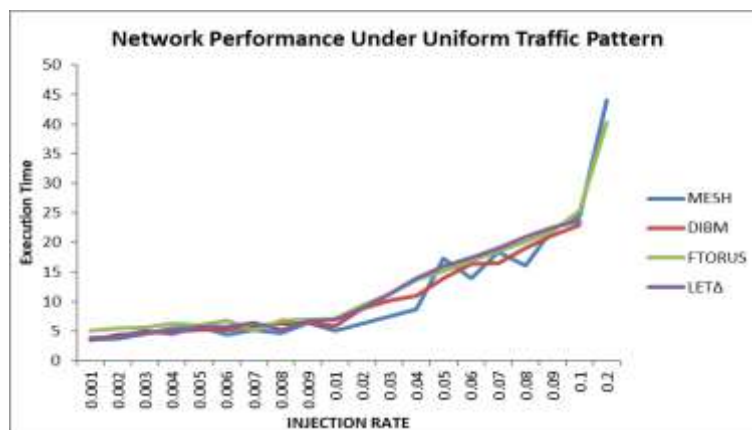


Fig 6. (a) Execution Time with Uniform Traffic

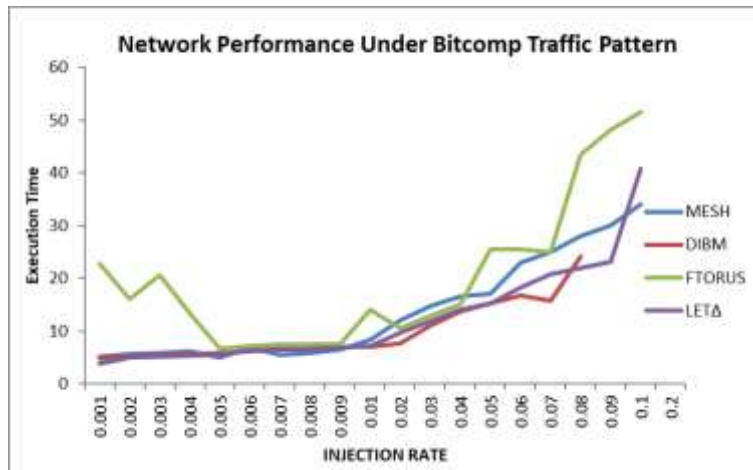


Fig 6. (b) Execution Time with Bitcomp Traffic

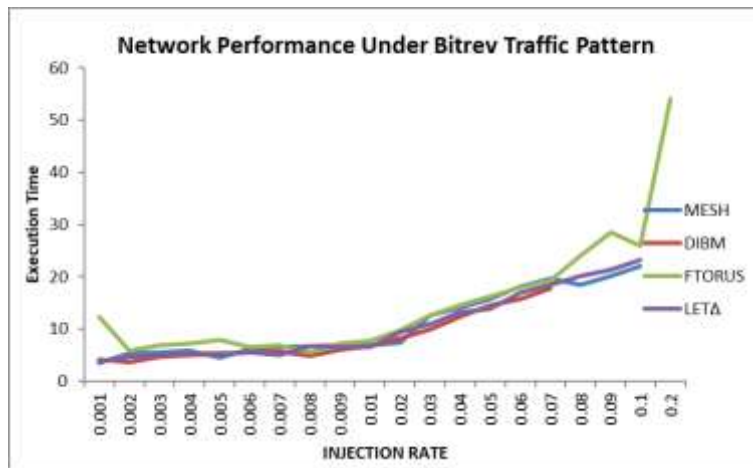


Fig 6. (c) Execution Time with Bitreversal Traffic

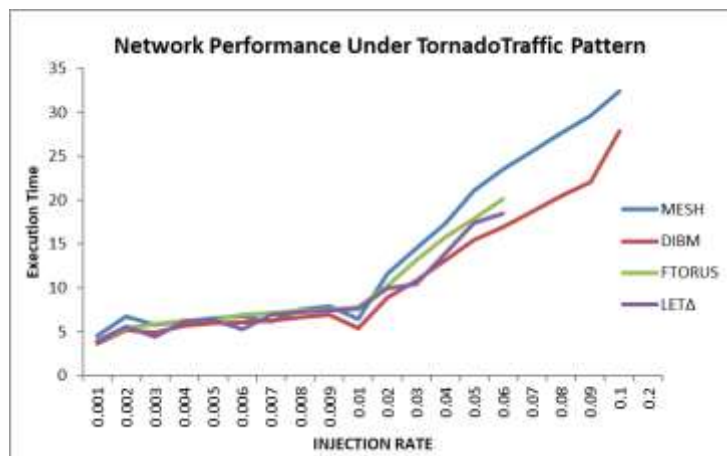


Fig 6. (d) Execution Time with Tornado Traffic

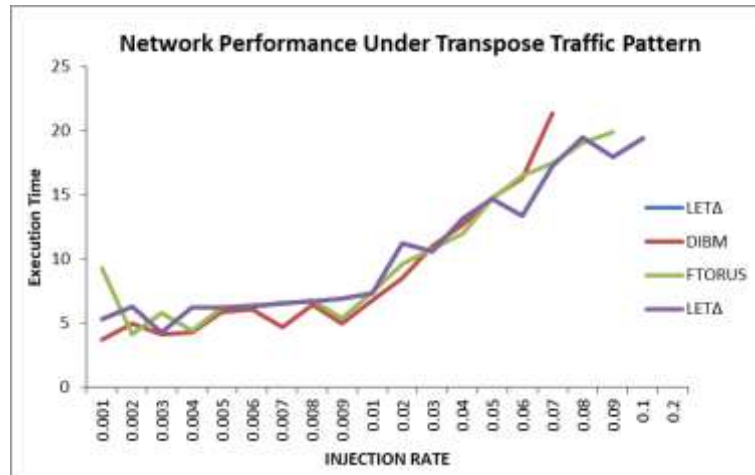


Fig 6. (e) Execution Time with Transpose Traffic

Fig. 6. Variation of Execution Time under different Traffic pattern

4. Discussion

As described in section 3 we have used five types of traffic patterns namely uniform, bit-complement, bit-reversal, tornado and transpose type of traffic patterns. According to the simulation results demonstrated in previous sections, the LEC Δ network architecture has a better performance in terms of latency as compared to the equivalent Mesh, Folded Torus and DIMB NoC architectures. The network only fails to cope up with latency under bit-complement traffic in which there is sharp rise in latency. The result is depicted in Fig. 4. (b). For other traffic patterns the LEC Δ network attaining outstanding performance. The reason is that the average distance a message travels in the network in a DIMB network is lower than that of a mesh.

When comparing the performance in terms of throughput there is significant improvement in case of LEC Δ architecture as compared to other considered network on chip architectures. This trend is obtained for all types of traffic patterns as shown in Fig. 5. Particularly, the LEC Δ and Mesh networks are producing similar results with an improvement of approximately 75% in throughput in case of LEC Δ network for almost every type of traffic pattern. Similar patterns are observed for mesh and folded torus as both of them are having equal number of cores with one hop routing distance. In DIMB, all links are not connected directly which restrict a limited improvement in throughput.

For the accurate estimates of the effectiveness of the proposed topologies, execution time is evaluated in terms of task completion or reaching to the saturation value of throughput at a particular stage of packet injection. The curve shown in Figure 6 indicate that similar behaviour is obtained for all the considered NoC architectures when different traffic patterns are applied to them. The performance of folded torus network is slightly lesser as compared to other networks in case of bit reversal traffic, on the other hand with the same traffic, LEC Δ producing best performance. For other traffic pattern also, LEC Δ network taking lesser execution time to achieve the desired performance. Fig. 6. (d) shows an improvement of approximately 50% reduction in execution time for folded torus and LEC Δ networks. As demonstrated in Table 1, LEC Δ network is having good parameters which are required for network NoC communication and also having lesser cost. Therefore, LEC Δ network could be considered a best choice in terms of topological characteristics as well as for good on-chip communication.

5. Conclusion

In this paper, a simulation model for core mapping using different traffic pattern is proposed. The regular mesh, folded torus, DIMB and LEC Δ topologies are considered as NoC networks. BookSim simulator is used to test and evaluate the performance of considered networks in terms of communication latency, execution time and throughput. Comparative study is carried out and curves are drawn for latency and throughput with packet injection rate. Compared with two latest network architectures namely LEC Δ topology and DIMB and two conventional architectures i.e., conventional mesh and folded torus networks, the simulation results show that significant improvement on the network latency is obtained for a variety of network traffic. Execution time of the task graph is achieving desired performance with similar pattern for all the considered NoC topologies. The

LECA topology producing 75% improvement in throughput and 50% reduction in execution time when compared to other similar considered networks. The better performance of LECA network may be attributed for its good topological properties such as small diameter and cost. Moreover, the proposed LECA network is scalable and even better results are obtained for larger size of networks and with varying values of VCs under different traffic patterns.

6. References

- [1] Romanov A. and Ivannikov A, (2018) SystemC Language Usage as the Alternative to the HDL and High-level Modeling for NoC Simulation. *International Journal of Embedded and Real-Time Communication Systems (IJERTCS)*, vol 9, no. 2, pp 18-31.
- [2] Romanov A. Y. and Romanova I. I. (2015) Use of irregular topologies for the synthesis of networks-on-chip. *IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)*, pp 445–449.
- [3] Ghalwash, H. and Huang. C. H. (2019) QoS for SDN-based fat-tree networks. *Future of Information and Communication Conference FICC 2019: Advances in Information and Communication*, pp 691–705.
- [4] Jantsch A. and Tenhunen H, (2003) *Network on Chips*, Kluwer Academic Publishers, Boston
- [5] O. He, S. Dong, W. Jang, J. Bian and D. Z. Pan, (2012). "UNISM: Unified Scheduling and Mapping for General Networks on Chip," in *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 20, No. 8, pp. 1496-1509, Aug. 2012, doi: 10.1109/TVLSI.2011.2159280.
- [6] Yu. H., Ha, Ya., and Veeravalli B. (2010). Communication-aware application mapping and scheduling for NoC-based MPSoCs, in *Proc. IEEE Int.Symp. Circuits Syst. (ISCAS)*, pp. 3232–3235
- [7] Mirza-Aghatabar, M. and Koochi, S. and Hessabi, S. and Pedram, M. (2007). An Empirical Investigation of Mesh and Torus NoC Topologies Under Different Routing Algorithms and Traffic Models. *10th Euromicro Conference on Digital System Design Architectures, Methods and Tools (DSD 2007)*. pp 19-26.
- [8] Toldan P. and Kumar M. (2013). Design and Implementation of (N X N) Folded Torus Architecture for Network on Chip with E-Cube Routing. *IJCSC* vol 4, no. 2, pp 145-152.
- [9] de Mesquita, J.W., da Cruz M.O, Pereira, M.M and Kreutz, M.E. (2016). Design Space Exploration Using UTNoCs and Genetic Algorithm. In *Proceedings of the Computing Systems Engineering (SBESC)*, pp 198–202.
- [10] Zhang, Weihua & Sun, Gengxin & Bin, Sheng. (2015). A Novel Task Communication and Scheduling Algorithm for NoC-based MPSoC. *International Journal of Smart Home*. 9. 179-188. 10.14257/ijsh.2015.9.10.20.
- [11] Ahmed, M., Gaur, M.S. and Laxmi, V. (2010). Adaptive Routing over the 2D Hexagonal NOC. *The International Conference on Embedded Systems (ICES 2010)*, Coimbatore, pp 1-5
- [12] Khosravi, A., Khorsandi, S. and Akbari, M.K. (2011). Hyper Node Torus: A New Interconnection Network for High-Speed Packet Processors. *International Symposium on Computer Networks and Distributed Systems (CNDS)*, pp 106-110
- [13] Ravindra Kumar Saini and Mushtaq Ahmed. (2015). 2D Hexagonal Mesh Vs 3D Mesh Network on Chip: A Performance Evaluation, *Int. J. Com. Dig. System*, vol. 4, no. 1, pp 33-41.
- [14] Gautam S., Samad A. (2019) Properties and Performance of Linearly Extensible Multiprocessor Interconnection Network. In: *Communication, Network and Computing, CNC 2018, Lecture Note in Communication in Computer and Information Science*, vol. 839, pp. 3-12, Springer, Singapore. <https://doi.org/10.1007/978-981-13-2372-0>
- [15] Sabbaghi-Nadooshan R. and Patooghy, A. (2015). Analytical performance modeling of de Bruijn inspired mesh-based network-on-chips, *Microprocessors and Microsystems*, vol. 39, pp 27–36.
- [16] K. K. Paliwal, V. Janyani, M. S. Gaur and V. Laxmi, (2009) Impact of Faulty Links on Quality-of-Service in Network-on-Chip under Different Traffic Patterns, *IJCSNS International Journal of Computer Science and Network Security*, Vol. 9 No. 3, pp. 108-117.
- [17] M. Tang and Jing Lin, (2018) A Comparative Study on NoC Transpose Traffic, *Advances in Intelligent Systems Research (AISR)*, Vol. 151, 2018 International Conference on Computer Modeling, Simulation and Algorithm (CMSA 2018). pp. 165-168.
- [18] Salem, Ridha & Salah, Yahia & Atri, Mohamed. (2016). Decentralized Scheduler for 3D NoC with QoS. *International Journal of Computer Science and Information Security*, Vol. 14, pp. 746-751.