



NC-G-SIM: A Parameterized Generic Simulator for 2D-Mesh, 3D-Mesh & Irregular On-chip Networks with Table-based Routing

By Kartika Vyas, Naveen Choudhary & Dharm Singh

Maharana Pratap University of Agriculture and Technology, India

Abstract - As chip density keeps doubling during each course of generation, the use of NoC has become an integral part of modern microprocessors and a very prevalent architectural feature of all types of SoCs. To meet the ever expanding communication challenges, diverse and novel NoC solutions are being developed which rely on accurate modeling and simulations to evaluate the impact and analyze their performances. Consequently, this aggravates the need to rely on simulation tools to probe and optimize these NoC architectures. In this work, we present NC-G-SIM (Network on Chip-Generic-SIMulator), a highly flexible, modular, cycle-accurate, configurable simulator for NoCs. To make NC-G-SIM suitable for advanced NoC exploration, it is made highly generic that supports extensive range of cores in any kind of topology whether 2D, 3D or irregular. Simulation results have been evaluated in terms of latencies, throughput and the amount of energy consumed during the simulation period at different levels.

Keywords : *NC-G-SIM, generic, simulator, 3D, 2D, irregular, distributed table-based routing, deadlock-free.*

GJCST-E Classification : *C.2.2*



NC-G-SIMA PARAMETERIZED GENERIC SIMULATOR FOR 2D-MESH, 3D-MESH IRREGULAR ON-CHIP NETWORKS WITH TABLE-BASED ROUTING

Strictly as per the compliance and regulations of:



NC-G-SIM: A Parameterized Generic Simulator for 2D-Mesh, 3D-Mesh & Irregular On-chip Networks with Table-based Routing

Kartika Vyas ^α, Naveen Choudhary ^σ & Dharm Singh ^ρ

Abstract - As chip density keeps doubling during each course of generation, the use of NoC has become an integral part of modern microprocessors and a very prevalent architectural feature of all types of SoCs. To meet the ever expanding communication challenges, diverse and novel NoC solutions are being developed which rely on accurate modeling and simulations to evaluate the impact and analyze their performances. Consequently, this aggravates the need to rely on simulation tools to probe and optimize these NoC architectures. In this work, we present NC-G-SIM (Network on Chip-Generic-SIMulator), a highly flexible, modular, cycle-accurate, configurable simulator for NoCs. To make NC-G-SIM suitable for advanced NoC exploration, it is made highly generic that supports extensive range of cores in any kind of topology whether 2D, 3D or irregular. Simulation results have been evaluated in terms of latencies, throughput and the amount of energy consumed during the simulation period at different levels.

Keywords : NC-G-SIM, generic, simulator, 3D, 2D, irregular, distributed table-based routing, deadlock-free.

I. INTRODUCTION

With the accelerated growth in calibration and denseness of integrated circuits in recent years, data communication amid the internal on-chip processing cores has become a hitch for SoC performance (Dally & Towles, 2004). This has led to a new paradigm on-chip communication in Ultra Large Scale Integrated systems. Traditional bus structure could not conciliate such a tight and high call of time-to-market as substantial amount of labor has to be enmeshed for engrafting these non-scalable architectures in such large systems, amplifying power consumption and stretching disposition time (Dally & Towles, 2004). NoC here takes the charge and provides an elegant approach to manage interconnect complexity (Heo & Asanovic, 2005), facilitating integration of multi-cores by abstracting computation from communication (Benini & Micheli, 2002) achieving better scalability and more predictable performances (Atienza, Angiolini, Murali, Pullini, Benini & Micheli, 2008). Until now, most of the researches were on making effective and potent 2D

NoC designs scrutinized against power, performance and reliability (Heo & Asanovic, 2005).

However, simply increasing the number of cores over a 2D plane is not efficient due to long interconnects. Emergence of viable 3D integration technologies has created opportunities for chip architectures that were prohibitive due to several constraints with the firm reliance that many of these 3D implementations can outperform their 2D counterparts (Feero & Pande, 2009). Research has also gained momentum in the domain of irregular networks (Silla & Duato, 2000) with the increased demand of application-specific SoCs as it saves significant power and area overhead, for most SoCs the mapping of tasks to processors and hardware cores is done on basis of static (or semi-static) prior knowledge and hence, has well characterized communication traffic characteristics are at design time of the SoC. Therefore, it has become excessively necessary to provide the architecture, methodology and simulation platform, which remains sustainable over different technology generations (Dally & Towles, 2004). With this perspective, NC-G-SIM simulator is developed that caters all the popular requirements of NoCs.

The rest of the paper is systematized as follows. Section 2 highlights background information about a few interconnection network simulators and the motivation behind developing generic NC-G-SIM simulator. Section 3 presents the simulator architecture and functionality. Section 4 describes the statistics collection process and depicts experimental results. Finally, Section 5 concludes the paper and highlights some directions for possible future work.

II. RELATED WORK

Varied NoC simulators have been instituted and research communities are currently operating on these. Noxim (Palesi et al.) is such a simulator that put up the experimentation with different traffic distributions using various routing algorithms, but it supports mesh topology only and also lacks the concept of virtual channels disabling physical channel's bandwidth sharing. Simulators gpNoCsim (Hossain, Ahmed, Al-Nayeem, Islam & Akbar, 2007) and BookSim (Jiang et al.) support evaluation of different existing 2D

Authors ^{α σ ρ} : Department of Computer Science & Engineering, College of Technology and Engineering, Maharana Pratap University of Agriculture and Technology, Udaipur, Rajasthan, India.
E-mails : katty.b08@gmail.com, naveenc121@yahoo.com, dharm@mpuat.ac.in

architectures but with confined traffic injection models. A simulator with similar features is NoC sim (Jones 2005), only support VC routers with synthetic traffics and do not report router power and energy. Ocim tsim (Prabhu, Grot, Gratz & Hu, 2009), GARNET (Agarwal, Krishna, Peh & Jha, 2009), SICOSYS (Puente, Gregorio & Bevide, 2002) and Darsim (Mieszko, Sup, Hyon, Pengju, Khan & Devadas, 2010) support only 2D NoCs and compute router power based on the ORION model (Glass & Ni, 1992). Nostrum-NNSE (Lu, Thid, Millberg, Nilsson & Jantsch, 2005) developed in SystemC, has packet switched simulation environment. It supports irregular topologies but has no tool to estimate communication energy requirement and power consumption. NIRGAM (Jain, Al-Hashimi, Gaur, Laxmi & Narayanan, 2007) provides extension flexibility and modularity. However, only supports Mesh and torus topologies with limited support to various routing methodologies. *IrNIRGAM* (Choudhary, Gaur & Laxmi, 2011) supports irregular topologies using deterministic as well as adaptive table-based routing to make routing deadlock free and performance oriented. Therefore, most of existing simulators are specific to 2D-NoC topology architecture and may not be able to provide flexibility and modularity for existing and novel architectures. The market is also flooded with numerous simulators that researchers have materialized themselves to simulate their distinct architectures which aren't applicable for other topologies. The fast pace of current trends of the market raises the need for a generic simulation tool to efficiently explore the wide design space (Dally & Towles, 2004) where design choices are primarily taken on the basis of simulation before resorting to implementation as this is more flexible and cheaper.

IrNIRGAM is therefore, selected and extended to develop NC-G-SIM, a discrete event-cycle accurate 3D NoC simulation framework aimed typically at the NoC research and development community wherein it provides them with convenient and efficacious mechanism to experiment with NoC design in terms of 2D, 3D, irregular topology, routing and applications, with the facility to easily plug-in their own routers and attaching to cores any user-specified application library and configuration of parameters as per the need. It thereby, proves to be a powerful and competent tool to model vivid and intricate systems having large number of cores (upto 500) providing accurate performance estimation.

III. SIMULATOR DESCRIPTION

a) NC-G-SIM Functionality

The flexibility of the structure of *IrNIRGAM* motivated to extend it in order to incorporate in it 3D NoC structure. NC-G-SIM simulator supports 2D regular Mesh and Torus topologies with a set of source, XY and

odd-even routings. Irregular topology support and its features remain existent as they were in *IrNIRGAM*.

The 3D mesh NoC in NC-G-SIM follows the footsteps and uses both one of the well-known Dimension Order Routing (DOR) routing scheme i.e. XYZ algorithm as well as distributed table-based routing scheme. XYZ is a simple traditional scheme, easy to implement and free of deadlock and live lock. However, every time a packet is routed between a particular source-destination pair the same path is always followed which might lead to a scenario where a set of paths are heavily used while other paths remain idle, affecting overall system performance. But, use of distributed routing scheme overcomes this bottleneck, exploiting the fact that there might be multiple alternative paths to reach every node. Hence, look-up tables corresponding to each node is generated which cumulate this information, facilitating diverging traffic uniformly among the paths. Routing table files follow the format [*sourceID*, *destinationID*, *nexttileID*, *pathID*] each path is assigned a unique id '*pathID*' to make paths easily distinguishable. Each node decides next routing direction on basis of the entries in the node's corresponding table file. At each intermediate node, the same procedure is repeated till the packet reaches its destination. These tables are filled offline, based on the networks' localized knowledge, in accordance to either of miscellaneous deadlock free routing algorithms such as Left-Right (Schroeder et al. 1991), up*/down* (Schroeder et al. 1991) etc for irregular topologies and for 3D topology any of the minimal non-minimal, DOR or adaptive routing could be used meeting negligible chances of any deadlock formation.

Using tables we have also deployed a generic methodology based on *escape paths* by (Silla & Duato, 2000) in NC-G-SIM which enforces to follow deadlock free route only when the shortest paths are clogged up by high amount of traffic, thereby offering greater adaptivity, reduced latency and better overall throughput. It presumes that every physical channel should be split into two sets; original virtual channels and the new virtual channels representing minimal paths and escape-paths respectively. Packets following new channels are routed by the router to any channel without any circumscriptions giving first preference to new channels. Due to congestion, when no new channels are accessible then, the original routing function is followed and once a packet acquires an original channel it is not permitted to transit to a new channel anymore to avoid deadlock situation. Format of the tables based on *escape paths* methodology is similar to the above format except that instead of the '*pathID*'s we have '*virtual_channel_no*.' that distinguishes between minimal paths and deadlock free paths, which is being followed by the packet.

The simulator facilitates the user to design one's own topology according to ones choice and

research scenario. This is made possible by providing an overview of the topology layout by the user itself in a topology configuration file which is read in at the beginning of the simulation. This facility is provided for both irregular topologies, since there is no regularity or particular trend in the connections among the tiles, as well as for 3D topology. The topology is in a matrix format $[tileID, Neighbour_tileID_1, Neighbour_tileID_2, \dots, Neighbour_tileID_n, -1]$ where $tileID$ is current tile identifier followed by the $tileIDs$ of its neighbouring nodes, list terminated by the delimiter “-1”. With this, the topological arrangement becomes wieldy descriptive and alterable as and when desired. This not only simplifies the experimentations on the traditional NoC architectures but also encourages further extension of these architectures. Therefore, the simulator can be used for one and all purposes hence, proving to be totally generic in nature.

NC-G-SIM simulator can support 6 neighbours of any tile, with maximum number of tiles to be 500 for irregular and 3D topologies. Along with the topology file, another file of format $[tileID\ next_tileID\ link_length_{(in\ mm)}]$ is provided which contains the length of the links in the topology in order to estimate the energy consumption.

Apart from these traffic characteristics are essentially required to run a simulation. Besides the well-known yet primitive CBR and Bursty traffic generator applications, NC-G-SIM additionally implements BWCBR traffic facility to generate concurrent traffic from a source node to different destinations by attaching multiple traffic generating processes thereby, making the simulation conditions more proximate to the actual

scenarios where heavy traffic is exchanged between the nodes. Another important feature of this simulator is that it allows dynamic attachment of routing algorithms and application cores at the run time, letting re-modification of routing and traffic, saving the user from re-compilation of the code.

b) *NC-G-SIM Architecture*

A comprehensive architecture of NC-G-SIM simulator is shown in Fig.1. The additional customizable parameters are the size of the topologies, simulation and traffic generation cycle, buffer size, packet and flit sizes, packet interval, adjusted for the simulation engine. Multiple numbers of virtual channels can be multiplexed per physical channel supporting fifo buffer for storing incoming traffic. It works at a default clock frequency of 1 GHz which can be altered. Wormhole switching mechanism is used, where packets are serialized into several flits therefore, smaller buffers can be utilized in contrast to the packet switching technique and thus, makes packet latency relatively insensitive to the path length.

When simulation ends, NC-G-SIM generates a log file that can be made detailed or concise by switching to different log levels. Apart from this the detailed results of the simulation can be read, checked and compared in terms of flit load, energy dissipation, throughput, latency and standard deviation from the result files generated on completion of the simulation. Using GnuPlot graphs can also be plotted of these output performance metrics.

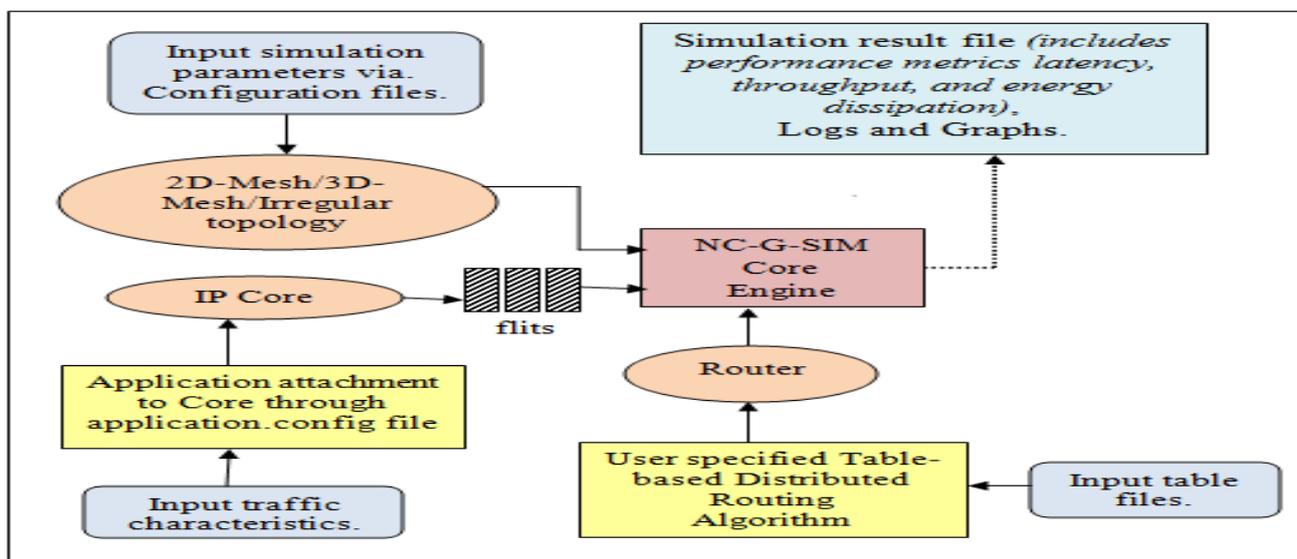


Figure 1 : NC-G-SIM Simulation Framework

Internal Architecture is implemented in the core where a module “NoC” models network of tiles and each tile is then prepared by the “NWTile” module. The preeminent components of each tile (“their module

names”) are as follows: **Input Channel Controller**[IC] (“InputChannel”), **Controller**: (“Controller”), **Virtual Channel Allocator** [VCA] (“VCAlocator”), **Output Channel Controller** [OC] (“OutputChannel”) and **ipcore**

("ipcore"). The number of IC and OC in each tile is equal to the number of its neighbor plus one for ipcore. IC manages the arrival and storage of the flits whereas OC manages the transfer of the flits to the next neighbor node. Controller handles the routing requests from all ICs as it implements the router and VCA serves virtual channel allocation requests from all ICs. Ipcore in each tile consists of an IP element to which an application or traffic generator can be attached if needed.

IV. EXPERIMENTAL RESULTS

NC-G-SIM simulator was tested to generate the various performance parameters. Flit interval rate is taken as 2 clock cycles which means that time taken for one flit injection at each core and its transmission into the network approximately takes 2 clock cycles. The evaluation is performed under medium congestion with varied traffic load ranges from 10 to 100 percent for a set of random sources and destinations decided randomly. These source-destination pairs are kept alike for fair comparison between 3D and 2D topologies in each run of simulation. For traffic generation BWCBR generator was chosen to send out traffic to multiple cores concurrently. For performance comparison, a period of 5000 simulation clock cycles was kept in NC-G-SIM with correlated packet injection interval to scrutinize network performance.

For the test cases of NoC, the energy required during communication is evaluated as per energy model proposed in (Hu 2005). The average energy consumption by router in transmitting a bit is calculated using analytical model such as Orion (Glass & Ni, 1992) and the dynamic bit energy consumption for inter-node links (E_{bit}) is calculated using equation (Hu 2005).

$$E_{bit} = (1/2) \times \alpha \times C_{phy} \times (V_{DD})^2 \quad (1)$$

Where α denotes average transition probability for a specific bit between two subsequent segments in the transit course. The value of α is assumed 0.5 for purely random data stream. C_{phy} is the physical capacitance of inter-node wire taken under consideration for any liable technology and V_{DD} is the supply voltage.

Within the vast range of supportable number of tiles 100 topologies were selected that possess same number of cores in both 2D and 3D for better comparison and for each topology 10 different test cases were prepared and executed. Having been experimented on innumerable number of topology sets, the average of obtained results is represented to show the fulfillment of our objectives.

Fig.2. proves three dimensional topologies exhibit better performance when experimented with constant average throughput. The average flit latency gets reduced by 8% to 21.8% as the architecture of 3D

is more compact compared to 2D, therefore needs less number of hops to traverse from the source node to the destination node, thereby taking lesser time to reach its desired node. The average total energy per flit, consumed in the whole simulation scenario for each test case also proves the above stated fact of 3D architectures. The average energy dissipation for the received flits at destination in 3D has showed tremendous improvement, as we have got 21% to 36.8% lesser consumption of energy in 3D topologies as compared to 2D topologies. The reason behind this is since the whole bigger topology in 3D is sliced into layers so it has shorter routes to traverse, this in return utilises lesser number of routes and much less switch arbitration saving the energy and the time too.

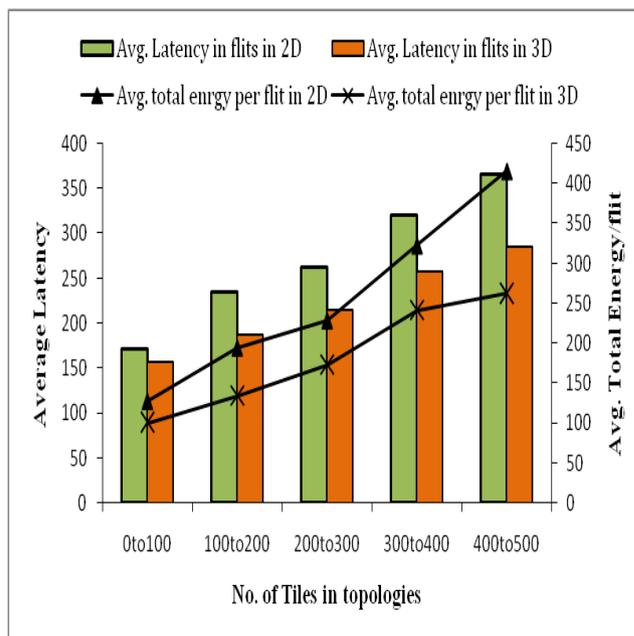


Figure 2 : Comparison of performance characteristics for both two dimensional and three-dimensional topologies, with topologies grouped in the range of hundreds

For testing the efficacy of irregular arrangement, a set of application specific irregular topologies were explored using table-based scheme supporting Left-Right routing. Our results in Fig.3. demonstrate that latency and energy dissipation of irregular NoC was lowered by approximately 23% and 7.6% respectively, in comparison to 3D NoCs. Irregular NoCs have achieved these results as they are more design specific according to the application requirements of the system. So, their property of being more system-oriented and targeted towards the fulfillment of precise system needs make them a suitable choice to those systems that require specific designs with maximum performance yield and in lesser time-to-market constraint as well as cost trade-offs. Using irregular interconnection systems will also prove beneficial if they are tried to be fused with the

three-dimensional interconnection feature of the upcoming network-on chips.

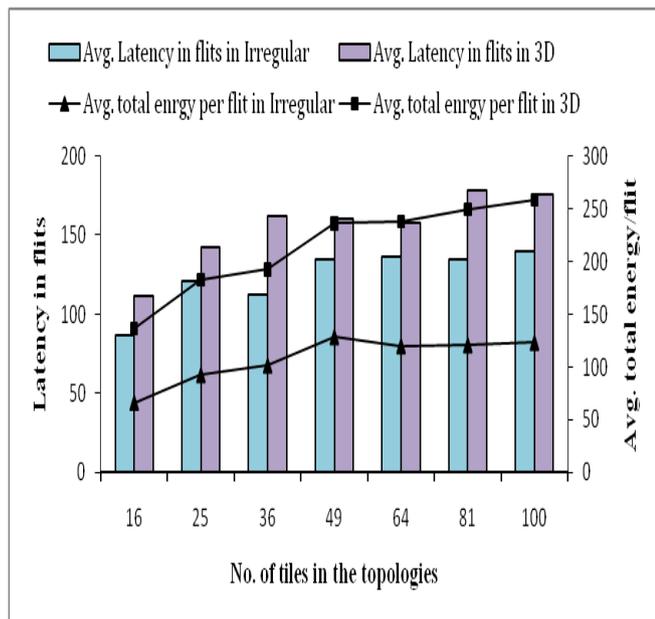


Figure 3 : Performance comparison between irregular and three-dimensional topologies in terms of latency and average energy consumed per flit

V. CONCLUSION AND FUTURE EXTENSIONS

In this paper, we presented a robust and generic simulation framework NC-G-SIM to study the performance of all the NoC domains 2D, 3D and irregular. System C is preferred as simulator design because it is meant for higher layer of NoC abstraction which leads to better understanding about the design enabling better system tradeoffs. NC-G-SIM features the use of distributed table-based routing in 3D & irregular interconnection networks with the potentiality of evaluating accurately and smartly the effectuation of multitudinous core designs, allowing the experimenter to migrate from one design archetype to some other effortlessly. For the future work we plan to support a wider range of NoC topologies, routing schemes and switching mechanisms. Fault tolerant feature is also seen as an upcoming prominent area of interest. Also, heterogeneous on-chip communication requirements will be soon addressed to make the simulator more advantageous for the researchers of NoC.

VI. ACKNOWLEDGMENTS

This work is supported by Department of Science and Technology, Jaipur, Rajasthan, India under the research project “Network-on Chip simulation framework for regular, irregular and 3D-Mesh Interconnection Architecture”.

REFERENCES RÉFÉRENCES REFERENCIAS

1. Agrawal, N., Krishna, T., Peh, L.S., & Jha, N.K. (2009). GARNET: A detailed on-chip network model inside a full system simulator. In IEEE International Symposium on Performance Analysis of Systems and Software, 33–42.
2. Atienza, D., Angiolini, F., Murali, S., Pullini, A., Benini, L., & Micheli, G.D. (2008). Network-on-chip design and synthesis outlook. In Integration-The VLSI Journal, 41(3), 340-359.
3. Benini, L., & Micheli, G.D. (2002). Networks on chips: A new SoC paradigm. In IEEE Computer, 35(1), 70–78.
4. Choudhary, N., Gaur, M.S., & Laxmi, V. (2011). Irregular NoC simulation framework: IrNIRGAM. In International Conference on Emerging Trends in Networks and Computer Communications (ETNCC), 1-5.
5. Dally, W.J., & Towles, B. (2004). Principles and practices of interconnection networks. Publication, Elsevier/Morgan Kaufmann, Amsterdam, London.
6. Feero, B.S., & Pande, P.P. (2009). Networks-on-Chip in a Three-Dimensional Environment: A Performance Evaluation. In IEEE Transactions on Computers, 58(1), 32-45.
7. Glass, C.J., & Ni, L.M. (1992). The turn model for adaptive routing. In Proceedings of 19th Annual International Symposium on Computer Architecture, 278– 287.
8. Heo, S., & Asanovic, K. (2005). Replacing global wires with an on-chip network: a power analysis. In Proceedings of the International Symposium on Low power electronics and design, 369–374.
9. Hossain, H., Ahmed, M., Al-Nayeem, A., Islam, T.Z., & Akbar, M.M. (2007). Gpnocsim – A General Purpose Simulator for Network-On-Chip. In International Conference on Information and Communication Technology, 254 – 257.
10. Hu, J. (2005). Design methodologies for application specific networks-onchip. PhD thesis, Carnegie Mellon University.
11. Jain, L., Al-Hashimi, B.M., Gaur, M.S., Laxmi, V., & Narayanan, A. (2007). NIRGAM: a simulator for noc interconnect routing application modeling. DATE-2007.
12. Jiang, N. et al. BookSim Interconnection Network Simulator. Online, <http://nocs.stanford.edu/booksim.html>.
13. Jones, M. (2005). NoCsim: a versatile network on chip simulator. <https://circle.ubc.ca/handle/2429/16550>.
14. Lu, Z., Thid, R., Millberg, M., Nilsson, E., & Jantsch, A. (2005). NNSE: Nostrum Network-on-Chip Simulation Environment. In Swedish System-on-Chip Conference, Stockholm, Sweden.

15. Mieszko, L., Sup, S.K., Hyon, C.M., Pengju, R., Khan, O., & Devadas, S. (2010). Darsim: A parallel cycle-level NoC simulator. Online, <http://dspace.mit.edu/handle/1721.1/59832>
16. Palesi, R. et al. Noxim: the NoC Simulator. Online, <http://noxim.sourceforge.net>.
17. Prabhu, S., Grot, B., Gratz, P.V., & Hu, J. (2009). Ocim tsim- DVFS aware simulator for NoCs.
18. Puente, V., Gregorio, J.A., & Beivide, R. (2002). SICOSYS: an integrated framework for studying interconnection network performance in multiprocessor systems. In Euromicro Workshop on Parallel, Distributed and Network based Processing, 15 –22.
19. Schroeder, M.D. et al. (1991). Autonet: A high-speed self-configuring local area network using point-to-point links. In Journal on Selected Areas in Communications, 9, 1318 – 1335.
20. Silla, F., & Duato, J. (2000). High-performance routing in networks of workstations with irregular topology. In IEEE Transactions on Parallel and Distributed Systems, 11, 699-719.