CLUSTERED TWO-DIMENSIONAL MESH TOPOLOGY FOR LARGE-SCALE NETWORK-ON-CHIP ARCHITECTURE

MEHDI BABOLI

UNIVERSITI TEKNOLOGI MALAYSIA

CLUSTERED TWO-DIMENSIONAL MESH TOPOLOGY FOR LARGE-SCALE NETWORK-ON-CHIP ARCHITECTURE

MEHDI BABOLI

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy (Electrical Engineering)

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

JUNE 2017

Dedication to my parents, my wife, and my kids.

ACKNOWLEDGEMENT

All praise is for GOD the beneficent, the merciful and lord of all the universes. I am grateful to almighty GOD for rewarding me good health, courage, patience and determination to complete this research work.

Foremost, I offer my sincerest gratitude and deepest appreciation to my supervisor Dr. Nasir Shaikh Husin and Assoc. Prof. Dr. Muhammad Nadzir Marsono for their guidance, patience, motivation, enthusiasm, immense knowledge and continuous support throughout my Ph.D study and also allowing me the room to work in my own way. His advices helped me in all the time during conducting experiment and writing of this thesis. Without his encouragement and effort this thesis would not have been completed or written. I would also like to thank Dr. Jasmine Hau Yuan Wen and Dr. Usman Ullah Sheikh for their valuable advises comments and guidance during my research work. I owe a very important debt to my family, especially my wife Fatemeh Asadi Amiri and sons Ermia and Arshia, who support and encourage me a lot during my Ph.D studies. I am very thankful and oblige to my family for their support, courage and help during completion of this task. I express my sincere appreciations to my parents amd my mother in low, whose hands always rose in prayers, for their sacrifices in growing me up and make me able to achieve the present level. In my daily work I have been blessed with a friendly and cheerful group of fellow students. Special thanks goes to all my local and international colleagues for their support, warm and friendly atmosphere they have created within the research groups especially Alireza, Yin Zhen, Huiru, Stephen, Kim, Jay, Jeevan, Monina, and Shaliza. A very special thanks goes to my former supervisors Assoc. Prof. Dr. Reza Ebrahimi Atani and Assoc. Prof. Dr. Sattar Mirzakuchaki for Iran University of Science and Technology (IUST) Tehran Iran, for their valuable advises comments and guidance during my previous research work. I would like to thank all my teachers from primary to present level, whose teaching and efforts make me able to achieve this level. At the end, a special thanks to my institute Universiti Teknologi Malaysia (UTM) for providing a very constructive and encouraging research environment to complete my research work. I am also grateful to UTM for sponsoring my research work.

ABSTRACT

Driven by the continuous scaling of Moore's law, the number of processing cores in chip multiprocessors and systems-on-a-chip are expected to grow tremendously in the near future. Connecting the different components of a multiprocessor chip in a scalable and efficient way has become increasingly challenging. Current network-on-chip (NoC) topologies are adequate for small-size networks but are not optimized for large-scale networks. Transmitted packets inside a large NoC require longer route to reach their destinations, resulting in an increase in certain performance parameters such as latency and power consumption. Thus, it is necessary to develop a new topology appropriate for large-size NoCs. In this research, we proposed a cost-effective network topology for large-size NoCs that improves performance in terms of end-to-end latency. The topology, called RaMesh, consists of clusters of mesh networks. A routing algorithm suitable for this topology was also proposed. The RaMesh architecture together with mesh, torus, and clustered 2D-mesh were simulated using Noxim (NoC simulator), C for software NoC models, and Altera ModelSim for Verilog hardware models. Simulations were conducted under different network traffic and for a variety of network sizes. Experimental results showed that RaMesh performed better than equivalent 2D-mesh and torus topologies. RaMesh topology was also benchmarked against a clustered mesh topology. Average hop count in the proposed topology was at least 22.7% lower compared to the mesh and torus. Average latency was also decreased by at least 24.66% as compared to the mesh and torus. Finally, the saturation point for the proposed topology increased by at least 15% as compared to mesh and torus.

ABSTRAK

Didorong oleh peningkatan berterusan dalam Hukum Moore, bilangan teras pemprosesan dalam multiprosesor cip dan sistem dalam satu cip dijangka berkembang dengan pesat dalam masa terdekat. Menyambung komponen yang berlainan dalam multiprosesor cip dengan cara yang cekap dan berskala tinggi telah menjadi semakin mencabar. Topologi Rangkaian-atas-Cip (NoC) semasa adalah cukup untuk rangkaian saiz kecil tetapi bukan teroptimum untuk rangkaian berskala besar. Paket-paket yang dihantar di dalam NoC besar mungkin mempunyai laluan yang panjang untuk sampai ke destinasinya. Ini menyebabkan peningkatan dalam parameter tertentu seperti kependaman dan penggunaan kuasa. Oleh itu, adalah perlu untuk menghasilkan topologi baharu sesuai untuk NoC bersaiz besar. Dalam kajian ini, kami mencadangkan satu topologi rangkaian kos-berkesan untuk NoC bersaiz besar yang memperbaiki prestasi dari segi pendaman hujung-ke-hujung. Topologi yang dinamakan RaMesh, terdiri daripada kelompok rangkaian jejaring. Algoritma penghalaan yang sesuai untuk topologi ini juga dicadangkan. Seni bina RaMesh bersama-sama dengan jejaring dan torus disimulasi menggunakan Noxim (NoC Simulator), C untuk model NoC perisian, dan Altera ModelSim untuk model perkakasan Verilog. Simulasi dilakukan di bawah lalu lintas rangkaian yang berbeza dan untuk aneka saiz rangkaian. Hasil uji kaji menunjukkan prestasi RaMesh lebih baik daripada topologi jejaring-2D setara dan torus. Topologi RaMesh juga ditanda aras dengan topologi jejaring berkelomplok. Kiraan hop purata dalam topologi yang dicadangkan adalah sekurang-kurangnya 22.7% lebih rendah berbanding dengan jejaring dan torus. Kependaman purata juga diturunkan sekurang-kurangnya 24.66% berbanding dengan jejaring dan torus. Akhirnya, titik tepu bagi topologi yang dicadangkan bertambah sekurang-kurangnya 15% berbanding dengan jejaring dan torus.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

bps	-	Bit per second
BPSF	-	Bypass Forward
BPSR	-	Bypass Reverse
CMP	-	Chip multiprocessors
CPU	-	Central Processing Unit
DNLY	-	Down Layer
DOR	-	Dimension Ordered Routing
DyAD	-	Dynamical Adaptive and Deterministic routing
FPGA	-	Field Programmable Gate Array
FRWD	-	Forward
GAD	-	Global Average Delay
GAT	-	Global Average Throughput
GS	-	Global Switch
HDL	-	Hardware Description Languages
IC	-	Integrated Circuit
IDE	-	integrated development environment
IP	-	Intellectual Property
ISM	-	Interface Switch Master
ISO	-	International Organization for Standardization
ISS	-	Interface Switch Slave
IVC	-	Input Virtual Channel
LS	-	Local Switch
MPSoC	-	Multiprocessor System-on-Chip
NoC	-	Network on Chip
OVC	-	Output Virtual Channel
PE	-	Processor Element
QoS	-	Quality-of-service
RaMesh	-	Ring-based Mesh
ROMM	-	Randomized Oblivious Multi-phase Minimal routing
RTL	-	Register transfer level

RVRS	-	Reverse
SoC	-	System on Chip
UCDB	-	Unified Coverage Database
UPLY	-	Up Layer
VC	-	virtual channel
VHDL	-	VHSIC Hardware Description Language
XML	-	Extensible Markup Language

LIST OF APPENDICES

CHAPTER 1

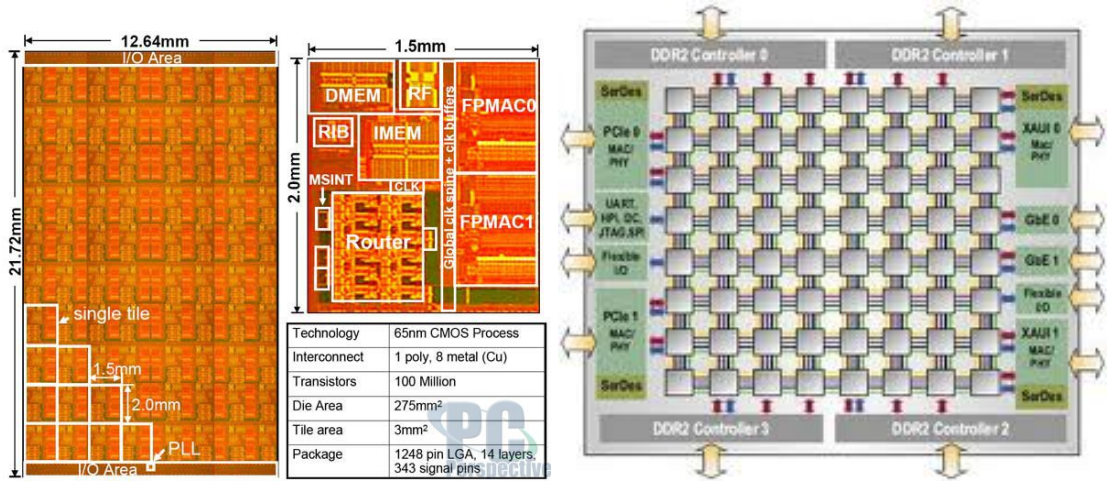
INTRODUCTION

1.1 Network-on-Chip

Multi-processor system-on-chip (MPSoC) is capable of accommodating many processing resources for high-performance computation [12, 13]. On-chip communication is the main bottleneck of MPSoC. Conventional bus-based on-chip interconnect cannot provide efficiency and scalability to connect many cores on one chip. Network-on-chip (NoC) has been proposed to meet on-chip interconnect challenges. NoC consists of interconnected routers based on certain topology (e.g., a mesh), that integrates memories, computational processors or the Intellectual Property (IP) components. The method of communication among IPs within an NoC-based system is through packet transmission via routers instead of circuit switching in bus-based interconnect.

Designing an efficient high performance and low latency NoC is still an open area of research. According to [14, 15], MPSoC size with hundreds or thousands of cores are likely to be common-place today. The increase of on-chip cores requires a high-bandwidth and scalable communication fabric [16, 17]. To satisfy these requirements, NoCs have been presented and has very quickly emerged as the preferred interconnection fabric.

As example, there exists real chips with 80 cores by Intel [1, 18], 100 cores by Tilera [19], and even a research prototype with 1000 cores by University of Glasgow [20]. While increased core count has allowed processor chips to scale without experiencing complexity and power dissipation problems inherent in larger individual cores, challenges still exist. NoC has been utilized to solve this problem. Figure 1.1 shows an example of a 80-core research prototype from Intel [1] (Figure 1.1a) and a commercial 64-core chip for embedded applications from Tilera (Figure 1.1b) that



(a) Intel Teraflops **(b)** Tilera TILE64

Figure 1.1: Examples of multi-core chips with on-chip networks [1, 2]

employs an on chip network for inter-tile communication [2].

1.2 Problem Statement

NoC topology defines how routers are connected together with network endpoints (i.e. IP cores). The performance and cost of NoC are greatly affected by the topology in large-scale MPSoC. Large-scale NoC topology is referred to as the one that has more than 100 IP cores [21]. An NoC topology is characterized by number of hop and network latency [22, 23]. The main issue with a large-scale NoC is the large number of hops that packets have to pass through to reach their final destination, hence creating significant network latency. A large number of hops also has a direct impact on the energy consumed in the interconnect for buffering, transmission, and control.

There are several critical outstanding problems of large-scale NoCs. Due to increasing number of the nodes inside the NoC and also the increase in the transaction between nodes, the rate of data transmission in the links rises. Thus, some links are used more excessively than other links which can lead to difficulty in load balancing inside the NoC. This imbalance makes some packets to take longer paths to reach the destination [24]. The long route results in increase latency, hop count, packet loss and power consumption, and decrease in throughput.

One of the ways to remove the aforementioned problems is to use routing

algorithm. Many routing algorithms were created to solve these problems, but perfect solution is still elusive. The topologies currently used are good for small size networks only. Thus it is necessary to design and develop a new topology which is appropriate for large size NoCs. Besides, an optimized routing algorithm suitable for the suggested topology must be developed. Nychis et al. in [25] have evaluated large NoCs of up to 4,096 cores, and they have shown two important issues with existing topologies in a large-scale NoC, which are high latency and low throughput.

Topologies are increasingly becoming the bottleneck that is limiting the performance of NoC [26, 27]. Indeed, for a large-scale NoC, the topology has a key impact on the performance and cost of the network [5, 7]. It is responsible for 60% to 75% of the miss latency [28].

The classical NoC topology is the two-dimensional mesh [17, 29]. It is preferred over other topologies. Since its simple implementation and the overall layout is very regular [22]. However, in spite of its advantage, the two-dimensional mesh topology is disadvantaged with congestion, high hop count, and high communication latency for large-scale NoC. Indeed, a significant disadvantage of the mesh topology is in its large communication radius which induces long path for packet delivery [7, 30, 24]. For small-scale network (up to 64 nodes [25]), mesh topology is proven to be efficient [23, 7, 25]. However, for large-scale NoC network, the performance of mesh topology degrades significantly [7, 31]. The performance of mesh topology does not scale well with network size.

The torus is also a favored topology for NoCs [32]. There are many long-range links in torus topology that may create problems in terms of performance and cost. A packet that uses a long-range link takes longer time to reach the next hop than when a packet uses a normal link [33]. In addition, each long link imposes a minimum latency and is a potential point of contention [3]. However, long-range links may improve performance by reducing number of hops [3].

Based on aforementioned disadvantages, there is a need to develop a topology with low network latency and hop count [3, 30]. The combination of mesh and ring topology have a potential to address latency and hop count and avoidance of congestion for large-scale networks [7, 8, 34].

1.3 Objective

The main goal of our research is to develop a topology with a suitable routing algorithm for large-scale NoC. The objectives in this thesis are:

- To propose a topology based on mesh clusters that reduces the number of switches, the number of hops, and latency in large-scale NoC. This thesis proposes a new NoC topology called RaMesh. RaMesh is designed based on clusters and it is suitable for large-scale NoCs that have more than 100 IP cores. Each cluster is a mesh topology. However, internal communication between IP cores inside the cluster uses the rule of ring topology. The target performance metric include a low hop count and low average network latency, and congestion avoidance.
- To propose a routing algorithm to cater for the proposed topology. The proposed routing algorithm is a combination of three existing routing algorithms, which are ring, XY, and TRANC [13] to avoid congestion and deadlock problems.

1.4 Scope

The proposed topology is a hierarchical network topology based on mesh clusters suitable for large scale NoCs with more than 100 IP cores. The structure of each cluster is the mesh topology, but the rule of the ring topology is used for internal communication among IP cores.

In this thesis, the proposed topology was evaluated for different NoC sizes, different traffic models, and different traffic ratios. The proposed topology is benchmarked in terms of average hop count and latency with clustered 2D-mesh [8], mesh, and torus under the same experimental conditions.

The proposed topology was implemented using Verilog and simulated using ModelSim. To characterize the proposed topology, we have used random, transpose, hotspot, and uniform distribution traffic models to obtain average hop count and average latency for four sizes of network with 144, 324, 576, and 900 IP cores.

The proposed routing algorithm is based on deterministic routing. We used XY

and TRANC [9] routing algorithms for mesh and torus topology respectively as they are deadlock free. The switching technique for these routing algorithms is wormhole switching.

The hardware evaluation to compare hardware cost (in terms of number of adaptive logic modules (ALM)) and estimate maximum hardware operating frequency was done based on Stratix V 5SEEBF45I4 FPGA using Quartus II software for 144 IP cores. The NoC code was written in C and translated to Verilog HDL and compiled using Quartus II version 13 software. The code was verified using Altera Modelsim. This process is explained in more detail in section 3.4.3.

1.5 Contribution of Study

This thesis proposed a topology called RaMesh, which is suitable for a large-scale NoC. A routing algorithm for RaMesh that minimizes congestion and deadlock is also proposed. The proposed topology is based on mesh clusters, is hierarchical, and has long-range links to help reduce the hop count. The performance of RaMesh is superior in terms of network latency compared to existing topologies such as clustered 2D-mesh, mesh and torus. In summary, the main contributions of this thesis are:

- The proposed NoC topology improves significantly the average hop count compared to clustered 2D-mesh, mesh and torus topologies. For example, RaMesh on average has 42.1% lower hop count compared to clustered 2D-mesh topology in tests done for various network sizes.
- In tests using RTL model for each topology, RaMesh also has superior end-to-end average latency compared to other topologies. Compared to clustered 2D-mesh, mesh, and torus topologies, Ramesh has 31.2%, 49.5%, and 41.5% lower average latency respectively.

1.6 Thesis Outline

The rest of the thesis is organized as follows.

- Chapter 2 contains literature survey on the studies of NoC, which includes topology and routing algorithm.
- Chapter 3 covers the methodology for the work done in this thesis. This also includes the general approach taken for the research done in this work, as well as tools and platform used.
- Chapter 4 presents the evaluation results of mesh topology. We also have simulated the topologies with appropriate routing algorithms under different ratios of traffic pattern.
- Chapter 5 describes the proposed topology that was designed based on hierarchical mesh topology for large-scale NoC called RaMesh. In addition, this chapter also presents a proposed routing algorithm for the proposed topology.
- Chapter 6 presents the results and analysis of the experimentations to compare average latency under different ratios of traffic and different traffic models, and hardware cost.
- Chapter 7 summarizes the thesis, re-stating contributions, and suggest directions for future research.

REFERENCES

1. Vangal, S. *et al.* (2007). An 80-tile 1.28 TFLOPS network-on-chip in 65nm CMOS. *Digest of Technical Papers, IEEE International Solid-State Circuits Conference (ISSCC 2007), San Francisco, CA, USA*. 98–99.

2. Howard, J. *et al.* (2010). A 48-core IA-32 message-passing processor with DVFS in 45nm CMOS. *Digest of Technical Papers, IEEE International Solid-State Circuits Conference (ISSCC 2010)*. 108–109.

3. Ogras, U. Y. and Marculescu, R. (2006). " It's a small world after all": NoC performance optimization via long-range link insertion. *Very Large Scale Integration (VLSI) Systems, IEEE Transactions on*, 2006. 14(7): 693–706.

4. Reza, A., Reshadi, M., Khademzadeh, A. and Bahmani, M. (2008). Norma: A hierarchical interconnection architecture for Network on Chip. *3rd International Design and Test Workshop (IDT 2008)*. 5–10.

5. Bahmani, M., Reshadi, M., Khademzadeh, A. and Reza, A. (2008). Corona: Ring-based interconnected topology for on-chip network. *3rd International Design and Test Workshop (IDT 2008)*. 199–204.

6. Samuelsson, H. and Kumar, S. (2004). Ring road NoC architecture. *Proceedings of IEEE Norchip Conference.* 16–19.

7. Bourduas, S. and Zilic, Z. (2007). A hybrid ring/mesh interconnect for network-on-chip using hierarchical rings for global routing. *First IEEE International Symposium on Networks-on-Chip, 2007. NOCS 2007.* 195–204.

8. Winter, M., Prusseit, S. and Gerhard, P. (2010). Hierarchical routing architectures in clustered 2D-mesh networks-on-chip. *Proceedings of IEEE SoC Design Conference (ISOCC 2010).* 388–391.

9. Rahmati, D., Sarbazi-Azad, H., Hessabi, S. and Kiasari, A. E. (2012). Power-efficient deterministic and adaptive routing in torus networks-on-chip. *Microprocessors and Microsystems*, 2012. 36(7): 571–585.

10. Monemi, A., Ooi, C. Y. and Marsono, M. N. (2015). Low latency Network-on-Chip router microarchitecture using request masking technique.

International Journal of Reconfigurable Computing, vol. 2015, Article ID 570836, 13 pages. doi:10.1155/2015/570836.

11. Jerger, N. E. and Peh, L.-S. (2009). On-chip networks. *Synthesis Lectures on Computer Architecture*, 2009. 4(1): 1–141.

12. Nicopoulos, C. *et al.* (2006). ViChaR: A dynamic virtual channel regulator for network-on-chip routers. *Microarchitecture, 2006. MICRO-39. 39th Annual IEEE/ACM International Symposium on*. IEEE. 333–346.

13. Seiler, L. *et al.* (2009). Larrabee: A many-core x86 architecture for visual computing. *IEEE micro*. (1): 10–21.

14. Borkar, S. (2007). Thousand core chips: a technology perspective. *Proceedings of ACM 44th annual Design Automation Conference*. 746–749.

15. Owens, J. D. *et al.* (2007). Research challenges for on-chip interconnection networks. *IEEE micro*. (5): 96–108.

16. Benini, L. and De Micheli, G. (2002). Networks on chips: a new SoC paradigm. *Computer, IEEE*. 35(1): 70–78.

17. Dally, W. J. and Towles, B. (2001). Route packets, not wires: On-chip interconnection networks. *IEEE Conference on Design Automation Conference*. 684–689.

18. Intel. Single-chip cloud computer. 2011. URL `http://goo.gl/SfgfN`.

19. Tilera. Announces the world's first 100-core processor with the new tile-gx family. 2011. URL `.http://goo.gl/K9c85`.

20. Vanderbauwhede, W. (2010). Scientists squeeze more than 1,000 cores on to computer chip. 2010. URL `http://goo.gl/KdBbW`.

21. Liu, Y., Liu, P., Jiang, Y., Yang, M., Wu, K., Wang, W. and Yao, Q. (2010). Building a multi-FPGA-based emulation framework to support networks-on-chip design and verification. *International Journal of Electronics*, 2010. 97(10): 1241–1262.

22. Chiu, G.-M. (2000). The odd-even turn model for adaptive routing. *Parallel and Distributed Systems, IEEE Transactions on*, 2000. 11(7): 729–738.

23. Taylor, M. B. *et al.* (2002). The Raw microprocessor: A computational fabric for software circuits and general-purpose programs. *Micro, IEEE*, 2002. 22(2): 25–35.

24. Zhang, Y., Lu, Z., Jantsch, A., Li, L. and Gao, M. (2009). Towards hierarchical cluster based cache coherence for large-scale network-on-

chip. *Proceedings of the 4th IEEE International Conference on Design & Technology of Integrated Systems in Nanoscale Era*.

25. Nychis, G. P., Fallin, C., Moscibroda, T., Mutlu, O. and Seshan, S. (2012). On-chip networks from a networking perspective: Congestion and scalability in many-core interconnects. *ACM SIGCOMM computer communication review*, 2012. 42(4): 407–418.

26. Bjerregaard, T. and Mahadevan, S. (2006). A survey of research and practices of network-on-chip. *ACM Computing Surveys (CSUR)*, 2006. 38(1): 1.

27. Salminen, E., Kulmala, A. and Hamalainen, T. D. (2008). Survey of network-on-chip proposals. *white paper, OCP-IP*, 2008: 1–13.

28. Sanchez, D., Michelogiannakis, G. and Kozyrakis, C. (2010). An analysis of on-chip interconnection networks for large-scale chip multiprocessors. *ACM Transactions on Architecture and Code Optimization (TACO)*, 2010. 7(1): 4.

29. Fu, W., Chen, T. and Liu, L. (2014). Energy-efficient Hybrid Optical-Electronic Network-on-Chip for Future Many-core Processors. *Elektronika ir Elektrotechnika*, 2014. 20(3): 83–86.

30. Lis, M., Ren, P., Cho, M. H., Shim, K. S., Fletcher, C. W., Khan, O. and Devadas, S. (2011). Scalable, accurate multicore simulation in the 1000-core era. *IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS 2011)*. 175–185.

31. Qian, Z., Bogdan, P., Wei, G., Tsui, C.-Y. and Marculescu, R. (2012). A traffic-aware adaptive routing algorithm on a highly reconfigurable network-on-chip architecture. *Proceedings of the eighth IEEE/ACM/IFIP international conference on Hardware/software codesign and system synthesis*. ACM. 2012. 161–170.

32. Ezhumalai, P., Chilambuchelvan, A. and Arun, C. (2011). Novel NoC Topology Construction for High-Performance Communications. *Journal of Computer Networks and Communications*. 2011.

33. Grecu, C., Pande, P. P., Ivanov, A. and Saleh, R. (2005). Timing analysis of network on chip architectures for MP-SoC platforms. *Microelectronics Journal*, 2005. 36(9): 833–845.

34. Johari, S., Kumar, A. and Sehgal, V. K. (2015). Heterogeneous and Hybrid Clustered Topology for Networks-on-Chip: 183–187.

35. Nicopoulos, C., Narayanan, V. and Das, C. R. *(2009). Network-on-Chip Architectures: A Holistic Design Exploration*. vol. 45. Springer Science &

Business Media.

36. Bell, S. *et al.* (2008). Tile64-processor: A 64-core soc with mesh interconnect. *Digest of Technical Papers, IEEE International Solid-State Circuits Conference, (ISSCC 2008)*. 88–98.

37. Maeurer, T. and Shippy, D. (2005). Introduction to the Cell multiprocessor. *IBM journal of Research and Development*, 2005. 49(4): 589–604.

38. Taylor, M. B. *et al.* (2002). The Raw microprocessor: A computational fabric for software circuits and general-purpose programs. *Micro, IEEE*, 2002. 22(2): 25–35.

39. Mutlu, O. and Witchel, E. *(2011). Network-on-Chip Architectures for Scalability and Service Guarantees*. Ph.D. Thesis. The University of Texas at Austin. 2011.

40. Bertozzi, D. *et al.* (2005). NoC synthesis flow for customized domain specific multiprocessor systems-on-chip. *IEEE Transactions on Parallel and Distributed Systems*, 2005. 16(2): 113–129.

41. Kim, J. *et al.* (2007). A novel dimensionally-decomposed router for on-chip communication in 3D architectures. 2007. 35(2): 138–149.

42. Mullins, R., West, A. and Moore, S. (2004). Low-latency virtual-channel routers for on-chip networks. 2004. 32(2): 188.

43. Kumar, A. *et al.* (2007). A 4.6 Tbits/s 3.6 GHz single-cycle NoC router with a novel switch allocator in 65nm CMOS. *25th IEEE International Conference on Computer Design, (ICCD 2007)*. 63–70.

44. Peh, L.-S. and Dally, W. J. (2001). A delay model for router microarchitectures. *Micro, IEEE*. 21(1): 26–34.

45. Peh, L.-S. and Dally, W. J. (2000). Flit-reservation flow control. *Proceedings IEEE Sixth International Symposium on High-Performance Computer Architecture (HPCA-6 2000)*. IEEE. 73–84.

46. Kim, J. and other. (2006). A gracefully degrading and energy-efficient modular router architecture for on-chip networks. *ACM SIGARCH Computer Architecture News*. IEEE Computer Society. vol. 34. 4–15.

47. Wang, H., Peh, L.-S. and Malik, S. (2003). Power-driven design of router microarchitectures in on-chip networks. *Proceedings of the 36th annual IEEE/ACM International Symposium on Microarchitecture*. IEEE Computer Society. 2003. 105.

48. Hu, J., Ogras, U. and Marculescu, R. (2006). System-level buffer allocation

for application-specific networks-on-chip router design. *Computer-Aided Design of Integrated Circuits and Systems, IEEE Transactions on*, 2006. 25(12): 2919–2933.

49. Saastamoinen, I., Alho, M. and Nurmi, J. (2003). Buffer implementation for Proteo network-on-chip. *Proceedings of the 2003 International Symposium on Circuits and Systems (ISCAS 2003).* vol. 2. II–113.

50. Lee, K., Lee, S.-J., Kim, D., Kim, K., Kim, G., Kim, J. and Yoo, H.-J. (2005). Networks-on-chip and networks-in-package for high-performance SoC platforms. *Asian Solid-State Circuits Conference, 2005*. IEEE. 485–488.

51. Choudhary, N. (2012). Bursty Communication Performance Analysis of Network-on-Chip with Diverse Traffic Permutations. *International Journal of Soft Computing and Engineering (IJSCE) ISSN*, 2012: 2231–2307.

52. Peh, L.-S. and Dally, W. J. (2001). A delay model and speculative architecture for pipelined routers. *High-Performance Computer Architecture, 2001. HPCA. The Seventh International Symposium on*. IEEE. 255–266.

53. Nousias, I. and Arslan, T. (2006). Wormhole routing with virtual channels using adaptive rate control for network-on-chip (NoC). *First IEEE NASA/ESA Conference on Adaptive Hardware and Systems, (AHS 2006)*. 420–423.

54. Pham, D. *et al.* (2005). The design and implementation of a first-generation CELL processor-a multi-core SoC. *IEEE International Conference on Integrated Circuit Design and Technology, (ICICDT 2005)*. 49–52.

55. Gratz, P., Kim, C., McDonald, R., Keckler, S. and Burger, D. Implementation and Evaluation of On-Chip Network Architectures. *International Conference on Computer Design, (ICCD 2006)*. 477–484. doi:10.1109/ICCD.2006.4380859.

56. Gratz, P. *et al.* (2006). Implementation and evaluation of on-chip network architectures. *Computer Design, 2006. ICCD 2006. International Conference on*. IEEE. 477–484.

57. Lee, K., Lee, S.-J. and Yoo, H.-J. (2006). Low-power network-on-chip for high-performance SoC design. IEEE. 2006, vol. 14. 148–160.

58. Kim, J., Dally, W. J. and Abts, D. (2007). Flattened butterfly: a cost-efficient topology for high-radix networks. *ACM SIGARCH Computer Architecture News*, 2007. 35(2): 126–137.

59. Ho, W. H. and Pinkston, T. M. (2003). A methodology for designing

efficient on-chip interconnects on well-behaved communication patterns. *The Ninth IEEE International Symposium on High-Performance Computer Architecture, (HPCA-9 2003)*. 377–388.

60. Ogras, U. Y. and Marculescu, R. (2005). Energy-and performance-driven NoC communication architecture synthesis using a decomposition approach. *Proceedings of the conference on Design, 2005. Automation and Test in Europe-Volume 1*. IEEE Computer Society. 352–357.

61. Pinto, A., Carloni, L. P. and Sangiovanni-Vincentelli, A. L. (2003). Efficient synthesis of networks on chip. *21st IEEE International Conference on Computer Design*. 146–150.

62. Srinivasan, K., Chatha, K. S. and Konjevod, G. (2006). Linear-programming-based techniques for synthesis of network-on-chip architectures. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2006. 14(4): 407–420.

63. Karim, F., Nguyen, A. and Dey, S. (2002). An interconnect architecture for networking systems on chips. *IEEE micro*, 2002. (5): 36–45.

64. Baboli, M., Husin, N. S. and Marsono, M. N. (2014). A Comprehensive Evaluation of Direct and Indirect Network-On-Chip Topologies. *Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management. Bali*. 2081–2090.

65. Kumar, S. *et al.* (2002). A network on chip architecture and design methodology. *Proceedings IEEE Computer Society Annual Symposium on VLSI*. 105–112.

66. Rantala, V., Lehtonen, T. and Plosila, J. *(2006). Network on chip routing algorithms*. 2006.

67. Dally, W. J. and Towles, B. (2001). Route packets, not wires: On-chip interconnection networks. *Design Automation Conference, 2001. Proceedings*. IEEE. 684–689.

68. Hu, W.-H., Lee, S. E. and Bagherzadeh, N. *(2011). Design and Application of Advanced Network-on-Chip Architecture*. Ph.D. Thesis. UNIVERSITY OF CALIFORNIA, IRVINE. 2011.

69. Kim, J., Balfour, J. and Dally, W. (2007). Flattened butterfly topology for on-chip networks. *Proceedings of the 40th Annual IEEE/ACM International Symposium on Microarchitecture*. IEEE Computer Society. 2007. 172–182.

70. Kim, J., Balfour, J. and Dally, W. Flattened Butterfly Topology for On-Chip

Networks. *Computer Architecture Letters*, 2007. 6(2): 37–40. ISSN 1556-6056. doi:10.1109/L-CA.2007.10.

71. Kim, J., Dally, W. J., Scott, S. and Abts, D. (2008). Technology-driven, highly-scalable dragonfly topology. *ACM SIGARCH Computer Architecture News*. IEEE Computer Society. 2008, vol. 36. 77–88.

72. Lee, K. and other. (2004) A 51mW 1.6GHz on-chip network for low-power heterogeneous SoC platform. *Digest of Technical Papers, IEEE International Solid-State Circuits Conference (ISSCC 2004)*. 152–518 Vol.1.

73. Coppola, M., Locatelli, R., Maruccia, G., Pieralisi, L. and Scandurra, A. (2004). Spidergon: a novel on-chip communication network. *International Symposium on System-on-Chip, 2004*. 15.

74. Grot, B., Hestness, J., Keckler, S. W. and Mutlu, O. (2009). Express cube topologies for on-chip interconnects. *IEEE 15th International Symposium on High Performance Computer Architecture (HPCA 2009)*. 163–174.

75. Das, R., Eachempati, S., Mishra, A. K., Narayanan, V. and Das, C. R. (2009). Design and evaluation of a hierarchical on-chip interconnect for next-generation CMPs. *IEEE 15th International Symposium on High Performance Computer Architecture (HPCA 2009)*. 175–186.

76. Ho, W. H. and Pinkston, T. M. (2003). A methodology for designing efficient on-chip interconnects on well-behaved communication patterns. *IEEE Ninth International Symposium on High-Performance Computer Architecture, (HPCA-9 2003)*. 377–388.

77. Guerrier, P. and Greiner, A. (2000). A generic architecture for on-chip packet-switched interconnections. *Proceedings of the conference on Design, automation and test in Europe*. ACM. 2000. 250–256.

78. Hossain, H., Akbar, M. M. and Islam, M. M. (2005). Extended-butterfly fat tree interconnection (EFTI) architecture for network on chip. *IEEE Pacific Rim Conference on Communications, Computers and signal Processing, (PACRIM 2005)*. 613–616.

79. Pande, P. P., Grecu, C., Ivanov, A. and Saleh, R. (2003). Design of a switch for network on chip applications. *Proceedings of IEEE International Symposium on Circuits and Systems, (ISCAS 2003)*. vol. 5. V–217.

80. Chang, M., Cong, J., Kaplan, A., Liu, C., Naik, M., Premkumar, J., Reinman, G., Socher, E. and Tam, S.-W. (2008). Power reduction of CMP communication networks via RF-interconnects. *Proceedings of the 41st annual IEEE/ACM International Symposium on Microarchitecture*. IEEE

Computer Society. 376–387.

81. BEIDU, S. K. *(2009). Formal Verification Of A Network On Chip*. Ph.D. Thesis. African University of Science and Technology. 2009.

82. Wang, C., Li, X., Zhang, J., Zhou, X. and Wang, A. (2012). A star network approach in heterogeneous multiprocessors system on chip. *The Journal of Supercomputing*, 2012. 62(3): 1404–1424.

83. Saneei, M., Afzali-Kusha, A. and Navabi, Z. (2006). Low-power and low-latency cluster topology for local traffic NoCs. *Proceedings IEEE International Symposium on Circuits and Systems, (ISCAS 2006)*. 1727–1730.

84. Jianmin XIE, G. L. S. L., Zhaoshan LIU. (2015). Hybrid connection-based mesh topology and pseudo pdaptive routing algorithm for network on chip. *Journal of Computational Information Systems*, 2015. 11(6): 1997–2005.

85. Kumar, R. (2015). Network-on-Chip Architecture Based on Cluster Method. *International Journal of Scientific Engineering and Research, (IJSER 2015)*. 3(3): 61–65.

86. Sanju, V., Chiplunkar, N., Khalid, M., Joshi, S. and Nirmala, J. (2013). A performance study of 2D mesh and torus for network-on-chip-based system. *Proc. Int. Conf. Emerging Research in Computing, Information, Communication and Applications, Bangalore, India (Elsevier, Amsterdam, 2013)*. 2013. 47–51.

87. Sheibanyrad, A. and other. *(2011). 3D integration for NoC-based SoC Architectures*. Springer.

88. Feero, B. S. and Pande, P. P. (2009). Networks-on-chip in a three-dimensional environment: A performance evaluation. *IEEE Transactions on Computers*, 2009. 58(1): 32–45.

89. Matsutani, H., Koibuchi, M. and Amano, H. (2007). Tightly-coupled multi-layer topologies for 3-D NoCs. *IEEE International Conference on Parallel Processing, (ICPP 2007)*. 75–75.

90. Park, D. and other. (2008). MIRA: A multi-layered on-chip interconnect router architecture. *ACM SIGARCH Computer Architecture News*. IEEE Computer Society. 2008, vol. 36. 251–261.

91. Ramanujam, R. S. and Lin, B. (2009). A layer-multiplexed 3D on-chip network architecture. *Embedded Systems Letters, IEEE*, 2009. 1(2): 50–55.

92. Ben Ahmed, A., Ben Abdallah, A. and Kuroda, K. (2010). Architecture

and design of efficient 3D network-on-chip (3D NoC) for custom multicore SoC. *IEEE International Conference on Broadband, Wireless Computing, Communication and Applications (BWCCA 2010)*. 67–73.

93. Mori, K., Abdallah, A. B. and Kuroda, K. (2009). Design and evaluation of a complexity effective network-on-chip architecture on FPGA. *Proc. of The 19th Intelligent System Symposium (FAN 2009)*. 318–321.

94. Li, F., Nicopoulos, C., Richardson, T., Xie, Y., Narayanan, V. and Kandemir, M. (2006). Design and management of 3D chip multiprocessors using network-in-memory. *ACM SIGARCH Computer Architecture News*. IEEE Computer Society. 2006, vol. 34. 130–141.

95. Richardson, T. D., Nicopoulos, C., Park, D., Narayanan, V., Xie, Y., Das, C. and Degalahal, V. (2006). A hybrid SoC interconnect with dynamic TDMA-based transaction-less buses and on-chip networks. *VLSI Design, Held jointly with IEEE 5th International Conference on Embedded Systems and Design*. 8–pp.

96. Daneshtalab, M. *et al.* (2011). Exploring Adaptive Implementation of On-Chip Networks. 2011.

97. Ye, T. T., Benini, L. and De Micheli, G. (2004). Packetization and routing analysis of on-chip multiprocessor networks. *Journal of Systems Architecture*, 2004. 50(2): 81–104.

98. Pande, P. P., Grecu, C., Jones, M., Ivanov, A. and Saleh, R. (2005). Performance evaluation and design trade-offs for network-on-chip interconnect architectures. *IEEE Transactions on Computers*, 2005. 54(8): 1025–1040.

99. Ingle, . V. V. and A.gaikwad, M. Article: Review of Mesh Topology of NoC Architecture using Source Routing Algorithms. *IJCA Special Issue on Recent Trends in Engineering Technology*, 2013. RETRET: 30–34. Full text available.

100. Baboli, M. and Husin, N. S. (2014). An Empirical Evaluation of Topologies for Large Scale NoC. *TELKOMNIKA Indonesian Journal of Electrical Engineering*, 2014. 12(12): 8133–8139.

101. Palesi, M. and Daneshtalab, M. *(2014). Routing algorithms in networks-on-chip*. Springer. 2014.

102. Kumar, A., Peh, L.-S., Kundu, P. and Jha, N. K. (2007). Express virtual channels: towards the ideal interconnection fabric. 2007. 35(2): 150–161.

103. Ramanujam, R. S. and Lin, B. (2008). Near-optimal oblivious routing on three-dimensional mesh networks. *Computer Design, 2008. ICCD 2008. IEEE International Conference on*. IEEE. 134–141.

104. Liu, H. and Zhang-Shen, R. (2005). On direct routing in the valiant load-balancing architecture. *Global Telecommunications Conference, (GLOBECOM'05)*. IEEE. vol. 2. 6–26.

105. Seo, D., Ali, A., Lim, W.-T., Rafique, N. and Thottethodi, M. (2005). Near-optimal worst-case throughput routing for two-dimensional mesh networks. 2005. 33(2): 432–443.

106. Zhang, W. and other. (2009). Comparison research between xy and odd-even routing algorithm of a 2-dimension 3x3 mesh topology network-on-chip. *IEEE WRI Global Congress on Intelligent Systems, (GCIS'09)*. vol. 3. 329–333.

107. Liu, H. and Zhang-Shen, R. (2005). On direct routing in the valiant load-balancing architecture. *Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE*. IEEE. vol. 2. 6–pp.

108. Dally, W. J. and Towles, B. P. *(2004). Principles and practices of interconnection networks*. Elsevier. 2004.

109. Duato, J., Yalamanchili, S. and Ni, L. M. *(2003). Interconnection networks: An engineering approach*. Morgan Kaufmann. 2003.

110. Dally, W. J. (1992). Virtual-channel flow control. *IEEE Transactions on Parallel and Distributed Systems*, 1992. 3(2): 194–205.

111. Taylor, M. B., Lee, W., Amarasinghe, S. P. and Agarwal, A. (2005). Scalar operand networks. *IEEE Transactions on Parallel and Distributed Systems*, 2005. 16(2): 145–162.

112. Lu, Z., Liu, M. and Jantsch, A. (2007). Layered switching for networks on chip. *Design Automation Conference, 44th ACM/IEEE (DAC 2007)*. 122–127.

113. Kumar, A., Peh, L.-S., Kundu, P. and Jha, N. K. (2007). Express virtual channels: towards the ideal interconnection fabric. 2007. 35(2): 150–161.

114. Andreas Hansson, K. G. and Radulescu, A. A Unified Approach to Mapping and Routing on a Network-on-Chip for Both Best-Effort and Guaranteed Service Traffic. *VLSI Design*, 2007. 2007. doi:10.1155/2007/68432. URL `http://doi:10.1155/2007/68432`.

115. Salminen, E. (2010). On design and comparison of on-chip networks.

Tampereen teknillinen yliopisto. Julkaisu-Tampere University of Technology. Publication; 872, 2010.

116. Tang, M. and Lin, X. (2011). Injection Level Flow Control for Networks-on-Chip (NoC). *J. Inf. Sci. Eng.*, 2011. 27(2): 527–544.

117. Navaridas, J. and other. (2012). Reservation-based Network-on-Chip timing models for large-scale architectural simulation. *Sixth IEEE/ACM International Symposium on Networks on Chip (NoCS)*. 91–98.

118. Grot, B., Hestness, J., Keckler, S. W. and Mutlu, O. (2011). Kilo-NOC: a heterogeneous network-on-chip architecture for scalability and service guarantees. *ACM SIGARCH Computer Architecture News*, 2011. 39(3): 401–412.

119. Mirza-Aghatabar, M., Koohi, S., Hessabi, S. and Pedram, M. (2007). An empirical investigation of mesh and torus NoC topologies under different routing algorithms and traffic models. *IEEE 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools (DSD 2007)*. 19–26.

120. Lee, J. W., Ng, M. C. and Asanovic, K. (2008). Globally-synchronized frames for guaranteed quality-of-service in on-chip networks. *IEEE 35th International Symposium on Computer Architecture (ISCA 2008)*. 89–100.

121. Rahmani, A.-M., Afzali-Kusha, A. and Pedram, M. (2009). A novel synthetic traffic pattern for power/performance analysis of network-on-chips using negative exponential distribution. *Journal of Low Power Electronics*, 2009. 5(3): 396–405.

122. Xu, Y., Jiang, T., Yang, M. and Chao, H. J. *(2010). Reservation Cut-through Switching Allocation for High-Radix Clos Network on Chip*. Technical report. Technical Report. http://eeweb. poly. edu/chao/publications/TechReports. html.

123. Palesi, M., Patti, D. and Fazzino, F. (2010). Noxim: Network on Chip (NoC) Simulator. *University of Catania (Italy)*. URL `http://www.noxim.org/`.

124. Puente, V., Gregorio, J. A. and Beivide, R. (2002). SICOSYS: an integrated framework for studying interconnection network performance in multiprocessor systems. *Proceedings. 10th Euromicro Workshop on Parallel, Distributed and Network-based Processing*. IEEE. 2002. 15–22.

125. Wang, D., Lo, C., Vasiljevic, J., Jerger, N. E. and Steffan, J. G. (2014). DART: a programmable architecture for NoC simulation on FPGAs. *IEEE Transactions on Computers*, 2014. 63(3): 664–678.

126. Becker, D. U. *(2012). Efficient microarchitecture for network-on-chip routers*. Ph.D. Thesis. Stanford University. 2012.

127. Lu, Y., McCanny, J. and Sezer, S. (2011). Exploring Virtual-Channel architecture in FPGA based Networks-on-Chip. *Proceedings of IEEE International, SOC Conference (SOCC)*. 302–307. doi:10.1109/SOCC.2011.6085089.

128. Papamichael, M. K. and Hoe, J. C. (2012). CONNECT: re-examining conventional wisdom for designing nocs in the context of FPGAs. *Proceedings of the ACM/SIGDA international symposium on Field Programmable Gate Arrays*. ACM. 37–46.

129. Reshadi, M., Khademzadeh, A., Reza, A. and Bahmani, M. (2013). A novel mesh architecture for on-chip networks. *http://www. design-reuse. com/articles/23347/on-chipnetwork*.

130. Teimouri, N., Modarressi, M., Tavakkol, A. and Sarbazi-Azad, H. (2011). Energy-optimized on-chip networks using reconfigurable shortcut paths. In: *Architecture of Computing Systems-ARCS 2011*. Springer. 231–242.

131. Tamir, Y. and Chi, H.-C. (1993). Symmetric crossbar arbiters for VLSI communication switches. *IEEE Transactions on Parallel and Distributed Systems*, 1993. 4(1): 13–27.

132. Becker, D. U. and Dally, W. J. (2009). Allocator implementations for network-on-chip routers. *Storage and Analysis, Proceedings of the Conference on High Performance Computing Networking*. 1–12.

133. Galles, M. (1997). Spider: A high-speed network interconnect. *Micro, IEEE*, 1997. 17(1): 34–39.

134. Peh, L.-S. and Dally, W. J. (2001). A delay model and speculative architecture for pipelined routers. *IEEE Seventh International Symposium on High-Performance Computer Architecture (HPCA 2001)*. 255–266.