HiWA: A Hierarchical Wireless Network-on-Chip Architecture

Amin Rezaei[1], Farshad Safaei[1], Masoud Daneshtalab[2,3], and Hannu Tenhunen[3]

[1] Department of Computer Engineering, Shahid Beheshti University (SBU), Tehran, Iran
me@aminrezaei.com, f_safaei@sbu.ac.ir
[2] Department of Information Technology, University of Turku (UTU), Turku, Finland
masdan@utu.fi
[3] Department of Electronic Systems, Royal Institute of Technology (KTH), Stockholm, Sweden
hannu@kth.se

Abstract— **Due to high latency and high power consumption in long hops between operational cores of NoCs, the performance of such architectures has been limited. In order to fill the gap between computing requirements and efficient communications, a new technology called Wireless Network-on-Chip (WNoC) has been emerged. Employing wireless communication links between cores, the new technology has reasonably increased the performance of NoC. However, wireless transceivers along with associated antenna impose considerable area and power overheads in WNoCs. Thus, in this paper, we introduce a hierarchical WNoC called Hierarchical Wireless-based Architecture (HiWA) to use the wireless resources optimally. In the proposed approach the network is divided into subnets where intra-subnet nodes communicate through wire links while inter-subnet communications are almost handled by single-hop wireless links. On top of that, we have also defined performance evaluation parameters. Simulation results show that the proposed architecture reduces average packet latency 16% and power consumption 14% in comparison with its conventional counterparts.**

Keywords— System-on-Chip; Network-on-Chip; Wireless Network-on-Chip; Architecture; Latency; Power Consumption

I. INTRODUCTION

Electronic industry is changing in so swift a fashion which is, of course, the result of the permanent demand for innovation and technological advancement. Many believe that emerging of System-on-Chip (SoC) has been an effective solution for advancement of the industry in the form of Moore's law [1-3]. As it is described in [4] and [5], SoC is a very compressed system with a complex function which is comprised of many independent operational units called IP Cores. However, according to [6] and [7] there are two major limitations in SoC. The first limitation is the latency of long wires to connect IP cores which by the decrease of technology scale, is changed into a bottleneck. The second limitation is integration of IP cores with different standards and various producers on a single chip. The method of designing based on information which is presented as a solution for the existing challenges in designing SoC [1], [6], [8] and [9] has a key feature according to which, more efficient and at the same time simpler systems are tried to be provided. This key feature includes separating information from communication. Network-on-Chip (NoC) is a very compressed system based on network technology in which IP cores are connected to each other according to a communication infrastructure consisted of a switch, router and communication links called Interconnection Network. According to [10] and [11], despite the fact that NoC has a lot of advantages because of the presence of long wires in two-sided metal interconnects, it suffers from high latency and high power consumption. At the present time, because of extremely high energy costs, single-core high-frequency processors are less considered and processors manufacturers are moving toward designing multi-core chips. Recent studies suggested that systems with hard-wired metal interconnects will finally face serious limitations and they can prevent the development of Ultra-Large Scale Integration (ULSI) systems [12]. Therefore, alternative technologies such as Wireless Network-on-Chip, 3D Network-on-Chip and Photonic Network-on-Chip were introduced [13-15]. In this paper, a hybrid wireless network-on-chip architecture called Hierarchical Wireless-based Architecture (HiWA) along with performance evaluation parameters is introduced. HiWA is based on a 2D mesh that is divided into square subnets. Then, in each subnet if necessary one of the baseline routers is replaced by a wireless router which has wireless connections with wireless routers in neighbors' subnets. Moreover, A wireless router placement algorithm as well as a routing algorithm is proposed for HiWA. Reducing number of wireless routers to obtain a trade-off between latency and power consumption has been targeted. The rest of the paper is organized as follows. Section II review backgrounds and related works. The HiWA architecture is proposed in section III. Section IV evaluates the proposed architecture based on average hop count, normalized latency and normalized power consumption parameters. Finally, some conclusions are given in section V.

II. BACKGROUNDS AND RELATED WORKS

Recent advances in silicon integrated circuit technology, has enabled the integration of small and low-cost transceivers antennas on a single chip, which results in introducing Wireless Network-on-Chip (WNoC). A low Terahertz (324

GHz) frequency generator is realized in 90 nm CMOS [16]. Moreover a signal source operating near 410 GHz that is fabricated using low-leakage transistors in a 6 M 45 nm digital CMOS technology is reported [17]. Based on these techniques, the output power level of the on-chip millimeter-wave generator is expected to be as high as -1.4 dBm in the 32 nm CMOS process, which is large enough for on-chip short distance communication [18]. Following the rule of thumb in RF design, the maximum available bandwidth is 10% of the carrier frequency. According to this experimental estimation, up to 16 channels can be available for WNoC in the range of 100 to 500 GHz. With recent developments of millimeter-wave circuits, bandwidths of hundred GHz will be reachable in near future. In addition to the bandwidth, WNoC requires low-power on-chip wireless transceivers. Silicon Mach-Zehnder electro-optic modulator at data rates up to 10 Gb/s with low RF power consumption of only 5 pJ/bit [19] is commercially available. Conventional NoCs utilize multihop communication. To improve performance, express virtual channels are introduced in [20]. Beyond traditional wired interconnect solutions, different emerging approaches including 3D Network-on-Chip and Photonic Network-on-Chip were proposed [14] and [15]. In addition, the design of a wireless NoC based on CMOS UWB technology is introduced [21]. The antennas used in [21] achieve a transmission range of 1mm; therefore, for a typically die area of 20mm×20mm, this architecture requires multihop communication. Using miniaturized on-chip antennas as an enabling technology, a hybrid wireless NoC (WiNoC) is designed [11]. Besides, design methodologies and technology requirements for scalable WiNoC architecture is discussed [10]. Complex design steps used in WiNoC require high hardware overhead to implement in micro-architecture level. In this paper, a hierarchical wireless network-on-chip architecture is introduced. Although several aspects of the proposed architecture are addressed in the paper, HiWA is a flexible platform that different placement and routing algorithms can be applied to it without changing the architectural structure. With optimal placement of wireless routers a trade-off between latency and power consumption parameters is obtained. The trade-off criteria are easily changeable according to the application where HiWA will be used.

III. PROPOSED ARCHITECTURE

A. Topology

The proposed Hierarchical Wireless-based Architecture (HiWA) is based on 2D mesh NoC. Topology is shaped in such a way that first 2D mesh is divided into square subnets. Then, in each subnet if necessary one of baseline routers is replaced by a wireless router which has wireless links with wireless routers in neighbors' subnets. Wireless routers are capable of transmitting packets in both wired and wireless channels. The purpose of wireless routers placement is decreasing number of mapped wireless routers as much as possible, while achieving high-performance.

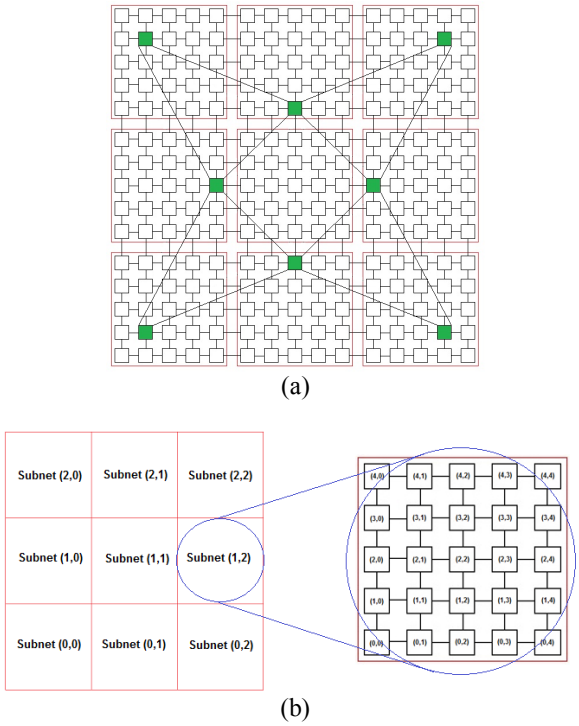


Fig. 1. (a) Illustration of a 225-node HiWA (Nodes without color are baseline routers and dark nodes are wirless routers) (b) 225-node HiWA physical structure and its addressing method

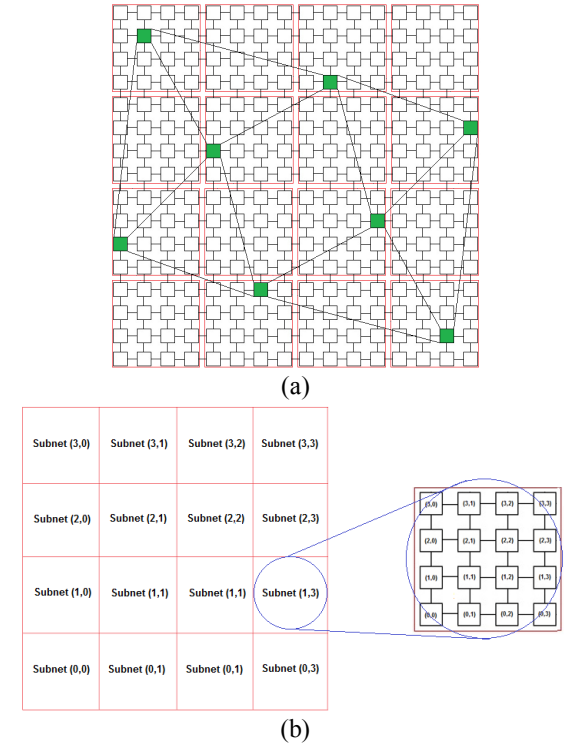


Fig. 2. (a) Illustration of a 256-node HiWA (Nodes without color are baseline routers and dark nodes are wirless routers) (b) 256-node HiWA physical structure and its addressing method

B. Adressing

Addressing method in HiWA is formed of X(subnet), Y(subnet), X(local), and Y(local) in which the first two fields specify the location of subnet and the next two fields determine the place of local node (router) in the subnet. Separating local and subnet address fields results in a simple design of hierarchical systems. Besides, it decreases hardware complexity of routers.

C. Wireless Router Placement Algorithm

The placement of the wireless routers is important as this is responsible for establishing high-speed, low-power interconnects on the network, which will eventually result in performance gains. The key task is minimizing the average distance between nodes. Once the network is initialized, an optimization by means of Ant Colony Optimization (ACO) heuristic is performed [22]. ACO is a population based approach for solving combinatorial optimization problems that is inspired by the foraging behavior of ants and their inherent ability to find the shortest path from a food source to their nest [23]. The optimization step is necessary as the random initialization might not produce the optimal network topology.

Fig. 1a shows a 225-node HiWA which is divided into 9 subnets of 5×5 nodes. 8 wireless routers are used and subnet (1,1) does not have any wireless router. According to an experimental estimation which is obtained through simulation results, for NoCs of 225 and 256 nodes, using 8 wireless routers is the best option to establish a trade-off between latency and power consumption parameters. Fig. 2a shows a 256-node HiWA which is divided into 16 subnets of 4×4 nodes. Again 8 wireless routers are used and half of the subnets don't have any wireless router. Fig. 1b and Fig. 2b represent addressing method of the HiWA. These architectures are easily scalable to 900 and 1024 nodes by quadruplicate each one and use 24 wireless routers.

ALGORITHM I. HIWA ROUTING ALGORITHM

input : Incoming packet and buffer usage of local wireless router
output: Routing path
s: source node
d: destination node
WR_s: source nearest wireless router
WR_d: destination nearest wireless router
H_B: traveling distance between s and d without using wireless links
H_W: traveling distance between s and d using both wireless and wired links

if *s and d in the same subnet* **then**
  Route packets from s to d using wired path
else
  if $H_w < H_B + \delta$ **then**
    Route packets in three step:
      a Route packets from s to WR_S using wired links
      b Route packets from WR_S to WR_d using wireless links
      c Route packets from WR_d to d using wired links
  else
    Route packets from s to d using wired path
  end
end

D. Routing Algorithm

In HiWA when one node send a packet to another node, it is possible to be transmitted using only wired paths or only wireless paths or a combination of wired and wireless paths; therefore, an efficient decision is required to select an appropriate path. This can be seen as a hybrid network that has been characterized by adding express paths (wireless links) to a 2D mesh NoC. One of the benefits of partitioning is that inta-subnet communications are handled through wire paths while inter-subnets communications are a function of hop counts and congestion. Algorithm 1 represents pseudo-code of the routing algorithm. Since each wireless router is shared by several nodes, there is a possibility of congestion. In order to balance utilization of wired and wireless networks, a balance parameter called δ is added in the 4th line of Algorithm 1. The values of δs depend on the network size and utilization of wireless network. In each router, there is a table that keeps δs based on different situations. In general, the larger the network size or the higher the link utilization are, the larger the δs. Dynamic allocation of wireless paths, ultimate to decrease congestion; Moreover, in a light traffic, more packets can utilize available wireless paths.

E. Deadlock Avoidance

Deadlock avoidance is one of the important issues should be addressed in such networks using wormhole switching. Although using XY routing in each wired and wireless network guarantees deadlock free packet transmission, when packets transmit through both wired and wireless paths, there is a possibility of graph dependency and deadlock as it is shown in Fig. 3.

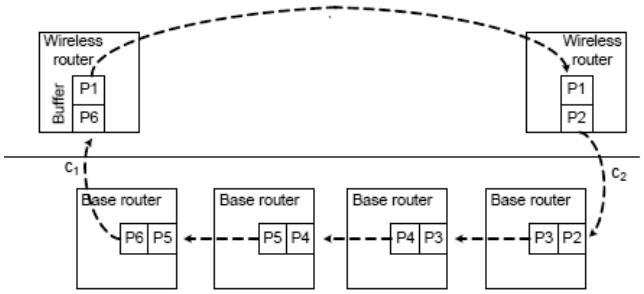


Fig. 3. A deadlock prone situation in HiWA [18]

In order to overcome the problem, virtual channels are taken into account. In each input port of the routers two sets of virtual channels are used. One of them is for traffic transmission using nearest wireless router (the 6th line of Algorithm 1) while the other one is for traffic transmission using wired links (the 2nd line of Algorithm 1) or traffic transmission of the wireless router to the destination node (the 8th line of Algorithm 1).

IV. SIMULATIN RESULTS

To evaluate the performance and feasibility of the proposed architecture, in this section 225-node HiWA (Fig. 1) and 256-node HiWA (Fig. 2) are compared with basic 2D mesh NoC based on average hop count, normalized latency and normalized power consumption parameters.

A. Simulator and Simulation Parameters

Experiments are perfumed using an open-source simulator called *XMulator* [24] a listener-based integrated simulation platform for interconnection networks. Moreover, for calculating the power consumption, *Orion* [25] library is added to the simulator. The wireless transceiver is designed with TSMC 65-nm standard CMOS process to obtain its power and delay characteristics and are fed in the simulator.

Table I shows a summary of simulation of parameters.

TABLE I. SIMULATION PARAMETERS

Parameter	Characteristic
Technology	65 nm
Clock Frequency	1 GHz
Number of IP Cores	225, 256
Switching	Wormhole
Routing	XY
Wireless Communication	mm-Wave Antennas
MAC	FDMA
Number of Virtual Channels	2
Number of Wireless Routers	8
Flit Length	64 bits

B. Traffic Patterns

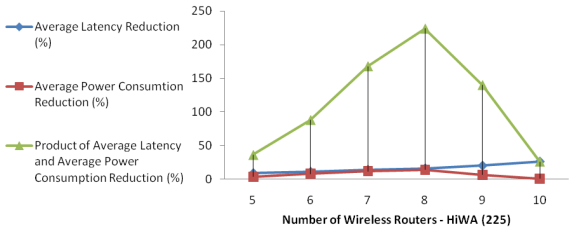
Evaluate each system with incorrect traffic patterns will likely lead to a wrong conclusion that would not be attributable; therefore, choosing suitable traffic patterns is vital. In simulations, six trace-driven traffic patterns taken from *Splash II* [26] and six synthetic traffic patterns as shown in Table II are used. Note that the trace-driven traffic patterns are repeated to cover all nodes of HiWA.

TABLE II. SYNTHETIC TRAFFIC PATTERNS

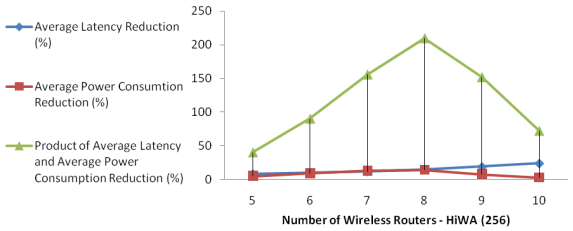
Traffic Pattern	Explanation
Uniform	Traffic distribution is random and nodes are likely to communicate with all other nodes equally.
Dataflow	Nodes are biased to communicate only with nodes within their subnets and with nodes in neighbors' subnets.
Hotspot	One or more nodes are transmitting 50% of total traffic. (Rest of traffics assumed to be Uniform)
Hot Dataflow	A dataflow pattern with one hot subnet which is transmitting 50% of total traffic.

C. Optimal Number of Wireless Nodes

Fig. 4 shows an experimental estimation which is obtained through simulation results under uniform traffic pattern, for different sizes of HiWA. According to this estimation, that is product of reduction in latency and power consumption, using 8 wireless routers is the optimal point to establish a trade-off between latency and power consumption parameters. In this experimental estimation, HiWA is configured with two different mesh sizes, 255 and 256 along with different number of wireless routers compared with conventional NoC. Considering only the latency reduction for evaluation, it is obvious that the more the number of wireless routers, the less the latency. On the other hand, increasing number of wireless routers leads to increased power consumption. The results show that using more than 10 wireless routers will significantly increases power consumption of HiWA.



(a)

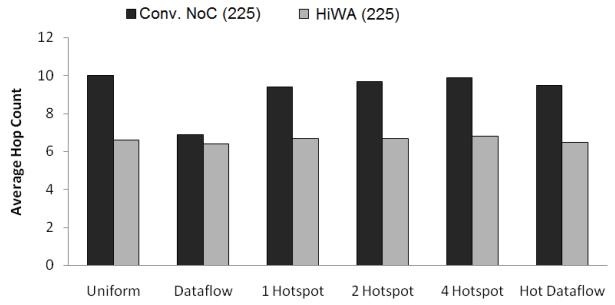


(b)

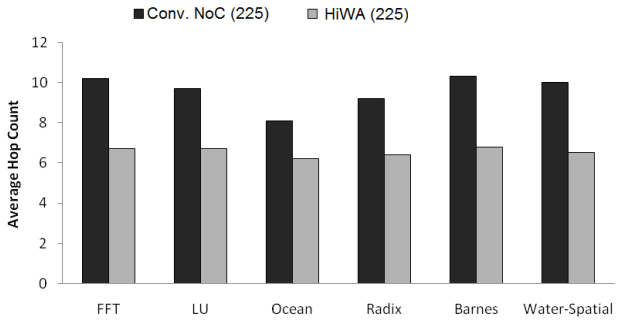
Fig. 4. Optimized number of wireless nodes under uniform traffic pattern (a) 225-node HiWA (b) 256-node HiWA

D. Hop Count

Fig. 5 and Fig. 6 compare the average hop counts between HiWA and conventional NoC. The results indicate that the proposed architecture reduces average hop counts 42% in comparison with conventional NoC. Since the wireless links act as shortcut paths, they decrease travel distance of packets from source to destination effectively. The result will show how much this reduction will affect the latency and power consumption.

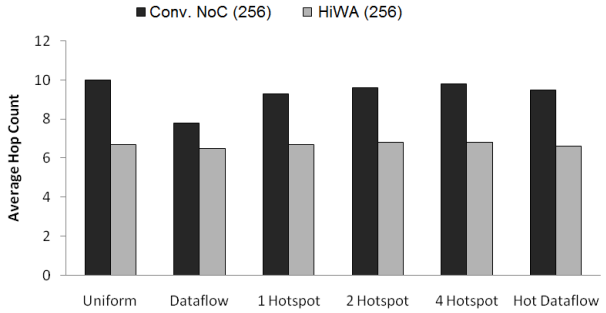


(a)

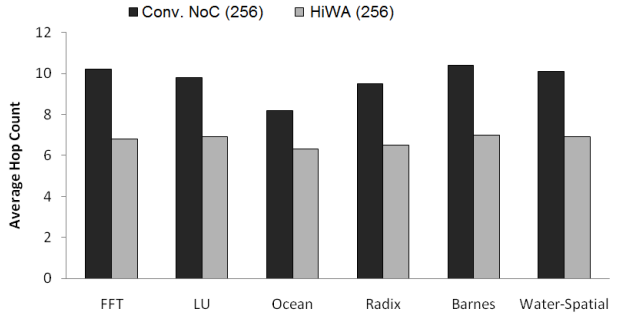


(b)

Fig. 5. Hop count comparision between 225-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns



(a)

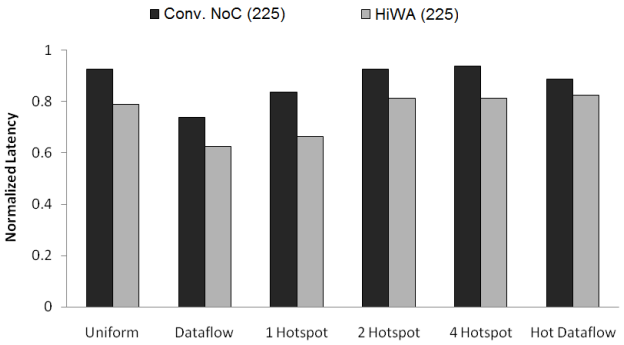


(b)

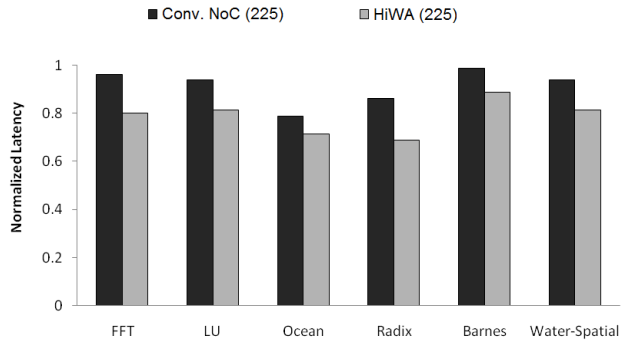
Fig. 6. Hop count comparision between 256-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns

E. Latency

Fig. 7 and Fig. 8 show the performance gains of the proposed approach under different traffic pattern and network sizes. The results indicate that the proposed architecture reduces average packet latency 16% in comparison with conventional NoC. This reduction is because of using wireless routers. Note that saving in hop counts is not directly convertible to latency reduction. In the current technologies bandwidth of wireless links is smaller than bandwidth of wired links. Consequently congestion occurs in wireless routers and overhead of packet blocking will be added to latency. Since carrier frequency increases in each technology generation, the number of available channels in the network will increase. According to *ITRS* [27] in 16nm CMOS technology gain frequency and power gain will reach in 600 GHz and 1 THz respectively; therefore, the proposed architecture will reach much more reduction in latency compared to conventional NoC.



(a)



(b)

Fig. 7. Latency comparision between 225-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns

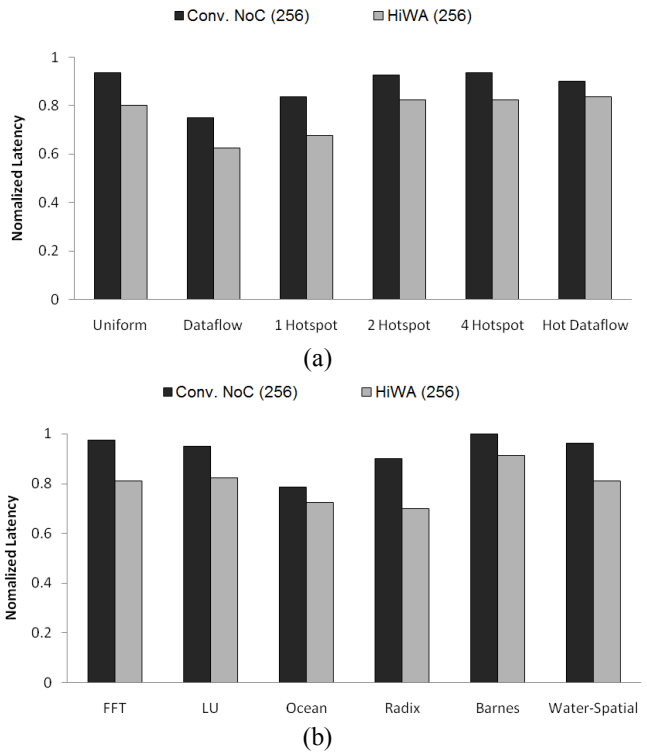


Fig. 8. Latency comparision between 256-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns

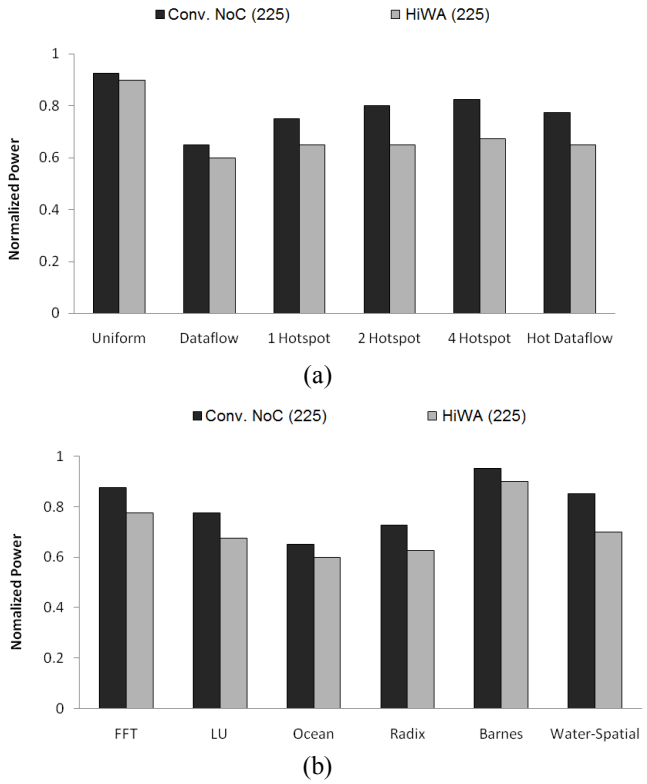


Fig. 9. Power consumption comparision between 225-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns

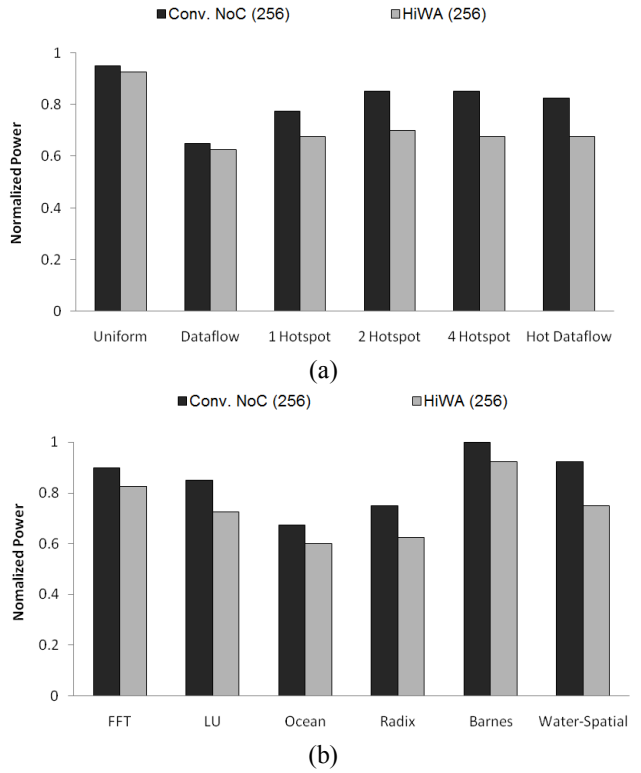


Fig. 10. Power consumption comparision between 256-node HiWA and conventional NoC (a) synthetic traffic patterns (b) trace-driven traffic patterns

F. Power Consumption

Fig. 9 and Fig. 10 show comparison between power consumption of proposed architecture and conventional NoC. 14% reduction is observable in total power consumption. Since the power consumption in a NoC originates from the operation of the IP cores and the interconnection components between those cores, it is proportional to the switching activity arising from packets moving across the network. In addition, router buffers have the greatest impact on power consumption. This improvement is caused by transmission speed up (reducing average hop counts) and reduction of average power consumption in buffers.

V. CONCLUSION

In this paper, a flexible wireless network-on-chip architecture called Hierarchical Wireless-based Architecture (HiWA) with wired and wireless communication was proposed. The proposed architecture is scalable to many IP cores and can be implemented with low hardware overhead at architecture level. In addition, simulation results shows that HiWA reduces average packet latency 16% and power consumption 14% in comparison with its conventional counterparts.

ACKNOWLEDGMENTS

This work is supported by VINNOVA (Swedish Agency for Innovation Systems) within the CUBRIC project.

REFERENCES

[1] L. Benini and G. D. Micheli, "Networks on chips: a new SoC paradigm," In *IEEE Computer,* Vol. 35, Issue 1, pp. 70-78, 2002.

[2] M. Horowitz and W. J. Dally, "How scaling will change processor architecture," In *IEEE International Solid State Circuits Conference*, Vol. 1, pp. 132-133, 2004.

[3] P. Magarshack and P. G. Paulin, "System-on-chip beyond the nanometer wall," In *Proceedings of Design Automation Conference*, pp. 419-424, 2003.

[4] R. K. Kamat, S. A. Shinde, and V. G. Shelake, *Unleash the System on Chip using FPGAs and Handel C*, Springer, 2009.

[5] V. K. Madisetti and C. Arpikanondt, A Platform-Centric Approach to System-on-Chip (SoC) Design, Springer, 2005.

[6] F. Fu, S. Sun, J. Song, J. Wang, and M. Yu, "A NoC performance evaluation platform supporting designs at multiple levels of abstraction," In *IEEE Conference on Industrial Electronics and Applications*, pp. 425-429, 2009.

[7] P. P. Pande, C. Grecu, M. Jones, A. Ivanov, and R. Saleh, "Performance evaluation and design trade-offs for network-on-chip interconnect architectures," In *IEEE Transactions on Computer*, Vol. 54, Issue 8, pp. 1025-1040, 2005.

[8] L. Benini and D. Bertozzi, "Network-on-chip architectures and design methods," In *Proceedings of Computers and Digital Techniques*, pp. 261-272, 2005.

[9] J. Sepulveda, M. Strum, W. Chau, and G. Gogniat, "A multi-objective approach for multi-application NoC mapping," In *IEEE Latin American Symposium on Circuits and Systems*, pp. 1-4, 2011.

[10] A. Ganguly, K. Chang, S. Deb, P. P. Pande, B. Belzer, and C. Teuscher, "Scalable hybrid wireless network-on-chip architectures for multicore systems," In *IEEE Transactions on Computers*, Vol. 60, Issue 10, pp. 1485-1502, 2011.

[11] P. P. Pande, A. Ganguly, K. Chang, and C. Teuscher, "Hybrid wireless network on chip: A new paradigm in multi-core design," In *Proceedings of International Workshop on Network on Chip Architectures*, pp. 71-76, 2009.

[12] M. F. Chang, V. P. Roychowdhury, L. Zhang, H. Shin, and Y. Qian, "RF/wireless interconnect for inter- and intra-chip communications," In *Proceedings of the IEEE*, Vol. 89, Issue 4, pp. 456-466, 2001.

[13] M. F. Chang, J. Cong, A. Kaplan, M. Naik, G. Reinman, E. Socher, and S. W. Tam, "CMP network-on-chip overlaid with multi-band RF-interconnect," In *Proceedings of IEEE International Symposium on High Performance Computer Architecture*, pp. 191-202, 2008.

[14] V. F. Pavlidis and E. G. Friedman, "3-D topologies for networks-on-chip," In *IEEE Transactions on Very Large Scale Integration*, Vol. 15, Issue 10, pp. 1081-1090, 2007.

[15] A. Shacham, K. Bergman, S. Member, and L. P. Carloni, "Photonic networks-on-chip for future generations of chip multiprocessors," In *IEEE Transactions on Computers*, Vol. 57, Issue 9, pp. 1246-1260, 2008.

[16] D. Huang, T. LaRocca, M.-C. Chang, L. Samoska, A. Fung,R. Campbell, and M. Andrews, "Terahertz CMOS frequency generator using linear superposition technique," In *IEEE Journal of Solid-State Circuits*, Vol. 43, Issue 12, pp. 2730-2738, 2008.

[17] E. Seok et al., "A 410GHz CMOS push-push oscillator with an on-chip patch antenna," In *IEEE International Solid-State Circuits Conference*, pp. 472-629, 2008.

[18] S. B. Lee, S. W. Tam, I. Pefkianakis, S. Lu, M. F. Chang, C. Guo, G. Reinman, C. Peng, M. Naik, L. Zhang, and J. Cong, "A scalable micro wireless interconnect structure for CMPs," In *Proceedings of the International Conference on Mobile Computing and Networking*, pp. 217-228, 2009.

[19] W. Green, M. Rooks, L. Sekaric, and Y. Vlasov, "Ultra-compact, low RF power, 10 Gb/s silicon Mach-Zehnder modulato," In *Optics Express*, Vol. 15, Issue 25, pp. 17106-17113, 2007.

[20] A. Kumar, L. S. Peh, P. Kundu, and N. K. Jha, "Toward ideal on-chip communication using express virtual channels," In *IEEE Micro*, Vol. 28, Issue 1, pp. 80-90, 2008.

[21] D. Zhao and Y. Wang, "SD-MAC: design and synthesis of a hardware-efficient collision-free QoS-aware MAC protocol for wireless network-on-chip," In *IEEE Transactions on Computers*, Vol. 57, no. 9, pp. 1230-1245, 2008.

[22] M. Dorigo and T. Stutzle, *Ant Colony Optimization*, MIT Press, 2004.

[23] N. M. A. AL-Salami, "System evolving using ant colony optimization algorithm," In *Journal of Computer Science*, Vol. 5, No. 5, pp. 380-387, 2009.

[24] A. Nayebi, S. Meraji, A. Shamaei, and H. Sarbazi-Azad, "XMulator: A listener-based integrated simulation platform for interconnection networks," In *First Asia International Conference on Modeling & Simulation*, pp. 128-132, 2007.

[25] A. Kahng, B. Li, L. Peh, and K. Samadi, "Orion 2.0: A power-area simulator for interconnection networks," In *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 20. Issue 1, pp. 191-196, 2012.

[26] S. Woo et al., "The SPLASH-2 programs: characterization and methodological considerations," In *International Symposium on Computer Architecture*, pp. 24-36, 1995.

[27] ITRS. International Technology Roadmap for Semiconductors, 2007 edition.

[28] M. Palesi and M. Daneshtalab (Eds.), *Routing Algorithms in Networks-on-Chip*, Springer, 2014.

[29] X. Wang, Z. Li, M. Yang, Y. Jiang, M. Daneshtalab, and T. Mak, "Agile frequency scaling for adaptive power allocation in many-core systems powered by renewable energy sources," In *Proceedings of Asia and South Pacific Design Automation Conference*, pp. 298-303, 2014.