Wireless Network-on-Chip Analysis of Propagation Technique for On-chip Communication

Vasil Pano, Isikcan Yilmaz, Yuqiao Liu, Baris Taskin, and Kapil Dandekar
Electrical and Computer Engineering
Drexel University, Philadelphia, Pennsylvania 19104
Email: vasilpano@gmail.com, iy43@drexel.edu, ting.yuqiao@gmail.com, taskin@coe.drexel.edu, dandekar@coe.drexel.edu

***Abstract*—Network-on-Chip (NoC) is a communication paradigm capable of facilitating a scalable interconnection infrastructure for multi core processors. Wireless NoCs have been introduced to improve the communication performance over long-distance processing nodes. Current on-chip antennas used in wireless NoCs communicate predominantly through surface waves, where the efficacy of the wireless nodes is partially determined by the radiation efficiency and transmission gain limited due to the conductivity loss of the silicon substrate. Recently, an on-chip propagation technique of radio waves was introduced, through the un-doped silicon layer as opposed to surface-waves prevalent in literature. The through-substrate propagation waves provide a unique solution to overcome the challenge of long-distance communication between processing nodes. In this work, overall improvements are shown compared to traditional wireless NoCs with the placement of antennas on undoped silicon (i.e. communicating through surface waves), simulated in NoC architectures across performance metrics of area, power consumption and latency.**

I. Introduction

Network-on-Chip (NoC) is a new communication paradigm targeting the communication infrastructure of chip multiprocessors (CMP) and multiprocessor System-on-Chip (MPSoC) in order to achieve increased performance and energy savings [1]. NoCs are particularly a preferred solution to support the increasing scalability of these devices [2]. One important enhancement to NoCs has been the inclusion of wireless communication through on-chip antennas [3]. Wireless NoCs have allowed for long-distance communication between far away processing elements, eliminating the need for multi-hop jumps between subnets or routers. This elimination of multi-hop jumps enable lower power consumption, lower latency and higher overall throughput of the NoC. The antenna implementation, placement and the communication medium of semiconductor technology are integral components that determine the quality (i.e. length) of the wireless, long-distance communication between nodes.

Antenna designs face the challenge of obtaining high radiation efficiency and transmission gain, both important factors in delivering packets or flits across the entire NoC topology. Current research has increasingly focused on strategically placing on-chip antennas in advantageous positions through novel heuristics and placement algorithms [4–6] so that multiple antennas would be able to communicate efficiently with each other. This approach increases area and power consumption prohibitively when considering the marginal performance benefits of additional, unnecessary antennas.

In this work, the impact of a recent propagation technique for on-chip communication [7] on hierarchical NoC system is analyzed. This recent antenna design utilizes substrate waves as opposed to the surface waves, where the substrate waves are provided with a un-doped silicon layer underneath — a placement that is fully compliant with CMOS manufacturing. This novel propagation mechanism allows geographically distant antennas to communicate directly and essentially eliminates the need for additional redundant antennas that are only present to bounce the information to the desired destination. By eliminating these additional antennas, latency and power consumption improve in addition to the overall reduction in area. It is also shown that the overall throughput of the system does not change by a large factor, demonstrating the overall superiority of the proposed wireless NoCs with surface-wave propagating antennas placed on undoped silicon.

The existing body of work related to wireless NoCs and their enhancements are presented in Section II-A. A top-down explanation of the propagation mechanism used in the NoC is shown in Section II-B. The overview for the hierarchical, multi-level Network-on-Chip is presented in Section III. Simulation setup, performance metric overview and simulation results for broadcasting antennas are discussed in Section IV-A, Section IV-B and Section IV-C, respectively. Finally, the conclusions of this work are presented in Section V.

II. Related Works

The wireless NoC literature is analyzed as relevant to this work in Section II-A. The recently proposed antenna with substrate-wave based propagation is briefed in Section II-B.

A. Wireless NoCs Literature Review

Multiple wireless network-on-chips have been introduced in recent years; For instance, Chang et al. [3], Deb at al. [8], and, DiTomaso et al. [6] have varying wireless NoC designs that show significant performance improvement over hierarchical or typical flat-mesh NoC.

In particular, Chang et al. [3] proposed a hybrid (wireless/wired) hierarchical NoC architecture with heterogeneous

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1305350 and ECCS-1232164. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

978-1-5090-5142-7/16/$31.00 ©2016 IEEE

subnets and small-world based upper level configuration. The architecture shows significant performance improvement but further analysis showed that incrementing the number of wireless antennas did not increase substantially the overall throughput of the system.

DiTomaso et al. [6] introduced a energy-efficient adaptive wireless NoC architecture which uses link utilization to adapt wireless channels accordingly to different traffic patterns. The results show great promise when considering adaptive algorithms that can lower power consumption but further analysis on variable wireless routers was not performed to measure the performance impact of decrementing the total number of antennas.

B. Antennas in Wireless Networks on Chip

Existing NoC literature mostly relies on variations of antennas presented in [9, 10]. O et al. [9, 10] presented meander dipole antennas operating at 15 GHz for wireless on-chip communication. These antennas mostly communicate through surface propagation of waves. Recently in [7], an alternative placement of antennas on the undoped silicon layer of a typical semiconductor technology was introduced, that predominantly propagate through the undoped substrate. In this new technique presented by Liu et al. [7], the chip layout is planned such that the antennas are placed over an undoped silicon layer as depicted in Figure 1.

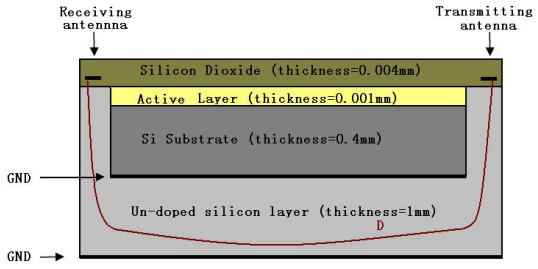


Fig. 1. Dominant propagation path after inserting a ground layer between Si substrate and un-doped silicon layer. [7]

3D finite element method based HFSS modeling and simulations are presented in [7] to demonstrate that:

1) The efficacy and directivity of the on-chip antennas was improved considerably,
2) The technique does not have an area overhead, and
3) The technique is fully compliant with CMOS manufacturing.

The propagation characteristics of the surface-wave propagating antenna placed over undoped silicon is shown to be solidified in [7] with strategically placed ground planes. In particular, the radiation efficiency can be improved from 2.8% to 28.78% and transmission gain improved from -30.3 dB to -15.4 dB by adding a ground plane between silicon substrates and the un-doped silicon layer, while removing the ground plane under antenna area. By doing so, the un-doped silicon layer acts as a wave guide that provides the main path for electromagnetic (EM) waves. Concurrently, the ground plane between the un-doped silicon layer and silicon substrate prevents the wave from leaking to the silicon substrates. These improvements in antenna efficiency and transmission gain allow for new NoC topologies that use less antennas but at a larger distance than previously possible for the same channel quality (e.g. BER, etc.).

III. Proposed Multi-Level Network-on-Chip

The network architecture proposed in this work is primarily based on a typical NoC flat-mesh design, with one router per processing element, 4 virtual channels per router (both input and output), router input and output buffer depth of 4 flits and a total number of 560 processing elements (PEs). The design choices and parameters sweep used in this work are detailed in Table I.

TABLE I. Network-on-Chip Topology Analysis

Topology	Design Point Chosen
# of PEs in Subnets	16 PEs (4x4)
# of Subnets	35 (5x7)
Total # of PEs	560 PEs
# of Wireless Antennas	Variable 0 – 10
# of Virtual Channels	4 VCs
# of Input/Output Buffers	4 flits
# of Antenna Buffer Depth	8 flits
Fixed Packet Size	10 flits

These design choices are comparable to the wireless NoCs from literature reviewed in Section II-A. A high number of processing elements is selected to estimate the performance of a next-generation CMP network that would presumably include hundreds of RISC cores and would require a very large NoC to operate efficiently. The main differences between a typical NoC and the one evaluated here is the hierarchical and wireless design choices that are implemented to further facilitate the main discussion points of the implementation of the novel propagation technique. The standard NoC is modified to implement a subnet structure that groups processing elements together in clusters of 16 PEs per subnet. These clusters are further directly connected to a hub which acts as a router on the top layer of the NoC. There are a total of 35 subnets (i.e. hubs) and in order for a packet to travel across the network to a different subnet, the packet has to go through its respective hub and any additional hubs required to reach the destination subnet.

In addition to the typical router/hub interface, some of the hubs are modified to be a wireless hybrid hub of sending packets either using an antenna or through the typical wire-based interconnect. These hybrid hubs are the basis of the overall analysis performed on the NoC. Depending on the configuration of the wireless antenna placement and which pair of hybrid hubs can communicate wirelessly, the performance of the NoC topology will vary greatly, effectively proving that a better performing NoC can be designed with the novel propagation technique.

IV. Performance Evaluation of Substrate-Propagation Antennas

The wireless NoC with antennas capable of broadcasting to different regions of proximity is detailed in Section IV-A. An overview and explanation of the performance metrics and traffic patterns used to evaluate the NoC is detailed in Section IV-B. The results of the varied configurations of

antenna placement on the NoC performance are discussed in Section IV-C.

A. Wireless NoC with Broadcasting Antennas

The test scenarios evaluated in this work show a comparison in performance between antennas,

- transmitting to their closest neighbors (N+1),
- their closest neighbors and the next ones (N+2), and
- broadcasting to all neighbors.

The antenna placement and an example of range of transmission are shown in Figure 2. In this example, hybrid hub #16 is capable of transmitting to subnets #7, 9, 21, 22 in the first test scenario (the inner circle), subnets #7, 9, 21, 22, 2, 14, 18, and 30 in the second test scenario (outer circle), and finally it can broadcast to every other wireless antenna in the last test scenario. In effect, test scenarios 1 and 2 are multi-cast whereas test scenario 3 is a broadcast operation.

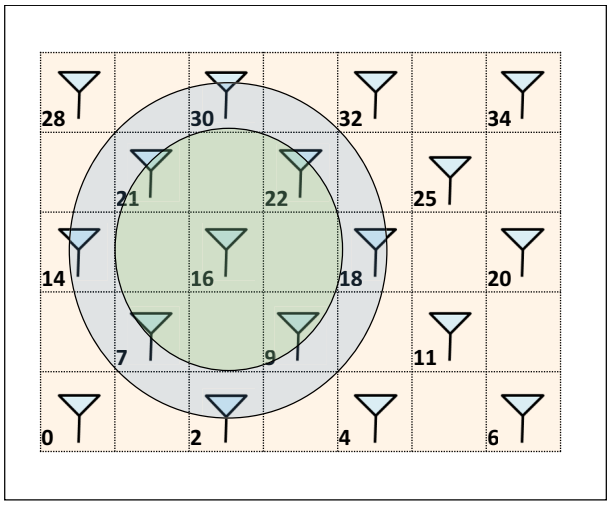


Fig. 2. Antenna placement on the wireless NoC (each tile represents a subnet, each circle shows the range of transmission).

B. Performance Metric and Traffic Patterns Overview

The NoC performance metrics are:

1) saturation throughput,
2) average energy per flit at saturation,
3) normalized average latency at saturation, and
4) normalized energy-delay per flit at saturation,

Saturation throughput is measured as total delivered flits per cycle per processing element. *Average energy per flit at saturation* is measured as pJ per flit and its primary use is to discuss the overall energy consumption over the total amount of flits delivered. *Normalized average latency at saturation* is measured and calculated in comparison to the baseline hierarchical NoC with no antennas. *Normalized energy-delay per flit at saturation* is the product of the overall energy per cycle per IP with the average delay normalized to the baseline hierarchical NoC.

The synthetic traffic patterns used in the analysis are:

1) *uniform* traffic,
2) *butterfly* traffic, and
3) *varying localization degree* traffic.

Though the selection of the *uniform* traffic seems counter-intuitive based on the current parallel programming paradigms, the exa-scale computing study in [11], reports that the future scientific workloads are expected to be highly non-localized. The *uniform* traffic is included in this list in part to address the vision in [11] of non-localized and heterogeneous CMPs.

Also taking into account current MPI based parallel programming paradigms where the communication between the nodes is to be localized as a facet of the task-to-core mapping. Synthetic workloads of localized traffic patterns with localization degrees in the range from 0.05 to 0.50 (the probability that the destination is within the subnet) are used to represent the varying degrees of task mapping.

C. NoC Simulation Results with Broadcasting Antennas

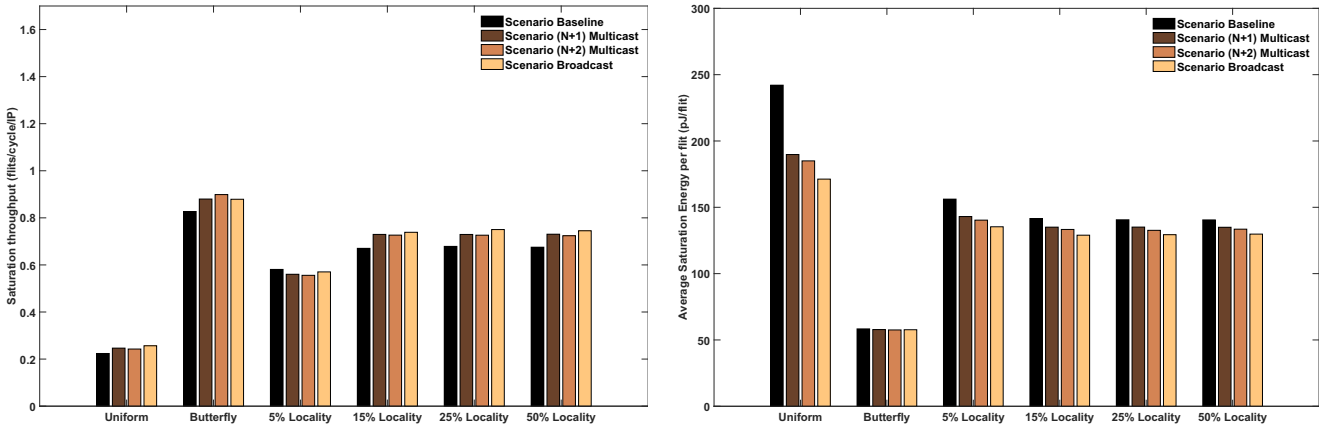
The performance differences between multicasting in a region of proximity and broadcasting in the entire wireless NoC are shown in Figure 3.

i. *Saturation throughput* results in Figure 3(a) show an improvement of up to 15% compared to the baseline NoC. Uniform random traffic benefits from the broadcasting capabilities of scenario Broadcast. Both multicast scenarios display similar saturation throughput across all traffic patterns and are not able to surpass the throughput achieved by the broadcast scenario.

ii. Results for *Average energy/flit at saturation* in Figure 3(b) show a decrease in energy per flit with the increase of the range of transmission. Packets are able to traverse longer distances with less hops and therefore the overall power consumption of the wireless NoC is lower compared to the baseline NoC.

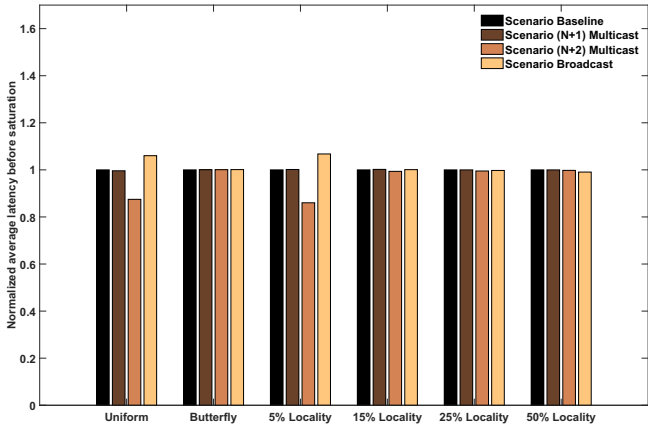
iii. *Normalized average latency* results detailed in Figure 3(c) show that the average latency is actually increased with the increase of the range of transmission. This is because network utilization is not even throughout the NoC and therefore is formed, increasing the delay of packet arrival. This behavior is more evident with the last scenario because the broadcasting capabilities of the antenna and the routing algorithm employed (XY routing) do tend to increase congestion as packets are routed through the shortest link possible without network utilization awareness. This is less clear in the other scenarios because the antennas cannot broadcast directly to the destination hubs and therefore the traffic is more uniformly distributed.

iv. *Normalized energy-delay/flit* is shown to be comparatively similar between scenarios. When the probability of localized traffic is increased, the broadcast scenario has a lower EDP compared to the multicast options. When using random uniform traffic and butterfly traffic, all scenarios have nearly identical EDP.

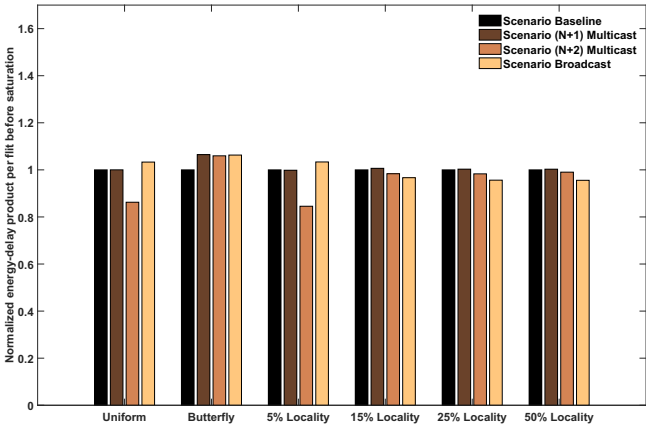


(a) Saturation throughput

(b) Average energy/flit at saturation



(c) Normalized average latency



(d) Normalized energy-delay/flit

Fig. 3. Experimental results for wireless NoCs with broadcasting surface-propagating antennas with varying synthetic traffic patterns.

V. Conclusions

In summary, the proposed wireless hierarchical NoC analyzing a new propagation technique capable of increased transmission gain and radiation efficiency performs favorably when compared to multicast antennas with a small range of transmission. The proposed wireless NoC implementations with surface-wave propagating antennas placed over undoped silicon with strategically placed ground planes permit NoC implementations with improved energy per flit, area and latency, with similar throughput with existing wireless NoC implementations with conventional antenna placement.

References

[1] W. J. Dally and B. Towles, "Route packets, not wires: on-chip interconnection networks," In ACM Proceedings of the Design Automation Conference (DAC), 2001, pp. 684-689.

[2] R. Marculescu, U. Y. Ogras, L. S. Peh, N. E. Jerger and Y. Hoskote, "Outstanding Research Problems in NoC Design: System, Microarchitecture, and Circuit Perspectives," In IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 28, no. 1, pp. 3-21, January 2009.

[3] K. Chang, S. Deb, A. Ganguly, X. Yu, S. P. Sah, P. P. Pande, B. Belzer, and D. Heo, "Performance evaluation and design tradeoffs for wireless network-on-chip architectures," ACM Journal of Emerging Technologies in Computer Systems, vol. 8, no. 3, pp. 23, 2012.

[4] A. Rezaei, F. Safaei, M. Daneshtalab and H. Tenhunen, "HiWA: A hierarchical Wireless Network-on-Chip architecture," In Proceedings of the International Conference on High Performance Computing and Simulation (HPCS), 2014, pp. 499-505.

[5] S. Abadal, M. Iannazzo, M. Nemirovsky, A. Cabellos-Aparicio, H. Lee and E. Alarcn, "On the Area and Energy Scalability of Wireless Network-on-Chip: A Model-Based Benchmarked Design Space Exploration," In IEEE/ACM Transactions on Networking, vol. 23, no. 5, pp. 1501-1513, October 2015.

[6] D. DiTomaso, A. Kodi, D. Matolak, S. Kaya, S. Laha and W. Rayess, "A-WiNoC: Adaptive Wireless Network-on-Chip Architecture for Chip Multiprocessors," In IEEE Transactions on Parallel and Distributed Systems, vol. 26, no. 12, pp. 3289-3302, December 2015.

[7] Y. Liu, V. Pano, D. Patron, K. Dandekar and B. Taskin, "Innovative propagation mechanism for inter-chip and intra-chip communication," In IEEE Transactions on Wireless and Microwave Technology Conference (WAMICON), 2015, pp. 1-6.

[8] S. Deb et al., "Design of an Energy-Efficient CMOS-Compatible NoC Architecture with Millimeter-Wave Wireless Interconnects," In IEEE Transactions on Computers, vol. 62, no. 12, pp. 2382-2396, 2013.

[9] K. Kim, K. K. O, "Integrated Dipole Antennas on Silicon Substrates for Intra-chip Communication," In IEEE Proceedings of the International Symposium on Antennas and Propagation Society (AP-S), 1999, pp. 1582-1585.

[10] K. K. O, K. Kim et al., "The feasibility of on-chip interconnection using antennas," In IEEE/ACM Proceedings of the International Conference on Computer-Aided Design (ICCAD), 2005, pp. 979-984.

[11] P. Kogge, K. Bergman, S. Borkar. 2008. ExaScale Computing Study: Technology Challenges in Achieving Exascale Systems. DARPA.