COMPARATIVE PERFORMANCE EVALUATION OF WIRELESS AND OPTICAL NOC ARCHITECTURES

Sujay Deb, Kevin Chang, Amlan Ganguly, Partha Pande
School of Electrical Engineering and Computer Science, Washington State University, USA
{sdeb,ganguly,jchang,pande}@eecs.wsu.edu

ABSTRACT

Network-on-Chip architectures with various new emerging interconnect technologies offer unprecedented performance gain compared to their more conventional planar metal interconnect-based counterparts. A comparative analysis of these emerging NoCs will provide a better understanding towards adopting them in the mainstream design. This paper compares performance of small-world NoC architectures having wireless and RF links with optical NoCs and presents design trade offs with system size scaling. Our analysis demonstrates that the small-world NoC architecture with THz wireless shortcuts provides the best performance-area overhead trade-off compared to the other NoCs for both uniform and non-uniform traffic across various system sizes.

I. INTRODUCTION

Networks-on-Chip (NoCs) have emerged as communication backbones to enable a high degree of integration in multi-core Systems-on-Chip (SoCs). Despite their advantages, an important performance limitation in traditional NoCs arises from planar metal interconnect-based multi-hop links, where the data transfer between two distant blocks causes high latency and power consumption. According to the International Technology Roadmap for Semiconductors (ITRS, 2007), "at 0.13um approximately 51% of microprocessor power was consumed by interconnect, with a projection that without changes in design philosophy, in the next five years up to 80% of microprocessor power will be consumed by interconnect". Consequently, both architectural and interconnect-level innovations are needed to address the performance bottleneck in conventional metal interconnect-based NoCs. The limitations of conventional NoCs can be addressed by drawing inspiration from the interconnection mechanism of natural complex networks [1]. Many networks, such as networks of neurons in the brain, the Internet, and social networks share the small-world (SW) property [2]. Compared to a purely locally and regularly interconnected network (such as a mesh interconnect), small-world networks have a very short average distance between any pair of nodes. This makes them particularly interesting for efficient communication in modern multi-core chips with increasing levels of integration. This small world topology can be incorporated in NoCs by introducing long-range, high bandwidth and low power links between distant cores. One of the possible ways to insert long-range links in a NoC is to strategically place wireless or RF transmission line-based links based on traffic distribution and the location of the cores. On the other hand optical NoC is emerging as a promising alternative to address the performance limitation of the conventional NoCs. In this paper we endeavor to establish a performance benchmark for a small-world NoC architecture designed with wireless or RF links with respect to emerging optical NoCs. As a representative case, we consider two recently proposed optical NoC architectures [3] [4] in this comparative study. We demonstrate the achievable bandwidths, energy dissipation profiles and the area overheads of all these novel NoC platforms with increase in system size.

II. RELATED WORK

Conventional NoCs use multi-hop packet switched communication. To improve performance, the concept of express virtual channels is introduced in [5]. It is shown that by using virtual express lanes to connect distant cores in the network, it is possible to avoid the router overhead at intermediate nodes, and thereby greatly improve NoC performance. NoCs have been shown to perform better by inserting long range wired links following principles of small world graphs [1].

The design principles of Photonic NoC are elaborated in various recent publications [2, 3]. It is estimated that a photonic NoC will dissipate significantly less power than its electronic counterpart.

Another alternative is NoCs with multi-band RF interconnects [6]. In this particular NoC, instead of depending on the charging/discharging of wires for sending data, electromagnetic (EM) waves are guided along on-chip transmission lines created by multiple layers of metal and dielectric stack. As the EM waves travel at the effective speed of light, low latency and high bandwidth communication can be achieved.

Recently, the design of a wireless NoC based on CMOS Ultra Wideband (UWB) technology was proposed [7]. In [8], the feasibility of designing on-chip wireless communication network with miniature antennas and simple transceivers that operate at the sub-THz range of 100-500 GHz has been demonstrated. If the transmission frequencies can be increased to THz/optical range then the corresponding antenna sizes decrease, occupying much less chip real estate. One possibility is to use nanoscale antennas based on CNTs operating in the THz/optical frequency range [9]. Consequently building an on-chip wireless interconnection network using THz frequencies for inter-core communications becomes feasible. Design of a wireless NoC with CNT antennas following the principles of small-world network has been demonstrated in [10, 11].

A comparative performance evaluation between these emerging NoCs is currently lacking. This paper aims to bridge that gap by evaluating the performance of small-world wireless NoCs with respect to a subset of recently proposed photonic NoCs.

III. NOC ARCHITECTURES UNDER CONSIDERATION

In this section we discuss the features of various NoC architectures considered for the performance evaluation in this paper.

A. Small-World NoC Architectures:

A conventional wired NoC suffers from multi-hop communication resulting in high energy dissipation and latency. To alleviate this problem, network topologies with small-world property having a very short average path length between any pair of nodes can be employed. This small-world network topology can be incorporated in NoCs by introducing long-range, high bandwidth and low power links between distant cores. Incorporating the small-world property will enable design of hierarchical NoC architectures, where closely spaced cores will communicate through traditional metal wires, but long distance communications will be predominantly achieved through high performance specialized links. In this work, we consider two types of small-world NoC architectures. In one, the long-range links are established using on-chip wireless interconnects and in the other on-chip RF interconnects are used as long-range links. Again, for the wireless interconnects, we consider two types of links, viz., THz links implemented with CNT antennas and sub-THz links as used in [8]. Due to the short range of the UWB antennas in [7], they are not suitable to create small-world architectures.

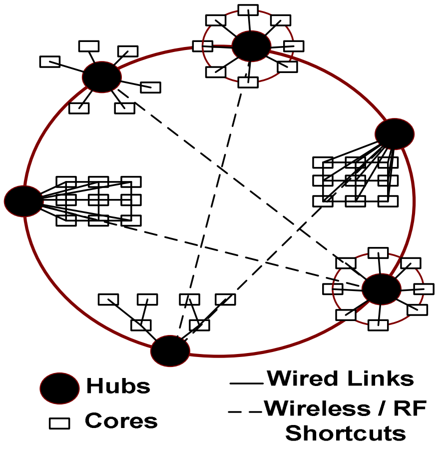

Figure 1. A conceptual hybrid (wireless/wired) NoC architecture with heterogeneous subnets following the small-world principle

i) Small-World Wireless NoC (WiNoC):

WiNoC is a hybrid wired/wireless NoC architecture. The whole system is first divided into multiple small clusters of neighboring cores. These smaller clusters are called *subnets*. As subnets are smaller networks, *intra-subnet* communication will have a shorter average path length than a single NoC spanning the whole system. Instead of a traditional uniform NoC, there may be subnets with varying architectures for different parts of the chip. Fig. 1 shows a hybrid (wireless/wired) NoC architecture with heterogeneous subnets. The cores in a subnet are connected to a centrally located hub through direct wireline links and the hubs from all subnets are connected in a 2nd level network forming a hierarchical architecture. This upper level of the hierarchy is designed to have characteristics of small-world graphs. To reduce the overhead arising due to the insertion of the wireless links, neighboring hubs are connected by traditional wired links forming a ring and a few wireless links are distributed between hubs separated by relatively long distances depending on their frequency of communication. The hubs with wireless links are equipped with wireless base stations (WBs) that transmit and receive data packets over the wireless channels. For inter-subnet and intra-subnet data transmission, wormhole routing is adopted [12].

Suitable on-chip antennas are necessary to establish wireless links for WiNoCs. Recent research has uncovered excellent emission and absorption characteristics leading to dipole like radiation behavior in carbon nanotubes (CNTs), making them promising for use as antennas for on-chip wireless communication [9]. By using multiband laser sources to excite CNT antennas, different frequency channels can be assigned to pairs of communicating subnets. This will require using antenna elements tuned to different frequencies for each pair, thus creating a form of Frequency Division Multiplexing (FDM). This is possible by using CNTs of different lengths, which are multiples of the wavelengths of the respective carrier frequencies. These CNT antennas principally operate in the range of tens of THz and consequently their sizes are in the order of few hundreds or tens of microns. High directional gains of these antennas, demonstrated in [9] [14], aid in creating directed channels between source and destination pairs. The different frequency channels are assigned to multiple wireless links in the WiNoC in such a way that a particular frequency is allocated only once to avoid signal interference. High-speed silicon integrated Mach-Zehnder optical modulator [15] can provide 10Gbps data rate per channel on these links. On-Off Keying (OOK) modulation scheme is adopted for WiNoC implementation.

The other wireless interconnect considered here is designed to operate at 100 − 500 GHz sub-terahertz frequency bands. As shown in [8], this available bandwidth can be divided into multiple frequency channels, each operating at 10-20 Gbps.

The wireless links are placed in the network using Simulated Annealing (SA) [16] based optimization technique. To perform SA, a metric μ has been established, which is closely related to the connectivity of the network. The metric μ is the average distance, measured in number of hops, between all source and destination hubs and weighted with a normalized frequency of communication between the particular pair of hubs. The optimization metric, μ can be computed as

$$\mu = \sum_{i,j} h_{ij} f_{ij} \, , \tag{1}$$

where h_{ij} is the distance (in hops) between the i^{th} source and j^{th} destination. The frequency f_{ij} of communication between the i^{th} source and j^{th} destination is the apriori probability of traffic interactions between the subnets. This is determined by the particular traffic pattern mapped onto the NoC. In this work we adopt the Cauchy annealing schedule where the temperature profile varies inversely with the number of iterations as

$$T = \frac{T_0}{k} \, , \tag{2}$$

where T is the temperature profile, T_0 is the initial temperature and k is the current annealing step. The convergence criterion is that the metric at the end of the current iteration differs by less than 0.1% from the metric of the previous iteration.

ii) Small-World NoC with RF Interconnects (RFNoC):

Another possible way to establish the long-range link is to use multi-band RF interconnects. The RFNoC is designed by replacing the wireless links of the WiNoCs by RF interconnects, maintaining the same hierarchical topology with shortcuts in the upper level. In 65nm technology it is possible to have 8 different frequency channels each operating with a data rate of 6 Gbps [6]. Like the wireless links, these RF links can be used as the long-range shortcuts in the hierarchical NoC architecture. These shortcuts are optimally placed using the same SA based optimization as used for the WiNoC.

B. NoCs with Optical Interconnects:

NoCs with high speed optical interconnects is another promising alternative to conventional wireline NoCs. We consider two such optical NoC architectures for the comparative study. The first one is adopted from [3] and is referred as Photonic NoC for the rest of this paper. This NoC requires an electrical control network to configure photonic switching elements which uses a flat wireline mesh. Photonic NoC architecture employs an optical circuit-switched network for bulk message transmission and an electronic packet-switched network for distributed control and short message exchange. The optical torus network is augmented with additional optical paths to provide path multiplicity, so that blocking can be avoided and path setup latency is accordingly reduced. Maximum path multiplicity in the Photonic NoC ensures that there is no blocking in the network. An illustrative 16 core Photonic NoC architecture with path multiplicity of 2 is presented in Fig. 2.

The second optical NoC considered in this paper is the Clos network proposed in [4]. Clos is a low-

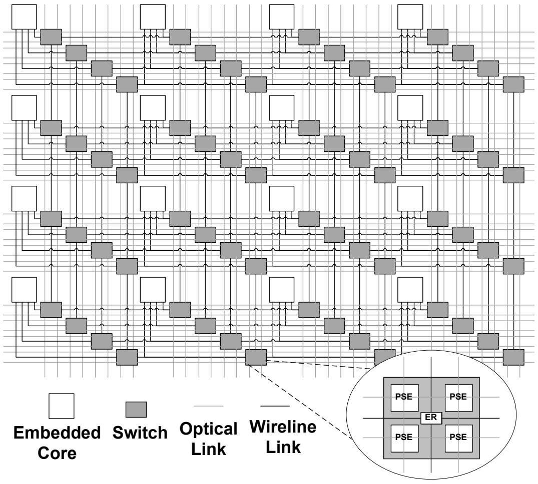

Figure 2. A representative 16 core Photonic NoC with path multiplicity of 2.

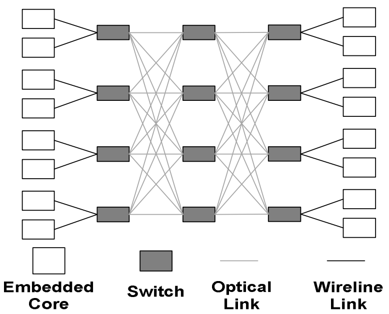

Figure 3. A 16 core Photonic Clos network

diameter non-blocking network. It uses multiple stages of routers to create a large non-blocking all-to-all connectivity. Clos network uses point-to-point channels instead of global shared channels resulting in significantly less overhead compared to global photonic crossbars. A 16 core photonic Clos network is shown in Fig. 3.

IV. PERFORMANCE EVALUATION

In this section, we present the experimental results to establish the relative performance benefits offered by the above mentioned emerging NoCs. To do this, we present detailed network level simulations with various system sizes and traffic patterns.

A. Characterization with uniform random traffic

As a first step, we analyze the characteristics of the different NoC architectures in presence of a uniform random traffic and study trends in their performance with scaling of system size. For our experiments we consider three different system sizes, namely 128, 256, and 512 cores. All the network architectures mentioned in section III are simulated using a cycle accurate simulator, which models the progress of data flits accurately per clock cycle accounting for flits that reach their destinations as well as those that are dropped. In this work we assume flit width of 32 bits and each packet consists of 64 flits. A self-similar traffic injection process is assumed.

For small world architectures the number of subnets and their sizes are chosen for optimum performance for different system sizes. 128 core system is divided into 16 subnets with each subnet consisting of 8 cores, 256 core system is divided into 16 subnets with each subnet consisting of 16 cores and 512 core system is divided into 32 subnets with each subnet consisting of 16 cores. For this exercise, we assumed that the cores within each subnet are connected in a mesh topology. Each core also has a direct path to the central hub. The upper level of the network is considered to be a ring with wireless shortcuts. The shortcuts are optimally deployed among the hubs following the SA methodology.

Table I: Packet Energy per Bandwidth (nJ/ TBps) for different NoC architectures with scaling of system size

System Size	Flat Mesh	WiNoC_ THz	WiNoC_ subTHz	RFNoC	Clos	Photonic NoC
128	560.0	21.0	31.3	34.1	16.6	34.1
256	721.7	13.3	19.6	25.8	9.0	29.6
512	1171	10.0	15.4	32.8	7.3	35.1

For small-world NoCs the subnet switches and the digital components of the hubs are synthesized using 65 nm standard cell library from CMP [17] at a clock frequency of 2.5GHz. The delays in flit traversals along all the wired interconnects that are introduced to enable the small-world NoC architecture were considered while quantifying the performance.

It is shown that 24 distinct frequency channels can be created for CNT antennas [13] used in WiNoC. In the WiNoC with sub-terahertz wireless links, 24 frequency channels each with a data rate of 10 Gbps is considered to provide same bandwidth as the CNT-based scheme, For WiNoC using CNT antennas the energy dissipation of the longest possible wireless link on the chip was found to be 0.33 pJ/bit. However, the sub-THz wireless link dissipates 4.5 pJ/bit in the same topology [8]. In case of RFNoC, the energy dissipation for RF interconnects in 65nm technology is considered to be 1pJ/bit [6].

For the Photonic NoC architecture, we have considered maximum path multiplicity for each system size for highest sustainable system throughput. The energy dissipation value for different optical components in the Photonic NoC is taken from [3].

The Clos NoC architecture uses multiple stages of small routers to create a larger non-blocking all-to-all network. As in [4], all of the Clos routers are implemented electrically and the inter-router channels are implemented with photonics. As assumed in [4], enough WDM optical channels are considered to enable a flit to be transmitted in a single cycle. The conservative energy projection values for all the optical components from [4] are used for energy calculations.

Fig. 4 shows the achievable overall network bandwidth for all the different NoCs under consideration. The bandwidth of flat mesh is also

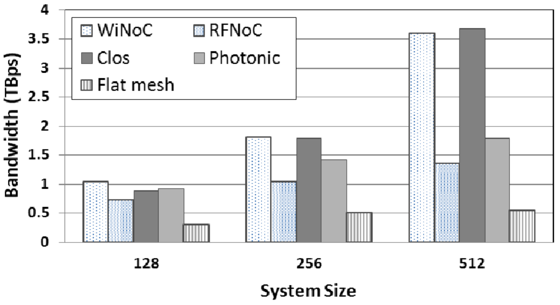

Figure 4. Achievable Bandwidth for different NoC architectures with scaling of system size

included for comparison. It is evident that all these emerging NoCs outperform the flat mesh. Among the small-world and optical NoCs, WiNoC with both types of wireless links and Clos perform significantly better than the others for all system sizes. As the different architectures produce different peak bandwidths, the packet energy per bandwidth for different NoC architectures is considered as a relevant metric and is presented in Table I. From the packet energy per bandwidth results it can be seen that Clos NoC performs best for all system sizes followed by THz WiNoC with CNT antennas.

Fig. 5 presents the area overhead arising out of the hubs and switches inclusive of necessary transceiver modules for the various NoC architectures. It is evident that Clos and the Photonic NoC have significantly higher area overhead compared to other NoCs. For 512 core system the area overheads for Clos and Photonic NoC even exceeds the limits of a 20mmX20mm die. The area overhead of Photonic NoC can be reduced by decreasing path multiplicity which will however correspondingly degrade the achievable bandwidth.

Fig. 6 shows the total wiring, RF and optical waveguide requirement of various lengths for the different NoC architectures considered in this work implemented in a 20mmx20mm die. The wiring requirement for flat mesh architecture is also shown for comparison. It is evident that all these emerging NoCs have extra interconnect overhead with respect to a flat mesh. WiNoC and RFNoC introduce less number of extra links compared to the Photonic and Clos NoCs. In the Photonic NoC high interconnect overhead arises mainly due to the path multiplicity and in the Clos this happens due to the non-blocking all-to-all nature of the network.

From all the above analysis, it is clear that among different NoC architectures compared in this paper, WiNoC with CNT antennas provides the best trade-off between performance and area overhead as system size scales up. This is because in WiNoC, the high-performance long range links are optimally placed to achieve a small-world network. WiNoC with sub-THz wireless links have higher packet energy per bandwidth and area overhead compared to WiNoC with CNT antennas. Though RFNoC uses the similar small-world network architecture, the less number of links (8 links compared to 24 available in WiNoC) and limited bandwidth of each link (6 Gbps compared to 10 Gbps of WiNoC) affect the performance, especially when the system size scales up. Both Clos and Photonic NoCs are benefited by the high bandwidth low power optical links. The Clos NoC dissipates least packet energy per bandwidth as it is non-blocking all-to-all network, whereas the Photonic NoC is affected by the electrical path set up latency. Though the clos network dissipates least packet energy per bandwidth, but it comes at a high price in real estate overheads compared to the small-world architectures.

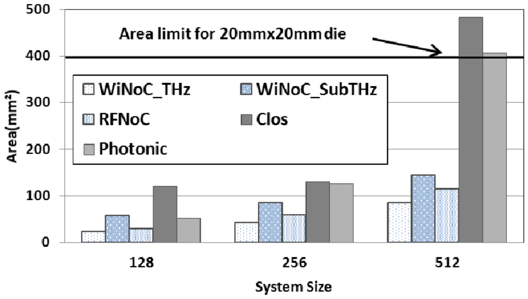

Figure 5. Area overhead for different NoC architectures with scaling of system size

B. Performance Evaluation in presence of application-specific traffic

In order to evaluate the performance of the different NoC architectures with non-uniform traffic patterns we considered both synthetic and application dependent traffic distributions. This result highlights the advantages of the Small World based WiNoC and RFNoC networks which have optimized topologies based on traffic distribution. In the following analysis, a 128 core system size is considered. We have principally considered inter-core communication patterns arising out of these traffic distributions.

We considered 2 types of synthetic traffic to evaluate the performance of the different NoC architectures. First, a transpose traffic pattern [5] is studied where a certain number of cores are

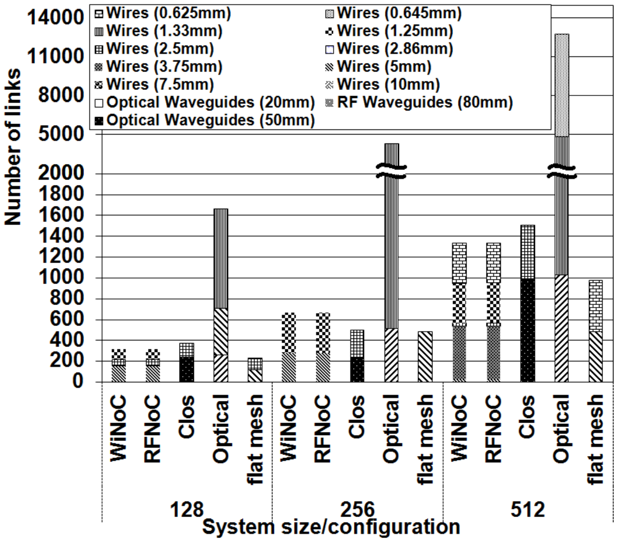

Figure 6. Total wiring, RF and optical waveguide requirement of various lengths for a 20mmx20mm die for different NoCs.

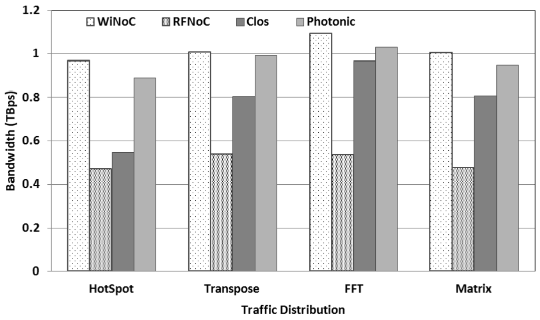

Figure 7. Achievable Bandwidth for non-uniform traffic distributions for different NoCs

considered to communicate more frequently with each other. We have assumed 3 such pairs and 50% of packets generated from one of these cores are targeted towards the other in the pair. The other synthetic traffic pattern considered is the hotspot [5], where each core communicates with a certain number of cores more frequently than with the others. We have assumed 3 such hotspot locations to which all other cores send 50% of the packets that originate from them. To represent a real application a 256-point Fast Fourier Transform (FFT) is considered on the 128 core system. Each core is assigned to perform a 2-point radix 2 FFT computation. The traffic pattern generated in performing multiplication of two 128x128 matrices is also used to evaluate the performance of the NoC architectures.

Fig. 7 shows the bandwidth for non-uniform traffic distributions for all the NoC architectures discussed in this paper. From the bandwidth results it is evident that WiNoC (both configurations with THz and sub-THz links) has a better bandwidth for all the different traffic patterns for 128-core system. This is precisely because WiNoC architecture is optimized taking the traffic distribution into account. The same traffic dependent optimization is carried out in RFNoC. But the achievable bandwidth is less due to less number of available shortcuts and their lower bandwidths. From this performance analysis it becomes clear that WiNoC provides better performance in both uniform and non-uniform traffic scenarios compared to the other alternative NoCs considered in this study.

V. CONCLUSIONS

In this paper we undertook a comparative performance evaluation between small-world wireless/RF NoC and optical NoC architectures. It is shown that small world network employing THz wireless shortcuts provide the best performance-overhead trade-off compared to the optical NoC architectures considered in this study for both uniform and non-uniform traffic distribution.

As part of this on-going investigation, we intend to undertake a more detailed and comprehensive performance benchmark for the small-world wireless NoC architecture with respect to a full spectrum of emerging optical NoCs. We also plan to evaluate the performance of the small-world wireless NoC with respect to other emerging NoCs like 3D NoCs.

VI. ACKNOWLEDGEMENT

This work was supported in part by the US National Science Foundation (NSF) CAREER grant (CCF-0845504).

REFERENCES

1. U. Y. Ogras and R. Marculescu, ""It's a Small World After All": NoC Performance Optimization Via Long-Range Link Insertion", *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, Vol. 14, No. 7, July 2006, pp. 693-706.
2. D. J. Watts and S. H. Strogatz. "Collective dynamics of 'small-world' networks." *Nature* 393, 1998, pp. 440–442.
3. A. Shacham et al., "Photonic Network-on-Chip for Future Generations of Chip Multi-Processors", *IEEE Transactions on Computers*, Vol. 57, no. 9, 2008, pp. 1246-1260.
4. Ajay Joshi et al., "Silicon-Photonic Clos Network for Global On-Chip Communication", *Proceedings of the 3rd International Symposium on Networks-on-Chip (NOCS-3)*, May 2009, pp. 124-133.
5. A. Kumar et al., "Toward Ideal On-Chip Communication Using Express Virtual Channels", *IEEE Micro*, Vol. 28, Issue 1, January-February 2008, pp. 80-90.
6. M. F. Chang et al., "CMP Network-on-Chip Overlaid With Multi-Band RF-Interconnect", *Proceedings of IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 16-20 February, 2008, pp. 191-202.
7. 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", *IEEE Transactions on Computers*, vol. 57, no. 9, September 2008, pp. 1230-1245.
8. S. B. Lee et al., "A Scalable Micro Wireless Interconnect Structure for CMPs", *Proceedings of ACM Annual International Conference on Mobile Computing and Networking (MobiCom),* September, 2009, pp. 20-25.
9. K. Kempa, et al., "Carbon Nanotubes as Optical Antennae," *Advanced Materials*, vol. 19, 2007, pp. 421-426.
10. Partha Pratim Pande, Amlan Ganguly, Kevin Chang, Christof Teuscher, "Hybrid Wireless Network-on-Chip: A New Paradigm in Multi-Core Design", *Proceedings of International Workshop on Network-on-Chip Architectures (NoCArc)*, December 12, 2009, pp. 71-76.
11. A. Ganguly et al., "Performance Evaluation of Wireless Networks on Chip Architectures", *Proceedings of the IEEE International Symposium on Quality Electronic Design (ISQED)*, 16th-18th March 2009.
12. J. Duato et al., Interconnection Networks – An Engineering Approach, Morgan Kaufmann, 2002.
13. B.G. Lee et al., "Ultrahigh-Bandwidth Silicon Photonic Nanowire Waveguides for On-Chip Networks," *IEEE Photonics Technology Letters*, vol. 20, no. 6, Mar. 2008, pp. 398-400.
14. Y. Huang et al., "Performance Prediction of Carbon Nanotube Bundle Dipole Antennas", *IEEE Transactions on Nanotechnology*, Vol. 7, No. 3, May 2008, pp. 331-337
15. W. M. J. Green et. al., "Ultra-compact, low RF power, 10Gb/s silicon Mach-Zehnder modulator", *Optics Express*, Vol. 15, No.25, pp. 17106-17113.
16. S. Kirkpatrick et al., "Optimization by Simulated Annealing". *Science*. New Series 220 (45978): 671-680.
17. Circuits Multi-Projects. http://cmp.imag.fr