TuE4.2 (Invited)
16:00 – 16:30

# Silicon Photonic Interconnection Networks for Data Centers

Madeleine Glick[1], Sébastien Rumley[2], Gouri Dongaonkar[2], Qi Li[2], Keren Bergman[2] and Raj Dutt[1]
1: APIC Corporation, 5800 Uplander Way, Culver City, CA 90230, USA
2: Department of Electrical Engineering, Columbia University, 500 W. 120th Street, New York, NY 10027, USA
glick@apichip.com

**Abstract:** To improve data center performance we seek to increase bandwidth and reduce power consumption. Integrated silicon photonics can help achieve both. System level simulations are required to demonstrate network level performance advantages (latency, execution time).
**OCIS codes:** (060.4250) Networks; (200.4650) Optical Interconnects

## 1. Introduction

Increasing bandwidth and reducing power consumption remain goals for the data center due to the ever increasing traffic demands. Low latency, previously only seen as a requirement for supercomputers, is also becoming a metric for the data center as its size increases and web search results gather information from more servers, the increasing "east-west" traffic. To meet these requirements, the current trend is towards "scale out"; including reducing bottlenecks by adding more low cost hardware [1]. However, with no end in sight to the increasing traffic demands we can envision that the scale out solution will also hit a limit. Considerable effort has been invested into exploring solutions to the bandwidth and energy challenges, through exploiting the high bandwidth capabilities of optics for switching and routing [2, 3]. Semiconductor optical amplifiers with their gain properties seemed an ideal solution for high speed reconfiguration and to overcome losses in switch fabrics [4, 5]. However, certainly as discrete devices, these require too much power to demonstrate a clear advantage. Architectures using MEMS optical switches have been employed for their commercial availability at relatively high port counts [6, 7, 8]. Their commercial availability is a major advantage enabling system level testing and demonstrations; however, there remain challenges due to the slow reconfiguration times. Silicon photonics has the potential to solve these challenges and offer improved network functionality.

## 2. Silicon photonics

Silicon photonics based device capabilities and manufacturability are advancing rapidly including significant advances in: ring resonator temperature stabilization [9], integrated switch fabrics [10, 11] and hybrid silicon technology [12]. Integrated silicon (and InP based) photonics reduce power consumption and cost in addition to their high speed switching capabilities. As the device technology matures, system simulations to the board, rack and cluster level are required to demonstrate the compelling advantages of their adoption in terms of improved application performance, i.e. execution time.

## 3. System Simulation Tools

To convince computer architects to incorporate photonics into their designs system level simulations are required. These need to compare network architectures involving electrical or optical components (or a mix of both), based on various design choices, for instance the topology. (bus, star, hierarchical, mesh). The optimum solution will depend on the specific traffic patterns and applications. For instance, real time video processing applications require considerable computational power and low latency, while search applications in a data center might sacrifice latency if power savings can be achieved. A simulation environment should give an accurate description of trade-offs for the different applications.

Many groups have explored simulation environments for optical networks on chip [13, 14, 15]. DSENT [13] incorporates simulations for timing, area and power. It can also be used to generate traffic dependent power traces. Reference 14 uses gem5 for cycle accuracy and a capability to model the network on chip up to the operating system level. In Reference 16, the authors study the signal degradation of interconnection networks based on micro ring resonator switching elements. The PhoenixSim simulator has been used to compare chip-scale electronic, photonic mesh and photonic crossbar of different network sizes and for different message sizes [15]. Hendry et al. [17] also used it to compare electronic and photonic solutions for Fast Fourier Transform, matrix multiply and projective transform applications for high performance embedded computing and analyze the results for power consumption, performance (operations per second) and efficiency.

Each of these simulators is the result of different modeling formulations, focuses on different aspect of the system and achieves a different trade-off between accuracy v. complexity and time required for execution. Each tool might also deliver metrics that others ignore. Cycle accurate simulators allow fine characterization of the sub-system performance. They also allow one to verify the correctness of the modeled system. Simulators embedding detailed power models capture the power consumption of the system under dynamic conditions. PhoenixSim ensures that the impact of each device is correctly taken into account in order to carefully reflect the optical signal quality.

However, simulators involving too fine an analysis of each component might be unable to simulate data-center sized networks. First, the amount of memory required to store the state of each component might be simply too large for the computing resources available. If memory is not a problem, the time required to conduct a simulation will likely be an obstacle. In [20], data center networks are reported to experience congestion duration of the order of 10s. Simulators should thus be expected to be able to simulate a real-sized network for at least this time scale.

With this in mind, we are extending the range of our studies to board and rack scale systems [18, 19] with LWSim (Lightweight Sim) which focuses on the traffic dynamics of the network only at simulation time, but incorporates the parameterized models for photonics devices of PhoenixSim at the model construction. More generally, we believe that rather than integrating all the possible models into one simulator, a preferable method consists in having a suite of simulators, each one focusing on a given time scale, level of detail and/or metric of interest

## 4. Conclusions

Silicon photonics has the potential to address power consumption and bandwidth challenges of the data center. System level modeling and simulation are necessary to evaluate performance trade-offs in high performance computing and data center applications. But as one tool cannot answer all questions, the right tool chain is still to be defined (how many tools, which tool and which model for what?). Recent advances and results will be reviewed.

## 5. References


[1] A.Vahdat, et al, “Scale-Out Networking in the Data Center,” in IEEE Micro, (July/August 2010), pp. 29-41.

[2] C. Kachris, I. Tomkos, “A survey on optical interconnects for data Centers,” IEEE Communication Surveys & Tutorials, pp. 1–16, 2012.

[3] RS Tucker, “Scalability and energy consumption of optical and electronic packet switching,” *IEEE Jour. Lightwave Technology.*, pp. 2011-2421, 2011.

[4] K.A. Williams, G.F. Roberts, T. Lin, R.V. Penty, I.H. White, M. Glick, D. McAuley, “Integrated optical 2x2 switch for wavelength multiplexed interconnects,” IEEE Journal Sel. Topics Quantum Electronics, vol. 11, pp. 78-85, Jan/Feb 2005.

[5] Assaf Shacham, Benjamin A. Small, Odile Liboiron-Ladouceur, and Keren Bergman, “A Fully Implemented 12x 12 Data Vortex Optical Packet Switching Interconnection Network”, IEEE Jour. Lightwave Technology, pp.3066-3075, 2005.

[6] M. Glick, D. G. Andersen, M. Kaminsky, L. Mummert,“Dynamically Reconfigurable Optical Links for High-Bandwidth Data Center Networks*,” OFC/NFOEC* 2009.

[7] N Farrington, G Porter, S Radhakrishnan, HH Bazzaz, V Subramanya, Y Fainman, Helios: a hybrid electrical/optical switch architecture for modular data centers, *ACM SIGCOMM Computer Communication Review*, pp. 339-350, 2010.

[8] L. Schares, X.J. Zhang, R. Wagle, D. Rajan, P. Selo,S.P. Chang, J. Giles, K. Hildrum, D. Kuchta, J. Wolf, E.Schenfeld, “A reconfigurable interconnect fabric with optical circuit switch and software optimizer for stream computing systems,” *OFC/NFOEC* 2009.

[9] B. Guha, K Preston, M Lipson, “Athermal silicon microring electro-optic modulator,” *Optics Letters*, pp. 2253-2255, 2012.

[10] B. Lee et al. “Four and Eight-Port Photonic Switches Monolithically Integrated with Digital CMOS Logic Drivers and Circuits,” *OFC/NFOEC* 2013.

[11] P. DasMahapatra, A. Rohit, R. Stabile, K. A. Williams, “Broadband 4x4 Switch Matrix using Fifth-order Resonators,” *OFC/NFOEC* 2013.

[12] M. J. R. Heck et al. “Hybrid Silicon Photonic Integrated Circuit Technology,” to appear in *IEEE Jour. Sel Topics Quantum Electronics*.

[13] Chen Sun et al., "DSENT -- A Tool Connecting Emerging Photonics with Electronics for Opto-Electronic Networks-on-Chip Modeling," *6th International Symposium on Networks-on-Chip (NOCS)*, Lyngby, Denmark, 2012.

[14] A. VanLaer, T. Jones, P. Watts, “Full System Simulation of Optically Interconnected Chip Multiprocessors Using gem5,” *OFC/NFOEC* 2013.

[15] J.Chan, G. Hendry, K. Bergman, L. P. Carloni, “Physical-Layer Modeling and System-Level Design of Chip-Scale Photonic Interconnection Networks*,” IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* 30, 1507-1520 (2011).

[16] A. Bianco et al., “Optical Interconnection Networks Based on Microring Resonators,” *J. Opt. Commun. Netw*., pp.546-556, 2012.

[17] G. Hendry, E. Robinson, V. Gleyzer, J. Chan, L. P. Carloni, N. Bliss, K. Bergman, “Circuit-Switched Memory Access in Photonic Interconnection Networks for High-Performance Embedded Computing,” *Supercomputing (SC*), 2010.

[18] Q. Li, R. Hendry, J. Chan, K. Bergman, M. Glick, R. Dutt, “Network Simulation of Passive Optical Broadcast-and-Select Network for Avionics Applications,” *Government Microcircuit Applications and Critical Technology Conference (38th Annual GOMACTech Conference)* 2013.

[19] G. Dongaonkar, S. Rumley, Q. Li, K. Bergman, M. Glick, “Ultra-low Latency Optical Switching for Short Message Sizes in Cluster Scale Systems,” *IEEE Optical Interconnects Conference 2013*, 2013.

[20] S. Kandula, S. Sungupta, A. Greenberg, P. Patel, R. Chaiken, “The Nature of Datacenter Traffic: Measurements & Analysis,” *ICM*, 2009.