A Multilayer Nanophotonic Interconnection Network for On-Chip Many-core Communications

Xiang Zhang and Ahmed Louri

Department of Electrical and Computer Engineering, The University of Arizona

1230 E Speedway Blvd.,Tucson, AZ, USA, 85721

{zxkidd, louri}@email.arizona.edu

ABSTRACT

Multi-core chips or chip multiprocessors (CMPs) are becoming the de facto architecture for scaling up performance and taking advantage of the increasing transistor count on the chip within reasonable power consumption levels. The projected increase in the number of cores in future CMPs is putting stringent demands on the design of the on-chip network (or network-on-chip, NOC). Nanophotonic interconnects have recently emerged as a viable alternate technology solution for the design of NOC because of their higher communication bandwidth, much reduced power consumption and wiring simplification. Several photonic NOC approaches have recently been proposed. A common feature of almost all of these approaches is the integration of the entire optical network onto a single silicon waveguide layer. However, keeping the entire network on a single layer has a serious implication for power losses and design complexity due to the large amount of waveguide crossings. In this paper, we propose MPNOC: a multilayer photonic networks-on-chip. MP-NOC combines the recent advances in silicon photonics and three-dimensional (3D) stacking technology with architectural innovations in an integrated architecture that provides ample bandwidth, low latency, and energy efficient on-chip communications for future CMPs. Simulation results show MPNOC can achieve 81.92 TFLOP/s peak bandwidth and an energy savings up to 23% compared to other proposed planar photonic NOC architectures.

Categories and Subject Descriptors

B.4.3 [**Hardware**]: Interconections—*Topology*
; C.1.2 [**Computer Systems Organization**]: Multiprocessors—*Interconnection architectures*

General Terms

Design, Performance

Keywords

Silicon photonics, Interconnection networks, 3D, CMP

1. INTRODUCTION

The ITRS Semiconductor roadmap [1] predicts that CMOS feature sizes will shrink from 45nm to sub-22nm regime within the next 5 years. Additionally, it has been projected that by 2017 [25], up to 256 general-purpose cores can be put on a single die. The proliferation of multiple cores on the same chip heralded the advent of a communication-centric system wherein the design of the on-chip network connecting various modules, namely the cores, the cache banks, the memory units, and the I/O devices has become extremely critical [3].

Nanophotonic interconnects are under serious consideration for providing the communication needs of future CMPs especially for long metallic wires [14, 19, 6]. Silicon waveguides can propagate end to end signals 70% faster than optimized and repeated global wires.[9]. A number of 2D planar nanophotonic on-chip interconnects have been proposed recently[20, 4, 7, 25, 24, 11, 10]. However, the design of planar nanophotonic on-chip networks is proving to be very challenging and may not be scalable due to power consumption and wiring complexity. For a large scale on-chip interconnect system, signal paths will have a large amount of waveguide crossings. This results in a significant optical signal power loss and back-reflection due to the changes in refractive index at the crossing points[9].

Recently, the semiconductor industry has proposed 3D stacking technology as the next growth engine for performance improvement [5]. The emerging 3D stacking technology has provided new design dimension for on-chip networks. Several 3D metallic-based interconnection network designs have been proposed and have shown a tangible improvement in performance and power savings over 2D interconnections[21, 13, 26]. A prevalent way to connect these layers vertically is using through silicon vias(TSVs)[22]. The pitch of these vertial vias is very small (4μm$\sim$$10\mu$m), and can be further reduced to 1μm [18]. The delay of these vertical lines is generally very small, only 20ps over a 20-layer stack. Unfortunately, 3-D metallic-based interconnection networks, still inherent the fundamental physical limits of electrical signaling and this will be compounded by the thermal and power challenges of 3-D stacking technologies.

In this paper, we leverage the advantages of two emerging technologies, namely, silicon photonics and 3D stacking with architectural innovations to design a high bandwidth, low latency, energy-efficient on-chip network called MPNOC: a
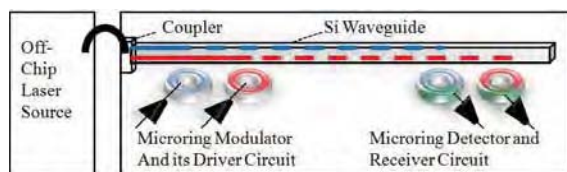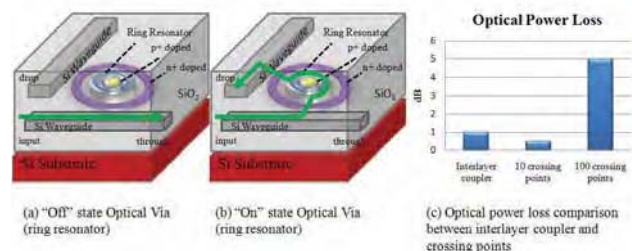
Figure 1: A typical on-chip nanophotonic link

Figure 2: Illustration of an optical via built using a microring resonator. (a) 'ON' state (b) 'OFF' state (c) Power loss comparison between optical via and planar waveguide with various number of waveguide crossings

multilayer photonic networks-on-chip. The proposed architecture targets 256 cores CMPs and 22nm CMOS technology. On the architecture side, MPNOC provides a global crossbar-like connectivity with much improved power efficiency and performance.

The remainder of the paper is organized as follows. In Section 2, we review the recent advances in 3-D silicon photonics as they apply to MPNOC. We provide a detailed description of the proposed MPNOC in Section 3. We evaluate its performance in Section 4, and we conclude the paper in Section 5.

2. 3-D SILICON PHOTONICS

A silicon photonic integrated circuit requires a laser source, a modulator and its driver circuit, medium (Si waveguide), a photodetector and on-chip off-chip interface (coupler) as shown in Figure 1. External laser source generates light with multiple wavelengths. Light is carried by the optical fiber and coupled to the silicon waveguide. The waveguide passes through an array of microring modulators. Each ring is tuned to a different wavelength to modulate the intensity of the light of that wavelength. At the receiver end, an array of tuned microring Ge-doped detectors absorbs the light and converts signal back to electrical domain.

Recent advances in silicon photonics have opened up the door to design 3D on-chip nanophotonic interconnects. Jalali group at UCLA has fabricated a SIMOX (Separation by IMplantation of OXygen) 3-D sculpting to stack optical devices in multiple layers[15]. Lipson group at Cornell has successfully buried active optical ring modulators in polycrystalline silicon[23]. Another interesting device is optical vias (Interlayer coupler) as shown in Figure 2. The basic function is to couple light from one silicon layer to another. According to[4, 9], interlayer coupler introduces a 1dB optical power loss, while each optical waveguide crossing undergoes a 0.05dB loss. If we stack the connections in 3-D as opposed to keeping all the waveguides in 2-D we can realize a very significant energy savings. For example, if we consider one hundred crossing points per waveguide for a 2-D implementation, then using a 3-D, where the waveguides are stacked vertically, we can realize an about 60% optical laser power savings.

3. MPNOC ARCHITECTURE

The proposed architecture comprises 256 cores in 64 tiles on a 400 mm^2 3D IC. As shown in Figure 3, 256 cores are mapped on an 8x8 network with a concentration factor of four. Since the performance and energy of the electrical interconnects are sufficient for short links (<2.5mm), small degree of concentration will significantly reduce system complexity[3]. Each tile is comprised of four cores. Each core has a private L1 cache, and four cores in the same tile share

a L2 cache. The bottom layer, adjacent to the heat sink, contains cores and local caches. One or more high level caches and memory layers in the middle provide the bulk of on-chip storage. The upper part of the chip contains four optical layers implementing the decomposed optical crossbar as will be described later. Silicon photonic devices, such as planar waveguides, couplers, microring resonators, and Ge photodectors, are combined to provide the photonic infrastructure for intra-chip and chip-to-chip communications. Inter-layer communications are realized by TSVs. As discussed in the previous section, such vertical wires occupy very small area and can transmit the signals from top layer to bottom layer in less than one clock cycle.
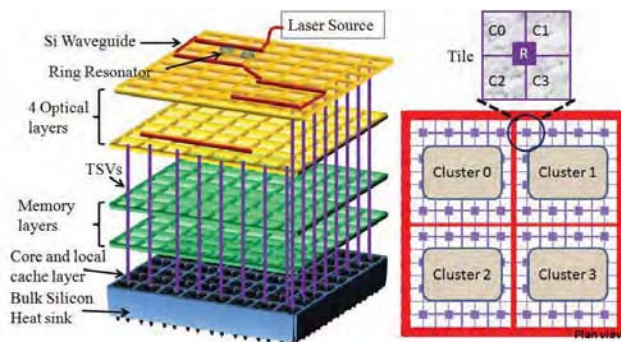

Figure 3: Proposed 256-core 3-D Chip Layout

In MPNOC, waveguides that have the potential to cross each other are laid out on different optical layers (cloverleaf intersection). Ring resonators are interleaved on different layers to minimize potential temperature variation. The proposed architecture takes the advantage of the unique properties of nanophotonic interconnects for global communication channels and switching capabilities of electronics at the router level. Such hybrid combination reduces the power dissipation on long inter-router communication while electrical switching provides flow control to regulate traffic and prevent buffer overflow.

In the proposed 3-D layout, we divide tiles into four clusters based on their physical location. Each cluster contains 16 tiles. Unlike the global 64x64 optical crossbar design in [25] and the hierarchical architecture in [20], MPNOC consists of 16 decomposed optical crossbar slices mapped on
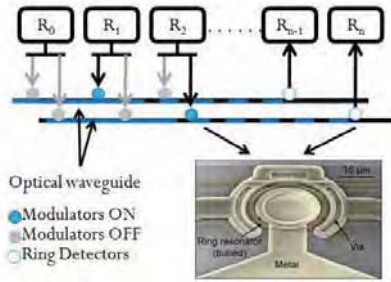
Figure 4: Optical MWSR bus implementation, deposited silicon ring resonators are courtesy of[23]

four optical layers. Each slice is a 16x16 optical crossbar connecting all tiles from one cluster to another (Inter-cluster communication), or all tiles from same clusters (Intra-cluster communication). Figure 4 shows the implementation of each slice in the decomposed optical crossbar. It is composed of a few Multiple-Write-Single-Read (MWSR) nanophotonic channels, which require much less power than Single-Write-Multiple-Read (SWMR) channels described in [20]. Token slot [25] is adopted to improve the arbitration efficiency (up to 100%)for the channel. Each wavelength in the waveguide operates at 10Gb/s. In MPNOC, we consider a 256 bit per phit size to achieve a 2.56Tb/s bandwidth, a 4 waveguide bundle with 64 wavelengths in each waveguide is required for each crossbar channel. Considering the total number of optical channels on the chip, MPNOC can achieve 81.92TFLOPS peak performance (81.92TB/s bandwidth).

Since each optical crossbar channel has multiple senders and a single receiver, we define each optical channel as the home channel for the receiver. A source tile sends packets to a destination tile by modulating the light on the home channel of the destination tile. Off-chip laser source generates 128 continuous wavelengths, $\Lambda = \lambda_0, \lambda_1, \lambda_2, ..., \lambda_{127}$. We divide these wavelengths into two groups. Figure 5 shows the detailed optical device floorplan of optical layer 1. The detailed decomposition and slicing of optical crossbar on four optical layers is shown in Figure 6.
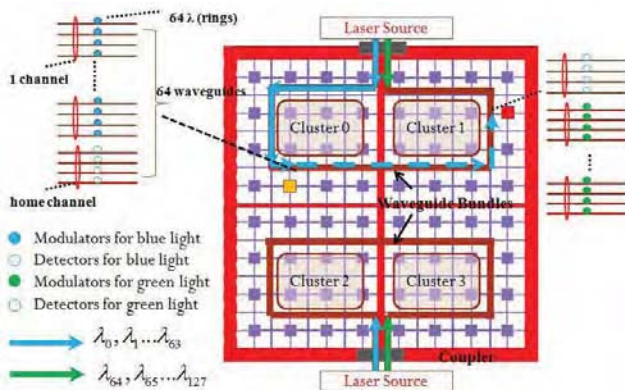

Figure 5: Plan view of optical layer one

Inter-cluster communication: The upper part of the chip in Figure 5 shows the waveguide bundle for inter-cluster communications between Cluster 0 and Cluster 1. Blue wavelengths ($\lambda_0, \lambda_1, ..., \lambda_{63}$) and green wavelengths ($\lambda_{64}, \lambda_{65},$
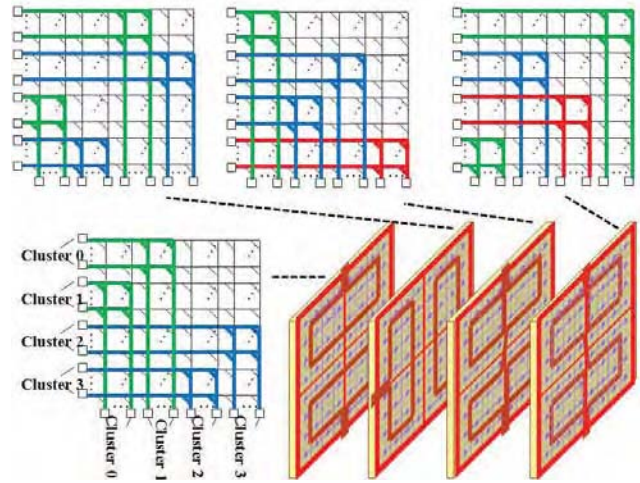

Figure 6: The decomposition, slicing and mapping of the optical crossbar. Color lines of the crossbar represent valid part of optical crossbar on each layer. The lines with same color mean they share the physical waveguides on each layer.

..., λ_{127}) are injected into both ends of the waveguide bundle. The waveguide bundle contains 32 crossbar channels (16 in each direction and physically 64 waveguides) corresponding to each tile of Cluster 0 and Cluster 1. Here the blue wavelengths are assigned as the communication channels for the tiles from Cluster 0 to Cluster 1, and the green wavelengths are assigned as the reversed communication channels (from Cluster 1 to Cluster 0). When the blue wavelengths are coupled onto the waveguide bundle, tiles of Cluster 0 will arbitrate and enable portion of the blue modulators to transmit the packets on these wavelengths. The blue microring photodiodes in the tiles of Cluster 1 will be passively tuned to resonant wavelengths and detect the signals on their home channel.

Intra-cluster communication: Intra-cluster communications are very similar to inter-cluster communications. The difference is that there are 64 wavelengths on each waveguide of the waveguide bundle for intra-cluster communications as opposed to 128 wavelengths for inter-cluster communications. This results in 50% reduction in the number of ring resonators required for intra-cluster communications. Consequently, there is considerable power savings for intra-cluster communications.

Each tile contains an electrical router as shown in Figure 7. The electrical router provides the proper interface to local cores/caches, on-chip nanophotonic interconnects and on-chip/off-chip memory/IO devices. In addition to having the same features as those routers in electrical NOC, the routers of MPNOC should provide interface to receive/generate tokens from/to optical waveguides. Tokens are used for optical crossbar arbitration and flow control.

Arbitration and Flow Control: Our proposed token slot arbitration is slightly different from[25], where only the destination tile can inject one-bit token every clock cycle. Such a modification can add flow control mechanism for the architecture without extra hardware overhead. Tokens are transmitted in arbitration waveguide through piggybacking. Tokens will be generated from the receiver end when
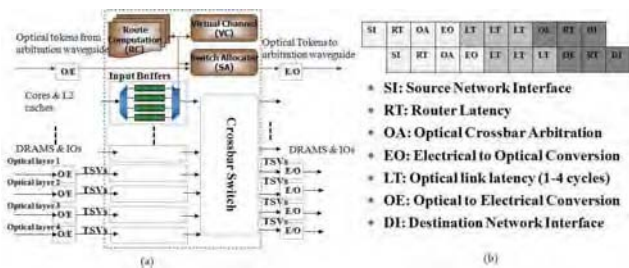
Figure 7: (a) On-chip Network Router architecture for MPNOC, (b) An example of 2-flit packet from source tile to destination tile, assuming the optical transversal latency is 3 clock cycles

there are enough buffers left in the input port. The receiver should reserve enough buffers for the worst case optical token round-trip latency, 12 clocks for inter-cluster communications and 8 for intra-cluster communications. Since the token is one-bit, it only carries the information whether there is an available buffer at the receiver. As a result, when a source router captures the token, it will have the privilege to send one flit (assume flit size = phit size) to the corresponding destination. Successive transmissions depend on whether the successive tokens are captured. Since there are four optical input/output pairs for each router, their tokens are maintained separately and send to different arbitration waveguides on different layers.

4. EVALUATION AND SIMULATION

In this section, we evaluate our proposed architecture and compare the performance, power-efficiency, device requirement and area against alternative architectures.

4.1 Simulation Setup

We first describe the simulation setup of the proposed architecture. A cycle-accurate simulator was modified and developed based on Booksim simulator[8] to support optical networks. The packet injection rate was varied from 0.1 to 0.9 of the network capacity. Since the delay of Optical/Electrical (O/E) and Electrical/Optical (E/O) conversion can be reduced to less than 100ps each, the total optical transmissions latency is determined by physical location of source/destination pair and two additional clock cycles for the conversion delay. Our simulation model includes the pipeline model, router arbitration and contentions, flow control and other overhead. The simulator is warmed up under load without taking measurements until steady-state is reached. An aggressive single cycle electrical router[17] is applied in each tile and the flit transversal time is 1 cycle from the local core to electrical router. A detailed simulation configuration is shown in Table 1.

Table 1: Simulation Configurations

Concentration (# cores of per router)	4
Buffer per input port	64 flits
Phit size (Flit size)	256 bits
Packet size	1 flit
V_{dd}	1.0V
CPU Frequency	5GHz

All the evaluated architectures are 256-core systems listed

in Table 2. They all implemented with a concentration degree of 4. We evaluate MPNOC architecture with two other crossbar-like architectures, CORONA[25] and FIREFLY[20] and one electrical architecture(CMESH)[3] using dimension-ordered routing (DoR). We assume token slot for both MP-NOC and CORONA to pipeline the arbitration process to increase the efficiency. Multiple requests can be sent from four local cores to optical channels to increase the arbitration efficiency. We use Fly_Src routing algorithm for Firefly architectures, which operates intra-cluster communications by electrical mesh link first and then operates inter-cluster communication through optical crossbar.

Table 2: Evaluated Architectures

NAME	Routing	VC#	Description
MPNOC	Token Slot	1	Multilayer optical crossbar
CORONA	Token Slot	1	a single layer optical crossbar
FIREFLY	Src_Fly	1	a hierachical architecture, local electrical mesh link, several global optical crossbars
CMESH	DoR	8	Concentrated mesh network

4.2 Performance

We simulate four architectures on seven sythetic traffic traces[8], including both random uniform traffic patterns and permutation patterns, such as bit-complement(bitcomp), bit-reversal(bitrev), transpose, tornado, neighbor and prefect shuffle. Figure 8 shows the throughput and average network latency per packet for the uniform traffic. The proposed architecture outperforms all the other networks on the uniform traffic. It improves zero load latency by 28%, 43%, and 51% as compared to CORONA, FIREFLY, and CMESH respectively. We observe MPNOC exceeds the throughput to 2.4x and 2x compared to Corona and Firefly.

The throughput study for all traffic traces is shown in Figure 9. The maximum throughput is normalized to 1. We observe MPNOC can achieve 100% throughput in bit-reversal traffic, because there is no contention in the network. MPNOC outperforms FIREFLY and CMESH in most of the traffic patterns and has the same throughput in bit-comp, transpose and shuffle as CORONA. While CMESH and FIREFLY has a better performance in neighbor traffic because such traffic pattern exploits the spatial locality and these two use electrical links for local traffic, MPNOC and Corona has a better performance in global traffic (non-neighbor) traffic patterns, where nanophotonic crossbar can dramatically reduce the hop counts and traverse from source tile to destination tile within one hop. On average, MPNOC provides a 55%, 109% and 233% improvement in throughput
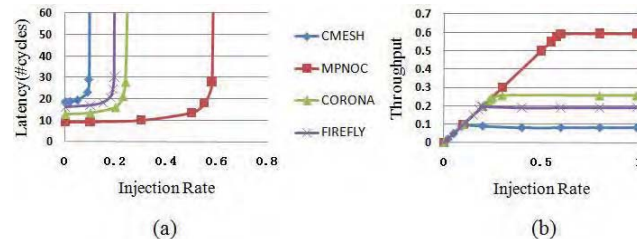

Figure 8: (a) Load-latency curve for uniform; (b) Load-throughput curve for uniform traffic
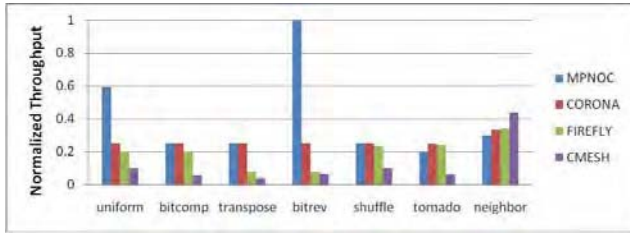
Figure 9: Simulation results showing normalized saturation throughput for seven traffic patterns

compared to CORONA, FIREFLY and CMESH on seven traffic patterns.

4.3 Energy Comparison

The energy consumption of a nanophotonic interconnection network can be divided into two parts, electrical energy and optical energy. Optical energy consists of the off-chip laser energy and on-chip microring resonator heating energy.

4.3.1 Electrical Energy Model

$$E_e = E_{link} + E_{router} + E_{O/E,E/O} \qquad (1)$$

Electrical power includes the energy of link, router and back-end circuit for optical transmitter and receiver. We use ORION 2.0[12] model and modified some parameters for 22nm technology according to [1]. We assume the injection rate of the electrical link is 0.1. The energy of electrical link include both planar links and vertical links (TSVs). The length of electrical planar links in Firefly and CMesh is determined to be 20mm/8=2.5mm. The energy for planar link is conservatively obtained as 0.15pJ/bit under low-swing voltage level. The length of vertical links is very small. For a 10-layer chip, the vertical via is determined as ∼100-200μm[18], which is much less than planar links. As a result, the power consumption of vertical links is very small. We neglect it when we calculate our electrical link power model. For the electrical router power, we assume a 8x8 router consumes 0.30pJ/bit/hop and a 5x5 router with the same buffer size requires 0.22pJ/bit/hop. For each optical transmitted bit, we need to provide electrical back end circuit for transmitter end and receiver end. We assume the O/E and E/O converter energy is 100fJ/b, as predicted in [16].

4.3.2 Optical Energy Model

$$P_{laser} = P_{rx} + C_{loss} + M_s \qquad (2)$$

The optical power budget is the result of the laser power and the power dissipation for the microring resonators. The laser power budget is determined by Equation (2). P_{laser} is the laser power requirement, P_{rx} is the receiver sensitivity, C_{loss} is the channel losses and M_s is the system margin. The ring power comes from the static power: fabrication error trimming and the heating power to keep the ring resonators in the resonance region, the dynamic power: direct data modulation power. In order to perform an accurate comparison with the other two optical architectures, we use the same optical device parameters and loss values from provided in [2, 4], as listed in Table 3.

Table 3: Laser and Ring Power Budget

Component	Value	Unit
Laser efficiency	5	dB
Coupler (Fiber to Waveguide)	1	dB
Waveguide	1	dB/cm
Splitter	0.2	dB
Non-Linearity	1	dB
Ring Insertion & scattering	1e-2 - 1e-4	dB
Ring drop	1.5	dB
Waveguide Crossings	0.05	dB
Photo Detector	0.1	dB
Ring Heating	26	μW/ring
Ring Modulating	500	μW/ring
Receiver Sensitivity	-26	dBm

4.3.3 Synthetic Workload Energy Comparison

Table 4: Power parameters of four architectures

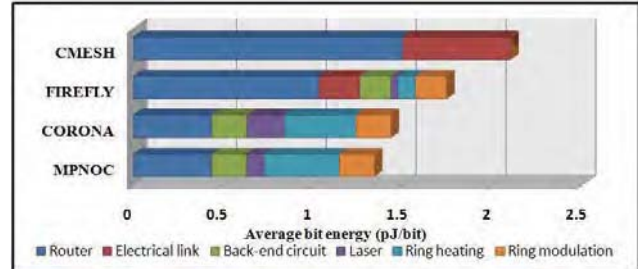|	CORONA	FIREFLY	MPNOC	CMESH
Electrical link	-	0.15pJ/b	-	0.15pJ/b
Router	0.22pJ/b	0.30pJ/b	0.22pJ/b	0.30pJ/b
O/E, E/O	100fJ/b	100fJ/b	100fJ/b	-
Optical channel loss	-25.2dB	-17.6dB	-16.0dB[1]	-
Optical power per λ	0.81mW	0.14mW	0.10mW	-
Laser requirement	13.6W	2.4W	6.1W	-
Ring heating	26W	6.5W	27.5W	-

Figure 10: Average per-bit energy consumption

Based on the energy model discussed in the previous section, we calculate the energy parameters of four architectures as shown in Table 4. We test uniform traffic with 0.1 injection rate to the four architectures and obtain energy per-bit comparison shown in Figure 10. Althrough Firefly has $\frac{1}{4}$ as much as the rings in CORONA and MPNOC, which results in $\frac{1}{4}$ energy consumption per bit on ring heatings, it still consumes more energy per bit than MPNOC and CORONA because of the energy consumption overhead on routers and electrical links. In general, MPNOC saves 6.5%, 23.1%, 36.1% energy per bit compared to CORONA, FIREFLY, and CMESH respectively. It should be noted that when the network injection rate increases, MPNOC becomes much more energy efficient than other three architectures.

4.3.4 Optical Device Requirement

In Figure 11(a), the contour line is the optical link power per wavelength budget in mWatts. The power budget of MPNOC requires further improvement of the ring devices, while FIREFLY requires the further improvement on waveguide propagation loss and CORONA requires both parame-

---

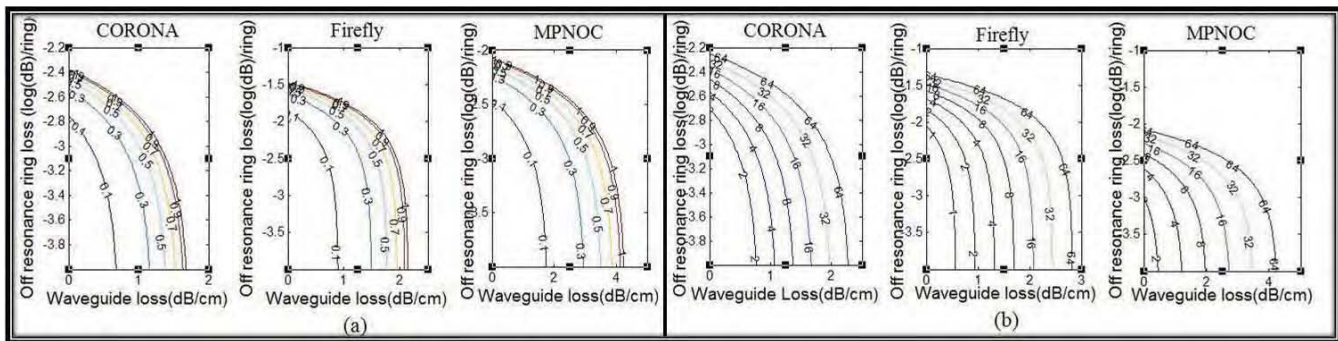[1]-16.0dB for inter-cluster comm. and -14.4dB for intra-cluster comm.

Figure 11: (a) Optical link power per wavelength, (b)Optical Laser Power requirement

ters. In Figure 11(b), we show optical laser power contour in Watts. The total laser power of MPNOC can be limited to 2W with the waveguide propagation loss to 0.3dB/cm and off resonance ring loss to 0.0003dB.

5. CONCLUSIONS

Recent advances in silicon photonics and 3D stacking technology have motivated us to explore multilayer nanophotonic interconnects to meet the performance and power requirements of future many-core CMPs. To this end, we propose MPNOC: a power-efficient multilayer nanophotonic network design for on-chip interconnects. MPNOC can achieve 81.92 TFLOP/s peak performances with reasonable power consumption. Simulation results show the 3D MPNOC approach outperforms 2D photonic designs both for performance and energy savings.

6. ACKNOWLEDGEMENT

7. REFERENCES

[1] http://www.itrs.net.
[2] J. Ahn and et al. Devices and architectures for photonic chip-scale integration. *Applied Physics A: Materials Science and Processing*, 95(4):989–997, June 2009.
[3] J. Balfour and W. Dally. Design tradeoffs for tiled cmp on-chip networks. In *ICS '06*, pages 187–198, Cairns, Queensland, Australia, 2006.
[4] C. Batten and et al. Building manycore processor-to-dram networks with monolithic silicon photonics. In *HOTI '08*, pages 21–30, Stanford, CA, USA, 2008.
[5] B. Black and et al. Die stacking (3d) microarchitecture. In *MICRO 39*, pages 469–479, 2006.
[6] G. Chen and et al. Predictions of cmos compatible on-chip optical interconnect. *Integr. VLSI J.*, 40(4):434–446, 2007.
[7] M. J. Cianchetti and et al. Phastlane: a rapid transit optical routing network. In *ISCA '09*, pages 441–450, Austin, TX, USA, 2009.
[8] W. Dally and B. Towles. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2003.
[9] R. K. Dokania and A. B. Apsel. Analysis of challenges for on-chip optical interconnects. In *GLSVLSI '09*, pages 275–280, 2009.
[10] H. Gu and et al. A low-power low-cost optical router for optical networks-on-chip in multiprocessor systems-on-chip. *ISVLSI '09*, 0:19–24, 2009.
[11] A. Joshi and et al. Silicon-photonic clos networks for global on-chip communication. In *NOCS '09*, pages 124–133, 2009.
[12] A. Kahng and et al. Orion 2.0: A fast and accurate noc power and area model for early-stage design space exploration. In *DATE*, pages 423–428, April 2009.
[13] J. Kim and et al. A novel dimensionally-decomposed router for on-chip communication in 3d architectures. *SIGARCH Comput. Archit. News*, 35(2):138–149, 2007.
[14] N. Kirman and et al. Leveraging optical technology in future bus-based chip multiprocessors. In *MICRO 39*, pages 492–503, 2006.
[15] P. Koonath and B. Jalali. Multilayer 3-d photonics in silicon. *Opt. Express*, 15(20):12686–12691, 2007.
[16] A. Krishnamoorthy and et al. Computer systems based on silicon photonic interconnects. *Proceedings of the IEEE*, 97(7):1337–1361, July 2009.
[17] A. Kumar and et al. A 4.6tbits/s 3.6ghz single-cycle noc router with a novel switch allocator in 65nm cmos. In *ICCD '07*, October 2007.
[18] G. H. Loh. 3d-stacked memory architectures for multi-core processors. *SIGARCH Comput. Archit. News*, 36(3):453–464, 2008.
[19] D. Miller. Device requirements for optical interconnects to silicon chips. *Proceedings of the IEEE*, 97(7):1166–1185, July 2009.
[20] Y. Pan and et al. Firefly: illuminating future network-on-chip with nanophotonics. In *ISCA '09*, pages 429–440, Austin, TX, USA, 2009.
[21] D. Park and et al. Mira: A multi-layered on-chip interconnect router architecture. In *ISCA '08*, pages 251–261, 2008.
[22] S. Pasricha. Exploring serial vertical interconnects for 3d ics. In *DAC '09*, pages 581–586, 2009.
[23] K. Preston and et al. Deposited silicon high-speed integratedelectro-optic modulator. *Opt. Express*, 17(7):5118–5124, 2009.
[24] A. Shacham and et al. On the design of a photonic network-on-chip. In *NOCS '07*, pages 53–64, 2007.
[25] D. Vantrease and et al. Corona: System implications of emerging nanophotonic technology. In *ISCA '08*, pages 153–164, Beijing, China, 2008.
[26] Y. Xu and et al. A low-radix and low-diameter 3d interconnection network design. In *HPCA '09*, pages 30–42, 2009.