LumiNOC: A Power-Efficient, High-Performance, Photonic Network-on-Chip for Future Parallel Architectures

Cheng Li, Mark Browning, Paul V. Gratz and Samuel Palermo
{seulc, mabrowning, pgratz, spalermo}@tamu.edu
Department of Electrical and Computer Engineering, Texas A&M University, College Station, Texas 77843-3128

ABSTRACT

Achieving scaling performance as core counts increase to the hundreds in future chip-multi-processors (CMPs) requires high performing, yet energy-efficient interconnects. Silicon nanophotonics is a promising replacement for electronic on-chip interconnect due to its high bandwidth and low latency, however, prior techniques have required high static power for the laser and ring thermal tuning. We propose a novel nano-photonic NoC architecture, LumiNOC, optimized for high performance and power-efficiency. In a 64-node NoC under synthetic traffic, LumiNOC enjoys 50% lower latency at low loads and 40% higher throughput per Watt on synthetic traffic, versus other reported photonic NoCs. LumiNOC reduces latencies 40% versus an electrical 2D mesh NoCs on the PARSEC shared memory, multithreaded benchmark suite.

Categories and Subject Descriptors

C.1.2 [**Multiple Data Stream Architectures**]: Interconnection arch.

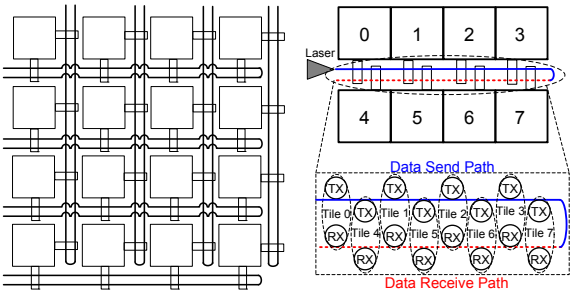
Keywords

CMP; NoC; Synthetic/Realistic Workload; Power Efficiency

1. INTRODUCTION

Monolithic silicon photonics have been proposed as a scalable alternative to meet future many-core systems bandwidth demands, however, current photonic NoC (PNoC) architectures suffer from high static power demands, high latency and low efficiency, making them less attractive than their electrical counterparts. For example, conventional photonic crossbars [1, 2], achieve nearly uniform latency and high bandwidth, however, channels are dedicated to each node and cannot be flexibly shared by the others. Due to the unbalanced traffic distribution in realistic workloads, channel bandwidth cannot be fully utilized. This leads to inefficient energy usage, since the static power is constant regardless of traffic load. Over-provisioned channels also implies higher ring resonator counts, which must be maintained at the appropriate trimming temperature, consuming on-chip power. In addition, many photonic NoCs globally route waveguides in a bundle, connecting all the tiles in the CMP. The long waveguide leads to significant laser power losses over the long distance. Moreover, the power and overhead introduced by the separated arbitration channels or networks in previous photonic NoCs can lead to further power efficiency losses. In future latency and power constrained CMPs, these characteristics will hobble the utility of photonic interconnect. In this paper we propose LumiNOC, a novel PNoC architecture which addresses the power wasted due to channel over-provisioning, while reducing latency and maintaining high bandwidth. LumiNOC makes three contributions: *First*, instead of conventional, globally distributed, photonic channels, requiring high laser power, we propose a novel channel sharing arrangement composed of sub-sets of cores in photonic subnets. *Second*, we propose a novel, purely photonic, distributed arbitration mechanism, dynamic channel scheduling, which achieves extremely low-latency without degrading throughput. *Third*, our photonic network architecture leverages the same wavelengths for channel arbitration and parallel data transmission, allowing efficient utilization of the photonic resources, lowering static power consumption.

PACT'12, September 19–23, 2012, Minneapolis, Minnesota, USA.
ACM 978-1-4503-1182-3/12/09.



(a) (b)

Figure 1: (a) LumiNOC interconnection of CMP with 16 tiles, (b) One-row subnet of eight nodes. Circles (TX and RX) represent groups of rings; one dotted oval represents a tile.

2. LUMINOC ARCHITECTURE

The LumiNOC design breaks the network into several smaller networks (subnets) with shorter waveguides. Fig. 1a shows an example of this subnet inter-connected 16-node CMP system. All tiles are interconnected by two different subnets, one horizontal and one vertical. If a sender and receiver do not reside in the same subnet, transmission requires a hop through an intermediate node's electrical router. In this case, transmission experiences longer delay due to the extra O/E-E/O conversions and router latency.

Fig. 1b details the shared channel for a LumiNOC one-row subnet design. Each tile contains W modulating "Tx rings" and W receiving "Rx Rings", where W is the number of wavelengths multiplexed in the waveguide. Since the optical signal uni-directionally propagates in the waveguide from its source at off-chip laser, each node's Tx rings are connected in series on the "Data Send Path", shown in a solid line from the laser, prior to connecting each node's Rx rings on the "Data Receive Path", shown in a dashed line. In this "double-back" waveguide layout, modulation by any node can be received by any other node; furthermore, the node which modulates the signal may also receive its own modulated signal, a feature that is leveraged in our collision detection scheme in the arbitration phase. The same wavelengths are leveraged for arbitration and parallel data transmission. At any given time a multi-wavelength channel with N nodes may be in one of three states: **Idle** - All wavelengths are un-modulated and the network is quiescent. **Arbitration** - One more sender nodes are modulating N copies of the arbitration flags; one copy to each node in the subnet (including itself) with the aim to gain control of the channel. **Data Transmission** - Once a sender has established ownership of the channel, it modulates all channel wavelengths in parallel for data transmission.

To send a packet, a node first waits for the channel to enter the **Idle** state. Then, it places the link in the **Arbitration** state and modulates a copy of the arbitration flags to the appropriate arbitration wavelengths for each of the N nodes. A "1-hot" source address serves as collision detection mechanism: since the subnet is operated synchronously, any time multiple nodes send overlapping arbi-

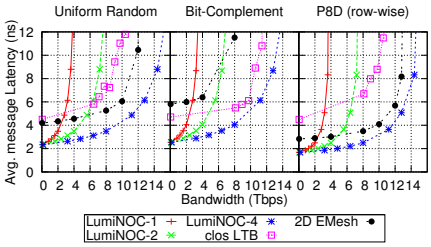


Figure 2: LumiNOC vs. Clos LTBw and electrical network under synthetic workloads.

Literature		ELP (W)	TTP (W)	ERP (W)	EO/OE (W)	ITP (Tbps)	RTP (Tpbs)	TP (W)	**RTP/W (Tbps/W)**
EMesh [3]		NA	NA	NA	NA	10	3.0	26.7	**0.1**
Corona [2]		26.0	21.00	0.52	4.92	160	73.6	52.4	**1.4**
FlexiShare [1]		5.80	11.00	0.13	0.60	20	9.0	17.5	**0.5**
Clos [4]		3.30	0.14	0.10	0.54	18	10.0	4.1	**2.4**
LumiNOC	1-Layer	0.35	0.33	0.13	0.30	10	4.0	1.1	**3.6**
	2-Layers	0.73	0.65	0.26	0.61	20	8.0	2.3	**3.4**
	4-Layers	1.54	1.31	0.52	1.22	40	16.0	4.6	**3.4**

Table 1: Power efficiency comparison of different photonic NoC architectures.

tration flags, the "1-hot" precondition is violated and all nodes are aware of the collision. We leverage self-reception of the arbitration flag: right after sending, the node monitors the incoming arbitration flags. If uncorrupted, then the sender succeeded arbitrating the channel and the channel proceeds to the **Data Transmission** phase. If the arbitration flags are corrupted (>1 is hot), then a conflict occurred. In this event, conflicting nodes back-off a specified number of cycles dependent upon their node id and the current clock cycle, and attempt arbitration again.

3. IMPLEMENTATION AND EVALUATION

We develop a baseline physical implementation of the general LumiNOC architecture for the evaluation of LumiNOC against competing PNOC architectures. We assume a 400 mm^2 chip implemented in a 22nm CMOS process and containing 64 square tiles that operate at 5GHz. Each tile contains a processor core, private caches, a fraction of the shared last-level cache, and a router connecting it to one horizontal and one vertical photonic subnet. The effective waveguide length is 4.0 cm, yielding a propagation delay of 2.7 cycles at 10GHz network frequency. The wavelengths per waveguide is limited to 32 to reduce the ring through loss. To evaluate this implementation's performance, we use a cycle-accurate, microarchitectural-level network simulator. The network was simulated under both synthetic and realistic workloads. LumiNOC designs with 1, 2, and 4 Network Layers are simulated to show results for different bandwidth design points.

Synthetic Workload Results: in Fig. 2, the LumiNOC design is compared against the electrical and Clos networks under *uniform random*, *bit complement*, and *P8D*. The figure shows the low-load latencies of the LumiNOC design are much lower than the competing designs. This is due primarily to the lower diameter of the LumiNOC topology, destinations within one subnet are one "hop" away while those in a second subnet are two. The 1-layer network saturates at 4Tbps realistic throughput as determined by analyzing the offered vs accepted rate.

The different synthetic traffic patterns bring out interesting relationships. On the *P8D* pattern, which is engineered to exploit the physical locality inherent in the LumiNOC and Clos designs, all have universally lower latency than on other patterns. However, while both the electrical and LumiNOC network have around 25% lower low-load latency than uniform random, Clos only benefits by a few percent from this optimal traffic pattern. At the other extreme, the electrical network experiences a 50% increase in no-load latency under the bit-complement pattern compared to uniform random while both Clos and the LumiNOC network are only marginally affected. This is due to the LumiNOC having a worst-case hop count of 2 and not all routes go through the central nodes as in the electrical network. Instead, the intermediate nodes are well distributed through the network under this traffic pattern. However, as the best-case hop count is also 2 with this pattern, the LumiNOC network experiences more contention and the saturation bandwidth is decreased as a result.

Realistic Workload Results: Fig. 3 shows the performance of the LumiNOC network in 1-, 2- and 4-layers, normalized against the performance of the baseline electrical NoC. Even with one layer, the average message latency is about 10% lower than the electrical

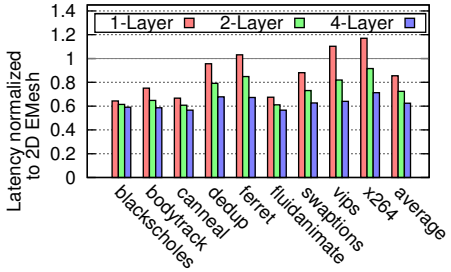


Figure 3: Message Latency in PARSEC benchmarks for LumiNOC compared to electrical network.

network. With additional network layers, LumiNOC has approximately 40% lower average latency. These results are explained by examining the bandwidth-latency curves in Fig. 2. The average offered rates for the PARSEC benchmarks are of the order of 0.5Tbps, so these applications benefit from LumiNOC's low latency while being well under even the 1-layer, LumiNOC throughput.

Power Efficiency: in order for a fair comparison versus other reported PNoC architectures, we refer to the photonic loss of various photonic devices reported by Joshi et al. [4] and Pan et al. [1]. Total power (TP) consists of the electrical laser power, thermal tuning power required to maintain microring resonant at the work wavelength, electrical router power, and electrical to optical/optical to electrical conversion power. Power efficiency is compared in Table 1. **ITP** is the ideal throughput of the design, while **RTP** is the maximum throughput of the design under a uniform random workload as shown in Fig. 2. A 2GHz electrical 2D-mesh [3] was scaled to 8×8 nodes operating at 5GHz, in a 22nm CMOS process, to compare against the photonic networks.

The table shows that LumiNOC has the highest power efficiency of all designs compared in RTP/Watt, increasing efficiency by ∼40% versus the nearest competitor, Clos [4]. By reducing wavelength multiplexing density, utilizing shorter waveguides, and leveraging the data channels for arbitration, LumiNOC consumes the least electrical laser power (ELP) among all the compared architectures. A 4-layer LumiNOC consumes ∼1/4th the ELP of a competitive Clos architecture, of nearly the same throughput.

4. CONCLUSIONS

In this work, LumiNOC addresses power inefficiencies issues of prior photonic NoCs by adopting a shared-channel, photonic on-chip network with a novel, in-band arbitration mechanism to efficiently utilize power, achieving a high performance and scalable interconnect with extremely low latency.

References

[1] Y. Pan et al., "FlexiShare: Channel Sharing for an Energy-Efficient Nanophotonic Crossbar," *HPCA*, pp. 1-12, 2010.

[2] D. Vantrease et al., "Corona: System Implications of Emerging Nanophotonic Technology," *ISCA*, pp. 153-164, 2008.

[3] J. Howard et al., "A 48-Core IA-32 Processor in 45 nm CMOS Using On-Die Message-Passing and DVFS for Performance and Power Scaling," *JSSC*, vol. 46, no. 1, pp. 173-183, Oct., 2011.

[4] A. Joshi et al., "Silicon-Photonic Clos Networks for Global On-Chip Communication," *NOCS*, pp. 124-133, 2009.