NORTHWESTERN
UNIVERSITY

Electrical Engineering and Computer Science Department

Technical Report
Number: NU-EECS-14-03

April, 2014

LaC: Integrating Laser Control in a Photonic Interconnect

Yigit Demir, Nikos Hardavellas

Abstract

The high-speed and low-cost modulation of light make photonic interconnects an attractive solution to the communication demands of manycore processors. However, the high optical loss of many nanophotonic components results in high power requirements for the laser source. Most of this power is wasted as optical interconnects stay always on, even during periods of system inactivity. We capitalize on this observation to propose LaC, a laser control mechanism that saves laser power by turning the laser off when not needed, while meeting high throughput. LaC runs on SWMR crossbars and saves between 66-92% (77% on average) of the laser energy. The power savings of LaC allow the cores to exploit a higher power budget and run faster, achieving speedups of 2-2.2x (2.1x average). Finally, a 64-core processor with Firefly interconnect topology using LaC attains 49-53% (52% on average) lower energy per instruction.

Keywords

Interconnection networks, nanophotonics, laser control, energy efficiency

LaC: Integrating Laser Control in a Photonic Interconnect

Yigit Demir
Northwestern University
Dept. of Electrical Engineering and Computer Science
Evanston, IL, USA
yigit@u.northwestern.edu

Nikos Hardavellas
Northwestern University
Dept. of Electrical Engineering and Computer Science
Evanston, IL, USA
nikos@northwestern.edu

ABSTRACT

The high-speed and low-cost modulation of light make photonic interconnects an attractive solution to the communication demands of manycore processors. However, the high optical loss of many nanophotonic components results in high power requirements for the laser source. Most of this power is wasted as optical interconnects stay always on, even during periods of system inactivity. We capitalize on this observation to propose LaC, a laser control mechanism that saves laser power by turning the laser off when not needed, while meeting high throughput. LaC runs on SWMR crossbars and saves between 66-92% (77% on average) of the laser energy. The power savings of LaC allow the cores to exploit a higher power budget and run faster, achieving speedups of 2-2.2x (2.1x average). Finally, a 64-core processor with Firefly interconnect topology using LaC attains 49-53% (52% on average) lower energy per instruction.

1. INTRODUCTION

Silicon photonics have emerged as a promising solution to meet the growing demand for high-bandwidth, low-latency, and energy-efficient communication in manycore processors. Silicon waveguides can be manufactured alongside CMOS logic on the same die by adding a few new steps in the manufacturing process [5], and they are more efficient for long-distance on-chip communication than electrical signaling [13,20]. However, the high optical loss of typical silicon waveguides, optical couplers, and on-ring resonators, together with the low efficiency of WDM-compatible lasers, dramatically increase the laser power consumption.

Typical silicon waveguides exhibit optical loss between 0.1-0.3 *dB/cm* [4], resulting in modest optical loss over short distances. However, replacing global wires with silicon photonics often requires long optical channels that traverse the entire chip in a serpentine form (for example, Firefly [20] on a 580 mm^2 chip would require a 16 *cm* waveguide, which increases the laser power by a factor of 1.5-3x). Aggressive technology can produce low-loss waveguides (0.05 *dB/cm* [13]) which allow the routing of long optical channels. However, these low-loss waveguides are much wider than conventional ones [13,15]. Their high area occupancy may force the use of narrow data paths (e.g., 2-bit links for an 8x8 array in the Oracle MacroChip [12,13]) which in turn impose significant serialization delays that degrade performance, and ultimately increase power consumption.

Additionally, WDM-compatible lasers are highly inefficient, with typical efficiencies in the range of 5-8% [24], and up to 10% [28]. Thus, the wall-plug laser power requirement is 10-20x higher than the required laser output power. Process variations impose additional losses, forcing designers to increase the laser power even higher, in order to maintain a safety margin. Sharing the optical path with other senders or receivers may also increase the laser power. While sharing is commonly employed to keep the hardware overhead manageable, it requires additional components which accumulate optical loss. While some optical interconnect topologies strike a better balance between power and performance [6,20,19], most of these costs are hard to avoid, and the laser power remains a considerable fraction of the total power budget. As all these factors are multiplicative, and not additive, it is easy for the wall-plug laser power to grow by more than one order of magnitude when all the losses and inefficiencies are factored in.

Unfortunately, the majority of this power is typically wasted. While the full laser power is required to support periods of high interconnect activity, most of it is wasted when activity is low because photonic interconnects are always on. In a typical setting, light is continuously injected into the waveguides and coupled onto several optical devices, regardless of whether packets are actively sent or not. By comparison, electrical interconnects stay idle consuming only a small amount of leakage power, until a packet attempts to traverse them. It is often the case that the interconnect stays idle for long periods of time, both in scientific computing (compute-intensive execution phases underutilize the interconnect), and in server computing (servers in Google-scale datacenters have a typical utilization of less than 30% [1]).

Motivated by these observations, we propose LaC, a laser control mechanism that reacts to the demands of the aggregate workload by opportunistically turning the laser off during periods of low activity to save energy, and leaving it on during periods of high activity in order to meet the high bandwidth demand. We then augment the initial LaC design with a technique that further increases laser energy savings when sending smaller coherence messages (Eco-LaC), and a design that attains higher performance by exploiting the correlations among the cache coherence messages (Smart-LaC). LaC capitalizes on recent advancements in Ge lasers [14,17], which enable energy-efficient on-chip laser sources that can be turned on or off within nanoseconds. More specifically, the contributions of this paper are:

1. We advocate laser control as a viable technique for saving power, and we quantify the maximum opportunity.

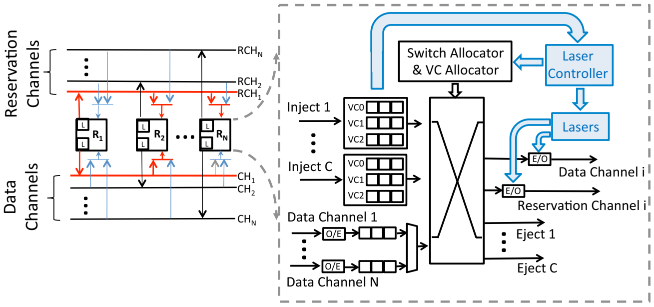

FIGURE 1. SWMR crossbar and router microarchitecture.

2. We propose LaC, a set of laser control mechanisms and policies that approximate the maximum possible savings.
3. We evaluate the impact of LaC on the performance and energy of a multicore running a range of synthetic and scientific workloads, under realistic physical constraints.

Our results indicate that LaC saves between 89-92% of the laser power for network topologies with SWMR at low injection rates (0.02-0.08). The potential energy savings drop as the injection rate increases, however LaC still provides high throughput, as the laser stays on almost all the time to service the high bandwidth demand. Performance- and energy-efficiency-focused Smart-LaC closely tracks closely tracks (within 2-4% on average) a perfect control mechanism with perfect knowledge of future interconnect requests. Thus, Smart-LaC harvests the vast majority of the energy benefits that can be achieved by controlling the laser source.

Moreover, the power savings of LaC allow for providing a higher power budget to the cores, which enables them to run faster. Employing Smart-LaC on a topology with SWMR crossbar (Firefly) allows the multicore chip to achieve 2-2.2x speedup (2.1x average). Similarly, Smart-LaC attains 49-53% lower energy consumption per instruction (52% on average) when employed on a 64-core processor with Firefly..

2. BACKGROUND

The nanophotonic devices used in this work include Ge-based lasers which generates light, ring modulators to modulate the light (electrical-to-optical conversion), waveguides routing optical signals to its destination, and resonant demodulators to demodulate the optical signal (optical-to-electrical conversion). By employing Dense Wavelength Division Multiplexing (DWDM), lasers of different wavelengths can be guided in the same waveguide without interfering with each other, which increases the bandwidth density.

The Ge-based laser introduced in [17] can be built within a standard-width (1um) waveguide, thereby incurring minimal area overhead, operates in room temperature, and is shown to be DWDM-compatible. Ge-based lasers can be turned on and off within 1ns [14,17]. DWDM-compatible on-chip lasers have been shown to have 5-10% energy efficiency [28,24], thus the total power delivered to the laser, and the total heat removed from the laser, are 10-20x higher than the energy of light required to maintain the communication. Silicon waveguides manufactured within SiO_2 cladding have high light confinement with transmission loss as low 0.3 *dB/cm* [4], with the width of 1 *um*. DWDM-compatible modulators and demodulators using resonant rings have been manufactured and characterized in [2], and can handle energy efficient signal conversions at speeds higher than 10 *GHz*.

Previously proposed interconnect topologies [11, 14, 20] employ SWMR crossbars. In Single-Writer-Multiple-Reader (SWMR) [11] crossbars, each router has its own dedicated data channel which delivers messages to all other routers (Figure 1). A sender first broadcasts on its reservation channel [20] a flit with the receiver's ID (in Figure 1, router R_1 broadcasts on RCH_1 a flit with ID=2). Upon receiving a reservation flit, the receiver (R_2) turns on its demodulators to receive the message from the sender's data channel (CH_1), which is now dedicated to transfer data from the sender to the receiver. Reservation channels are narrow because reservation flits only carry the receiver ID and message type information. However, the laser power required to broadcast increases exponentially with the number of readers, making it impractical to broadcast at high-radix crossbars (e.g., radix-64). Instead of having a single broadcast link with many readers, slicing [3] limits the number of readers per link, keeping the power low at the cost of deploying more waveguides. Reservation channels and slicing enable SWMR to scale to high radix counts (e.g., 64).

3. LASER CONTROL SCHEMES

The objective of the laser control (LaC) is to save laser energy by turning off the lasers whenever the bus (i.e., data channel) is idle. The laser should be turned back on when the bus will be used. The Ge-based laser [17] assumed in this work turns on in 1 *ns*, during which period it consumes the same power as when it is lasing. To control the lasers quickly, we place the laser for each router's dedicated channel within the router.

3.1 Laser Control for SWMR Crossbar

The shaded components in the router microarchitecture in Figure 1 correspond to components added by LaC. The laser controller turns the laser on if there is a message at any of the injection buffers, and turns it off only when all injection buffers are empty. The laser controller keeps the switch allocator waiting while the laser turns on. After the turn-on delay, the laser is ready and the switch allocator moves messages to the modulators. If a message is queued while the laser is on (servicing a previous message), the laser stays on and the message is sent out as soon as possible. This way, LaC provides high throughput under heavy utilization, while maintaining high laser energy savings at low injection rates.

3.2 Eco-LaC: An Energy Saving Optimization for LaC

LaC saves laser energy by turning the optical bus off when it is idle. However, LaC still wastes some laser energy, because it activates the whole optical bus to send small coherence messages (data-less) which don't occupy the whole bus. As photonic links provide high bandwidth, they offer wide busses which can send a data message in one cycle. A data message is 600-bits wide, and contains a 64 byte cache block and 64-bit address and 20-bit ID and 4-bit message type. However, an optical bus is 300-bit (or 300 wavelengths) wide, because the optical links runs at double the processor frequency. On the other hand, small coherence messages are transmitted in two 44-bit wide flits (64-bit address, 20-bit ID and 4-bit message type), which means 256 bits of this optical bus are used only when sending data messages and remain idle otherwise. Eco-LaC capitalizes on this by activating only 44 wavelengths of the optical bus (keeping the remaining 256 wavelengths off) when sending small coherence messages. The whole bus (300 wavelengths) will only be activated to send data messages.

Eco-LaC promises high laser energy savings, because it keeps a big fraction of the optical bus turned-off while servicing the majority of coherence messages. Although it requires independent control of the data-only portion (256 wavelengths) of the bus, that shouldn't increase the total laser power consumption, as the optical link loss, and the total number of wavelengths remain the same. Eco-LaC may require an additional waveguide at high DWDM levels (or may require an additional independent laser source) depending on how wavelength generation, splitting and waveguide assignment is done, but this potential area overhead won't be significant, as waveguides have small pitch and lasers are built within the waveguides.

3.3 Smart-LaC: A Performance Focused Optimization for Eco-LaC

Eco-LaC activates the whole optical bus only for data messages, and keeps the data-only portion (256 wavelengths) of the bus switched off most of the time. This lowers the data messages' likelihood of finding the whole data bus turned on. Therefore, data messages suffer from higher message latencies, which degrades performance. Smart-LaC turns the laser on proactively for the data messages to reduce the latency overhead. Smart-LaC anticipates the early laser activation for data messages by correlating cache coherence request messages to replies. In a directory-based cache coherence protocol, every data message is generated upon receipt of a read or write request. The request results in a lookup in the local L2 cache slice, which resolves quickly if it misses as only the tags are accessed first. However, if the tag hits, an additional long latency access to the data array is required to get the data. Smart-LaC recognizes that and turns on the laser 1ns before the data become available.

Similar to Eco-LaC, Smart-LaC keeps a portion of the optical bus inactive while sending small coherence messages, so it achieves high laser energy savings. It also achieves high performance because it proactively turns the laser on for data messages. We expect to see rare early or false laser activations with Smart-LaC, because it relies on L2 access latencies for timing, and race conditions in the cache coherence protocol can alter the L2 behavior. Thus, a slight increase in laser energy consumption compared to the Eco-LaC is expected.

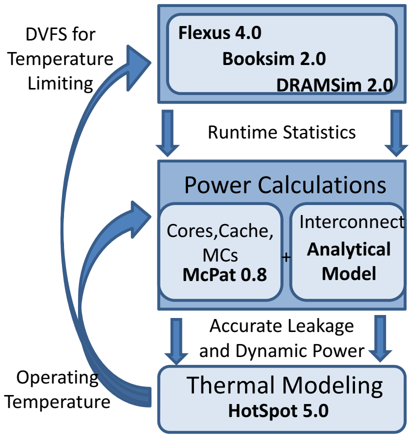

FIGURE 2. Simulation flow chart.

3.4 The Perfect Laser Control

A perfect control scheme has perfect and complete knowledge of future interconnect accesses. The perfect scheme saves the maximum laser energy without incurring any performance overhead by turning the laser on ahead of time, so the light reaches the writer at the exact time the writer attempts to transmit. Also it keeps the laser on for a message which will need it in the immediate future, if the energy consumed by keeping the laser on is lower than the energy if the laser was allowed to turn off and on again. Similar to Eco-LaC, Perfect can control data-only portion of the bus independently. Thus, the perfect scheme demonstrates the limit of energy savings with the given laser technology.

4. EXPERIMENTAL METHODOLOGY

4.1 Interconnect Performance and Energy Analysis

To evaluate the performance and energy consumption of LaC in isolation from the interference of other system components or application characteristics, we employ a cycle-accurate network simulator based on Booksim 2.0 [7], which models a radix-16 SWMR crossbar servicing random uniform traffic. The simulator models a single-cycle router, with 1-cycle E/O and O/E conversions. We assume a 480 mm^2 chip, which employs a 10 cm waveguide with a round trip time of 5 cycles. The link latency (1-5 cycles) is calculated based on the traversed waveguide length. The buffers are 20-flits deep, with a flit size of 300 bits. The maximum core frequency is 5 GHz, and the optical interconnect runs at 10 GHz. Latency is measured as the time required for the network to process a sample of injected packets. We evaluate the load-latency and energy-per-flit of LaC and Eco-LaC, and compare it against a baseline without laser control (*No-Ctrl*), and a perfect control scheme with full knowledge of future messages (*Perfect*).

TABLE 1. Architectural Parameters.

CMP Size	64 cores, 480mm^2
Processing Cores	ULTRASPARC III ISA, up to 5Ghz, OoO, 4-wide dispatch/retirement, 96-entry ROB
L1 Cache	Split I/D, 64KB 2-way, 2-cycle load-to-use, 2 ports, 64-byte blocks, 32 MSHRs, 16-entry victim cache
L2 Cache	Shared, 512 KB per core, 16 way, 64-byte blocks, 14 cycle-hit, 32 MSHRs, 16-entry victim cache
Memory Controllers	One per 4 cores, 1 channel per Memory Controller Round-robin page interleaving
Main Memory	Optically connected memory [2], 10ns access
Networks	SWMR_XBAR and Firefly

TABLE 2. Nanophotonic Parameters and Laser Power.

		SWMR_XBAR	Firefly
	per Unit	**Total**	**Total**
DWDM		64	64
WG Loss	0.3 *dB/cm*	3 *dB*	3 *dB*
Nonlinearity	1 *dB*	1 *dB*	1 *dB*
Modulator Ins.	0.5 *dB*	0.5 *dB*	0.5 *dB*
Ring Through	0.01 *dB*	10.24 *dB*	10.24 *dB*
Filter Drop	1.2 *dB*	1.2 *dB*	1.2 *dB*
Photodetector	0.1 *dB*	0.1 *dB*	0.1 *dB*
Total Loss		**16.04 *dB***	**16.04 *dB***
Detector		**-20 *dBm***	**-20 *dBm***
Laser Power Per Wavelength		**0.401 *mW***	**0.401 *mW***
Total LaserPower		**20.1 *W***	**80.4*W***

4.2 Multicore System Performance and Energy Analysis

To evaluate the impact of laser control schemes on a realistic multicore system, we model a 64-core processor on a full-system cycle-accurate simulator based on Flexus 4.0 [9,26] integrated with Booksim 2.0 [7] and DRAMSim 2.0 [21]. Table 1 details the architectural modeling parameters. The power consumption of the electrical interconnect is calculated using DSENT [23]. We target a 16 *nm* technology, and have updated our tool chain accordingly based on ITRS projections [8]. The simulated system executes a selection of SPLASH-2 benchmarks and other scientific workloads. All systems we model employ a throttling mechanism to keep the chip within safe operational temperatures (below 90*C).* Without loss of generality, we use Dynamic Voltage and Frequency Scaling (DVFS) as the throttling mechanism.

We collect runtime statistics from full-system simulations, and use them to calculate the power consumption of the system using McPAT [16], and the power consumption of the optical networks using the analytical power model by Joshi *et al.* [10]. We estimate the temperature of the chip using HotSpot 5.0 [22]. The estimated temperature is then used to refine the leakage power estimate. We adjust DVFS based on the stable-state power and temperature estimates (Figure 2).

To put LaC schemes' performance and energy consumption into perspective, we include in our evaluation a traditional all-electrical on-chip interconnect: a 2D-concentrated mesh with express links (***CMesh***). For CMesh we model routers with 8 input and output ports and a 3-cycle routing delay. Routers are connected through 150-bit bi-directional links with 1-cycle local and 3-cycle global delay. To show the range of LaC's impact, we evaluate its application on two optical network topologies that are at the opposite ends of the spectrum, a radix-16 SWMR crossbar (***SWMR_XBAR***) and a topology which uses 4 optical SWMR crossbars similar to [20] (***Firefly***). The SWMR_XBAR approximates a worst case scenario for LaC. It has low power consumption (similar to the power consumption of CMesh) and its high concentration factor (4) creates heavier traffic. The low power consumption and heavy traffic limit LaC's opportunity. The Firefly connects 16 local clusters (4 routers each) using 4 SWMR crossbars. The local clusters use an electrical ring to communicate locally, and each of the routers in a local cluster is connected to a different SWMR crossbar. Local electrical ring has 150-bit bi-directional links with 1-cycle delay. Firefly has high laser power consumption and a low concentration factor (1), which results in light traffic, thus giving ample opportunity to LaC to conserve laser power. The modeling of the optical SWMR crossbars is described in Section 4.1. Laser turn-on delay of 1 *ns* is included in LaC implementations. Finally, we contrast LaC to a power-equivalent optical interconnect design with no laser control (*Power_Eq*). Power_Eq is similar to the No-Ctrl baseline, but its interconnect width has been scaled down to approximate best LaC's average energy savings.

We compare the performance (user instructions per sec), energy per instruction (EPI) of CMesh, the baseline scheme without laser control (No-Ctrl), Power_Eq, LaC, Eco-LaC, Smart-LaC, and perfect laser control (Perfect).

4.3 Laser Power Consumption Calculation

Table 2 shows the optical loss parameters for the modulators, demodulators, drop filters, and detectors introduced in [2] and assumed in this work. The modulation and demodulation energy is 150 *fJ/bit* at 10 *GHz* [2]. The laser power per wavelength and total laser power are calculated in Table 2 using the analytical models introduced in [10]. The total laser power in Table 2 includes the laser power for both data and reservation channels, plus the laser efficiency of 10%, so it is the wall plug power for the laser. The data bus is 300-bits wide, so it can push a data message in one processor cycle (both edges of a 5 *GHz* clock).

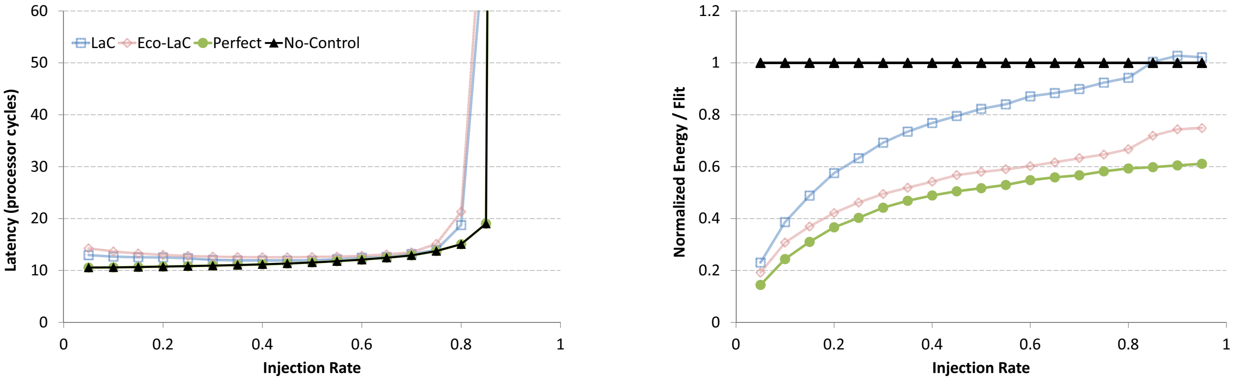

FIGURE 3. Load-Latency (left) and Energy-per-Flit (right) for radix-16 and SWMR crossbar.

4.4 Sensitivity to Optical Parameters

Unfortunately, there is little consensus on the optical loss parameters used or projected in literature. In same cases, parameters exhibit a variance over 10x across publications. However, we observe that the design of an optical interconnect highly depends on the losses of the optical components used. For example, if the off-ring through loss on the radix-16 crossbar was 10x higher (i.e., $0.1dB$) the interconnect wouldn't employ 64-way DWDM, as this would increase the laser power to unsustainable levels. Rather, the interconnect would be optimized with a lower 6-way DWDM and it would employ more waveguides, resulting in a total optical loss (and hence laser power) similar to the interconnect modeled in our work. In the extreme case where the off-ring loss were to increase by 10x, and on top of that the modulator insertion, drop loss, detection and non-linearity losses were to double, a 4-way DWDM would accommodate the increased losses and keep the total laser power at the same level.

In either case, the fraction of laser energy that LaC saves depends on the network utilization, not on the optical loss parameters. Moreover, the higher the total optical loss, the more power in absolute terms LaC would save, which would have a higher impact on the performance of the processor if this power is given back to the cores. Thus, in this work, we remain conservative in our estimates of optical losses.

4.5 Resonant Ring Heater Modeling

To calculate the total ring heating power we extend the method by Nitta *et al.* [18] by additionally accounting for the heating of the photonic die by the operation of the cores. We model the thermal characteristics of a 3D-stacked architecture where the photonic die sits underneath the logic die. We use the 3D-chip extension of HotSpot [22] to model the transient temperature changes in the optical die. After we execute a workload and collect transient temperature traces, we calculate the ring heating power required to maintain the entire photonic die at the constant micro-ring trimming temperature during the entire execution. In addition, we account for the individual ring trimming power required to overcome process variations, as described in [10]. The Individual ring trimming power is less significant when using smaller-radix crossbars.

5. EXPERIMENTAL RESULTS

5.1 Network Performance

LaC opportunistically turns the laser off when the optical bus is idle, and saves energy. When a message needs to be sent, if the laser is off, the message waits for the laser to turn on (1 *ns*); if the laser is on, the message is sent immediately. The LaC controller saves energy at the potential cost of latency overhead. We investigate the energy savings and performance trade-off LaC using synthetic traffic patterns by comparing it to a SWMR crossbar where the laser is always on (*No-Ctrl*).

Figure 3 shows the load-latency and energy per flit estimations under uniform random traffic pattern on a radix-16 SWMR nanophotonic crossbar with LaC and Eco-LaC. The latency overhead of LaC is just over 3 cycles (even though the laser turn-on latency is 5 cycles) at very low injection rates, because some of the messages find the laser on and transmit. This overhead decreases slightly as the injection rate grows, because the laser gets turned off less frequently and higher fraction of messages catch the laser on. As the injection rate increases laser energy savings decrease because the laser stays on longer. Therefore, LaC's laser energy consumption is close to No-Ctrl's at high injection rates. Energy per flit of LaC is higher than No-Ctrl under very heavy utilization, because LaC provides lower throughput than No-Ctrl.

Eco-LaC aims to achieve higher energy savings by keeping a portion of optical data bus inactive while sending smaller (dataless) messages. The message latency of Eco-LaC is slightly higher than LaC at low injection rates, because data messages can't catch the laser on, as the data-only portion of bus is turned off more frequently. The laser energy consumption of Eco-LaC doesn't grow as fast as LaC under heavy utilization, because the data-only portion of the bus can be kept inactive. Eco-LaC saves 56% of the laser energy on average across injection rates (minimum 25%) compared to No-Ctrl.

Perfect laser control achieves maximum laser energy savings with no latency overhead. Even though the total laser energy savings of Eco-LaC is similar to Perfect, its energy per flit consumption is 15% higher than Perfect on average, because Eco-LaC imposes high laser turn-on latency overhead.

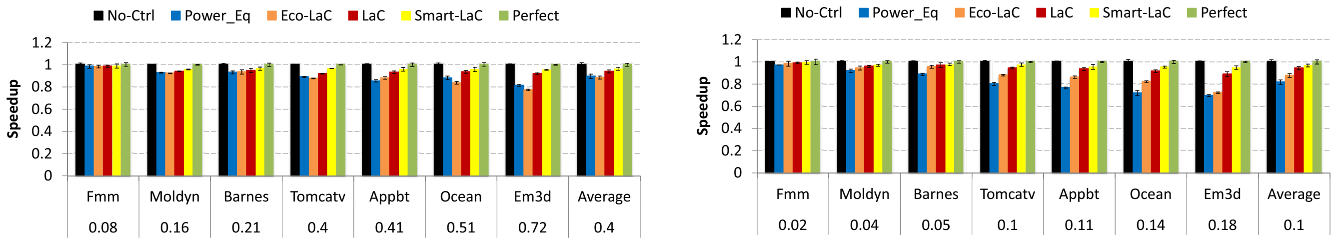

FIGURE 4. Speedup for SWMR_XBAR (left) and Firefly (right) with No-Ctrl, Eco-LaC, LaC, Smart-LaC and Perfect Control.

5.2 The Performance Cost of Laser Control

LaC and Eco-LaC saves laser energy at the cost of increased network latency. Smart-LaC is an advanced version of Eco-LaC, which achieves high energy savings with low latency overhead, because it proactively turns the laser on for data messages. The correlation between coherence messages help Smart-LaC to anticipate when to turn the laser on (Section 3.3), therefore we had to measure its performance using messages generated by real-world workloads running on a multicore processor (instead of a synthetic traffic pattern). We analyzed the effect of the latency overhead on the system performance by running real-world workloads on a multicore processor which is not subject to thermal constraints, and always runs at maximum frequency (5 GHz). Our suite includes both memory-intensive workloads that are sensitive to the interconnect latency (em3d, ocean, appbt, tomcatv), as well as compute-intensive workloads whose performance is less dependent on message latency (fmm, moldyn, barnes).

Figure 4 shows the performance of a multicore processor with two different topologies SWMR_XBAR and Firefly (detailed in Section 4.2) implementing laser control schemes LaC, Eco-LaC and Smart-LaC. The performance of laser controlled topologies is normalized against the performance of topology that keeps the lasers always on (No-Ctrl). The SWMR_XBAR topology has concentration of 4 cores per router. As a result, the message injection rate per router (shown under the workload name in Figure 4) is higher than the injection rate in the Firefly topology (concentration of 1).

Eco-LaC has the highest laser turn-on latency overhead, therefore it causes highest performance decrease when running memory intensive workloads. Eco-LaC has the minimum slowdown while running compute intensive workloads because their performance is less sensitive to the network latency. Eco-LaC saves energy by keeping the laser off when the optical bus is idle, therefore, its energy savings is highest while running compute intensive workloads with low injection rates. On average Eco-LaC saves 59% of the laser energy (81% maximum) compared to No-Ctrl on SWMR_XBAR. Eco-LaC on Firefly saves 77% of the laser energy on average (92% maximum) compared to No-Ctrl. Firefly, with lower core to router concentration ratio, manages to save higher fraction of the laser energy.

LaC is faster compared the Eco-LaC because it has lower laser turn-on latency overhead, however, LaC saves less energy as it activates the whole bus for any type of message, even the ones with no data. Compared to No-Ctrl, LaC manages to save 29% of the laser energy on average (63% maximum) for SWMR_XBAR. On the other hand, LaC on Firefly saves 64% of the laser energy on average (89% maximum).

Smart-LaC is the fastest laser control scheme, because it causes minimal laser turn-on delay overhead to the data messages. Because it turns off the data-only portion of the data bus while sending data-less messages, it achieves high laser energy savings too. Smart-LaC's laser energy savings are within 2-4% vicinity of the Eco-LaC. The decrease in savings is because of rare early turn on or false turn on occasions. However, Smart-LaC has the lowest laser energy consumption per flit compared to all other laser control schemes, because it is faster. Smart-LaC's laser energy savings is in the 4% vicinity of what Perfect scheme can achieve.

Power_Eq keeps the lasers always on similar to the No-Ctrl baseline, but its interconnect width (flit size) has been scaled down to approximate best LaC's average energy savings. Power_Eq with SWMR_XBAR uses 120-bit wide flits, with Firefly it uses 75-bit wide flits. Power_Eq is slower than laser control schemes, because it imposes high message serialization delays and provides lower throughput.

Overall, laser control schemes save more laser energy when running real-world workloads than when running a synthetic random traffic pattern (Section 5.1), because real workloads have a bursty (and sparse) memory access patterns.

5.3 Impact of Laser Control on a Realistic Multicore

Mechanisms like DVFS keep a chip within safe operating temperatures (under 90C) by throttling the cores during the execution. When the laser power consumption is high, more power is dissipated on the chip which increases its temperature. As a result, DVFS throttles the cores more in order to keep the chip within safe operating temperatures when there is high laser power consumption (No-Ctrl). Thus, even though LaC trades off network latency for laser energy savings, a realistic power-limited system will exhibit higher performance with LaC, because the cores will not be throttled as much as without laser control (No-Ctrl). For example, for Firefly running fmm, the LaC runs at 3.3 GHz, Smart-LaC

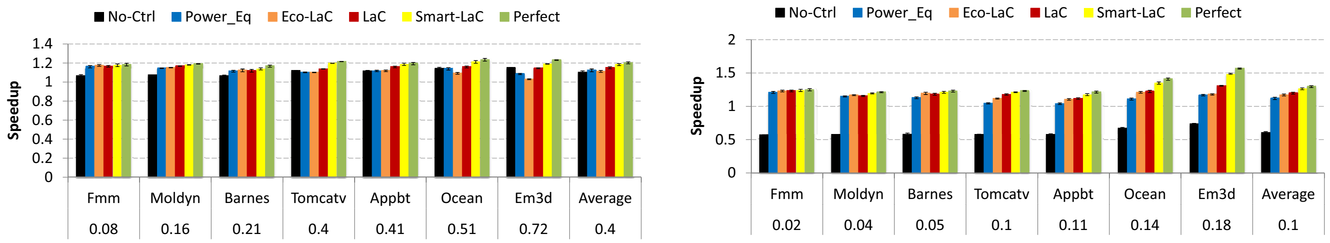

FIGURE 5. Speedup for SWMR_XBAR (left) and Firefly (right) with No-Ctrl, Eco-LaC, LaC, Smart-LaC, Perfect with DVFS.

and Eco-LaC runs at 3.35 GHz, while No-Ctrl runs at 1.5 GHz (its cores are throttled more).

Figure 5 shows the speedup for a realistic multicore system with LaC, Smart-LaC, Eco-LaC, Power_Eq, No-Ctrl and Perfect control schemes in SWMR_XBAR and Firefly normalized against CMesh. SWMR_XBAR with No-Ctrl is faster than CMesh because it provides higher throughput. Architectures with laser control schemes aren't throttled as much as No-Ctrl, this is why they are faster except for Eco-LaC running memory intensive workloads. Eco-LaC doesn't provide enough throughput and slowdown the execution under memory intensive workloads. Eco-LaC consumes more energy per instruction (EPI) compared to No-Ctrl, because it is slower (Figure 6, EPI is normalized against CMesh). Power_Eq is also less energy-efficient under memory intensive workloads for the same reason. The fastest and energy-efficient laser control scheme is Smart-LaC, which outperforms No-Ctrl by 1.08x, while consuming 8% less EPI on average. Smart-LaC is only 1.5% slower and has 2% higher EPI compared to Perfect on average.

The Firefly topology provides much higher bandwidth than SWMR_XBAR and CMesh, however it consumes over 80 Watts of laser power with No-Ctrl, so it performs poorly as its cores are throttled the most. The amount of energy saved by any laser control scheme running on Firefly is significant, therefore all control schemes outperform No-Ctrl. Firefly with No-Ctrl is 40% slower and has 40% higher EPI compared to CMesh, however, Smart-LaC improves system's performance by 2.1x and help lower its EPI by 2.11x. Average laser power consumption of Firefly with Smart-LaC is similar to the power consumption of CMesh. Smart-LaC improves the energy efficiency of the Firefly, making it feasible and more attractive alternative to conservative optical interconnects (SWMR_XBAR) and electrical interconnects (CMesh).

6. RELATED WORK

Different on-chip interconnect networks have been proposed that exploit CMOS-compatible photonics for future multicore processors. The hierarchical Firefly architecture [20] advocates the use of partitioned nanophotonic SWMR crossbars to connect clusters of electrically-connected mesh networks. Firefly improves power efficiency and provides uniform global bandwidth between all clusters. Batten *et al.* [2,3] propose to connect a many-core processor to DRAM memory using monolithic silicon nanophotonics, and present energy-efficient and scalable implementations of SWMR crossbars. All of these network topologies, which use SWMR LaC schemes to achieve higher laser energy efficiency while maintaining their performance.

Previously, Thonnart *et al.* [25] proposed power regulation techniques to reduce the static power consumption in electrical interconnects. Powering down the unused asynchronous units results in substantial energy savings. Zhou *et al.* [27] identify the constant laser power consumption when channel utilization is low as an inefficiency, and they propose a prediction-based mechanism to increase average channel utilization. Their mechanism controls active splitters, to tune channel bandwidth on a binary tree network. Kurian *et al.* [14] propose an optical SWMR crossbar and electrical hybrid interconnection network, and improve performance by utilizing the coherence protocol. [14] mentions that a Ge-based laser can be controlled to improve the laser energy efficiency, but they do not present nor evaluate a detailed laser-control

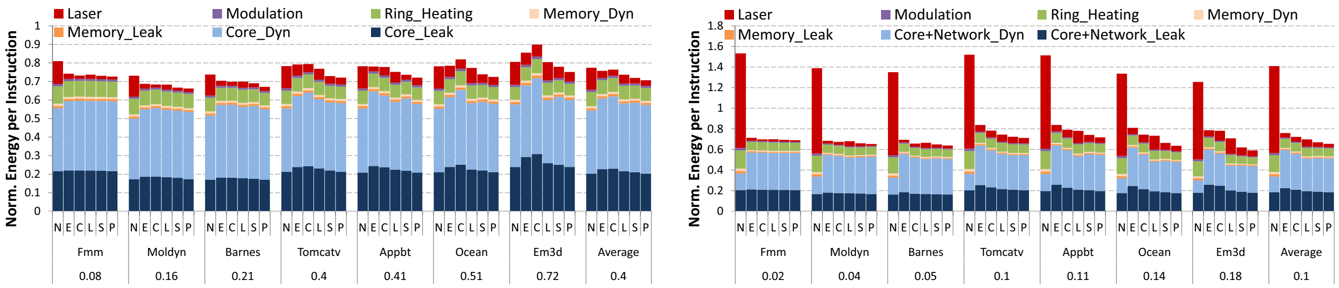

FIGURE 6. Energy per Instruction in SWMR_XBAR (left) and Firefly (right). The evaluated designs are from left to right: No-Ctrl (N), Power_Eq (E), Eco-LaC (C), LaC (L), Smart-LaC (S), and Perfect (P).

scheme. In contrast to previous work, we propose LaC, an set of on-chip laser control schemes that improve the laser energy efficiency significantly for SWMR crossbars, while providing high bandwidth and performance.

7. CONCLUSION

In this paper we propose LaC, a laser-control mechanism that turns the laser off during periods of inactivity to save energy. We improve upon the initial LaC design by proposing Eco-LaC, a laser-control mechanism that can also keep the majority of the data bus off while sending small (data-less) messages. Finally, we introduce Smart-LaC, a technique that improves the performance of laser control schemes by turning the laser on proactively for data messages. Compared to keeping the lasers always on (No-Ctrl), LaC saves up to 50-89% (64% on average), and Eco-LaC and Smart-LaC save up to 66-92% (77% on average) of the laser energy while running real-world applications. The energy savings and performance of Smart-LaC stays within 2-4% of a Perfect scheme with future knowledge of interconnect accesses. More importantly, the power savings of our laser control schemes allow the cores to exploit a higher power budget and run faster, achieving speedups up to 2-2.2x (2.1x average) on a 64-core processor with Firefly interconnect topology. On that multicore system, Smart-LaC attains 49-53% (52% on average) lower energy per instruction.

8. ACKNOWLEDGEMENTS

This work was partially supported by National Science Foundation award CCF-1218768, an ISEN booster award, and the June and Donald Brewer Chair in EECS at Northwestern University.

9. REFERENCES

[1] L. A. Barroso and U. Holzle. The case for energy-proportional computing. *IEEE Computer*, 40(12):33- 37, 2007.

[2] C. Batten, A. Joshi, J. Orcutt, A. Khilo, B. Moss, C. W. Holzwarth, M. A. Popovic, H. Li, H. I. Smith, J. L. Hoyt, F. X. Kartner, R. J. Ram, V. Stojanovic, and K. Asanovic. Building many-core processor-to-dram networks with monolithic cmos silicon photonics. *IEEE Micro*, 29(4):8- 21, 2009.

[3] C. Batten, A. Joshi, V. Stojanovic, and K. Asanovic. Designing chip-level nanophotonic interconnection networks. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2(2), 2012.

[4] J. Cardenas, C. Poitras, J. Robinson, K. Preston, L. Chen, and M. Lipson. Low loss etchless silicon photonic waveguides. *Optics Express*, 17(6):4752- 4757, 2009.

[5] G. Chen, H. Chen, M. Haurylau, N. Nelson, P. M. Fauchet, E. Friedman, and D. Albonesi. Predictions of cmos compatible on-chip optical interconnect. In *7th International Workshop on System-Level Interconnect Prediction (SLIP)*, San Francisco, CA, 2005.

[6] M. J. Cianchetti, J. C. Kerekes, and D. H. Albonesi. Phastlane: a rapid transit optical routing network. In *Proceedings of the 36th Annual International Symposium on Computer Architecture*, ISCA 09, 2009.

[7] W. J. Dally and T. B. *Principles and Practices of Interconnection Networks*. Morgan Kaufmann Publishing Inc., 2004.

[8] European, Japan, Korean, Taiwan, and United States Semiconductor Industry Associations. The international technology roadmap for semiconductors (ITRS). http://www.itrs.net/, 2012 Edition.

[9] N. Hardavellas, S. Somogyi, T. F. Wenisch, R. E. Wunderlich, S. Chen, J. Kim, B. Falsafi, J. C. Hoe, and A. G. Nowatzyk. SimFlex: a fast, accurate, flexible full-system simulation framework for performance evaluation of server architecture. *SIGMETRICS Performance Evaluation Review, Special Issue on Tools for Computer Architecture Research*, 31(4):31- 35, April 2004.

[10] A. Joshi, C. Batten, Y.-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, and V. Stojanovic. Silicon-photonic clos networks for global on-chip communication. In *Proceedings of the IEEE International Symposium on Networks-on-Chip (NOCS)*, pages 124- 133, 2009.

[11] N. Kirman, M. Kirman, R. K. Dokania, J. F. Martinez, A. B. Apsel, M. A. Watkins, and D. H. Albonesi. Leveraging optical technology in future bus-based chip multiprocessors. In *Proc. of the 39th IEEE/ACM Annual International Symposium on Microarchitecture*, 2006.

[12] P. Koka, M. McCracken, H. Schwetman, X. Zheng, R. Ho, and A. Krishnamoorthy. Silicon-photonic network architectures for scalable, power-efficient multi-chip systems. In *Proc. of the 37th Annual International Symposium on Computer Architecture*, 2010.

[13] A. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, and J. Cunningham. Computer systems based on silicon photonic interconnects. *Proceedings of the IEEE*, 97(7):1337 - 1361, july 2009.

[14] G. Kurian, C. Sun, C.-H. Chen, J. Miller, J. Michel, L. Wei, D. Antoniadis, L.-S. Peh, L. Kimerling, V. Stojanovic, and A. Agarwal. Cross-layer energy and performance evaluation of a nanophotonic manycore processor system using real application workloads. In *26th IEEE International Parallel Distributed Processing Symposium (IPDPS)*, pages 1117- 1130, 2012.

[15] G. Li, J. Yao, H. Thacker, A. Mekis, X. Zheng, I. Shubin, Y. Luo, J. hyoung Lee, K. Raj, J. E. Cunningham, and A. V. Krishnamoorthy. Ultralow-loss, high-density soi optical waveguide routing for macrochip interconnects. *Optics Express*, 20(11):12035- 12039, May 2012.

[16] S. Li, J. H. Ahn, R. D. Strong, J. B. Brockman, D. M. Tullsen, and N. P. Jouppi. Mcpat: an integrated power, area, and timing modeling framework for multicore and manycore architectures. In *42nd IEEE/ACM Annual International Symposium on Microarchitecture*, 2009.

[17] J. Liu, X. Sun, R. Camacho-Aguilera, L. C. Kimerling, and J. Michel. Ge-on-si laser operating at room temperature. *Opt. Lett.*, Mar 2010.

[18] C. Nitta, M. Farrens, and V. Akella. Addressing system-level trimming issues in on-chip nanophotonic networks. In *17th IEEE International Symposium on High Performance Computer Architecture*, 2011

[19] Y. Pan, J. Kim, and G. Memik. Flexishare: Channel sharing for an energy-efficient nanophotonic crossbar. In *17th IEEE International Symposium on High Performance Computer Architecture*, 2010.

[20] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, and A. Choudhary. Firefly: Illuminating future network-on-chip with nanophotonics. In *Proceedings of the 36th Annual International Symposium on Computer Architecture*, ISCA 09, Austin, TX, 2009.

[21] P. Rosenfeld, E. Cooper-Balis, and B. Jacob. Dramsim2: A cycle accurate memory system simulator. *Computer Architecture Letters*, 10(1):16- 19, 2011.

[22] K. Skadron, M. R. Stan, W. Huang, S. Velusamy, K. Sankaranarayanan, and D. Tarjan. Temperature-aware microarchitecture. In *Proceedings of the Annual International Symposium on Computer Architecture (ISCA)*, ISCA 03, pages 2- 13, 2003.

[23] C. Sun, C.-H. O. Chen, G. Kurian, L. Wei, J. Miller, A. Agarwal, L.-S. Peh, and V. Stojanovic. Dsent - a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling. In *6th IEEE/ACM Symposium on Networks-on-Chip*, 2012.

[24] S. Tanaka, S.-H. Jeong, S. Sekiguchi, T. Kurahashi, Y. Tanaka, and K. Morito. Highly-efficient, low-noise si hybrid laser using flip-chip bonded soa. In *IEEE Optical Interconnects Conference*, 2012.

[25] Y. Thonnart, E. Beigne, A. Valentian, and P. Vivet. Automatic power regulation based on an asynchronous activity detection and its application to anoc node leakage reduction. In *14th IEEE International Symposium on Asynchronous Circuits and Systems*, 2008.

[26] T. F. Wenisch, R. E. Wunderlich, M. Ferdman, A. Ailamaki, B. Falsafi, and J. C. Hoe. SimFlex: statistical sampling of computer system simulation. *IEEE Micro*, 26(4):18- 31, Jul-Aug 2006.

[27] L. Zhou and A. Kodi. Probe: Prediction-based optical bandwidth scaling for energy-efficient nocs. In *Seventh IEEE/ACM International Symposium on Networks on Chip (NoCS)*, pages 1- 8, 2013.

[28] A. Zilkie, B. Bijlani, P. Seddighian, D. C. Lee, W. Qian, J. Fong, R. Shafiiha, D. Feng, B. Luff, X. Zheng, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari. High-efficiency hybrid III-V/Si external cavity DBR laser for 3um SOI waveguides. In *9th IEEE International Conference on Group IV Photonics (GFP)*, pages 317- 319, 2012.