Towards Energy-Efficient Photonic Interconnects

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

ABSTRACT

Silicon photonics have emerged as a promising solution to meet the growing demand for high-bandwidth, low-latency, and energy-efficient on-chip and off-chip communication in many-core processors. However, current silicon-photonic interconnect designs for many-core processors waste a significant amount of power because (a) lasers are always on, even during periods of interconnect inactivity, and (b) microring resonators employ heaters which consume a significant amount of power just to overcome thermal variations and maintain communication on the photonic links, especially in a 3D-stacked design. The problem of high laser power consumption is particularly important as lasers typically have very low energy efficiency, and photonic interconnects often remain underutilized 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%). We address the high laser power consumption by proposing EcoLaser+, which is a laser control scheme that saves energy by predicting the interconnect activity and opportunistically turning the on-chip laser off when possible, and also by scaling the width of the communication link based on a runtime prediction of the expected message length. Our laser control scheme can save up to 62 - 92% of the laser energy, and improve the energy efficiency of a many-core processor with negligible performance penalty. We address the high trimming (heating) power consumption of the microrings by proposing insulation methods that reduce the impact of localized heating induced by highly-active components on the 3D-stacked logic die.

Keywords

Nanophotonic Interconnection Networks; Adaptive Laser Control; Microring Trimming; Thermal Insulation

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 [1], and they are more efficient for long-distance on-chip communication than electrical signaling [2]. However, the high optical loss of typical silicon waveguides, optical couplers, and on-ring resonators dramatically increase the laser power consumption. While some optical interconnect topologies better balance power and performance [3, 2, 4] most of these costs are hard to avoid, and the laser power remains a considerable fraction of the total power budget. These costs together with the low efficiency of WDM-compatible lasers (in the range of 5-10% [5]), result in wall-plug laser power that is 10-20x higher than the required laser output power.The majority of this power is typically wasted when activity is low because photonic interconnects are always on. By comparison, electrical interconnects stay idle until a packet traverses them. It is common for the interconnect to stay idle often for long periods of time, both in scientific computing (compute-intensive execution phases underutilize the interconnect), and in server computing (Google-scale datacenters have a typical utilization of 30% [6]).

At the same time, integrating photonics with CMOS devices is typically done through 3D-die stacking, where a die with nanophotonic devices is stacked underneath a typical multicore processor logic die. However, fluctuations in the activity of the logic components (e.g., cores) induce significant temporal and spatial thermal variations [7] that can exceed 30^{o}C. To counteract these thermal variations, microring resonators employ heaters which consume a significant amount of power to maintain communication on the photonic links.

Motivated by these observations, in prior work we proposed EcoLaser [8], a collection of laser control mechanisms that react to the demands of the aggregate workload and opportunistically turn the laser off during periods of low activity to save energy, and leave it on during periods of high activity to meet the high bandwidth demand. In this paper we propose EcoLaser+, a laser control technique that improves upon EcoLaser [8] by keeping the majority of the data bus off while sending small (data-less) messages to minimize wasted energy, and provides better performance by turning the laser on proactively when possible. EcoLaser+ addresses the high trimming power consumption of the microrings by introducing insulation methods that reduce the impact of localized heating by highly-active components on the 3D-stacked logic die.

Our results indicate that EcoLaser+ saves between 62-92% of the laser power for radix-16 SWMR crossbar on real-world workloads. EcoLaser+ closely tracks (within 2-6% on average) a perfect controller with full knowledge of future intercon-

nect requests. Thus, EcoLaser+ harvests the vast majority of the energy benefits that can be achieved by controlling the laser source. Moreover, the power savings of EcoLaser+ leave a higher power budget to the cores, which allows a multi-core chip with EcoLaser+ to outperform a baseline scheme with no control by 1.5-1.7x. Finally, EcoLaser+ reduces the microring trimming power up to 5.9x by exploiting the photonic-die insulation.

2. BACKGROUND

Previous works [9, 10, 11, 12, 2] typically use off-chip laser sources because of their temperature stability, easy replacement, and energy efficiency (30% for gaussian comb lasers [13]). However, recent work [14] shows that output spectrum power variations and laser-to-fiber and fiber-to-chip coupling losses add 7-8 *dB* to the total optical loss, thus off-chip lasers are in reality only 6% efficient. In comparison, on-chip laser sources [15] attain wall-plug efficiencies up to 15%, while enabling wavelength-division multiplexing (WDM) to increase the bandwidth density. WDM can be implemented by feeding a set of wavelengths generated by a single-wavelength laser array into an optical bus. On-chip laser sources offer energy-efficiency and easy packaging, but their wall-plug power consumption is counted against the chip's overall power budget. In either case, the laser power consumption remains a considerable overhead, especially when accounting for realistic optical loss parameters and laser efficiencies, emphasizing the need for power gating the laser source. Power gating on-chip lasers can increase the energy efficiency of a photonic interconnect by up to 4x [14].

Laser power-gating has been overlooked, because of the high turn-on latency (~0.1 *us* [14]) of the traditional distributed feedback comb lasers that are widely assumed for photonic interconnects [9, 10, 11, 12, 2]. The traditional comb lasers use diffraction grating to form the optical cavity. Temperature affects the diffraction grating pitch and the active region's refractive index, which alter the diffraction grating's wavelength selection, and hence the laser's emission wavelength. Thus, when comb lasers are turned on they need time to reach a certain temperature and lock at the designated wavelength. This high turn-on delay hampers power gating. In contrast, Fabry-Perot lasers are suitable for power gating. They use two discrete mirrors to form the optical cavity, and their emission wavelength depends not on temperature but on the n-type doping level and the strain applied during the optical cavity development. Thus, when they are turned on (pumped to the lasing threshold), they lase at the designated wavelength without requiring time for temperature stabilization/locking.

EcoLaser+, and laser power gating in general, strongly depends on fast lasers. While such technology is still in its infancy and highly experimental, it is important to note that fast lasers with *ns*-scale turn-on times have been manufactured and their turn-on delay has been characterized on real hardware prototypes, and is in agreement with theoretically-derived results. To turn the laser on, a supply current is applied to the laser. When the carrier density exceeds the threshold density, laser oscillation starts and light output increases drastically (laser turn-on). The time it takes from the current injection to the laser turn-on is the "laser turn-on delay" which is governed by the carrier life time and is in the order of *ns* [16]. The turn-on delay of Fabry-Perot lasers is highly tunable by design parameters, and nanosecond or sub-nanosecond laser turn-on delays are both theoretically predicted [17, 18, 16] and achievable in real implementations [19,20,21,16].

For example, Ge-based Fabry-Perot on-chip lasers have been manufactured [20] and the turn-on delay of real hardware prototypes was measured to be at most 1.5 *ns*, both for optically- and for electrically-pumped implementations [20, 19]. Similarly, InP-based diode lasers [21] have been manufactured and shown to emit light with a 2 *ns* long electrical pulse excitation (so the laser turn-on latency is at most 2 *ns*). Besides their fast turn-on time, Ge-lasers [19] are suitable for on-chip photonic interconnects because they can be built within a standard-width (1 *um*) waveguide, thereby incurring minimal area overhead, operate in room temperature, and are DWDM-compatible as they exhibit gain spectrum over 200 *nm* [19]. Likewise, InP-lasers [21] have high peak power, and their emission wavelength is tunable in a wide range and highly stable with temperature, which makes them DWDM compatible. InP-lasers can be integrated on Si [22, 23] so they can also be used as an on-chip laser source. While EcoLaser+ requires a fast laser, it does not depend on a singular laser technology; any of these technologies [19, 21, 22, 23] is suitable for laser power gating with EcoLaser+.

It is important to offer the interested reader an additional perspective on laser turn-on times: VCSELs can turn-on with sub-100 *ps* delay [16], thus they can be directly modulated at speeds exceeding 35 GHz [24]. However, there is no clear path forward on how a multi-wavelength VCSEL array can be implemented, as VCSEL's operating wavelengths are defined by the epitaxial growth [14], which makes them unsuitable for on-chip applications with WDM. On top of that, it is hard to protect the integrity of messages with direct laser modulation due to chirping and the pattern effect [16]. For these reasons, in this paper we consider direct-modulated VCSELs inappropriate for EcoLaser+, and assume laser sources like Ge-lasers, InP-lasers, or any other fast WDM-compatible continuous-wave laser that can be integrated on chip.

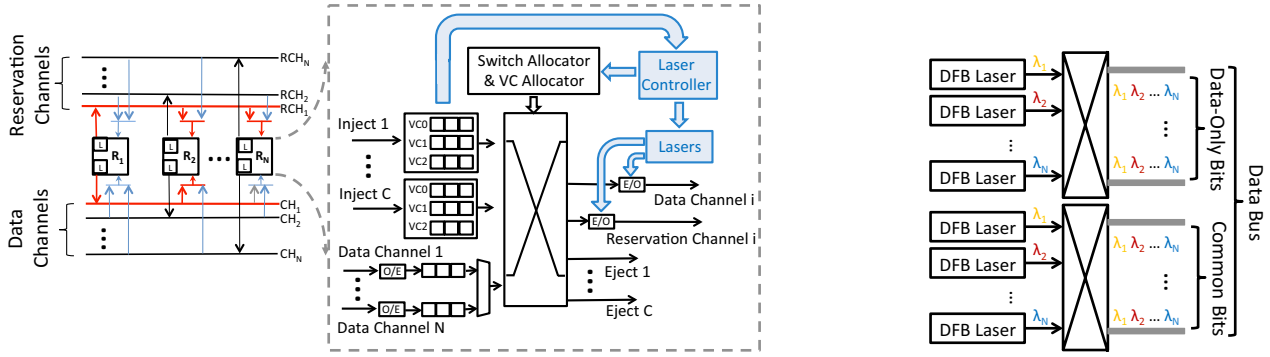

FIGURE 1. SWMR Crossbar and router microarchitecture(a), On-Chip laser array configuration (b).

3. LASER CONTROL SCHEMES

The laser control schemes aim to save laser energy by turning the lasers off whenever the bus (i.e., data channel) is not utilized. Energy savings come at the cost of potential increase in the message latency, because messages have to wait for the laser to turn on before transmission, when they find the laser off. The previously proposed adaptive EcoLaser scheme [8] controls on-chip lasers to achieve laser energy savings at low utilization levels while providing high performance under higher utilization. The adaptive EcoLaser scheme reacts to the demands of the aggregate workload and opportunistically turn the laser off during periods of low activity to save energy, and leave it on during periods of high activity in order to meet the high bandwidth demand. In this work, we propose EcoLaser+, a novel laser control scheme, which achieves higher laser energy savings for all utilization levels while minimizing the additional laser turn-on delay overhead of laser control. EcoLaser+ achieves energy efficiency with high performance by keeping the majority of the data-bus inactive while sending small (dataless) messages, and anticipating upcoming messages to turn the lasers on ahead of time.

EcoLaser saves laser energy by turning the optical bus off when it is idle. However, EcoLaser 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. A 300-bit (or 300 wavelengths) optical bus sends a cache line in one processor cycle, because the optical links runs at double the processor frequency (narrower bus causes serialization delay). 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 (data-only bits) and remain idle otherwise. Remaining 44 bits of the optical bus is activated for all messages (common bits). Figure 1.b illustrates the separation of the data bus into two independent sections as common bits and data-only bits. In order to achieve higher laser energy savings EcoLaser+ activates the data-only portion of the to send data messages only.

EcoLaser+ 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. EcoLaser+ turns the laser on proactively for most of the data and control messages to reduce the latency overhead. EcoLaser+ anticipates the early laser activation for data messages by correlating cache coherence request messages to replies, and turns the laser on proactively which reduces the latency overhead. 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. EcoLaser+ detects that and turns on the laser 1.5 *ns* (i.e., one laser turn-on delay) before the data become available.

EcoLaser+ does not turn off the lasers unless (a) there is no message at the injection buffers, and (b) the laser has stayed on for the minimum laser stay-on time *"K"*. EcoLaser+ tends to turn off the laser quicker, which saves more laser energy when the crossbar is not heavily utilized. However, under heavier traffic, turning the laser off quickly results in lost opportunities to quickly send, and increases the number of times the laser has to be turned on anew. The frequent laser turn-on delays decrease performance. On the other hand, when *K* is high, the laser tends to stay on for longer, which increases per-

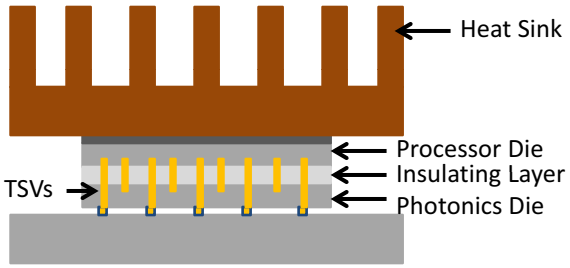

FIGURE 2. A 3D-Stacked Architecture with the Thermal Insulating Layer.

formance under heavier traffic, but wastes more laser energy when the utilization is low. EcoLaser+ observes the amount of laser turn-on requests to adjust the laser stay-on time *K* at run time. Frequent laser turn-on requests hint to lost opportunities to transmit opportunistically, and EcoLaser+ increases *K* to keep the laser on for longer. A low number of laser turn-on requests hints at potentially wasted laser energy, so EcoLaser+ decreases *K* to save more laser energy, by turning the laser off more quickly. To prevent oscillation or overshooting *K* from its ideal setting, we employ a hysteresis counter which robustly captures the laser turn-on request trends. The hysteresis counter decrements on every cycle on which there is no other counter activity. Upon sensing a laser turn-on signal, the counter increments by adding some value to it. Whenever the counter reaches its upper threshold, *K* increases by 1; whenever the counter reaches its lower threshold, *K* decreases by 1. The hysteresis counter controls the value of *K* in a stable manner, because increasing *K* results in a reduction of laser turn-on requests, as the likelihood of a writer finding an available data slot with light increases, and vice versa. The threshold settings and the increment and decrement values of the hysteresis counter change its reactive behavior (making it more lazy or aggressive).

3.1 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 EcoLaser+, 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. MICRORING RESONATOR INSULATION

The basic building block of photonic interconnects is the microring resonators, which are designed to resonate at a specific wavelength to realize add/drop filters and modulators. Microring resonators are very susceptible to temperature-related changes to their resonance wavelength, because the refractive index of Si changes with temperature. In a multicore processor there is a potential for significant variations in temperature, so the microring resonance wavelength shift is a problem that needs to be addressed. *Trimming* is a technique to correct the shift in the resonant wavelength either by using heating (shifts the resonant wavelength towards the red) or current injection (shifts the resonant wavelength towards the blue). Previous work [25] shows that trimming by current injection causes instability and thermal runaways and should be avoided, thus microrings need to be trimmed and maintained at a constant temperature using the heaters only. Such a trimming methodology consumes significant amount of power (up to 40W [25]), because the microrings need to be heated up above the maximum temperature the microprocessor is allowed to operate in, at all times.

The temperature variations in a microprocessor may occur rapidly depending on the workload [7], are both spatial and temporal variations in nature, and can exceed 30^{o}C difference. As the microprocessor die is 3D-stacked and directly coupled to the photonics die, these temperature variations result in high trimming power consumption. EcoLaser+ addresses this problem by thermally decoupling the 3D-stacked logic die from the 3D-stacked photonics die by introducing an insulating layer between them. The insulation layer (a) reduces the temperature variation in the photonics layer and (b) allows for easier trimming by trapping the heat within the photonics layer. An oxidized macro porous Si layer [26] or an air cavity etched between layers [27] can be exploited to realize the thermal insulation. In both cases, high ratio TSV's, introduced in [27, 28], should be routed through the insulating layer to maintain the communication between the 3D-stacked logic die and

the photonics die. Figure 2 illustrates an example of the thermal insulation. In this case, the insulating layer decouples the photonics die (bottom chip) from the logic die (top chip) which is cooled by a forced-air cooled heatsink.

5. EXPERIMENTAL METHODOLOGY

5.1 Interconnect Performance and Energy Analysis

To evaluate the performance and energy consumption of EcoLaser+ in isolation from the interference of other system components or application characteristics, we employ a cycle-accurate network simulator based on Booksim 2.0 [29], 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 EcoLaser+ and EcoLaser schemes, and compare them 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 [9], 10ns access
Networks	radix-16 SWMR and Firefly

5.2 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 [30, 31] integrated with Booksim 2.0 [29] and DRAMSim 2.0 [32]. Table 1 details the architectural modeling parameters. We assume a shared and physically distributed L2 cache and directories. The memory controllers are uniformly distributed on the chip, and they use the same physical interconnect with VCs to avoid deadlock. All messages below L1 cache traverses the interconnect. The power consumption of the electrical interconnect is calculated using DSENT [33]. We target a 16 *nm* technology, and have updated our tool chain accordingly based on ITRS projections [34]. The simulated system executes a selection of benchmarks from SPLASH-2, PARSEC 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 [35], and the power consumption of the optical networks using the analytical power model by Joshi *et al.* [36]. We estimate the temperature of the chip using HotSpot 5.0 [37]. The estimated temperature is then used to refine the leakage power estimate. We adjust DVFS based on the stable-state power and temperature estimates.

To put EcoLaser+'s performance and energy consumption into perspective, we include in our evaluation a high-performance all-electrical on-chip interconnect: a 4-ary 4-flat flattened butterfly (derived by combining 4 level butterfly network) [38] with concentration ratio of 4 (*Flat-Butterfly*). For Flat-Butterfly we model routers with 10 input and output ports and a 3-cycle routing delay. Routers are connected through 88-bit bi-directional links (1-flit control, 7-flit data messages) with 1-cycle local, 2-cycle semi-local and 3-cycle global link delay. To show the range of EcoLaser+'s impact, we evaluate its application on an optical network topology that uses 4 optical SWMR crossbars similar to (*Firefly* [2]). The Firefly connects 16 local clusters (4 routers each) using 4 SWMR crossbars. The local clusters use an electrical ring to route packets within the destination cluster, 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. The modeling of the optical SWMR crossbars is described in Section 5.1. Laser turn-on delay of 1.5 *ns* is included in the EcoLaser+ and EcoLaser scheme's implementations. Finally, we contrast EcoLaser+ 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 EcoLaser+'s average energy savings.We compare the performance (user instructions per sec), energy per instruction (EPI) of Flat-Butterfly, the baseline scheme without laser control (No-Ctrl), Power_Eq, EcoLaser [8], EcoLaser+, and perfect laser control (Perfect).

5.3 Laser Power Consumption Calculation

Table 2 shows the optical loss parameters for the modulators, demodulators, drop filters, and detectors introduced in [9] and assumed in this work. The modulation and demodulation energy is 150 *fJ/bit* at 10 *GHz* [9]. The laser power per wavelength and total laser power are calculated in Table 2 using the analytical models introduced in [36]. The total laser power in Table 2 includes the laser power for both data and reservation channels, plus the laser efficiency of 15%, 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).

TABLE 2. Nanophotonic Parameters and Laser Power.

		Firefly
	per Unit	**Total**
DWDM		64
Splitter	0.2 *dB*	0.6 *dB*
WG Loss	0.3 *dB/cm*	3 *dB*
Nonlinearity	1 *dB*	1 *dB*
Modulator Ins.	0.5 *dB*	0.5 *dB*
Ring Through	0.01 *dB*	10.24 *dB*
Filter Drop	1.2 *dB*	1.2 *dB*
Total Loss		**16.64 *dB***
Detector		**-20 *dBm***
Laser Power Per Wavelength		**0.461 *mW***
Total LaserPower	*15% Eff.*	**59.2 *W***

5.4 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., 0.1*dB*) 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 EcoLaser+ 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 EcoLaser+ 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.

5.5 Resonant Ring Heater Modeling

To calculate the total ring heating power we extend the method by Nitta *et al.* [25] 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 [37] to model the transient temperature changes in the optical die. After we execute a workload and collect transient temperature traces (300 samples with 10K processor cycle interval), we calculate the average ring heating power required to maintain the photonic die at the constant micro-ring trimming temperature during the entire execution. We model 3D-stacked silicon dies (logic die and photonics die) with 150 *um* thickness, separated by a 100 *um* porous silicon insulation layer or a 10 *um* air cavity. The thermal resistivity for Si is 0.01 *m-K/W*, for the porous Si insulation layer is 1 *m-K/W* [26], and for air is 40 *m-K/W* [27]. The TSVs running through the insulation layer are highly conductive, so we also include the thermal impact of TSVs in our model. The ambient temperature is fixed at *45C*. In addition, we account for the individual ring trimming power required to overcome process variations, as described in [36].

6. EXPERIMENTAL RESULTS

6.1 Network Performance

Laser control saves energy by turning off the lasers whenever the data bus is idle. Energy savings come with the potential cost of increased message latency, because messages may have to wait for the laser to turn back on. EcoLaser+ turns the

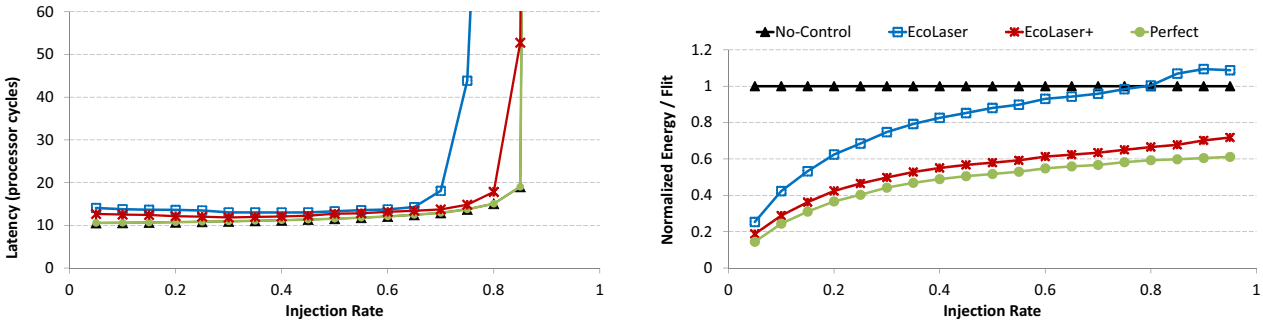

FIGURE 3. Load-Latency (left) and Energy-per-Flit (right)

laser on proactively by predicting upcoming messages, so the majority of the messages do not have to wait. Moreover, by keeping the data-only portion of the data bus inactive for small (dataless) messages, EcoLaser+ achieves higher energy savings than EcoLaser. We investigate the trade-off between the laser energy savings and the network performance of EcoLaser+ on a Firefly topology using random traffic, and compare it against No-Ctrl, EcoLaser [8] and Perfect.

Messages in EcoLaser exhibit a 6-cycle average delay at low injection rates, because some of the messages find the laser active and transmit immediately (Figure 3). This overhead decreases slightly for higher injection rates as more messages find the laser active. In comparison, EcoLaser+ incurs only a 1-cycle overhead at low injection rates, because it predicts the majority of the messages and activates the laser ahead of time. EcoLaser provides 12% lower throughput than EcoLaser+, as it is more susceptible to exposing the laser turn-on delay.

EcoLaser's energy savings disappear as the injection rate grows, as it keeps the laser on most of the time. However, EcoLaser+ exhibits high energy savings at high injection rates because it can keep the data-only portion of the data bus inactive (Figure 3-left). On average over injection rates, EcoLaser+ consumes 34% lower laser energy per flit than EcoLaser, and is within 4% of Perfect.

6.2 The Performance Cost of Laser Control

EcoLaser and EcoLaser+ trade off increased message latency for laser energy savings, therefore they are expected to lower the performance of the system compared to No-Ctrl. However, in a realistic system with on-chip lasers, the power saved by EcoLaser+ may decrease the thermal emergencies and decrease the need for core throttling, and thus increase performance. Previous work showed that EcoLaser [8] lowers the on-chip laser power consumption, which allows better cooling, reduces core throttling, and increases the performance of a realistic multicore that employs a thermal management system like DVFS. Similarly, EcoLaser+ lowers the power consumption by a significant margin and exposes less of the laser turn-on delay than EcoLaser. Therefore, a realistic power-limited system is expected to attain higher performance with EcoLaser+ because the cores will be throttled less.

We aim to analyze separately the effects of increasing the effective message latency and reducing the need for core throttling. First, we analyze the performance cost of EcoLaser+ by evaluating it on a multicore that is not subject to thermal constraints, thus cores run at their maximum frequency (5 *GHz*). Our workload suite includes both memory-intensive workloads that generate high traffic and are sensitive to interconnect latency (bodytrack, em3d, ocean, appbt, tomcatv), as well as compute-intensive workloads that have low injection rates and are less sensitive to message latency (fmm, moldyn,

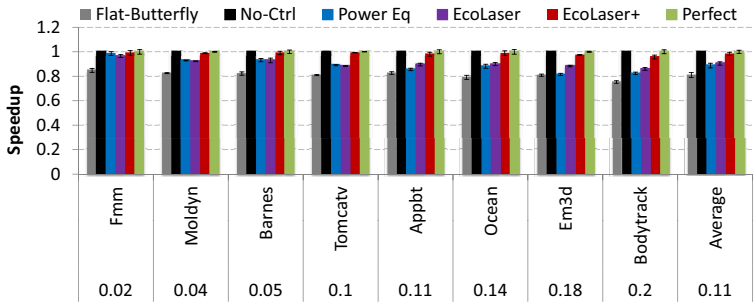

FIGURE 4. Speedup for Firefly on a hypothetical multicore without thermal constraints.

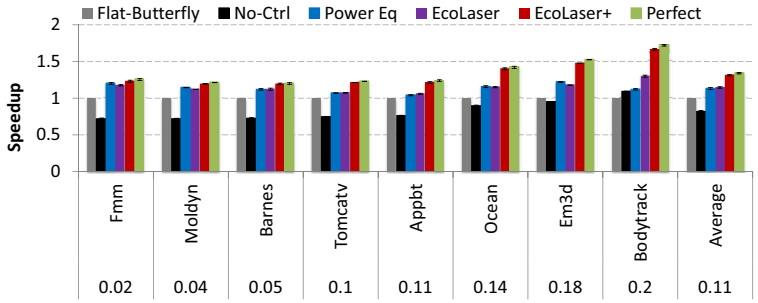

FIGURE 5. Speedup over Flat-Butterfly for Firefly under realistic thermal constraints.

barnes). Figure 4 summarizes our findings. The injection rate of each application appears below its name. For reference, we also present the performance of a multicore with a traditional electrical network (Flattened-Butterfly).

EcoLaser exposes the laser turn-on delay overhead, causing 10% slowdown while saving only 30% of laser energy on average (Figure 4). With the support of partial activation of the data bus and proactive laser turn on, EcoLaser+ saves 62% of the laser energy (92% maximum) while causing only 1.9% slowdown compared to No-Ctrl. Furthermore, the laser energy savings of EcoLaser+ are in the vicinity of 3-6% of Perfect when running real-world workloads. Overall, EcoLaser+ exhibits higher energy savings on real-world workloads than on synthetic random traffic patterns, because real-world workloads typically have bursty (and sparse) memory access patterns.

Power_Eq approximates EcoLaser+'s energy savings by scaling down its width (100-bit flits instead of 300-bit flits), but otherwise is similar to No-Ctrl. While achieving similar energy savings, Power_Eq suffers from high serialization delays and under-performs both EcoLaser and EcoLaser+. Thus, saving laser energy by reducing the width of the interconnect is not a good alternative to laser control.

6.3 The Impact of EcoLaser+ on a Realistic Multicore

The wall-plug power consumption of on-chip laser sources can be a major contributor to a multicore's power budget, due to the low laser efficiency levels (15% [15]). Under realistic thermal and power constraints, DVFS in No-Ctrl will have to throttle the cores to keep the chip within a safe temperature. EcoLaser+, however, reduces the laser power and results in a cooler chip with less core throttling and higher performance.

It is important to note that the impact of EcoLaser+'s energy savings on system performance and energy efficiency depends on the total laser power consumption of the photonic network.Firefly consists of 4 radix-16 SWMR crossbars, therefore the wall plug power consumption for on-chip lasers is 56.5W. All laser control schemes outperform No-Ctrl on all workloads, because the laser energy savings are a considerable fraction of the chip's power budget (Figure 5). EcoLa-

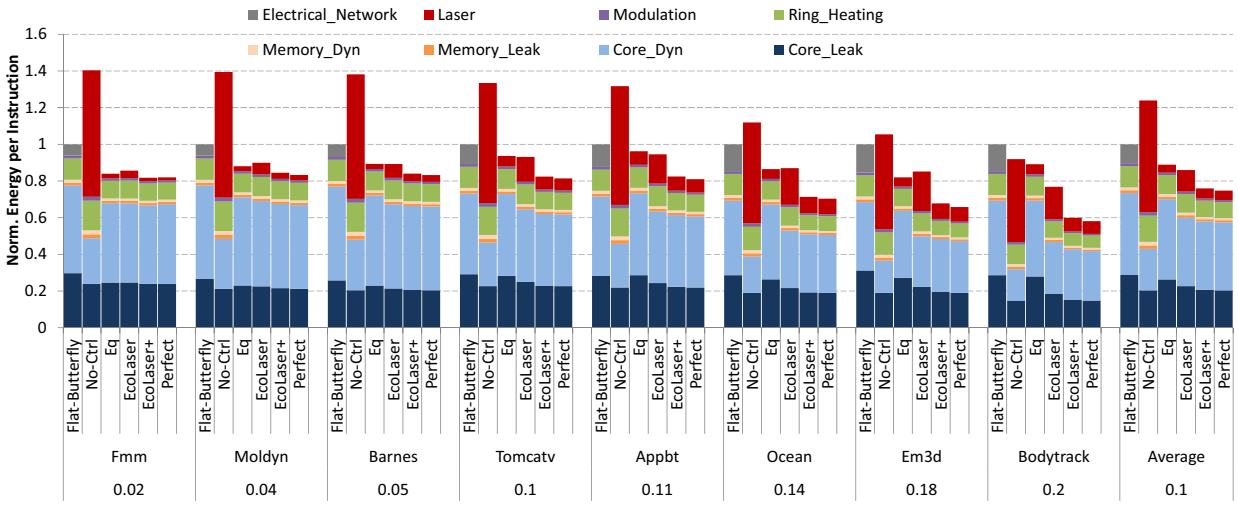

FIGURE 6. Energy Per Instruction for Firefly. The evaluated designs are from left to right: Flattened-Butterfly, No-Ctrl, Power_Eq (Eq), EcoLaser, EcoLaser+, and Perfect.

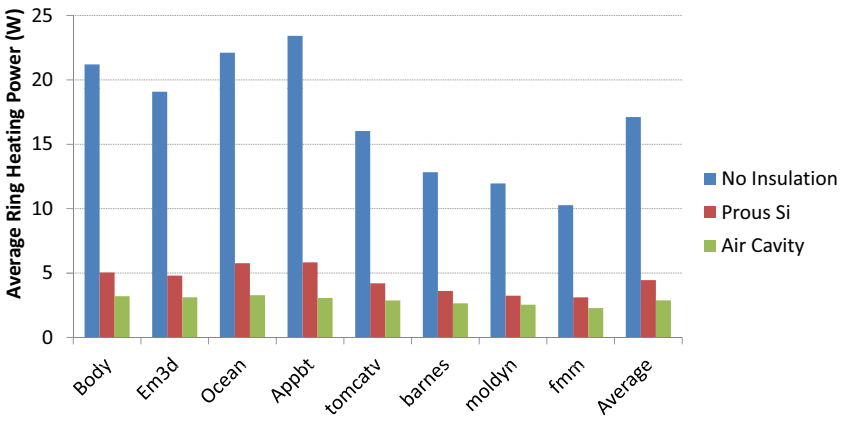

FIGURE 7. Average Microring Trimming Power Consumption (W) for No Insulation, and Insulation with Porous Silicon, and Insulation with Air cavity.

ser+ outperforms No-Ctrl for all workloads (1.6x on average) and achieves 39% lower EPI (Figure 6). Similarly, EcoLaser+ outperforms EcoLaser by 1.15x and has 12% lower EPI (Figure 6). Firefly with no laser control (No-Ctrl) is unable to deliver high performance on a realistic multiprocessor as more than half of the chip's power budget it consumed by the interconnect. However, EcoLaser+ enables Firefly to deliver on its performance promise by making the interconnect energy proportional. It is important to note that EcoLaser+'s performance and EPI are within 5-6% of Perfect's, which indicates that EcoLaser+ is harvesting the majority of the possible laser energy savings.

6.4 The Impact of Microring Resonator Insulation

The microring heating power is a significant contributor to the multicore processor's power budget, because the microrings need to be heated at all times above the maximum temperature that the microprocessor is allowed to operate in (Section 4). EcoLaser+ aims to reduce microring heating power consumption by thermally decoupling the 3D-stacked logic die from the 3D-stacked photonics die by introducing an insulating layer between them. The insulation layer (a) reduces the temperature variation in the photonics layer and (b) allows for easier trimming by trapping the heat within the photonics layer. In order to observe the impact of the insulating layer, we calculate the required microring heating power consumption when (a) there is no insulating layer, (b) with a porous Si insulating layer, and (c) with an air cavity as an insulating layer (Section 5.5). We observe that, when there is no insulation between the 3D-stacked logic die and the photonics die, the thermal fluctuations in the logic die directly reflect on the photonics die and force the ring heaters to consume 17W on average across our application suite. The porous Si insulating layer shields the photonics die from the high temperature fluctuations, and increases the thermal resistivity between dies, so it reduces the microring heating power consumption by 3.8x. Even though its feasibility and practicality is not clear, an air cavity formed between the dies acts as a better insulator, and can reduce the microring heating power consumption by up to 5.9x compared to no insulation.I

7. RELATED WORK

Different on-chip interconnect networks have been proposed that exploit CMOS-compatible photonics for future multicore processors. The hierarchical Firefly architecture [2] 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.* [9, 10] 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 that use SWMR can utilize EcoLaser+ to achieve higher laser energy efficiency while maintaining their performance.

Previous work has explored segregating the interconnect used for core communication from the interconnect used for communication with the cache [39, 40] to lower the network cost or to optimize for data placement and partitioning. However, such designs have not been proposed or evaluated in the context of photonic interconnects. EcoLaser+ segregates the data portion of the photonic interconnect from the control portion and manages them separately, in order to maximize power savings without hurting performance.

Previously, Thonnart *et al.* [41] 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.* [42] identify the constant laser power consumption when channel utilization is low as an inefficiency, and 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.* [12] propose an optical SWMR crossbar and electrical hybrid interconnection network, and improve performance by utilizing the coherence protocol. Kurian *et al.* [12] mention 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 scheme. Nitta et al. [43] shows the energy inefficiency of photonic interconnects under low utilization, and proposes to improve efficiency by recapturing the energy of photons which are not used for communication. In contrast to previous work, we propose EcoLaser+, a laser control mechanism that improves the laser energy efficiency for SWMR crossbars, while providing high bandwidth and performance.

8. CONCLUSION

In this paper we propose EcoLaser+, a laser-control mechanism that turns the laser off during periods of inactivity to save energy, and meets high bandwidth demands by turning the laser on for as long as necessary, while at the same time reduces the power consumption of the microring heaters by thermally insulating the photonic die from the logic die. EcoLaser+ improves upon EcoLaser [8] by keeping the majority of the data bus off while sending small (data-less) messages and provides better performance by turning the laser on proactively. Our results indicate that EcoLaser+ saves between 62-92% of the laser power for the Firefly topology on real-world workloads. EcoLaser+ harvests the vast majority of the energy benefits, as it closely tracks (within 2-6% on average) a perfect controller with full knowledge of future interconnect requests. Moreover, the power savings of EcoLaser+ allow for providing a higher power budget to the cores, which enables a multicore chip with EcoLaser+ to achieve speedups of 1.5-1.7x over a multicore with a traditional optical interconnect that keeps the lasers always on. Finally, by thermally insulating the photonic die from the logic die, EcoLaser+ reduces the power consumed by the microring heaters by a factor of 3.8x to 5.9x, depending on the insulating material used.

9. ACKNOWLEDGEMENTS

This work was partially supported by NSF award CCF-1218768, an ISEN booster award, and the June and Donald Brewer Chair at Northwestern.

10. REFERENCES

[1] Chen, G., Chen, H., Haurylau, M., Nelson, N., Fauchet, P. M., Friedman, E., and Albonesi, D., "Predictions of cmos compatible on-chip optical interconnect," in *7th International Workshop on System-Level Interconnect Prediction (SLIP)*, 13- 20 (2005).

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

[3] Cianchetti, M. J., Kerekes, J. C., and Albonesi, D. H., "Phastlane: a rapid transit optical routing network," in *Proceedings of the 36th Annual International Symposium on Computer Architecture*, *ISCA 09*, 441- 450 (2009).

[4] Pan, Y., Kim, J., and Memik, G., "Flexishare: Channel sharing for an energy-efficient nanophotonic crossbar," in *Proceedings of the IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 1- 12 (2010).

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

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

[7] Lee, J. S., Skadron, K., and Chung, S. W., "Predictive temperature-aware dvfs," *IEEE Transactions on Computers*, 59(1):127- 133, (2010).

[8] Demir, Y. and Hardavellas, N., "Ecolaser: An adaptive laser control for energy efficient on-chip photonic interconnects," in *Proceedings of the International Symposium on Low-Power Electronics and Design*, *ISLPED 14* (2014).

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

[10] Batten, C., Joshi, A., Stojanovic, V., and Asanovic, K., "Designing chip-level nanophotonic interconnection networks," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2(2):137- 153, (2012).

[11] Kirman, N., Kirman, M., Dokania, R. K., Martinez, J. F., Apsel, A. B., Watkins, M. A., and Albonesi, D. H., "Leveraging optical technology in future bus-based chip multiprocessors," in *Proceedings of the 39th IEEE/ACM Annual International Symposium on Microarchitecture*, *MICRO 39*, 492- 503 (2006).

[12] Kurian, G., Sun, C., Chen, C.-H., Miller, J., Michel, J., Wei, L., Antoniadis, D., Peh, L.-S., Kimerling, L., Stojanovic, V., and Agarwal, A., "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)*, 1117- 1130 (2012).

[13] Duan, G.-H., Shen, A., Akrout, A., Dijk, F. V., Lelarge, F., Pommereau, F., LeGouezigou, O., Provost, J.-G., Gariah, H., Blache, F., Mallecot, F., Merghem, K., Martinez, A., and Ramdane, A., "High performance inp-based quantum dash semiconductor mode-locked lasers for optical communications," *Bell Labs Technical Journal*, 14(3):63- 84, (2009).

[14] Heck, M. and Bowers, J., "Energy efficient and energy proportional optical interconnects for multi-core processors: Driving the need for on-chip sources," *Selected Topics in Quantum Electronics, IEEE Journal of*, 20(4):1- 12, July (2014).

[15] Koch, B. R., Norberg, E. J., Kim, B., Hutchinson, J., Shin, J.-H., Fish, G., and Fang, A., "Integrated silicon photonic laser sources for telecom and datacom," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013* , PDP5C.8, Optical Society of America (2013).

[16] Petermann, K., [*Laser Diode Modulation and Noise*], , vol. 3 of *Advances in Optoelectronics (ADOP*), Springer (1988).

[17] Hisham, H., Mahdiraji, G., Abas, A., Mahdi, M., and Adikan, F., "Characterization of turn-on time delay in a fiber grating fabry-perot lasers," *IEEE Photonics Journal*, 4(5):1662- 1678, (2012).

[18] Hisham, H., Mahdiraji, G., Abas, A., Mahdi, M., and Adikan, F., "Characterization of transient response in fiber grating fabry-perot lasers," *IEEE Photonics Journal*, 4(6):2353- 2371, (2012).

[19] Camacho-Aguilera, R. E., Cai, Y., Patel, N., Bessette, J. T., Romagnoli, M., Kimerling, L. C., and Michel, J., "An electrically pumped germanium laser," *Optics Express*, 20(10):11316- 11320, (2012).

[20] Liu, J., Sun, X., Camacho-Aguilera, R., Kimerling, L. C., and Michel, J., "Ge-on-si laser operating at room temperature," *Opt. Lett.*, 35(5):679- 681, (2010).

[21] Kotelnikov, E., Katsnelson, A., Patel, K., and Kudryashov, I., "High-power single-mode ingaasp/inp laser diodes for pulsed operation," *Proceedings of SPIE*, 8277:827715- 827715- 6, (2012).

[22] Paniccia, M. and Bowers, J., "First electrically pumped hybrid pumped hybrid silicon laser silicon laser," http://www.intel.com/content/dam/www/public/us/en/documents/technology-briefs/intel-labs-hybrid-silicon-laser-announcement.pdf (2006).

[23] Fang, A. W., Park, H., Cohen, O., Jones, R., Paniccia, M. J., and Bowers, J. E., "Electrically pumped hybrid AlGaInAs-silicon evanescent laser," *Optics Express*, 14(20):9203- 9210, (2006).

[24] Wolf, P., Moser, P., Larisch, G., Hofmann, W., Li, H., Lott, J., Lu, C.-Y., Chuang, S., and Bimberg, D., "Energy-efficient and temperature-stable high-speed VCSELs for optical interconnects," in *15th International Conference on Transparent Optical Networks (ICTON)*, 1- 5 (2013).

[25] Nitta, C., Farrens, M., and Akella, V., "Addressing system-level trimming issues in on-chip nanophotonic networks," in *17th IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 122- 131 (2011).

[26] Mondal, B., Basu, P., Reddy, B., Saha, H., Bhattacharya, P., and Roychoudhury, C., "Oxidized macro porous silicon layer as an effective material for thermal insulation in thermal effect microsystems.," in *International Conference on Emerging Trends in Electronic and Photonic Devices Systems*, 202- 206 (2009).

[27] Zhang, Y., Oh, H., and Bakir, M., "Within-tier cooling and thermal isolation technologies for heterogeneous 3d ics," in *2013 IEEE International 3D Systems Integration Conference (3DIC)*, 1- 6 (2013).

[28] Fischer, A. C., Bleiker, S. J., Haraldsson, T., Roxhed, N., Stemme, G., and Niklaus, F., "Very high aspect ratio through-silicon vias (tsvs) fabricated using automated magnetic assembly of nickel wires," *Journal of Micromechanics and Microengineering*, 22(10):105001, (2012).

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

[30] Hardavellas, N., Somogyi, S., Wenisch, T. F., Wunderlich, R. E., Chen, S., Kim, J., Falsafi, B., Hoe, J. C., and Nowatzyk, A. G., "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, (2004).

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

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

[33] Sun, C., Chen, C.-H. O., Kurian, G., Wei, L., Miller, J., Agarwal, A., Peh, L.-S., and Stojanovic, V., "Dsent - a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *6th IEEE/ACM International Symposium on Networks-on-Chip*, 201- 210 (2012).

[34] European Semiconductor Industry Association (ESIA), Japan Electronics and Information Technology Industries Association (JEITA), Korean Semiconductor Industry Association (KSIA), Taiwan Semiconductor Industry Association (TSIA), and United States Semiconductor Industry Association (SIA), "The international technology roadmap for semiconductors (itrs).," http://www.itrs.net/ (2012).

[35] Li, S., Ahn, J. H., Strong, R. D., Brockman, J. B., Tullsen, D. M., and Jouppi, N. P., "Mcpat: an integrated power, area, and timing modeling framework for multicore and manycore architectures," in *Proceedings of the 42nd IEEE/ACM Annual International Symposium on Microarchitecture*, *MICRO-42*, 469- 480 (2009).

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

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

[38] Kim, J., Dally, W. J., and Abts, D., "Flattened butterfly: A cost-efficient topology for high-radix networks," in *Proceedings of the 34th Annual International Symposium on Computer Architecture*, *ISCA 07*, 126- 137 (2007).

[39] Huh, J., Kim, C., Shafi, H., Zhang, L., Burger, D., and Keckler, S. W., "A nuca substrate for flexible cmp cache sharing," in *Proceedings of the 19th Annual International Conference on Supercomputing*, *ICS 05*, 31- 40, ACM, New York, NY, USA (2005).

[40] Lotfi-Kamran, P., Grot, B., and Falsafi, B., "Noc-out: Microarchitecting a scale-out processor," in *Proceedings of the 2012 45th Annual IEEE/ACM International Symposium on Microarchitecture*, *MICRO-45*, 177- 187, IEEE Computer Society, Washington, DC, USA (2012).

[41] Thonnart, Y., Beigne, E., Valentian, A., and Vivet, P., "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*, 48- 57 (2008).

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

[43] Nitta, C., Farrens, M., and Akella, V., "Dcof: An arbitration free directly connected optical fabric," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, 2(2):169- 182, June (2012).