Run-Time Laser Power Management in Photonic NoCs with On-Chip Semiconductor Optical Amplifiers

Ishan G Thakkar, Sai Vineel Reddy Chittamuru, Sudeep Pasricha
Department of Electrical and Computer Engineering
Colorado State University, Fort Collins, CO, U.S.A.
{ishan.thakkar, sai.chittamuru, sudeep}@colostate.edu

***Abstract*— Photonic network-on-chip (PNoC) architectures are projected to achieve very high bandwidth with relatively small data-dependent energy consumption compared to their electrical counterparts. However, PNoC architectures require a non-trivial amount of static laser power, which can offset most of the bandwidth and energy benefits. In this paper, we present a novel low-overhead technique for run-time management of laser power in PNoCs, which makes use of on-chip semiconductor amplifiers (SOA) to achieve traffic-independent and loss-aware savings in laser power consumption. Experimental analysis shows that our technique achieves 31.5% more laser power savings with 12.8% less latency overhead compared to another laser power management scheme from prior work.**

***Index Terms*—Photonic Networks-On-Chip, Power Management, Laser Source, Optical Signal Loss**

I. Introduction

In the many-core era, processors with hundreds of cores on a single chip are gradually becoming a reality. The performance of these many-core processors is driven by the effectiveness of the underlying electrical network-on-chip (ENoC) fabrics that are becoming increasingly crosstalk- and energy-limited [1]. To this end, due to the recent developments in the area of silicon photonics, photonic network-on-chip (PNoC) fabrics have been projected to supersede ENoCs. PNoCs offer a multitude of benefits over ENoCs such as: 1) higher bandwidth density; 2) distance-independent bit-rate; and 3) smaller data-dependent energy. However, irrespective of network traffic and utilization, PNoCs dissipate a non-trivial amount of static power in their laser source. The high laser power overheads can offset the bandwidth and energy advantages of PNoCs. Therefore, it is imperative to forge new techniques that can reduce the static power consumed in the laser sources of future PNoC architectures.

Several techniques have been proposed in prior works, e.g. [1]-[6], that aim to reduce static power in the laser sources of PNoCs. All of these techniques leverage temporal and spatial variations in network traffic and opportunistically adjust the power in laser sources by tuning or distributing the available network bandwidth. These methods tend to notably reduce the power in laser sources during low network load conditions. However, if the losses encountered by optical signals in the network between the source and destination are high, these methods would still require excessive laser power to compensate for the high losses, even under low network load conditions. *This observation motivates the need for a technique that can provide traffic-independent and loss-aware savings in laser power.*

In this paper, we present a novel low overhead technique for run-time management of laser power in PNoCs, which makes use of on-chip semiconductor amplifiers (SOA) to achieve traffic-independent and loss-aware savings in laser power. We refer to our SOA-enabled laser power management technique as *SOA_LPM*. Unlike the techniques proposed in prior works [1]-[6], *SOA_LPM* draws minimum power from the off-chip laser source and offloads the burden of loss-aware run-time laser power management to on-chip SOAs, which in turn enables significant savings in laser power with minimal overheads. Moreover, *SOA_LPM* is orthogonal to all the other laser power management techniques reported in prior works, and can be used in combination with any of them. Our novel contributions in this paper are summarized below:

- We propose a low overhead, SOA-enabled, and loss-aware technique (*SOA_LPM*) to manage and optimize the laser power overhead in PNoC architectures at run-time;
- We implement our *SOA_LPM* technique on a multiple-write -multiple-read (MWMR) photonic waveguide architecture;
- We present device-level analytical models of on-chip SOAs, and based on this models, we analyze the energy, area and performance overhead of our *SOA_LPM* technique;
- We evaluate *SOA_LPM* by implementing it on a well-known PNoC architecture Flexishare [7] and compare it with another laser power management technique from prior work [3].

II. Background

PNoC architectures (e.g., [17]-[19]) employ multiple high-bandwidth photonic links, each of which connects two or more nodes (e.g., cores). Typically, a large number of wavelengths are dense wavelength division multiplexed (DWDM) in a single photonic link. Each wavelength corresponds to a channel that is used for serial data transfers. Additionally, a photonic link employs microring resonator (MR) modulators (that are in resonance with the utilized wavelengths) at the source node to modulate electrical signals onto photonic signals that travel through the link, and MR detectors at the destination node to detect photonic signals and recover electrical signals. In general, the use of multiple wavelengths (or channels) in a photonic link enables high bandwidth parallel data transfers across the link.

The amount of laser power required by a source node to transfer data across the parallel wavelength channels of a photonic link to a destination node can be expressed as:

$$P_{Laser} - S \geq P_{Loss} + 10\log_{10}(N_\lambda), \tag{1}$$

where, P_{Laser} is the required laser power in dBm, N_λ is the number of wavelength channels in the link, P_{Loss} (in dB) is the total optical loss faced by a photonic signal along the link from the source to the destination, and S is the sensitivity of the detector (assumed to be -20dBm [8]). P_{Loss} includes optical signal losses such as through loss in MR modulators and detectors, modulating losses in modulator MRs, detection loss in detector MRs, propagation and bending loss in waveguides,

978-1-4673-9030-9/16/$31.00 (c) 2016 IEEE

and splitting loss in splitters. Overall, P_{Laser} thus depends on two main factors: 1) link bandwidth in terms of $N_λ$, which controls the network utilization and traffic, and 2) the total optical loss P_{Loss} during photonic signal propagation [20]-[21].

As implied from Eq. (1), if a link is underutilized (due to low traffic), its laser power consumption P_{Laser} can be reduced by decreasing $N_λ$ associated with the link. This is exactly what is done in prior works [1], [2], and [3] to reduce the total laser power consumption in PNoCs. Some other prior works also propose laser power management techniques, e.g., [4], [5], wherein they achieve significant power savings using dynamic reconfiguration of photonic networks. In spite of requiring periodic evaluation of network traffic and expensive run-time decision-making, these methods (presented in [1]-[5]) achieve notable savings in laser power. But the savings are highly contingent on information related to network traffic and losses. In contrast to the approaches from [1]-[5], Wang et al. [6] proposed a technique that achieves loss-aware savings in laser power. However, this technique requires the network to function in the TDM communication paradigm only, which incurs architecture specific overheads, limiting the scope of this technique.

In this paper, we present a complementary, low overhead, and SOA-enabled technique called SOA_LPM that can provide traffic-independent and loss-aware laser power savings for PNoC architectures. For transmitting a packet between source and destination nodes, *SOA_LPM* first allocates only the minimum amount of laser power to the source node that is enough for correct detection at the destination node. It then accounts for losses to be faced by the packet on its path from the source to the destination and enables the source to amplify the allocated laser power to the necessary level by using an on-chip SOA. As will be evident in the following sections, *SOA_LPM* proves to be more energy efficient than previously proposed techniques.

Table 1: Definitions and values of parameters for our SOA model

Parameter	Definition	Value
a_1	Constant	6.7×10^{-16} cm^2
n_0	Transparency carrier concentration	1.2×10^{18} cm^{-3}
$α$	Loss in SOA active region	10 cm^{-1}
$Γ$	Light confinement factor	0.4
L	Length of SOA active region	10μm
$Δλ$	SOA gain linewidth	95nm
I_0	Threshold input current	5μA

III. Semiconductor Optical Amplifiers: Overview

We first give a brief explanation of the structure and behavior of SOAs, before presenting analytical models for SOA gain (Section IV.A) and overheads (Section IV.B).

A detailed description of the structure, functionality, and modeling of SOAs is given in [10]. Briefly, an SOA is an optoelectronic device, which can be heterogeneously integrated with a silicon-on-insulator (SOI) based silicon photonic chip, and under suitable operating conditions can amplify an input broadband light signal. The structure of an SOA consists of an active waveguide region made of an intrinsic narrow bandgap material (e.g., AlInGaAs, InGaAsP), which is sandwiched between n-type and p-type cladding materials (e.g., n-InP, p-InP) with wider bandgaps. Free carriers are injected into the active waveguide region from the applied bias current, which in turn causes population inversion in the active region. The population inversion of free carriers in the active region results in stimulated emission of light, which imparts "gain" to the input optical signal. The operational characteristics and the gain spectrum of an SOA depend on its structure and materials used.

A. Analytical Model for SOA Gain

As mentioned earlier, an SOA can provide broadband optical gain in its active region through stimulated emission. This gain obtained in the active region of unity length is called material gain (g_m). The effect of g_m on SOA output power can be exponentially increased by increasing the length of the active region to provide a very high value of single-pass bulk SOA gain (G). Both g_m and G are functions of wavelength ($λ$) and input bias current (I), which can be expressed as [13]:

$$g_m(\lambda, I) = \left[\left(\Gamma a_1 n_0 \left\{\frac{I}{I_0} - 1\right\}\right) - \alpha\right]\left[1 - \frac{2(\lambda - 1570)^2}{\Delta\lambda^2}\right], \quad (2)$$

$$G(\lambda, I) = 10\log_{10}\left(e^{L * g_m(\lambda, I)}\right), \quad (3)$$

Here, g_m and G are in cm^{-1} and dB, respectively; and a_1, $α$, n_0, $Γ$, L, I_0 and $Δλ$ are constants that depend on the structure and operating conditions of the SOA. We took the typical values of these constants from [13], as shown in Table 1. We modeled the SOAs used in this work using Eq. (1), (2) and the constants in Table 1.

B. Overhead Analysis

From the description in the previous sections, the SOA gain is proportional to SOA input current (I). As a result, use of SOA to amplify a DWDM signal comes with a non-zero power overhead corresponding to the non-zero SOA input current. The power consumed by an SOA can be calculated by multiplying the target I with the operating voltage. Based on guidelines for bulk SOA design [13], we design the length (L) of the active region and operating voltage of our SOAs to be 15μm and 1.5V, respectively. Thus, we assess the power overhead of an SOA by multiplying its target I with 1.5V. Moreover, an SOA takes about 20-50ps to adjust to the target gain [13]. In *summary*, an SOA incurs power overhead equal to input current (I) × 1.5V, and latency overhead of at most 50ps (0.25 cycles at 5GHz).

Note that the broadband gain profile of an SOA is subject to fluctuations due to temperature variations and noise. To compensate for these fluctuations, our *SOA_LPM* technique operates the SOAs at such input current levels that can yield 3dB more SOA gain than the desired gain.

IV. SOA Enabled Laser Power Management

Our proposed *SOA_LPM* technique uses SOAs in combination with comb switches [9] and lookup tables to enable loss-aware run-time laser power management in PNoC architectures. *SOA_LPM* can be easily ported to different PNoCs and its implementation depends on the type of bus waveguides (BWGs) used in PNoC architecture. To evaluate our proposed *SOA_LPM* in this work, we implement it on a crossbar based PNoC architecture Flexishare [7]. The Flexishare PNoC uses multiple write multiple read (MWMR) type of BWGs. Each node (e.g., a processing element) connected to a crossbar requires both read and write access to the other nodes. A detailed analysis of *SOA_LPM* implementation for MWMR BWGs is presented in the following sub-section.

A. Implementation for MWMR Bus Waveguide

In an MWMR BWG, multiple nodes are capable of sending and receiving data using their modulating and detecting MR banks respectively. Therefore, MWMR BWGs require arbitration among multiple sender nodes, and also receiver selection among multiple receiver nodes [7][11].

Fig. 1 illustrates the implementation of *SOA_LPM* on a typical MWMR BWG based PNoC. As shown in the figure, the PNoC is comprised of multiple MWMR BWGs. In general, MWMR BWG based PNoCs with *N* nodes have *N* sender nodes and *N* receiver nodes (i.e., all *N* nodes can send as well as receive), with implementation-specific *K* MWMR BWGs. Each MWMR BWG employs a comb switch [9] (a broadband MR switch that can switch the entire DWDM spectrum) and an on-chip SOA [10]. Thus, the PNoC requires *K* SOAs and *K* comb switches as there are *K* BWGs in the PNoC. The SOA of each MWSR BWG can be controlled by any of the *N* sender nodes depending on which sender node wins the arbitration.

As shown in the figure, each sender node of each MWMR BWG has an SRAM based lookup table that stores *N-1* values of loss (corresponding to *N-1* receivers). After arbitration, the authorized sender node initiates a receiver selection phase, at the end of which the sender node accesses a corresponding entry from the lookup table to determine the total loss value that the signal will face on its way to the target receiver. Then, the sender node adjusts the gain of the SOA to be equal to the loss value. Next, the sender node controls the comb switch to extract only the minimum laser power equal to S=-20dBm (detector sensitivity) from the power waveguide and provide it as an input to the SOA. The SOA amplifies the allocated laser power by a value that is equal to the accessed loss value and provides it to the sender node, which then modulates it for communication with the target receiver. Note that we assume each entry of the lookup table to be of 8 bits, therefore, each sender has *N-1* number of 8-bit entries in its SRAM based lookup table.

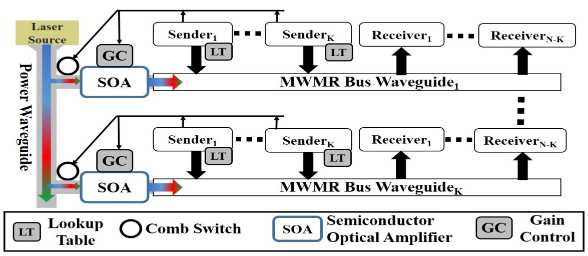

Fig. 1: Implementation of SOA_LPM on MWMR BWG based PNoC.

V. Experimental Evaluation

A. Evaluation Setup

We target a 256-core many-core system for evaluating our SOA enabled laser power management (*SOA_LPM*) technique. We evaluate *SOA_LPM* on a well-known crossbar-based PNoC architecture Flexishare [7]. Flexishare uses 32 groups of MWMR BWGs with a 2-pass token stream arbitration. Each MWMR BWG in Flexishare architecture is capable of transferring 512 bits of data from a source node to a destination node.

We modeled and simulated the architectures at cycle-accurate granularity with a SystemC-based NoC simulator. We used real-world traffic from applications in the PARSEC benchmark suite [12]. Full-system gem5 simulation of parallelized PARSEC benchmarks [13] was used to generate traces that were fed into our cycle-accurate NoC simulator. We set a "warm-up" period of 100M instructions and captured traces for 1B instructions. We targeted a 22nm process node and 5GHz clock frequency for the 256-core system. We considered BWGs with 64 DWDM wavelengths sharing the working band 1510nm–1590nm.

The static and dynamic energy consumption of electrical routers and concentrators in Flexishare architecture is based on results from the DSENT tool [16]. We model and consider area and performance overheads for *SOA_LPM* enabled laser power management. We estimated electrical area and power overhead using gate-level analysis and the open-source CACTI tool [15] for the SRAM-based lookup tables. The electrical area overhead, the electrical power overhead, and the photonic area overhead is estimated to be 0.8mm^2, 0.24mW, and 237.5μm^2 respectively for the Flexishare PNoC architecture.

To compute laser power, we considered detector sensitivity of -20dBm, MR through loss of 0.02 dB, waveguide propagation loss of 1dB/cm, waveguide-bending loss of 0.005dB/90^0, and waveguide coupler/splitter loss as 0.5dB [14]. We calculated photonic loss in components using these values, which sets the photonic power budget and correspondingly the electrical power for the off-chip laser source. Moreover, we take the energy/power and latency overhead values of SOAs and comb switches as discussed in Section III.A.

B. Comparison with Prior Work

We compared *SOA_LPM* with a dynamic laser power management technique (*BW_LPM*) from prior work [3]. *BW_LPM* performs a weighted time-division multiplexing of the photonic network bandwidth, and leverages the temporal fluctuations in network bandwidth to opportunistically save laser power. *BW_LPM* is designed to perform laser power management in MWMR BWGs [3]. Therefore, we focus on the Flexishare PNoC architecture with MWMR BWGs for our evaluation.

We analyzed power consumption and average packet latency for the *SOA_LPM* and *BW_LPM* techniques when they were integrated into the Flexishare PNoC architecture. For a fair comparison with *BW_LPM*, it is important to enable weighted time-division multiplexing of the network bandwidth in the Flexishare PNoC. Therefore, we enhanced the baseline Flexishare PNoC to enable weighted time-division multiplexing of the network bandwidth using token stream arbitration (TS) in its MWMR BWGs through the laser controller. We refer to this enhanced Flexishare PNoC as *Flexishare-TS*.

We implemented the *BW_LPM* technique on *Flexishare-TS,* and refer to the resulting PNoC configuration as *Flexishare-TS-BW_LPM*. Similar to the *BW_LPM* enhanced PNoC architecture presented in [3], the *Flexishare-TS-BW_LPM* configuration also has four laser sources along with a laser source controller, which is capable of switching ON/OFF the laser sources based on the executing application bandwidth requirements.

We implemented *SOA_LPM* on *Flexishare-TS* to obtain the *Flexishare-TS-SOA_LPM* PNoC configuration, and compared it with the *BW_LPM* enhanced *Flexishare-TS-BW_LPM* PNoC configuration. We also implemented *BW_LPM* and *SOA_LPM* together and show the combined benefits for the resulting *Flexishare-TS-BW_LPM-SOA_LPM* PNoC configuration.

Fig. 2 shows average latency results of this comparison study, with all results normalized with respect to the baseline Flexishare PNoC. As evident from the figure, it can be observed that on average, *Flexishare-TS-SOA_LPM* has 7.7% higher latency compared to *Flexishare-TS*. The inferior values of latency are due to the additional time/cycles required for switching of laser power using comb switches (0.5ns), for accessing the lookup table entries (1ns), and for amplification using SOAs (50ps). Further, *Flexishare-TS-SOA_LPM* has 12.8% lower average latency compared to *Flexishare-TS-BW_LPM*. The increase in average latency for *Flexishare-TS-BW_LPM* is because of additional cycles required for periodic computation of average packet latency in the *BW_LPM* technique. *Flexishare-TS-BW_LPM-SOA_LPM* also has 25.8% higher latency compared to *Flexishare-TS*. The latencies contributed by *BW_LPM* and *SOA_LPM* cumulate to render worse average latency for *Flexishare-TS-BW_LPM-SOA_LPM*.

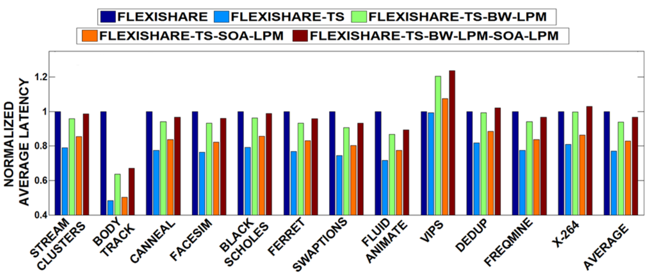

Fig. 2: Average latency for different variants of the Flexishare PNoC architecture. Results are normalized w.r.t baseline Flexishare PNoC.

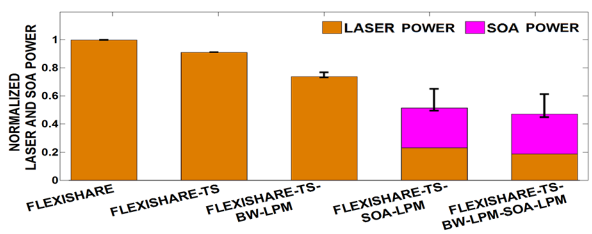

Fig. 3: Average laser and SOA power consumption comparison for different configurations of the Flexishare PNoC architecture. Results are normalized w.r.t. the baseline Flexishare PNoC architecture.

Fig. 3 presents a comparison of the laser and SOA power consumption values averaged across 12 applications of the PARSEC benchmark suite for the baseline *Flexishare, Flexishare-TS, Flexishare-TS-BW_LPM, Flexishare-TS-SOA_LPM* and *Flexishare-TS-BW_LPM-SOA_LPM* PNoC configurations. The error bars in the figure represent maximum and minimum power values across the 12 applications. *Flexishare-TS-SOA_LPM* has 50.8%, 43.4%, and 30.1% lower total power consumption compared to Flexishare baseline, *Flexishare-TS*, and *Flexishare-TS-BW_LPM* respectively. For *Flexishare-TS-SOA_LPM* and *Flexishare-TS-BW_LPM-SOA_LPM*, more than 50% of the total power consumption is due to the power overhead of SOAs. Even then, these PNoC configurations possess less total power consumption than all other variants. These results corroborate the excellent capabilities of our *SOA_LPM* technique in reducing laser power and hence total power consumption. Moreover, it can be inferred that *SOA_LPM* can be combined with other laser power management techniques such as *BW_LPM* to save even more laser power.

In *summary*, our *SOA_LPM* laser power management technique saves significant power with nominal latency overheads.

VI. Conclusions

We presented a low overhead, run-time laser power management technique called *SOA_LPM*, which made use of on-chip semiconductor amplifiers (SOA) to achieve traffic-independent and loss-aware savings in laser power consumption. Experimental analysis shows that our technique achieves 31.5% more laser power savings with 12.8% less latency overhead compared to another laser power management scheme from prior work. Thus, *SOA_LPM* represents an attractive solution to reduce laser power consumption in emerging PNoCs.

ACKNOWLEDGMENT

This research is supported by grants from SRC, NSF (CCF-1252500, CCF-1302693), and AFOSR (FA9550-13-1-0110).

References

[1] L. Zhou and A. K. Kodi, "PROBE: Prediction-based optical bandwidth scaling for energy-efficient NoCs," in *NOCS*, 2013.

[2] B. Neel et al., "Dynamic Power Reduction Techniques in On-Chip Photonic Interconnects," in *GLSVLSI*, 2015.

[3] C. Chen and A. Joshi, "Runtime Management of Laser Power in Silicon-Photonic Multibus NoC Architecture," *IEEE JSTQE,* 2013.

[4] C. Chen et al., "Managing Laser Power in Silicon-Photonic NoC Through Cache and NoC Reconfiguration," *IEEE TCAD,* 2015.

[5] Y.H. Kao and H. J. Chao, "Design of a Bufferless Photonic Clos Network-on-Chip Architecture," *IEEE TC*, 2014.

[6] Z. Wang et al., "Power Allocation Method for TDM-Based Optical Network on Chip," *IEEE PTL,* 2013.

[7] Y. Pan et al., "FlexiShare: Channel sharing for an energy-efficient nanophotonic crossbar," in *ISCA*, 2010.

[8] A. Biberman et al., "Photonic Network-on-chip Architectures Using Multilayer Deposited Silicon Materials for High-performance Chip Multiprocessors," *JETC, 2011.*

[9] A. Biberman et al., "Broadband Silicon Photonic Electrooptic Switch for Photonic Interconnection Networks," *IEEE PTL,* 2011.

[10] *Semiconductor Optical Amplifiers*, 2002 edition. Boston; London: Springer, 2002.

[11] S. V. R. Chittamuru, S. Pasricha, "Reconfigurable Silicon-Photonic Network with Improved Channel Sharing for Multicore Architectures," in *GLSVLSI, 2015.*

[12] C. Bienia et al., "The PARSEC Benchmark Suite: Characterization and Architectural Implications," in ICPACT, 2008.

[13] N. Binkert et al., "The Gem5 Simulator," *Comput Arch. News*, 2011.

[14] P. Grani and S. Bartolini, "Design Options for Optical Ring Interconnect in Future Client Devices," *JETC* 2014.

[15] K. Chen et al., "CACTI-3DD: Architecture-level modeling for 3D die-stacked DRAM main memory," in *DATE, 2012.*

[16] C. Sun et al., "DSENT - A tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling," in *NOCS, 2012.*

[17] S. Bahirat, S. Pasricha, "METEOR: Hybrid Photonic Ring-Mesh Network-on-Chip for Multicore Architectures", *ACM TECS, Mar 2014.*

[18] S. Pasricha, S. Bahirat, "OPAL: A Multi-Layer Hybrid Photonic NoC for 3D ICs", *in IEEE/ACM ASPDAC, Jan 2011.*

[19] S. Bahirat, S. Pasricha, "Exploring Hybrid Photonic Networks-on-Chip for Emerging Chip Multiprocessors", in *CODES+ISSS 2009.*

[20] S. V. R. Chittamuru, I. Thakkar, S. Pasricha, "PICO: Mitigating Heterodyne Crosstalk Due to Process Variations and Intermodulation Effects in Photonic NoCs", *in IEEE/ACM DAC, Jun. 2016.*

[21] I. Thakkar, S. V. R. Chittamuru, S. Pasricha, "A Comparative Analysis of Front-End and Back-End Compatible Silicon Photonic On-Chip Interconnects," *in ACM/IEEE SLIP, Jun 2016.*