# Dynamic Power Reduction Techniques in On-Chip Photonic Interconnects

Brian Neel
School of EECS
Ohio University
Athens, OH, 45701
bn179706@ohio.edu

Matthew Kennedy
School of EECS
Ohio University
Athens, OH, 45701
mk140409@ohio.edu

Avinash Kodi
School of EECS
Ohio University
Athens, OH, 45701
kodi@ohio.edu

## ABSTRACT

Photonic interconnects is a disruptive technology solution that can overcome the power and bandwidth limitations of traditional electrical Network-on-Chips (NoCs). However, the static power dissipated in the external laser may limit the performance of future optical NoCs by dominating the stringent network power budget. From the analysis of real benchmarks for multi-cores, it is observed that high static power is consumed due to the external laser even for low channel utilization. In this paper, we propose runtime power management techniques to reduce the magnitude of laser power consumption by tuning the network in response to actual application characteristics. We scale the number of channels available for communication based on link and buffer utilization. The performance on synthetic and real traffic (PARSEC, Splash-2) for 64-cores indicate that our proposed power scaling technique can reduce optical power by about 70% with less than 1% throughput penalty for real traffic.

## Categories and Subject Descriptors

C.1.2 [**Computer Systems Organization**]: Multiprocessors—*Interconnection architectures*

## Keywords

Interconnection Architecture, Nanophotonics, Power Reduction

## 1. INTRODUCTION

There have been several nanophotonic architectures and circuit optimizations related to reducing the external laser power. In [1], multi-bus NoC architecture between private L1 caches and distributed L2 caches which use weighted time division multiplexing to distribute the laser power across multiple buses. The proposed multi-bus NoC switches off laser sources at runtime during low bandwidth requirements to minimize the laser power consumption. In PROBE [2], multiple channels between source and destination are switched off using table-based prediction method. With a lightweight prediction technique, PROBE scales the bandwidth adaptively to the changing traffic demands while maintaining reasonable performance. This work differs from prior work as we evaluate the tree-based splitter to reduce the amount of optical intensity required by evaluating the network load. This provides a mechanism to increase or decrease the amount of optical intensity by enabling or disabling optical light into different levels of the tree splitter.

The focus of this paper is to design a power efficient reconfigurable optical power tree architecture that can consume energy in proportion to the application load. The proposed design allows the number of optical carrier signals to be reconfigured and time multiplexed across the entire optical network. Since off-chip communication to the laser is slow, the laser is designed to have a number of states corresponding to the number of optical carrier signals that it needs to provide. Then based on buffer and link utilization, a control unit will signal the laser to change states. During laser state changes, the network can still be used at the lower of the two states (current state and destination state) until the laser has stabilized, preventing significant performance degradation. This results in a system where the available bandwidth is tuned to match the network load providing a degree of energy proportionality to the optical network.


*GLSVLSI'15,* May 20–22, 2015, Pittsburgh, PA, USA.

http://dx.doi.org/10.1145/2742060.2742118.

## 2. ARCHITECTURE

### 2.1 Reconfigurable Optical Power Tree

By strategically tuning the tree ring resonators we can activate and deactivate links. We can then tune the laser source according to the amount of power required by only the active links. For example, by tuning only the ring resonator denoted $R_0^0$, we can direct all of the laser power to the left half of the power tree, i.e. only the row links. However, as the individual link optical power requirements remain constant, we can simultaneously conserve energy by tuning the laser to source approximately only half of the optical power. The power can be further reduced when only one or two active links are necessary by tuning second and third levels of ring resonators. Figure 2 presents example configurations where only a partial number of links are activated and the optical power is reduced to only half and a quarter of the original optical power respectively. The number of activated links required can be determined at runtime based on fluctuating bandwidth requirements. Since the tree is controlled by MRRs, the active channels can be quickly

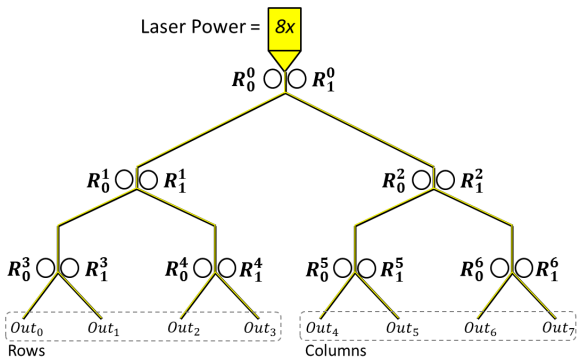


**Figure 1: The proposed optical power tree with ring notation. Without any rings activated, the optical laser power is halved at each branch of the tree. This requires approximately 8 times the power required to activate a single link. The left half of the tree corresponds to the row-based links while the right half corresponds to column-based links.**

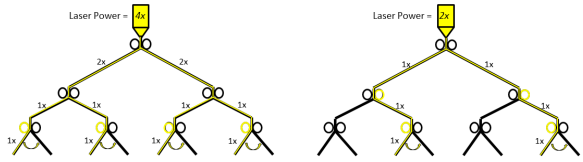


**Figure 2: Example cases showing only 4 of the 8 channels (left) and 2 of the 8 channels (right) in the tree active.**

reconfigured to any of the physical channels connected to the tree.

## 2.2 Reconfiguration Algorithm

There are a number of ways to determine when and how to change the laser states. A simple and effective approach is to set thresholds below the saturation point in each state. While the saturation point will vary across application traffic patterns, we chose fixed network load thresholds according to the worst-case traffic pattern, i.e. the pattern with the lowest saturation load. This was done to ensure adequate power is provided in each trial to avoid saturating the network. Methods could be employed to determine appropriate thresholds at runtime, however we save this for future work as it extends beyond the scope of this paper.

We define the laser power states as $S_i$, where $0 \leq i < log_2(N)$ and $N$ is the number of carrier signals. The number of threshold points required is one less then the number of laser states, where threshold count, $TC$, is $TC = log_2(N)$. We simulated eight synthetic traffic patterns (explained in Performance Evaluation section) and determined the thresholds used in our evaluation. After determining the thresholds, we estimated the network load based on aggregated buffer and link utilization measurements over a rolling reconfiguration window. We used the aggregated link utilization measurements to determine the total network load. We then compared the estimated network load to the thresholds at the end of the laser reconfiguration window, $RW_L$, to determine if a state change is necessary. If a state change is necessary, then the laser is signalled to change states and

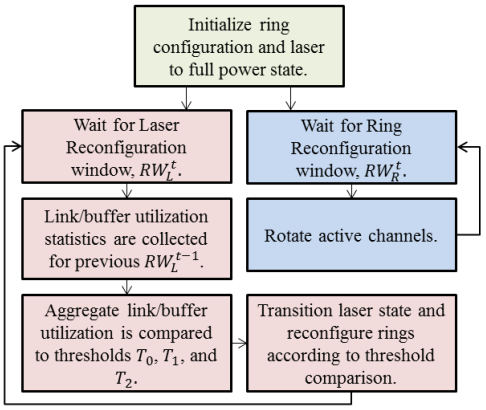


**Figure 3: Laser (red) and ring (blue) reconfiguration flow.**

wait for a response that the laser is stable at the new state. While waiting for a response back from the laser, the network can continue operating at the lower of the two states. For example, if the laser state is currently $S_2$ and we are going to transition to $S_3$, then we can continue operating in $S_2$. While moving to $S_3$, the laser will still be providing more light than what is necessary for $S_2$. However, as we transition down a state then we can immediately operate in the new state as the optical intensity will be less than what is needed for the current state, but always higher than what is needed for the new, lower state. Therefore, even during laser state transitions, the network is still active. Figure 3 depicts the laser and ring reconfiguration flow.

## 2.3 Active Channel Reconfiguration

By using MRRs to control light within the power tree it is possible to reconfigure the channels available in state $S_i$ to any of the output channels at the leaves of the tree. The ring reconfiguration window, $RW_R$, can be very small compared to the laser state reconfiguration window, $RW_L$, as tuning the resonators can be done at 10 Ghz or higher speeds.

In our simulations we set the $RW_R$ to 4, enough cycles to send a packet across each of the active channels, but only if the packet was ready to be sent in the first cycle of the new window. Otherwise, the packet must wait for the next time the channel is available. While round robin provides acceptable performance, it may be possible to further improve performance by skipping channels that do not have packets ready to send on the channel and allow another channel the use of the otherwise wasted bandwidth.

# 3. CASE STUDY

While an optical flattened butterfly has been chosen as a test case, other photonic architectures with multiple channels may also be compatible with our proposed reconfiguration mechanism. In our flattened butterfly, we combine optical transceivers and electronic switches to design a power-efficient high-performance network as shown in Figure 4. The proposed off-chip broadband light source will generate $W_N$ wavelengths, $\Lambda = \lambda_0, \lambda_1, \lambda_2, \lambda_3, ...\lambda_{W_N-1}$. By transmitting the continuous off-chip carrier signal in both x- and y-directions simultaneously, we modulate the signals at the optical transmitters. Figure 4 shows 4 cores and a shared L2 cache combined together to form a *tile*. This grouping

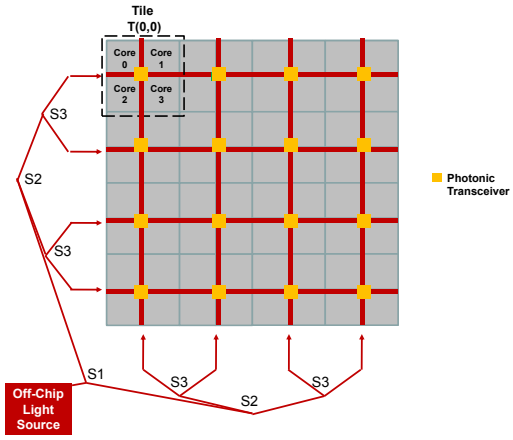


**Figure 4: Proposed layout of flattened butterfly architecture for 64 tile architecture. Four tiles are combined into a super-tile. Each waveguide is considered a single optical channel connected to the optical power tree**

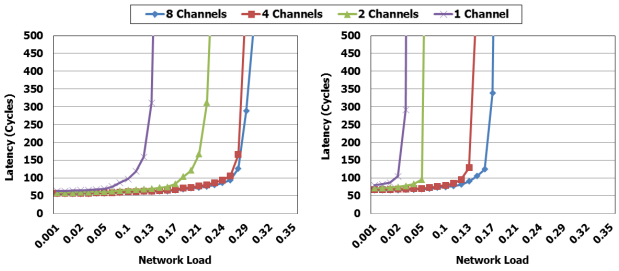


**Figure 5: Shows the saturation throughput for each state for a tree with N=3 for synthetic traffic patterns Uniform Random (left) and Complement (right).**

reduces the cost of the interconnect as every core does not require lasers attached and more importantly, facilitates local communication through cheaper electronic switching [3]. Each tile consists of dual-set (x and y) photonic transceivers and an electronic switch. Optical interconnects are used in two dimensions along the grid similar to an electronic 2D mesh or torus. It takes a maximum of 2 hops to traverse between any two tiles, one hop in the x-dimension and one hop in the y-dimension.

# 4. PERFORMANCE EVALUATION

## 4.1 Simulation Methodology

In this section we discuss our simulation methodology and assumptions leading to the results presented in the next section. The proposed optical power tree was simulated using a 64-core implementation of an optical flattened butterfly as the base network architecture. We also assume packets are composed of four 128-bit flits, requiring 4 network cycles to transmit a complete packet between two routers. With this in mind we chose a ring splitter reconfiguration period of 4 network cycles to allow the transmission of a complete network packet before deactivating the channel and reactivating a new channel. This optimizes network performance by preventing congestion of partial packets from building up in the network routers. We assume a 512 network clock cycle window for laser state reconfiguration based on Complement traffic analysis; further analysis is left for future work.

The SPLASH-2 [4], PARSEC [5], and SPEC CPU2006 [6] workloads were used to evaluate the performance of 64-core networks. The results acquired from these two implementations are included in the following section of this paper along with a comparison highlighting the energy savings achieved using the proposed reconfigurable laser architecture.

## 4.2 Results

Figure 5 shows the latency curves for each individual state for two select traffic patterns. For these simulations, the laser state is fixed and the network load is increased until the latency spikes at the saturation point. The saturation point is the point where further network demand no longer increases throughput as the network is already working at maximum capacity. Latency is plotted on the y-axis and network load (as a function of network capacity) is on the x-axis. A round robin ring reconfiguration method is used to distribute the active signal across the available channels.

Figure 5 (right) shows the latency curves for Complement traffic pattern which has the lowest saturation point for each state out of all the traffic patterns tested. Complement is used to determine the thresholds used throughout the rest of our testing as we are accounting for the worst case latency in each state. The thresholds are set at $T_0 = 0.01$, $T_1 = 0.05$, and $T_3 = 0.12$. The laser state is determined by estimating network load from buffer and link utilization and comparing it to the thresholds. Therefore, this plot is an important factor in determining the performance of the network.

Figure 6 presents the laser state transitions and network injection rate against elapsed time in network cycles for the Uniform Random synthetic traffic pattern. The network injection rate is stepped up nearly to the network saturation point for the corresponding traffic pattern, and then stepped back down to a near idle network load. The blue network load curve in the figure represents the simulation injection rate as a function of the total network capacity. As this is simply a controlled simulation parameter, this information would not normally be available to the reconfiguration controller without previous knowledge of the application traffic. Reconfiguration decisions must be made according to the estimated network load calculated using buffer and link utilization statistics collected by the laser reconfiguration controller. This calculated network load is included in the plot as the red color coded line. Only when the calculated network load crosses a threshold, denoted by the horizontal dotted lines, will the laser transition to a new power state. We can observe from the figures the laser reconfiguration reacts very rapidly to increasing network load, however, the laser tends to lower it's state at a lagging pace relative to the network injection rate. This is because the reconfiguration algorithm also accounts for buffer utilization which requires several network cycles to drain packets from the router buffers.

While the synthetic traffic simulations help to demonstrate the operation of the proposed design, real traffic simulations were also performed to exhibit the optical power savings achieved by the architecture. Figure 7 presents the laser state transitions for a real traffic trace acquired using

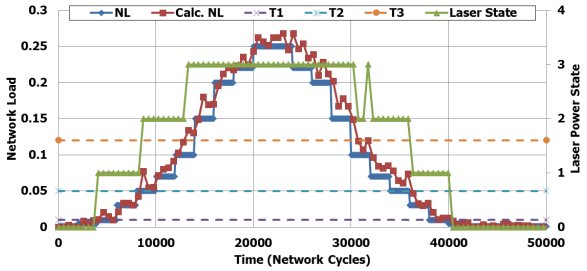


**Figure 6: Laser state transitions, network load (NL), and calculated network load (Calc. NL) versus elapsed simulation cycles using Uniform Random synthetic traffic.**

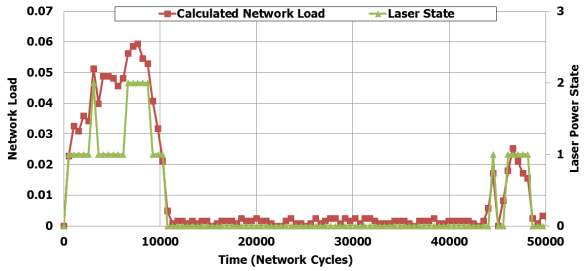


**Figure 7: Laser state transitions and calculated network load versus elapsed simulation cycles using PARSEC-Ferret traces.**

GEM5 [7]. We simulated the laser reconfigurable architecture using several real traffic traces. On average, optical power consumption was reduced by more than 70% as shown in Figure 8 with virtually no penalty to network throughput compared to the standard optical flattened butterfly architecture as shown in Figure 9.

## 5. CONCLUSIONS

This work has shown that by using the light source in lower intensity states and time multiplexing the available carrier signals, it is possible to trim the power and provide energy proportionality to the laser. An important aspect of this design is controlling the optical power tree with micro-ring resonators that can provide quick reconfiguration within a single power state. This paper just touches the surface of many aspects of the design such as light source design and reconfiguration algorithms leaving many ideas (laser control, electrical drivers) for future work. However, even with simple reconfiguration algorithms, the simulation results show significant reduction in power for real traffic benchmark traces. More than 70% reduction in power is demonstrated, while providing nearly no throughput penalty at less than 1%.

## 6. ACKNOWLEDGEMENT

This research was supported by the US National Science Foundation (NSF) awards CCF-1054339 (CAREER), ECCS-1129010, CCF-1318981, ECCS-1342657, and CCF-1420718.

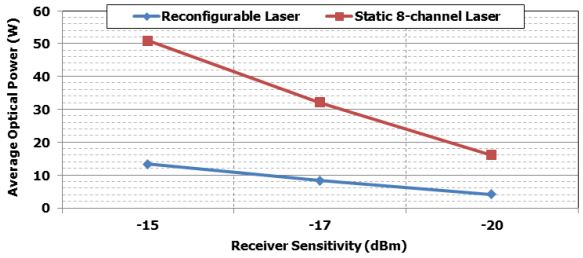


**Figure 8: Average laser power consumption using real traffic traces for reconfigurable and static 8-channel laser implementations vs. several receiver sensitivities. Power consumption was simulated for networks with optical receiver sensitivities of -15dBm (conservative), -17dBm, and -20dBm (aggressive).**

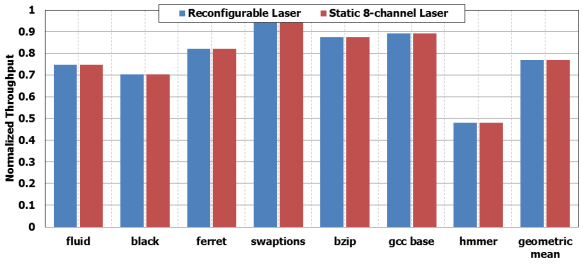


**Figure 9: Comparison of normalized network throughput using various real traffic traces for the reconfigurable laser implementation vs. static 8-channel laser implementation.**

## 7. REFERENCES

[1] C. Chen and A. Joshi, "Runtime management of laser power in silicon-photonic multibus noc architecture," *Journal of Selected Topics in Quantum Electronics*, vol. 2, March/April 2013.

[2] L. Zhou and A. K. Kodi, "Probe: Prediction-based optical bandwidth scaling for energy-efficient nocs," in *IEEE/ACM 7th International Symposium on Networks-on-Chip (NoCs)*, 2012.

[3] J. Balfour and W. J. Dally, "Design tradeoffs for tiled cmp on-chip networks," in *Proceedings of the 20th ACM International Conference on Supercomputing (ICS)*, Cairns, Australia, June 28-30 2006, pp. 187–198.

[4] S. Woo, M. Ohara, E. Torrie, J. Singh, and A. Gupta, "The splash-2 program: Characterization and methodological considerations," 1995, pp. 24–36.

[5] C. Bienia, S. Kumar, J. P. Singh, and K. Li, "The parsec benchmark suite: Characterization and architectural implications," in *Proceedings of the 17th International Conference on Parallel Architectures and Compilation Techniques*, October 2008.

[6] J. L. Henning, "Spec cpu suite growth: an historical perspective," *SIGARCH Comput. Archit. News*, vol. 35, pp. 65–68, March 2007.

[7] M. Martin, D. Sorin, B. Beckmann, M. Marty, M. Xu, A. Alameldeen, K. Moore, M. Hill, and D. Wood, "Multifacet's genreal execution-driven multiprocessor simulator (gems) toolset," *ACM SIGARCH Computer Architecture News*, no. 4, pp. 92–99, November 2005.