CLAP-NET: Bandwidth Adaptive Optical Crossbar Architecture

Matthew Kennedy and Avinash Karanth Kodi

Department of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701, United States

Abstract

While the number of processing cores placed on individual silicon dies climbs towards hundreds, and even thousands of cores, there is growing demand for efficient and scalable on-chip interconnects. Offering many advantages over metallic interconnects, nanophotonic interconnects enable new design possibilities, however, nanophotonic interconnects also require an off-chip laser source, which is often wasted to insertion losses and periods of low network activity. We present a new optical crossbar architecture that leverages capabilities of nanophotonics to provide high-performance inter-core communication while maximizing utilization of the laser power by means of dynamic bandwidth and laser power reconfiguration schemes. We compare our architecture with other proposed optical crossbar designs according to power consumption, throughput, and latency. We evaluate the network using synthetic patterns to show approximately a 13% improvement in throughput can be achieved compared to other optical crossbar designs, and a 92% improvement compared to a conventional electrical flattened butterfly architecture.

Keywords: Network-on-Chip, optical interconnects, optical crossbar, multicore, bandwidth reallocation, power reconfiguration

1. Introduction

On-chip network power consumption for multicore processors is a growing concern for computer architects as hundreds and even thousands of processing

Preprint submitted to Journal of Parallel and Distributed Computing January 15, 2016

cores [1] are being packed on to single silicon dies. While industry and consumer
 5 demand for faster and more efficient devices is constantly growing, computer ar-
 chitects must seek new technologies to enable the next generation of computing
 devices. With an already established presence in the communications industry,
 photonic technology is gaining acceptance as a potential enabler for efficient
 on-chip networks. Developments in nanophotonics promise several advantages
 10 over conventional electrical networks including low latency, high-bandwidth, and
 greater power efficiency for on-chip interconnects.

High performance, low latency topologies such as the optical crossbar, other-
 wise impractical using only electrical components, become possible with nanopho-
 tonics. Scaling an electrical crossbar network to support many cores requires
 15 N additional wires for each core added to the network, where N is the total
 number of cores. This $\mathbf{O}(N^2)$ growth in wire area overhead limits network
 scalability. Implementing crossbar topologies using nanophotonics is made pos-
 sible with innovations such as dense wavelength division multiplexing (DWDM)
 [2]. DWDM enables multiple network channels across only a single waveguide.
 20 Photonic links are also unburdened by increasing wire capacitances inherent in
 scaling electrical networks which increases signal latency and power dissipation.
 Optical link latencies are near constant regardless of traveling distance, how-
 ever, increasing the number of network participants contribute to higher optical
 insertion losses and laser power costs.

25 We present **C**lockwise/**C**ounter-Clockwise **A**daptive **P**hotonic **N**etwork (CLAP-
 NET), a unique decomposed optical crossbar architecture. The CLAP-NET
 design leverages shared waveguides to enable dynamic bandwidth allocation for
 improved network throughput, and a decomposed crossbar with reduced optical
 insertion losses and power consumption. Implementing a clockwise/counter-
 30 clockwise optical routing scheme avoids the common communication pitfall
 found in most optical crossbar designs where modulated signals must loop
 around the entire distance of waveguides, passing each tile on the network to
 reach their destination. This incurs unnecessary optical insertion losses due to
 extended communication distance and through-losses from adjacent ring mod-

ulators. Although our design requires additional waveguide area, we show reduced optical insertion losses and power consumption can be achieved without hindering network throughput. Simulations show our proposed CLAP-NET architecture with dynamic bandwidth allocation achieves a 13% improvement in network throughput versus a full optical crossbar design and a 92% improvement in throughput versus a conventional electrical network using synthetic traffic. Additionally, using a simple power reconfiguration technique we show laser power savings up to 48.9%. The major contributions highlighted in this paper are as follows:

- We introduce a unique optical crossbar network architecture, CLAP-NET, that mitigates power consumption through a decomposed crossbar design while maintaining single hop network communication.
- A dynamic bandwidth reconfiguration scheme is implemented to achieve significantly higher network throughput and improved utility of the static laser source.
- A dynamic laser power reconfiguration technique is explored to conserve power dissipation during periods of low link utilization.
- Our proposed CLAP-NET network architecture is evaluated by comparison with several competing network designs on power, network throughput, and latency metrics.

In this paper, Section 2 provides an overview of nanophotonic technology, Section 3 discusses related work in optical crossbar-based networks, Section 4 discusses the proposed CLAP-NET network architecture and reconfiguration, Section 5 evaluates power, area and performance, and Section 6 concludes the paper.

2. Nanophotonics

Developments in silicon nanophotonics have presented vast opportunities for improvements in the computing industry. Researchers are constantly finding

new methods to scale optical devices to smaller dimensions, enabling many future prospects for on-chip optical devices. Already researchers at companies
65 like IBM have successfully demonstrated on-chip optical communication links [3].

There are several necessary components to form a photonic communication link as depicted in Figure 1. Narrow optical waveguides can be fabricated by etching shallow trenches on silicon wafers. The highly reflective waveguides
70 transport photons emitted by an off-chip laser source. Waveguides can be fabricated on multiple silicon layers to reduce insertion losses from waveguide crossings and increase on-chip bandwidth [4]. Photons with varying wavelengths may be emitted from a single broadband laser or multiple single wavelength laser sources coupled to the optical waveguides using optical fibers.

On-chip ring resonators can be used to send, receive, and route photons be-
75 tween network devices. Ring resonators are small circular waveguide structures which are precisely fabricated to couple light from nearby waveguides. Changes in the refractive index of ring resonators can be induced by changing the thermal state of the resonator using ring heaters or by applying an electric field to
80 the resonator structure [5]. Having dynamic control over the resonators refractive index enables computer architects to implement dynamic optical routing schemes during runtime of a nanophotonic chip. Electro-optic modulation can be used to modulate the optical carrier signal at high data rates. Researchers have demonstrated modulators capable of effectively sending data at over 10
85 Gbps [6].

On the receiving end of photonic links, germanium photodetectors capable of receiving signals up to 40Gbps collect photons and produce an electrical output current [3]. Photodetectors with responsivities of 0.5A/W have been demonstrated [3]. Integrating the photodetector with a ring resonator allows the
90 destination device to actively listen to or ignore specific wavelength channels by thermally tuning the resonators refractive index. A transimpedance amplifier, or TIA, is used to convert the output of the germanium photodetector to a voltage output. Limiting amplifiers are used to normalize the voltage signal before

Figure 1: A basic nanophotonic link consists of an off-chip laser coupled to an on-chip silicon waveguide. Electrically controlled ring resonators are used to modulate the carrier signal. The optical signal is detected using Germanium photodetectors and is converted back to a digital signal through a series of amplifiers.

output as a usable signal by digital circuitry such as a router or switch [3]. Ring heaters are used at both transmitter and receiver circuitry to ensure stabilization of the ring resonant wavelength to avoid any drift due to thermal variations [7]. The low-energy and high-bandwidth characteristics makes nanophotonic links a promising solution for future many-core interconnects [8, 9].

3. Related Works

Many optical network designs have been proposed [10, 11, 12, 13, 14, 15]. PROPEL [10], a two-hop network leverages nanophotonics to implement a power-efficient interconnect. Another architecture, 3D-NoC [11], explores implementing an optical interconnect across multiple nanophotonic layers. However, there has been significant interest focused towards nanophotonic crossbars for their ultra-low-latency, single-hop communication capabilities. One such design, Corona, winds bundles of waveguides past each tile on the network, dedicating a bundle for receiving messages from all other tiles [14]. This provides high dedicated bandwidth and low latency traversal, however due to the unidirectional nature of the laser power, messages frequently need to lap around the entire chip depending on the location of the source and destination tiles. Not only does the required laser power scale directly with every tile added to the network, but bandwidth may also be wasted during inactive periods for individual tiles.

Another topology, Firefly, reduces the optical insertion losses and power

consumption by reducing the number of writers and readers allowed access to
 115 signals transmitted through the waveguides [15]. This is made possible with
 a two tiered network. Messages are required to first traverse electrical links
 between adjacent local tiles before (or after) the message is transmitted across
 the global optical crossbar. Tiering the network reduces the load on the global
 crossbar by routing short distance traffic across nearby electrical links. Limiting
 120 only a single writer from each quadrant of the chip to access each waveguide
 also reduces the number of modulators and optical insertion losses, resulting
 in reduced laser power consumption. Although optical power losses and global
 crossbar contention is reduced, transmission latency is increased as extra hops
 across electrical links may be necessary to reach the destination tile.

125 A common disadvantage among optical crossbar topologies is high static
 laser power consumption. Because channel wavelengths are assumed to be gen-
 erated by off-chip laser sources, reducing laser power at runtime can be difficult,
 and static laser power is simply wasted during periods of low network activity.
 Attempts to maximize static laser power utility through dynamic bandwidth
 130 allocation and wavelength stealing have been proposed [11, 16]. Bandwidth can
 easily be reallocated using nanophotonics as wavelengths can be reassigned to
 network tiles for reading and writing by retuning appropriate ring resonators.
 Several papers have also proposed dynamically toggling laser sources at runtime
 to mitigate unnecessary laser power consumption [17, 18].

135 4. CLAP-NET Architecture

The architecture proposed in this paper will maintain the low latency bene-
 fits typically found in optical crossbars while also attempting to mitigate many
 of the drawbacks such as poor scalability due to insertion losses and wasted
 static laser power. The design leverages shared optical waveguides to have
 140 the flexibility to shift bandwidth between tiles. Two optical layers are used
 to avoid waveguide crossing insertion losses at larger network scales. A new
 clockwise/counter-clockwise optical routing method is also explored to ensure

Figure 2: (a) The network connects a 64 core grid using shared waveguide bundles. Each network routes laser power in either a clockwise or counter-clockwise fashion. (b) Each tile consists of a concentration of four cores along with individual L1 caches, a shared L2 cache, and a router connecting the tile to the optical network. (c) Individual wavelengths are 'peeled' from the left side common waveguides onto the assigned modulating group waveguide. (d) Ring resonators are tuned to modulate and detect designated wavelengths on the shared waveguides according to the bandwidth allocation algorithm.

full communication capability exists between all tiles on the network.

Table 1: Layout Comparison

	Corona-64	Firefly-64	3D-NoC-64	PROPEL-64	CLAP-NET-64
Waveguide Distance (mm)	6175	2880	5266	3200	5920
MRRs	64k	24k	40k	32k	104k
Total Optical Loss	39.60dB	28.59dB	29.54dB	21.25dB	30.86dB
Network Diameter	1	3	1	2	1
Bisection Bandwidth	2.5TB/s	2.5TB/s	2.5TB/s	1.25TB/s	2.5TB/s

4.1. Data Network

145 The proposed architecture as shown in Figure 2(a)(b) is for a 64 core, 16
 tile system implementation on a 400mm^2 chip with four cores concentrated
 to a single tile. As seen in previous nanophotonic crossbars, a laser source is
 typically coupled to a bundle of looping waveguides that snake by all tiles on the
 chip. Due to the unidirectional nature of the laser source, this usually requires
 150 modulated signals to travel the entire distance of the waveguide loops, incurring
 significant insertion losses from passing ring resonators, splitters, and waveguide
 cladding. By splitting the full optical crossbar into smaller loops, we can reduce
 the waveguide and ring resonator insertion losses incurred by communication
 signals between the source and destination tiles. The contention rate for sending
 155 data to a target tile is also reduced, however, this benefit is a trade-off in the
 router radix as each crossbar a tile is affiliated must be accommodated for
 with an additional network input port as shown in Figure 3. Although there
 are many ways to partition the crossbar, we have chosen to separate the tiles

into row-based groups, where each row of tiles is logically connected to every
 160 other row by means of 4x4 optical crossbars. Because each group is also able
 to connect to tiles within itself through an intra-group crossbar, the resulting
 network topology consists of 16 total logical crossbars.

In order to facilitate communication between tiles in a manner that also flex-
 ibly supports the reallocation of bandwidth, shared waveguides are used instead
 165 of receiver or sender dedicated waveguides. As shown in Figure 2(c), common
 power waveguides with broadband laser inputs run vertically across the left
 side of the chip. Any combination of wavelengths may be 'peeled' off from the
 common power waveguides using dynamically tuned ring resonators to source
 outgoing bandwidth to each of the row-groups. The number of wavelengths
 170 provided to each group may vary according to bandwidth demand and is al-
 located using the Control Network discussed in the next section. Wavelengths
 may be modulated by tiles of each group along the bidirectional waveguides
 (Figure 2(d)) according to the token-based arbitration scheme implemented as
 part of the control network. Ring resonators along the right side of the chip are
 175 tuned according to the destination tiles of each wavelength channel as shown in
 Figure 2(e). Because of this, tiles may not send and receive signals on the same
 wavelength channel and network waveguide to avoid interfering signals.

Clockwise and counter-clockwise network pairs are utilized to successfully
 facilitate communication between any two tiles on the network. This is because
 180 the common waveguides must be avoided when routing messages between tiles
 as the raw laser power would wipe over any modulated signals. Consequently,
 signals cannot complete a full loop, e.g. the third tile in row-group 3 cannot use
 the clockwise network to send a message to the second tile in the group without
 the signal being lost to the laser power in the common waveguide. By intro-
 185 ducing the counter-clockwise network the path between the tiles is mirrored,
 providing a channel for the third tile to send messages to the second tile.

For the reader's convenience we denote the tiles using cartesian coordinates.
 The tiles discussed in the following examples are represented in this text by their
 row and column coordinates, i.e (x,y) where $0 \leq x \leq 3$ and $0 \leq y \leq 3$. Figure

190 4(a) depicts the optical paths of a sample **intra-group** network transaction
between tiles (1,3) and (2,3) of the CLAP-NET network architecture. In the
example provided either tile (1,3) or tile (2,3) of the group could have initiated
the transaction with a request, but the response is always echoed on the network
opposite of the request. The actual wavelengths used to carry the messages may
195 be the same, as the signals are conveyed on two separate halves of the network
pair, i.e. the clockwise network and the counter-clockwise network. Figure 4(b)
depicts the optical paths of a sample **inter-group** network transaction between
tiles (2,1) and (1,1) of the second and third row-groups respectively.

While the network is designed to support reconfigurable bandwidth alloca-
200 tion through dynamic wavelength assignment, a static communication schedule
must exist to ensure communication is always possible. Only one CLAP-NET
network pair must be dedicated to static communication wavelength assign-
ments. The row-group is sourced wavelengths λ_{0-15} of the 64 wavelengths
available for sending messages. Wavelengths λ_{0-3} are dedicated to intra-group
205 communication, where the first tile in the row-group always receives on λ_0 , the
second tile receives on λ_1 , and so on. Wavelengths λ_{4-15} are reserved for inter-
group communication from the first row-group. The rest of the wavelengths,
 λ_{16-63} are split amongst the other three row-groups respectively for sending
messages. If two tiles from the same group contend for the same destination
210 wavelength, their transmissions are regulated according to the token slot arbi-
tration scheme discussed in the Control Network section of this paper.

Figure 3: Router Microarchitecture.

Figure 4: Optical signal paths for sample (a) intra-group and (b) inter-group network transactions.

4.2. Control Network

In this section an optical control network is proposed for managing network arbitration and bandwidth reallocation. The major advantage of utilizing shared waveguides across the network instead of dedicated point to point links is having the ability to dynamically allocate bandwidth according to trends in application traffic and maximize utility of the static laser source. A means to quickly collect network link utilization statistics and signal retuning of resonators at the pace of the data network must be implemented to avoid bottlenecking network performance. The CLAP-NET network architecture accomplishes this by layering

wavelength assignments as determined by the bandwidth allocation algorithm.

4.2.1. Bandwidth Reconfiguration

System performance can vary depending on application traffic patterns. While some applications may generate balanced traffic loads between many
225 cores, others may have concentrated traffic patterns between only a few cores. Using shared waveguides to transport messages between tiles enables the data network to shift bandwidth to priority tiles simply by rerouting wavelengths and retuning ring resonators to ignore or actively listen on certain wavelength chan-
230 nels. The bandwidth allocation algorithm proposed for R-3D-NoC is adopted for use with the CLAP-NET architecture [11]. In this algorithm a reconfiguration controller in each group requests link and buffer utilization statistics over a reconfiguration period. Each reconfiguration controller then classifies the groups' links as being 'not-utilized', 'under-utilized', 'normal-utilized', or
235 'over-utilized' where link utilization is 0%, less than 25%, between 25% and 50%, and greater than 50% respectively. Link utilization is calculated as the ratio of flits sent to the number of cycles in the reconfiguration window. The reconfiguration controllers share link classification information with the other reconfiguration controllers corresponding to 90%, 50%, 25%, and 0% available
240 bandwidth. For example, a tile that sends a flit 30% of the cycles over the duration of the reconfiguration window would be classified as 'normal-utilized' and would be willing to release 25% of its bandwidth to any requesting tiles. These requests for available bandwidth from other controllers are relayed across the network to acknowledge the bandwidth shift and the appropriate ring res-
245 onators are retuned according to the new bandwidth allocation. No more than 90% of a tile's bandwidth may be reallocated to avoid completely orphaning a tile from the rest of the network.

Figure 5: Flow chart detailing the reconfiguration algorithm adopted from 3D-NoC.

4.2.2. Power Reconfiguration

While bandwidth reconfiguration improves utility of laser power by shifting
250 bandwidth to high demand links, bandwidth reconfiguration does little to improve laser utility during periods of low network demand where all links have excess bandwidth. In this state, the ideal adjustment would be to simply turn off optical channels at the laser source. However, this is no simple task as there are several complications. If a broadband laser is generating the carrier signal,
255 it may be impossible to disable specific optical channels without turning off the entire laser and disabling all channels. Perhaps, an array of narrowband lasers could potentially be used instead which could allow individual channels to be disabled, yet there is still the challenge of communicating to the off-chip lasers which channels should be toggled, as this must be done very quickly to avoid
260 depriving the network of bandwidth. Instead, we propose an approach similar to the design discussed in [19] where laser power is distributed to links in time

Figure 6: CLAP-NET with adaptive reconfigurable laser power distribution.

slots by actively tuning broadband ring resonators at run-time. Figure 6 show the CLAP-NET layout with reconfigurable laser power. We can reduce the laser power consumption of CLAP-NET by using a single laser source and alternating
 265 power between the clockwise and counter-clockwise networks. When the network is under high demand, the laser power is tuned to support both networks. Under low network demand, the laser power is reduced such that just enough optical power is provided to support one network. The power is then alternated every 4 clock cycles between the networks such that neither network is starved
 270 of power. A 4 clock cycle period was chosen as this is the time required to send a typical packet. Figure 7 shows a flowchart of the adapted power reconfiguration algorithm.

This power reconfiguration method can be implemented alongside bandwidth reconfiguration, adding an additional layer of laser power management.
 275 The benefits of this method, much like the adaptive bandwidth reconfiguration, depends heavily on application traffic characteristics. In the best case, required laser intensity could be reduced by 50% for low activity networks, however this does not imply a 50% reduction in power consumption as the laser output is not linearly proportional to laser input power.

Figure 7: Flow chart detailing the power reconfiguration algorithm adopted from [19].

280 4.2.3. Network Arbitration

We adopt a destination-based optical token slot arbitration scheme similar to the one used in Corona [14]. However, as the CLAP-NET architecture utilizes several split crossbars, a similarly decomposed arbitration scheme must also be implemented. Each row-group requires its own arbitration mechanism

285 to prevent tiles from simultaneously transmitting to the same destination tile and modulating the same wavelength. Just like the data network, instead of a single full loop for optical token arbitration, we propose four isolated arbitration loops for the 64 core CLAP-NET system. As shown in Figure 9 optical tokens are propagated through the loops and may be coupled from the waveguides to claim a destination link on the data network. Once the optical token is

290 secured, a bandwidth allocation table is used to direct flits to the appropriate modulation circuits for deserialization. Figure 8 presents the static wavelength allocation table. Each group is allocated four wavelengths for sending packets to destination nodes. Nodes must contend within their group for write access

295 to the corresponding wavelengths.

Figure 8: (a) Table showing static wavelength allocation amongst groups. (b) Table showing assignments of the first four wavelengths to corresponding destination nodes. (c) Logical topology diagrams for intra- and inter-group communication.

4.3. Network Scaling

We assume 64 wavelength channels may be modulated and received at 10Gbps effectively on a single waveguide. This limits a single clockwise/counter-clockwise network pair to only 1280Gbps of total available bandwidth. Bandwidth can be increased by using 3D-stacking to layer additional network pairs without introducing optical power losses from waveguide crossings. Only two optical layers are necessary to scale the network to 32 network pairs, providing up to 5TB/s of network bandwidth or 640Gbps per core.

The network could be scaled to support a larger number of cores by increasing the number of groups, and group to group crossbars. For example a 256 core network may be split into 8 row groups creating a total of 64 logical crossbars. Maintaining the same amount of per-node-bandwidth as achieved with

Figure 9: The optical token arbitration scheme for a single row of tiles. Ring resonators couple optical tokens from the arbitration waveguide to resolve network contention between tiles of the same row. Separate arbitration networks must accompany each row of the network. An optical token for every possible tile destination is propagated through each arbitration network.

the 64 core design would scale the number of required waveguides linearly with respect to the number of cores, i.e. 4 times the number of waveguides would be needed for the 256 core design. We could expect the number of waveguides in other competing designs to scale linearly with respect to the number of cores also, however the CLAP-NET routing scheme mitigates the incurred optical power insertion losses as signals are not required to travel the entire length of additional waveguides.

5. Performance Evaluation

5.1. Power Analysis

Network power consumption was modeled using MIT’s DSENT [24] and compared to several competing architectures. Several assumptions were made to provide a balanced comparison. All networks are assumed to be implemented using 3D-stacking, reducing insertion losses from potential waveguide crossings. The number of waveguides and ring resonators used in each design is scaled to match a common total network bandwidth of 5TB/s. Table 2 contains the optical device parameters assumed in the power analysis [20], [21], [22], [23].

Table 2: Optical Device Parameters [20], [21], [22], [23].

Parameter	Value
Waveguide Loss	1.0 dB/cm
Pass-by Ring Resonator Loss	0.0001 dB
Drop Ring Resonator Loss	1.0 dB
Waveguide Crossing Loss	0.05 dB
Receiver Sensitivity	-17 dBm
Laser Efficiency	15%
Ring Heating Power	26 μ W/ring
Ring Modulating Power	500 μ W/ring
Ring Modulation Frequency	10 Ghz
Wavelengths/Waveguide	64
Waveguide Pitch	4 μ m
Ring Area	100 μ m ²

Figure 10: Required laser power for various laser efficiencies.

A breakdown of per bit energy consumption between the proposed CLAP-
 325 NET architecture, Corona [14], Firefly [15], 3D-NoC [11], and PROPEL [21] has
 been charted in Figure 11. This was estimated using the following formula:

$$Energy/bit = TotalEnergy/TotalNetworkBandwidthPerSec$$

where *TotalEnergy* is calculated as the sum of the total router leakage and dynamic power, total ring modulating and heating power, and the laser power over a 1 second duration. We assumed the worst-case scenario of max utilization of all components in our energy/bit model. We estimate required laser power of around 5.19 W for the CLAP-NET architecture. Assuming a more conservative waveguide cladding loss of 2 dB/cm, we estimate a required laser power of 21.90 W. We have also included an analysis of required laser power over various laser efficiencies in Figure 10. Our CLAP-NET network architecture requires less laser power than Corona, but requires more than the other other three networks. CLAP-NET is able to achieve lower laser power compared to Corona because the signal propagation distance is significantly shorter. Unlike Corona, where the signal may need to propagate up to twice around the chip due to the unidirectional nature of the laser carrier signal, CLAP-NET signals have short propagation distances between communicating tiles. Having short propagation distances reduces the signal insertion losses due to waveguide cladding and requires less laser power. Compared to Firefly, there is a latency trade-off where more power is sacrificed for a lower diameter network with one hop instead of three. PROPEL also requires up to two hops between source and destination tiles. Our power consumption is fairly close to the power consumption of 3D-NoC. The excess power consumption for the CLAP-NET architecture may be partially attributed to significantly higher ring heating power. Decreasing the number of ring resonators may be possible by assuming lower flexibility in dynamic bandwidth allocation. 25% of the total network bandwidth can be allocated to a single group-to-tile link under the network assumptions of this paper, but realistically the total network bandwidth needed by a single tile may be drastically lower.

Dynamic power reconfiguration as discussed in the Power Reconfiguration section was also simulated for several real traffic traces. The CLAP-NET architecture with dynamic power was compared to vanilla CLAP-NET with laser power statically set to support both, the clockwise network and the counter-clockwise network simultaneously. When the laser power is throttled to state

Figure 11: Energy per bit comparison of different networks in nanojoules.

0, only the clockwise network or the counter-clockwise network is powered at a time. The power is toggled between the two networks every 4 clock cycles to prevent link power starvation. The average link utilization is calculated for periods of 4000 clock cycles and compared to a threshold value. The threshold value was determined by averaging link utilization for several applications. More advanced methods for determining the reconfiguration threshold may be implemented, however we do not consider these methods in this paper. Figures 12-15 show plots of average link utilization and laser state over the duration of an application execution period. The bar charts next to each plot show a comparison of the average laser power consumption for the dynamic power reconfigurable CLAP-NET and vanilla CLAP-NET without any power reconfiguration, i.e. full laser power. The simulation results show power savings up to 48.9% for the SPEC CPU2006 GCC benchmark where average link utilization remains low throughout most of the application execution.

5.2. Throughput and Latency

Network saturation throughputs and latencies were simulated using the cycle-accurate OPTISIM simulator [25]. The simulator was warmed up under load without taking measurements until steady state was reached. Then a sample of injected packets was labeled during a measurement interval. The simulation was allowed to run until all the labeled packets reached their destinations. All designs were tested with different synthetic traffic patterns such as

Figure 12: Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPLASH-2 Barnes traffic traces.

Figure 13: Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPEC CPU2006 GCC traffic traces.

uniform random, bit reversal, butterfly, matrix transpose, complement, perfect
 380 shuffle and neighbor traffic pattern for the network under test. We compared
 the network performance to Corona, Flattened Butterfly, Concentrated mesh
 (CMesh), CLAP-NET and reconfigured R-CLAP-NET networks.

Figure 20 shows a comparison of network throughputs for several topolo-
 gies, including our proposed bandwidth-reconfigurable CLAP-NET topology,
 385 standard CLAP-NET, Corona, an electrical concentrated flattened butterfly,
 and an electrical concentrated mesh using various traffic patterns. For uniform
 random traffic, performance is improved marginally by R-CLAP-NET by tak-

Figure 14: Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using PARSEC Freemine traffic traces.

Figure 15: Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPLASH-2 LU traffic traces.

ing advantage of the phases in which communication between cores is helped by reconfiguration. For permutation traffic such as complement, butterfly, matrix transpose, the performance gains due to reconfiguration is more impressive. The communication patterns are static and the reconfiguration algorithm can take advantage of the traffic pattern. The geometric mean of the throughputs for all traffic patterns shows approximately a 13% improvement in throughput is achieved with the bandwidth-reconfigurable CLAP-NET topology over the standard CLAP-NET topology and Corona. R-CLAP-NET demonstrates a 92% throughput improvement over the conventional electrical flattened but-

terfly network. As most real applications have phases in which the behavior resembles the synthetic traffic pattern, we expect similar performance.

Figures 16, 17, 18, and 19 shows average network packet latencies against
400 network congestion rates for several traffic patterns. The bandwidth-reconfigurable
R-CLAP-NET network exhibits slightly higher traffic congestion tolerances compared to the standard CLAP-NET and Corona topologies. For each traffic simulation R-CLAP-NET demonstrates modest performance gains in network latency over the standard CLAP-NET architecture. The network saturation
405 point is also increased in each case. Significant improvements to network performance are noticeable especially against the Complement traffic pattern. The Perfect Shuffle and Butterfly latency curves for the R-CLAP-NET network show abnormal variations at higher injection rates. This is likely due to fluctuations in the adaptive algorithms allocation of bandwidth. The average R-CLAP-NET
410 network latency declines for the Butterfly and Complement traffic patterns for injection rates of 0.03 to 0.06. This is likely due to the reconfiguration process allocating more bandwidth to high-demand links, reducing the overall average packet latency. Injection rates lower than 0.03 may not provide enough network traffic to trigger bandwidth reconfiguration.

415 Network throughput simulations were also conducted using PARSEC [26], SPEC [27], and SPLASH2 [28] benchmark traces. The results are presented in Figure 21 including the geometric mean. We chose the geometric mean to summarize the throughput results as outlier simulations will not significantly skew the final results. Simulating with real benchmark traces did not yield significant
420 throughput gains for R-CLAP-NET compared to CLAP-NET or Corona. This is because, unlike the synthetic traffic trials, the real application traffic did not saturate the network and so bandwidth was not a limiting factor in network performance. However, the (R-)CLAP-NET architectures performed on par with Corona and better than Firefly with respect to network throughput.

Figure 16: Average network latencies in cycles for Uniform traffic pattern at varying packet injection rates (flits/cycle).

Figure 17: Average network latencies in cycles for Perfect Shuffle traffic pattern at varying packet injection rates (flits/cycle).

Figure 18: Average network latencies in cycles for Butterfly traffic pattern at varying packet injection rates (flits/cycle).

Figure 19: Average network latencies in cycles for Complement traffic pattern at varying packet injection rates (flits/cycle).

Figure 20: Comparison of network throughputs at saturation for various traffic patterns. The geometric mean is shown in the last column.

Figure 21: Comparison of network throughputs simulated using PARSEC, SPEC, and SPLASH2 benchmark traces. The geometric mean is shown in the last column.

425 6. Conclusions

In this paper, we explore a new approach to implementing nanophotonic crossbars with the goal of maximizing laser power utility. By utilizing shared waveguides we are able to flexibly adjust bandwidth allocation according to trends in application traffic. The R-CLAP-NET network is capable of delivering high network throughput performance by maximizing utility of the static laser power source through dynamic bandwidth reconfiguration and also by implementing a dynamic laser power reduction technique. Network scalability is also improved by splitting crossbar functionality into multiple smaller crossbars with split arbitration for reduced network contention. Signal travel distance is shortened through the use of clockwise and counter-clockwise network pairs, resulting in lower optical insertion losses and increased power efficiency compared to competing architectures.

Acknowledgment

The authors would like to thank the anonymous reviewers for their feedback and Brian Neel for his insight into optical interconnects for networks-on-chip. This research was supported by NSF awards CCF-1054339 (CAREER), ECCS-1129010, ECCS-1342657, CNS-1318981.

References

- [1] S. Borkar, Thousand core chips: A technology perspective, in: Proceedings of the 44th Annual Design Automation Conference, DAC '07, ACM, New York, NY, USA, 2007, pp. 746–749.
- [2] C. Brackett, Dense wavelength division multiplexing networks: principles and applications, Selected Areas in Communications, IEEE Journal on 8 (6) (1990) 948–964.

- 450 [3] H. Pan, S. Assefa, W. M. J. Green, D. M. Kuchta, C. L. Schow, A. V. Rylyakov, B. G. Lee, C. W. Baks, S. M. Shank, Y. A. Vlasov, High-speed receiver based on waveguide germanium photodetector wire-bonded to 90nm soi cmos amplifier, *Opt. Express* 20 (16) (2012) 18145–18155.
- [4] A. Biberman, K. Preston, G. Hendry, N. Sherwood-Droz, J. Chan, J. S. Levy, M. Lipson, K. Bergman, Photonic network-on-chip architectures using multilayer deposited silicon materials for high-performance chip multi-
455 processors, *ACM Journal on Emerging Technologies in Computing Systems* 7 (2011) 1–25.
- [5] K. Padmaraju, J. Chan, L. Chen, M. Lipson, K. Bergman, Thermal stabilization of a microring modulator using feedback control, *Opt. Express*
460 20 (27) (2012) 27999–28008.
- [6] Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, M. Lipson, 12.5 gbit/s carrier-injection-based silicon micro-ring silicon modulators, *Opt. Express* 15 (2) (2007) 430–436.
- 465 [7] M. Georgas, J. Leu, B. Moss, C. Sun, V. Stojanovic, Addressing link-level design tradeoffs for integrated photonic interconnects, in: *CICC*, 2011, pp. 1–8.
- [8] A. Shacham, K. Bergman, S. Member, L. P. Carloni, Photonic networks-onchip for future generations of chip multiprocessors, *IEEE Trans. Computing* (2008) 1260.
470
- [9] R. G. Beausoleil, P. J. Kuekes, G. S. Snider, S.-Y. Wang, R. S. Williams, Nanoelectronic and nanophotonic interconnect, *Proceedings of the IEEE* 96 (2) (2008) 230–247.
- 475 [10] A. Kodi, R. Morris, Design of a scalable nanophotonic interconnect for future multicores, in: *Proceedings of the 5th ACM/IEEE Symposium on Architectures for Networking and Communications Systems, ANCS '09*, ACM, New York, NY, USA, 2009, pp. 113–122.

- [11] R. Morris, A. Kodi, A. Louri, 3d-noc: Reconfigurable 3d photonic on-chip interconnect for multicores, in: Computer Design (ICCD), 2012 IEEE 30th International Conference on, 2012, pp. 413–418.
- [12] A. Joshi, C. Batten, Y.-J. Kwon, S. Beamer, I. Shamim, K. Asanovic, V. Stojanovic, Silicon-photonic clos networks for global on-chip communication, in: Proceedings of the 2009 3rd ACM/IEEE International Symposium on Networks-on-Chip, NOCS '09, IEEE Computer Society, Washington, DC, USA, 2009, pp. 124–133.
- [13] A. V. Krishnamoorthy, R. Ho, X. Zheng, H. Schwetman, J. Lexau, P. Koka, G. Li, I. Shubin, J. E. Cunningham, Computer systems based on silicon photonic interconnects, in: Proceedings of the IEEE, Vol. 97, 2009, pp. 1337–1361.
- [14] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N. P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R. G. Beausoleil, J. H. Ahn, Corona: System implications of emerging nanophotonic technology, in: Proceedings of the 35th Annual International Symposium on Computer Architecture, ISCA '08, IEEE Computer Society, Washington, DC, USA, 2008, pp. 153–164.
- [15] Y. Pan, P. Kumar, J. Kim, G. Memik, Y. Zhang, A. Choudhary, Firefly: Illuminating future network-on-chip with nanophotonics, in: Proceedings of the 36th Annual International Symposium on Computer Architecture, ISCA '09, ACM, New York, NY, USA, 2009, pp. 429–440.
- [16] A. Zulfiqar, P. Koka, H. Schwetman, M. Lipasti, X. Zheng, A. Krishnamoorthy, Wavelength stealing: An opportunistic approach to channel sharing in multi-chip photonic interconnects, in: Proceedings of the 46th Annual IEEE/ACM International Symposium on Microarchitecture, MICRO-46, ACM, New York, NY, USA, 2013, pp. 222–233.
- [17] C. Chen, A. Joshi, Runtime management of laser power in silicon-photonic

multibus noc architecture, Selected Topics in Quantum Electronics, IEEE Journal of 19 (2) (2013) 3700713–3700713.

- [18] Y. Demir, N. Hardavellas, Ecolaser: An adaptive laser control for energy-efficient on-chip photonic interconnects, in: Proceedings of the 2014 International Symposium on Low Power Electronics and Design, ISLPED '14, 510 ACM, New York, NY, USA, 2014, pp. 3–8.
- [19] B. Neel, M. Kennedy, A. Kodi, Dynamic power reduction techniques in on-chip photonic interconnects, in: Proceedings of the 25th Edition on Great Lakes Symposium on VLSI, GLSVLSI '15, ACM, New York, NY, USA, 515 2015, pp. 249–252.
- [20] V. Stojanovic, A. Joshi, C. Batten, Y.-J. Kwon, K. Asanovic, Manycore processor networks with monolithic integrated cmos photonics, Conference on Lasers and Electro-Optics.
- [21] R. Morris, E. Jolley, A. Kodi, Extending the performance and energy-efficiency of shared memory multicores with nanophotonic technology, Parallel and Distributed Systems, IEEE Transactions on 25 (1) (2014) 83–92. 520
- [22] V. Stojanovic, A. Joshi, C. Batten, Y.-J. Kwon, K. Asanovic, Manycore processor networks with monolithic integrated cmos photonics, in: Conference on Lasers and Electro-Optics/International Quantum Electronics Conference, Optical Society of America, 2009, p. CTuC3. 525
- [23] G. Kurian, C. Sun, C.-H. O. Chen, J. E. Miller, J. Michel, L. Wei, D. A. Antoniadis, L.-S. Peh, L. Kimerling, V. Stojanovic, A. Agarwal, Cross-layer energy and performance evaluation of a nanophotonic manycore processor system using real application workloads, in: Proceedings of the 2012 IEEE 530 26th International Parallel and Distributed Processing Symposium, IPDPS '12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 1117–1130.
- [24] C. Sun, C.-H. O. Chen, G. Kurian, L. Wei, J. Miller, A. Agarwal, L.-S. Peh, V. Stojanovic, Dsent - a tool connecting emerging photonics with electron-

- ics for opto-electronic networks-on-chip modeling., 2012 IEEE/ACM Sixth
535 International Symposium on Networks-on-Chip (n.d).
- [25] A. Kodi, A. Louri, A system simulation methodology of optical intercon-
nects for high-performance computing (hpc) systems, *Journal of Optical
Networking* 6 (2007) 1282–1300.
- [26] C. Bienia, Benchmarking modern multiprocessors, Ph.D. thesis, Princeton
540 University (January 2011).
- [27] J. L. Henning, Spec cpu2006 benchmark descriptions, *SIGARCH Comput.
Archit. News* 34 (4) (2006) 1–17.
- [28] S. C. Woo, M. Ohara, E. Torrie, J. P. Singh, A. Gupta, The splash-2
programs: Characterization and methodological considerations, in: *Pro-
ceedings of the 22Nd Annual International Symposium on Computer Ar-
545 chitecture, ISCA '95*, ACM, New York, NY, USA, 1995, pp. 24–36.