Abstract—Future processors are anticipated to have hundreds or even thousands of processing cores placed entirely on a single silicon chip. The increasing number of cores placed on a single chip presents new challenges, pushing researchers to explore opportunities in emerging technologies such as on-chip silicon nanophotonics. Implications of nanophotonic technology have created a unique landscape for new interconnect designs. Among the many architectures made possible by nanophotonics, there has been notable interest in crossbar topologies that were previously impractical using only electrical components. In this paper, we present a new nanophotonic crossbar interconnect architecture with the aim of retaining the low latency, single-hop characteristic of the crossbar topology, while also improving the networks utility of the static laser source which is often wasted to insertion losses and unused bandwidth. We compare our architecture design to other proposed architectures according to area, power consumption, throughput, and latency. Approximately a 13% improvement in throughput is achieved compared to other optical crossbar topologies and a 92% improvement is achieved compared to a conventional electrical flattened butterfly topology on synthetic traffic patterns.

I. INTRODUCTION

While dual-core, quad-core, and even octa-core processors have become common in consumer devices, computer architects are already anticipating the challenges of many-core processors scaling up to thousands of cores per chip [1]. As clock scaling is pushed to its limitations, and the industry struggles to maintain the rise in computing performance as predicted by Moore’s law, researchers have turned to emerging technologies to propel computing to new levels. Among these technologies, nanophotonic interconnects have garnered considerable attention from academic and industry researchers [2]–[7]. Increasing the number of cores placed on a single chip to hundreds or thousands requires a scalable and efficient interconnect architecture to facilitate inter-cache and memory communication. Compared to electrical interconnects, optical interconnects offer several advantages including low latency, high-bandwidth, and low power transmission over extended distances [8]. These characteristics make nanophotonics an ideal solution for the scalability challenges faced by interconnect designers. Narrow optical waveguides can be fabricated by etching shallow trenches on silicon wafers. The highly reflective waveguides transport photons emitted by an off-chip laser source. 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. 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 [9] as shown in Figure 1.

In this paper we propose Clockwise/Counter-Clockwise (CW/CCW), a unique decomposed optical crossbar architecture. The CW/CCW 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-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 modulators. 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 CW/CCW 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. The major contributions highlighted in this paper are as follows:

- We introduce a unique optical crossbar network architecture, CW/CCW, 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.
- Our proposed CW/CCW network architecture is evaluated by comparison with several competing network designs on power, area, network throughput, and latency metrics.

II. RELATED WORKS

Although many optical network designs have been proposed, there has been a growing interest focused towards nanophotonic crossbars for their ultra-low-latency, single-hop communication capabilities [4], [5], [7]. Among these designs include Corona, a single full optical crossbar with bundles of
waveguides that snake by each tile on the chip. Corona boasts a simple and intuitive architecture with high bandwidth and low latency communication, however, the design incurs significant optical insertion losses as signals must sometimes travel almost twice around the chip to complete a transaction. Such insertion losses must be countered with increased laser power to ensure effective signal modulation and detection, reducing power efficiency. Firefly, another optical crossbar architecture, reduces optical insertion losses using a concentration of tiles which communicate locally across electrical mesh networks, and globally across an optical network. This design requires shorter optical transmission distance, reducing insertion losses and required laser power, yet increases the total number of network hops as messages must traverse the local and global networks.

A common challenge among optical crossbar topologies is high static laser power consumption. Because channel wavelengths are assumed to be generated by an off-chip laser source, there is currently no convenient means of reducing laser power during runtime; static laser power is simply wasted during low traffic loads. Attempts to maximize static laser power utility through dynamic bandwidth allocation and wavelength stealing have been proposed [7], [10]. Bandwidth can easily be reallocated using nanophotonics as wavelengths can be reassigned to network tiles for reading and writing by retuning appropriate ring resonators. The design presented in this paper will leverage this capability of nanophotonics to improve utility of the laser power.

III. CW/CCW ARCHITECTURE

The architecture proposed in this paper will maintain the low latency benefits 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 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 full communication capability exists between all tiles on the network.

A. Data Network

The proposed architecture is for a 64 core, 16 tile system implementation on 400mm² area chips. 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 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 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 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 flexibly supports the reallocation of bandwidth, shared waveguides are used instead of receiver or sender dedicated waveguides. As shown in Figure 2c, 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 sourced to each group may vary according to bandwidth demand and is allocated using the Control Network discussed in the next section. Wavelengths may be modulated by tiles of each group along the bidirectional waveguides according to the token-based arbitration scheme implemented as part of the control network. Ring resonators along the right side of the chip are tuned according to the destination tiles of each wavelength channel as shown in Figure 2e. Modulation and detection of signals is performed on the same waveguides (Figure 2d). 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 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 0 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 introducing 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.

Figure 4a depicts the optical paths of a sample intra-group network transaction between the second and third tiles of the CW/CCW network architecture. In the example provided either the second or third node of the group could have initiated the transaction with a request, but the response is
Fig. 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) Four cores are concentrated to each tile 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. (e) Resonators at the ends of two middle row-groups are tuned according to outgoing and incoming wavelength allocations.

### TABLE I

<table>
<thead>
<tr>
<th></th>
<th>Corona-64</th>
<th>Firefly-64</th>
<th>3D-NoC-64</th>
<th>PROPEL-64</th>
<th>CW/CCW-64</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waveguide Distance (mm)</td>
<td>6175</td>
<td>2880</td>
<td>5266</td>
<td>3200</td>
<td>5920</td>
</tr>
<tr>
<td>MRRs</td>
<td>64k</td>
<td>24k</td>
<td>40k</td>
<td>32k</td>
<td>104k</td>
</tr>
<tr>
<td>Total Optical Loss</td>
<td>39.60dB</td>
<td>28.59dB</td>
<td>29.54dB</td>
<td>21.25dB</td>
<td>30.86dB</td>
</tr>
<tr>
<td>Network Diameter</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Bisection Bandwidth</td>
<td>2.5TB/s</td>
<td>2.5TB/s</td>
<td>2.5TB/s</td>
<td>1.25TB/s</td>
<td>2.5TB/s</td>
</tr>
</tbody>
</table>

always made on the network opposite of the request. The actual wavelengths used to carry the messages may 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 4b depicts the optical paths of a sample inter-group network transaction between tiles of the second and third row-groups.

### B. 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 CW/CCW network architecture accomplishes this by layering an additional optical network responsible for notifying routers of changes to the wavelength assignments as determined by the bandwidth allocation algorithm.

System performance can vary depending on application traffic patterns. While some applications may generate balanced traffic loads between many 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 channels. The bandwidth allocation algorithm proposed for R-3D-NoC is adopted for use with the CW/CCW architecture [7]. 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 '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 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 resonators are re-tuned 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.

### C. 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 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 CW/CCW routing scheme mitigates the incurred optical power insertion losses as signals are not required to travel the entire length of additional waveguides.

### IV. Performance Evaluation

#### A. Power and Area Analysis

Network area and power consumption was modeled using MIT’s DSENT [15] 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 II contains the optical device parameters assumed in the area and power analysis [11], [12], [13], [14].

A breakdown of power consumption between the proposed CW/CCW architecture, Corona [4], Firefly [5], 3D-NoC [7],

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Waveguide Loss</td>
<td>1.0 dB/cm</td>
</tr>
<tr>
<td>Pass-by Ring Resonator Loss</td>
<td>0.0001 dB</td>
</tr>
<tr>
<td>Drop Ring Resonator Loss</td>
<td>1.0 dB</td>
</tr>
<tr>
<td>Waveguide Crossing Loss</td>
<td>0.05 dB</td>
</tr>
<tr>
<td>Receiver Sensitivity</td>
<td>-20dBm</td>
</tr>
<tr>
<td>Laser Efficiency</td>
<td>30%</td>
</tr>
<tr>
<td>Ring Heating Power</td>
<td>26uW/ring</td>
</tr>
<tr>
<td>Ring Modulating Power</td>
<td>500uW/ring</td>
</tr>
<tr>
<td>Ring Modulation Frequency</td>
<td>10Ghz</td>
</tr>
<tr>
<td>Wavelengths/Waveguide</td>
<td>64</td>
</tr>
<tr>
<td>Waveguide Pitch</td>
<td>4um</td>
</tr>
<tr>
<td>Ring Area</td>
<td>100um²</td>
</tr>
</tbody>
</table>

### C. 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 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 CW/CCW routing scheme mitigates the incurred optical power insertion losses as signals are not required to travel the entire length of additional waveguides.

### IV. Performance Evaluation

#### A. Power and Area Analysis

Network area and power consumption was modeled using MIT’s DSENT [15] 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 II contains the optical device parameters assumed in the area and power analysis [11], [12], [13], [14].

A breakdown of power consumption between the proposed CW/CCW architecture, Corona [4], Firefly [5], 3D-NoC [7],
and PROPEL [12] has been charted in Figure 5. Our CW/CCW network architecture requires less laser power than Corona, but requires more than the other other three networks. Compared to Firefly, this 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 CW/CCW architecture may be partially attributed to significantly higher ring heating power. Decreasing the number of ring resonators 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.

Figure 6 shows a comparative area breakdown of our CW/CCW network architecture against several other designs. Our design is comparable to Corona in terms of area footprint, but requires far more waveguides and ring resonators than the other networks. Again, as explained previously in the power comparison, the number of ring resonators may be reduced by sacrificing bandwidth allocation flexibility. Innovations such as 3D-stacking may also be used to further reduce area footprint in addition to reducing waveguide crossings and optical insertion losses.

**B. Throughput and Latency**

Network saturation throughputs and latencies were simulated using the cycle-accurate OPTISIM simulator [16]. 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 uniform random, bit reversal, butterfly, matrix transpose, complement, perfect shuffle and neighbor traffic pattern for the network under test. We compared the network performance to Corona, Flattened Butterfly, Concentrated mesh (CMesh), CWCCW and reconfigured R-CWCCW networks.

Figure 8 shows a comparison of network throughputs for several topologies, including our proposed bandwidth-reconfigurable CW/CCW topology, standard CW/CCW, 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-CW/CCW by taking 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 CW/CCW topology over the standard CW/CCW topology and Corona. R-CW/CCW demonstrates a 92% throughput improvement over the conventional electrical flattened butterfly network. As most real applications have phases in which the behavior resembles the synthetic traffic pattern, we expect similar performance gains.

Figure 7 shows average network packet latencies against network congestion rates for several traffic patterns. The bandwidth-reconfigurable R-CW/CCW network exhibits slightly higher traffic congestion tolerances compared to the standard CW/CCW and Corona topologies. For each traffic simulation R-CW/CCW demonstrates modest performance gains except for the Complement traffic pattern which demonstrates significant improvements in latency and network saturation tolerance.

Network throughput simulations were also conducted using PARSEC, SPEC, and SPLASH2 benchmark traces. The results are presented in Figure 9 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 throughput gains for R-CWCCW compared to CWCCW 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.

**V. CONCLUSIONS AND FUTURE WORK**

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. Dynamic bandwidth reconfiguration achieves significant improvements in network throughput. Network scalability is also improved by splitting crossbar functionality into multiple smaller crossbars with separate 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 power consumption.

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