J. Parallel Distrib. Comput. 100 (2017) 130–139

Contents lists available at ScienceDirect

J. Parallel Distrib. Comput.

journal homepage: www.elsevier.com/locate/jpdc

CLAP-NET: Bandwidth adaptive optical crossbar architecture

Matthew Kennedy *, Avinash Karanth Kodi

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

HIGHLIGHTS

- A unique optical network architecture with a decomposed crossbar design is proposed.
- Dynamic bandwidth reconfiguration achieves higher network throughput.
- Dynamically reconfiguring the laser source at runtime reduces power dissipation.
- The proposed architecture excels on multiple metrics of power, area, and throughput.

ARTICLE INFO

Article history:
Available online 7 May 2016

Keywords:
Network-on-chip
Optical interconnects
Optical crossbar
Multicore
Bandwidth reallocation
Power reconfiguration

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.

© 2016 Elsevier Inc. All rights reserved.

1. Introduction

On-chip network power consumption for multicore processors is a growing concern for computer architects as hundreds and even thousands of processing cores [4] are being packed on to single silicon dies. While industry and consumer demand for faster and more efficient devices is constantly growing, computer architects 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 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, otherwise impractical using only electrical components, become possible with nanophotonics. Scaling an electrical crossbar network to support many cores requires 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 possible with innovations such as dense wavelength division multiplexing (DWDM) [5]. DWDM enables multiple network channels across only a single waveguide. 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, however, increasing the number of network participants contribute to higher optical insertion losses and laser power costs.

We present **Cl**ockwise/Counter-Clockwise **A**daptive **P**hotonic **Net**work (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-clockwise optical routing scheme avoids the

* Corresponding author.
E-mail address: mk140409@ohio.edu (M. Kennedy).

http://dx.doi.org/10.1016/j.jpdc.2016.05.004
0743-7315/© 2016 Elsevier Inc. All rights reserved.

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 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 like IBM have successfully demonstrated on-chip optical communication links [19].

There are several necessary components to form a photonic communication link as depicted in Fig. 1. 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. Waveguides can be fabricated on multiple silicon layers to reduce insertion losses from waveguide crossings and increase on-chip bandwidth [2]. 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 between 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 the resonator structure [18]. 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 Gbps [27].

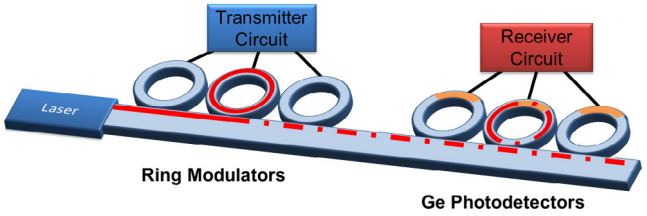


Fig. 1. A basic nanophotonic link consists of an off-chip laser coupled to an on-chip silicon waveguide. Electrically controlled ring resonators are used to modulate the carrier signal. The optical signal is detected using Germanium photodetectors and is converted back to a digital signal through a series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

On the receiving end of photonic links, germanium photodetectors capable of receiving signals up to 40 Gbps collect photons and produce an electrical output current [19]. Photodetectors with responsivities of 0.5 A/W have been demonstrated [19]. Integrating the photodetector with a ring resonator allows the 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 output as a usable signal by digital circuitry such as a router or switch [19]. 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 [8]. The low-energy and high-bandwidth characteristics makes nanophotonic links a promising solution for future many-core interconnects [21,1].

3. Related works

Many optical network designs have been proposed [12,16,10,13,25,20]. PROPEL [12], a two-hop network leverages nanophotonics to implement a power-efficient interconnect. Another architecture, 3D-NoC [16], 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 [25]. 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 signals transmitted through the waveguides [20]. 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 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.

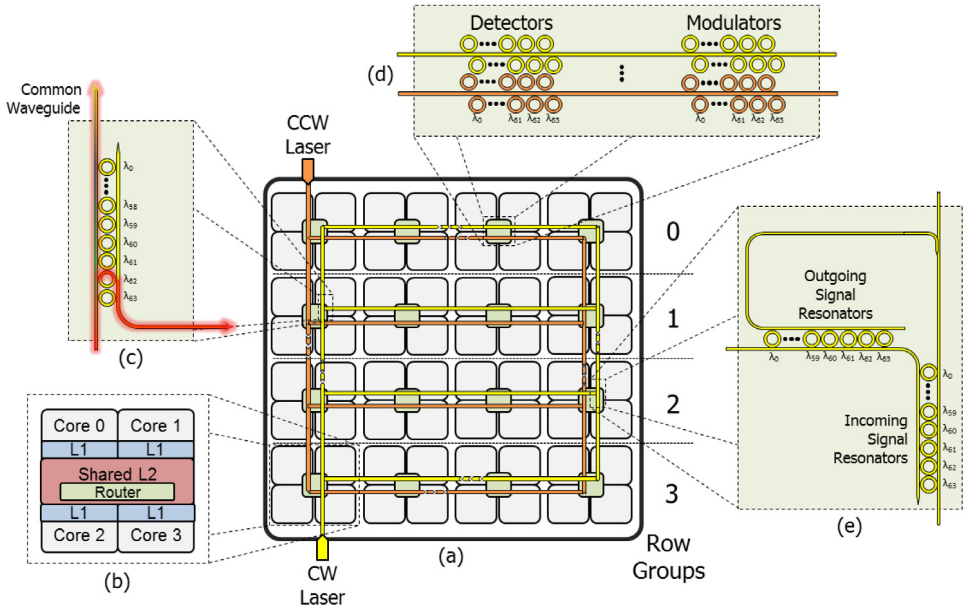


Fig. 2. (a) The network connects a 64 core grid using shared waveguide bundles. Each network routes laser power in either a clockwise or counter-clockwise fashion. (b) Each tile consists of a concentration of four cores along with individual L1 caches, a shared L2 cache, and a router connecting the tile 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.

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	39.60 dB	28.59 dB	29.54 dB	21.25 dB	30.86 dB
Network diameter	1	3	1	2	1
Bisection bandwidth	2.5 TB/s	2.5 TB/s	2.5 TB/s	1.25 TB/s	2.5 TB/s

A common disadvantage among optical crossbar topologies is high static laser power consumption. Because channel wavelengths are assumed to be generated 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 allocation and wavelength stealing have been proposed [16,28]. 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 [6,7].

4. CLAP-NET 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 (see Table 1).

4.1. Data network

The proposed architecture as shown in Fig. 2(a)(b) is for a 64 core, 16 tile system implementation on a 400 mm^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 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 Fig. 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 4×4 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 Fig. 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 provided to each group may vary according to bandwidth demand and is

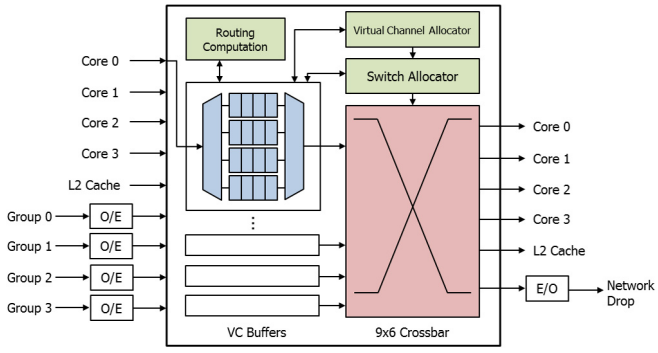


Fig. 3. Router microarchitecture.

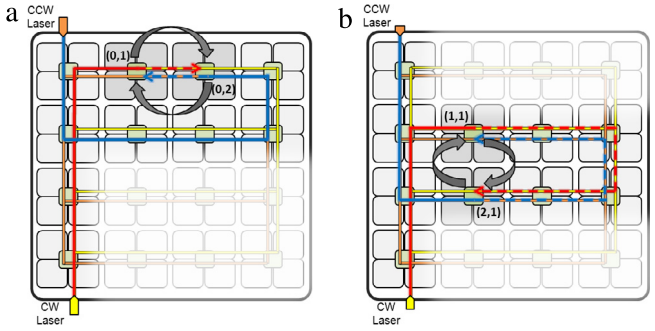


Fig. 4. Optical signal paths for sample (a) intra-group and (b) inter-group network transactions.

allocated using the Control Network discussed in the next section. Wavelengths may be modulated by tiles of each group along the bidirectional waveguides (Fig. 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 tuned according to the destination tiles of each wavelength channel as shown in Fig. 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 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 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.

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 $0 \leq y \leq 3$. Fig. 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 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. Fig. 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 allocation 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 assignments. 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 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 wavelength, their transmissions are regulated according to the token slot arbitration scheme discussed in the Control Network section of this paper.

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 an additional optical network responsible for notifying routers of changes to the 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 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 CLAP-NET architecture [16]. 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 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 (see Fig. 5).

4.2.2. Power reconfiguration

While bandwidth reconfiguration improves utility of laser power by shifting bandwidth to high demand links, bandwidth

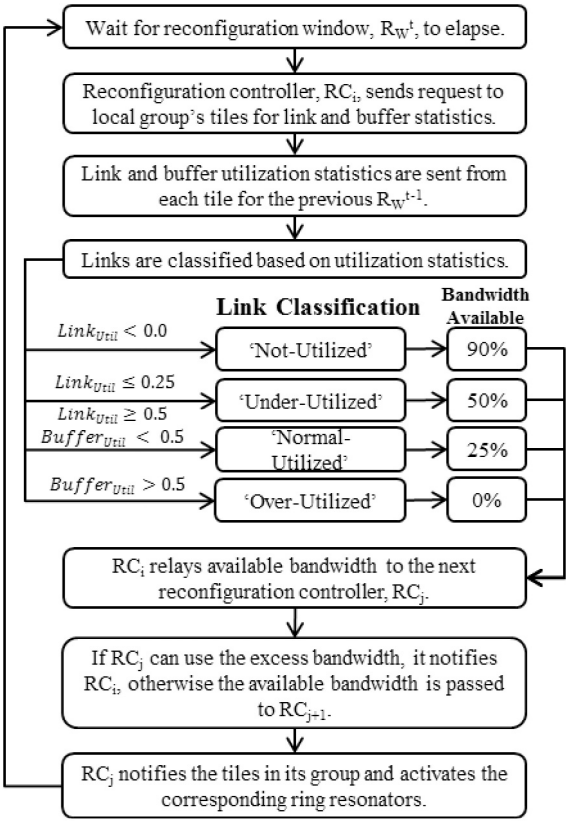


Fig. 5. Flow chart detailing the reconfiguration algorithm adopted from 3D-NoC.

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, 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 depriving the network of bandwidth. Instead, we propose an approach similar to the design discussed in [17] where laser power is distributed to links in time slots by actively tuning broadband ring resonators at run-time. Fig. 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 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 of power. A 4 clock cycle period was chosen as this is the time required to send a typical packet. Fig. 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. 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.

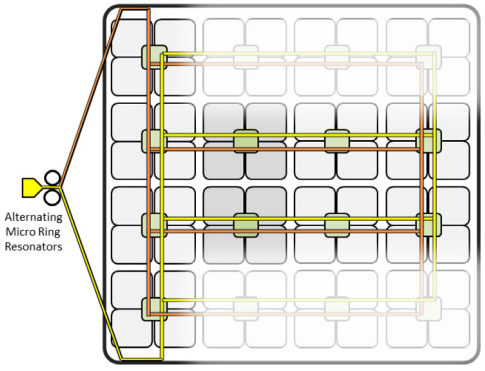


Fig. 6. CLAP-NET with adaptive reconfigurable laser power distribution.

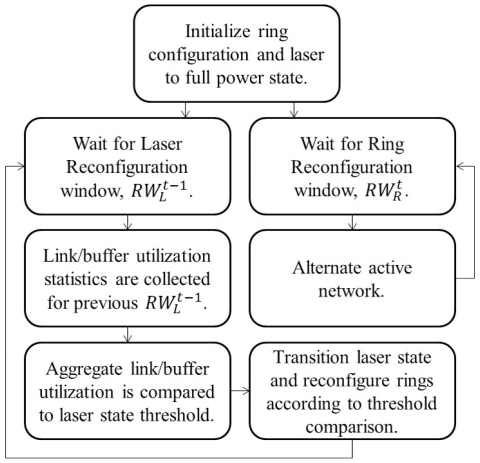


Fig. 7. Flow chart detailing the power reconfiguration algorithm.
Source: Adopted from [17].

4.2.3. Network arbitration

We adopt a destination-based optical token slot arbitration scheme similar to the one used in Corona [25]. 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 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 Fig. 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 secured, a bandwidth allocation table is used to direct flits to the appropriate modulation circuits for deserialization. Fig. 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 to the corresponding wavelengths.

4.3. Network scaling

We assume 64 wavelength channels may be modulated and received at 10 Gbps effectively on a single waveguide. This limits a single clockwise/counter-clockwise network pair to only 1280 Gbps of total available bandwidth. Bandwidth can

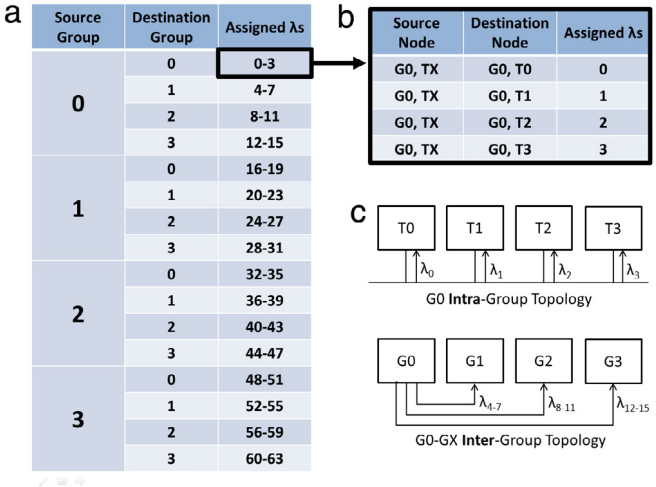


Fig. 8. (a) Table showing static wavelength allocation amongst groups. (b) Table showing assignments of the first four wavelengths to corresponding destination nodes. (c) Logical topology diagrams for intra- and inter-group communication.

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 5 TB/s of network bandwidth or 640 Gbps 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 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 5 TB/s. Table 2 contains the optical device parameters assumed in the power analysis [22,15,23,14].

Table 2
Optical device parameters [22,15,23,14].

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 μW/ring
Ring modulating power	500 μW/ring
Ring modulation frequency	10 GHz
Wavelengths/waveguide	64
Waveguide pitch	4 μm
Ring area	100 μm^2

A breakdown of per bit energy consumption between the proposed CLAP-NET architecture, Corona [25], Firefly [20], 3D-NoC [16], and PROPEL [15] has been charted in Fig. 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 s 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 Fig. 10. Our CLAP-NET network architecture requires less laser power than Corona, but requires more than the 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

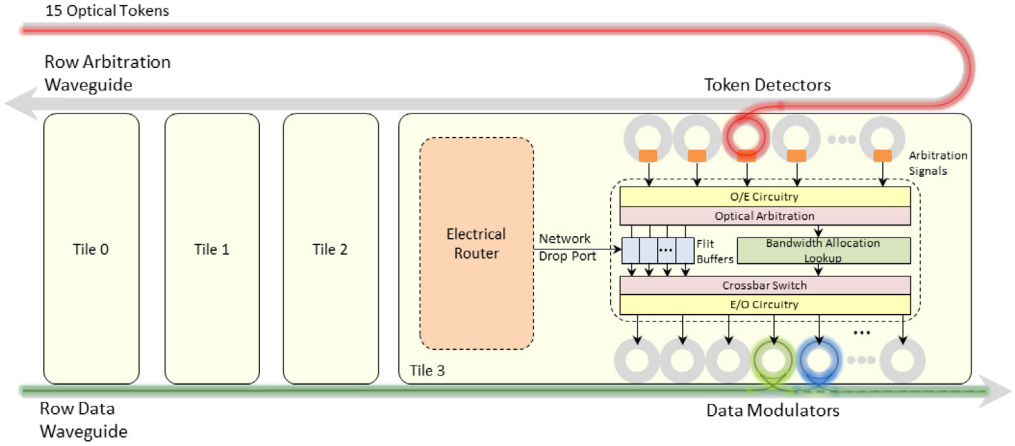


Fig. 9. The optical token arbitration scheme for a single row of tiles. Ring resonators couple optical tokens from the arbitration waveguide to resolve network contention between tiles of the same row. Separate arbitration networks must accompany each row of the network. An optical token for every possible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

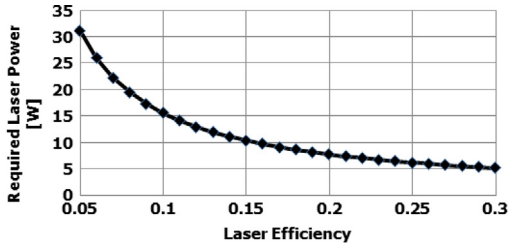


Fig. 10. Required laser power for various laser efficiencies.

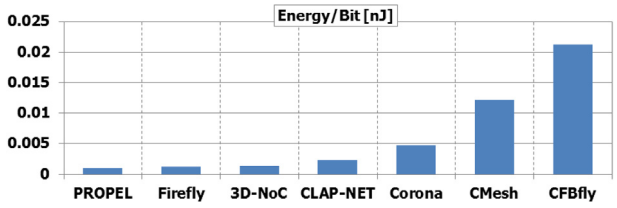


Fig. 11. Energy per bit comparison of different networks in nanojoules.

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 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 no not consider these methods in this paper. Figs. 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 [11]. 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), CLAP-NET and reconfigured R-CLAP-NET networks.

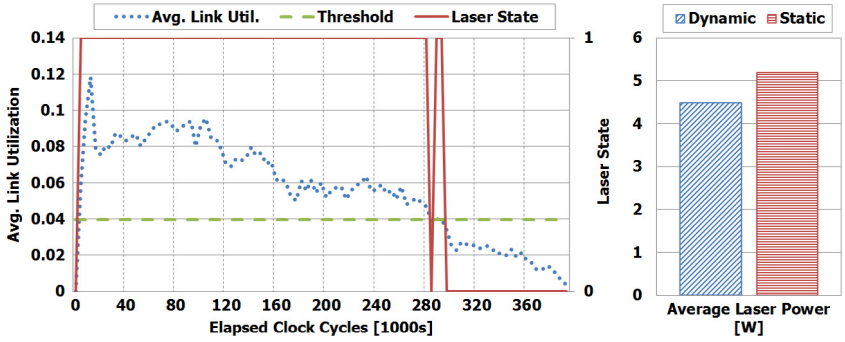


Fig. 12. Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPLASH-2 Barnes traffic traces.

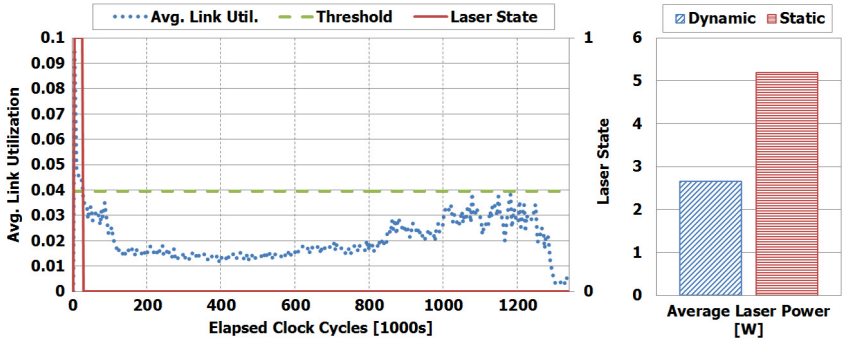


Fig. 13. Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPEC CPU2006 GCC traffic traces.

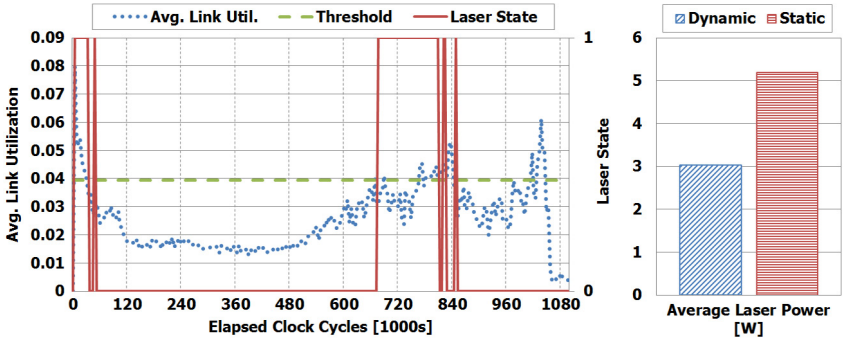


Fig. 14. Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using PARSEC Freqmine traffic traces.

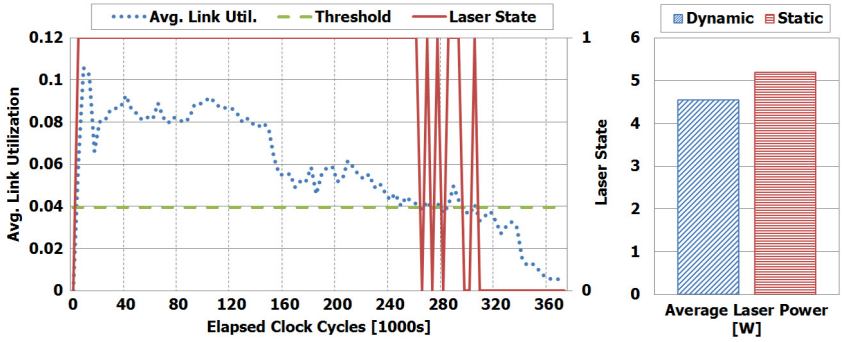


Fig. 15. Average link utilization and laser state plotted against elapsed clock cycles for dynamic power reconfiguration simulations using SPLASH-2 LU traffic traces.

Fig. 20 shows a comparison of network throughputs for several topologies, including our proposed bandwidth-reconfigurable CLAP-NET topology, 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 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 CLAP-NET topology over the standard CLAP-NET topology and Corona. R-CLAP-NET 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.

Figs. 16, 17, 18, and 19 shows average network packet latencies against 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 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 network latency declines for the Butterfly and Complement traffic patterns for injection rates of 0.03–0.06. This is likely due to the reconfiguration process allocating more

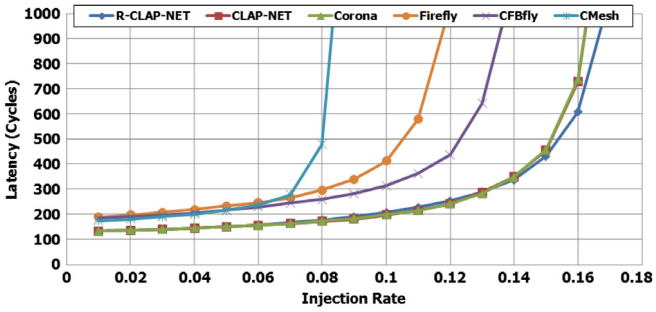


Fig. 16. Average network latencies in cycles for uniform traffic pattern at varying packet injection rates (flits/cycle).

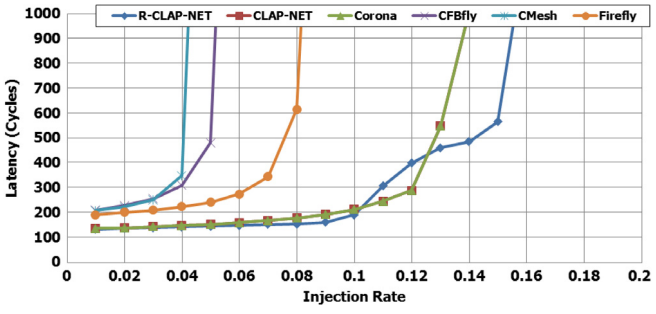


Fig. 17. Average network latencies in cycles for Perfect Shuffle traffic pattern at varying packet injection rates (flits/cycle).

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.

Network throughput simulations were also conducted using PARSEC [3], SPEC [9], and SPLASH2 [26] benchmark traces. The results are presented in Fig. 21 including the geometric mean. We chose the geometric mean to summarize the throughput

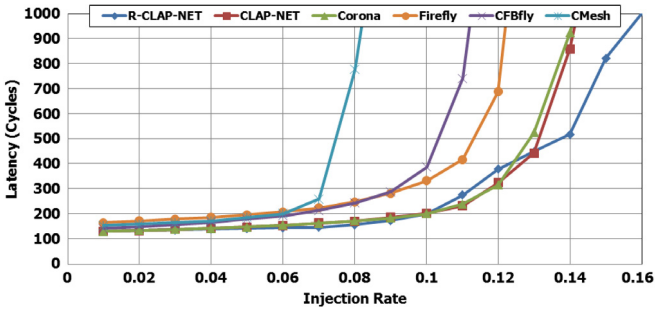


Fig. 18. Average network latencies in cycles for Butterfly traffic pattern at varying packet injection rates (flits/cycle).

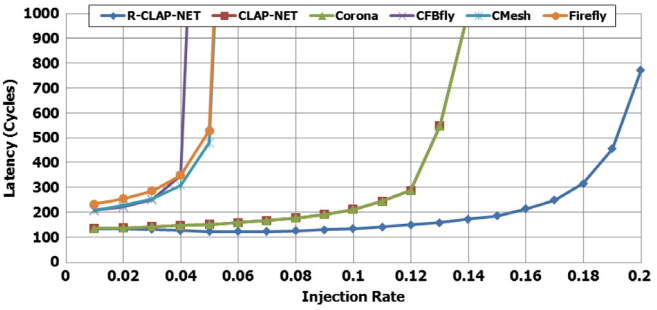


Fig. 19. Average network latencies in cycles for Complement traffic pattern at varying packet injection rates (flits/cycle).

results as outlier simulations will not significantly skew the final results. Simulating with real benchmark traces did not yield significant 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.

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.

Acknowledgments

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 National Science Foundation awards CCF-1054339 (CAREER), CCF-1420718, CCF-1318981, ECCS-1342657, and CCF-1513606.

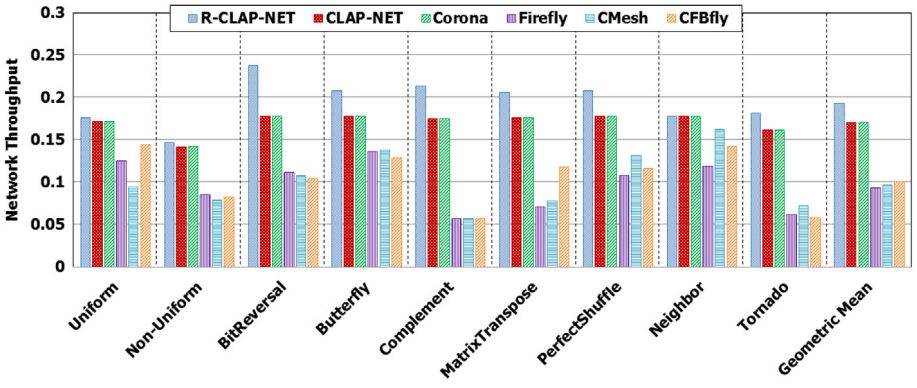


Fig. 20. Comparison of network throughputs at saturation for various traffic patterns. The geometric mean is shown in the last column.

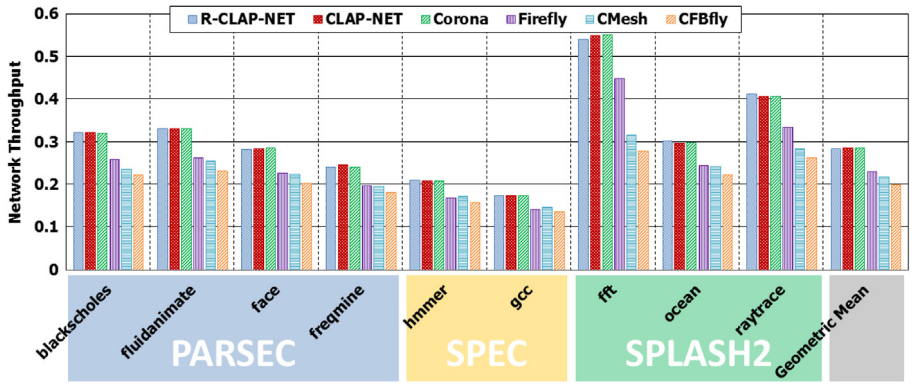


Fig. 21. Comparison of network throughputs simulated using PARSEC, SPEC, and SPLASH2 benchmark traces. The geometric mean is shown in the last column.

References

[1] R.G. Beausoleil, P.J. Kuekes, G.S. Snider, S.-Y. Wang, R.S. Williams, Nanoelectronic and nanophotonic interconnect, Proc. IEEE 96 (2) (2008) 230–247.

[2] 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 multiprocessors, ACM J. Emerg. Technol. Comput. Syst. 7 (2011) 1–25.

[3] C. Bienia, Benchmarking modern multiprocessors (Ph.D. thesis), Princeton University, 2011.

[4] 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.

[5] C. Brackett, Dense wavelength division multiplexing networks: principles and applications, IEEE J. Sel. Areas Commun. 8 (6) (1990) 948–964.

[6] C. Chen, A. Joshi, Runtime management of laser power in silicon-photonic multibus noc architecture, IEEE J. Sel. Top. Quantum Electron. 19 (2) (2013) 3700713.

[7] 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, ACM, New York, NY, USA, 2014, pp. 3–8.

[8] 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.

[9] J.L. Henning, Spec cpu2006 benchmark descriptions, SIGARCH Comput. Archit. News 34 (4) (2006) 1–17.

[10] 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.

[11] A. Kodi, A. Louri, A system simulation methodology of optical interconnects for high-performance computing (hpc) systems, J. Opt. Netw. 6 (2007) 1282–1300.

[12] 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.

[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, Proc. IEEE 97 (2009) 1337–1361.

[14] 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 26th International Parallel and Distributed Processing Symposium, IPDPS'12, IEEE Computer Society, Washington, DC, USA, 2012, pp. 1117–1130.

[15] R. Morris, E. Jolley, A. Kodi, Extending the performance and energy-efficiency of shared memory multicores with nanophotonic technology, IEEE Trans. Parallel Distrib. Syst. 25 (1) (2014) 83–92.

[16] 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.

[17] 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, 2015, pp. 249–252.

[18] K. Padmaraju, J. Chan, L. Chen, M. Lipson, K. Bergman, Thermal stabilization of a microring modulator using feedback control, Opt. Express 20 (27) (2012) 27999–28008.

[19] 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 90 nm SOI CMOS amplifier, Opt. Express 20 (16) (2012) 18145–18155.

[20] 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.

[21] A. Shacham, K. Bergman, S. Member, L.P. Carloni, Photonic networks-onchip for future generations of chip multiprocessors, IEEE Trans. Comput. (2008) 1260.

[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.

[23] 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.

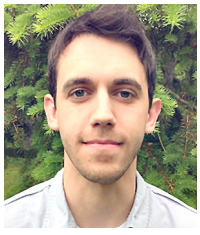
[24] C. Sun, C.-H.O. Chen, G. Kurian, L. Wei, J. Miller, A. Agarwal, L.-S. Peh, V. Stojanovic, Dsent—a tool connecting emerging photonics with electronics for opto-electronic networks-on-chip modeling, in: 2012 IEEE/ACM Sixth International Symposium on Networks-on-Chip (n.d).

[25] 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.

[26] S.C. Woo, M. Ohara, E. Torrie, J.P. Singh, A. Gupta, The splash-2 programs: Characterization and methodological considerations, in: Proceedings of the 22nd Annual International Symposium on Computer Architecture, ISCA'95, ACM, New York, NY, USA, 1995, pp. 24–36.

[27] 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.

[28] 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.



Matthew Kennedy received his B.S. in Electrical Engineering and Computer Science in 2012 from Ohio University, Athens. He is currently pursuing his M.S. degree in Electrical Engineering and Computer Science. His research interests include optical interconnects, network-on-chips (NoCs) and computer architecture.

Avinash Karanth Kodi received the Ph.D. and M.S. degrees in Electrical and Computer Engineering from the University of Arizona, Tucson in 2006 and 2003 respectively. He is currently an Associate Professor of Electrical Engineering and Computer Science at Ohio University, Athens. He is the recipient of the National Science Foundation (NSF) CAREER award in 2011. His research interests include computer architecture, optical interconnects, chip multiprocessors (CMPs) and network-on-chips (NoCs).