RCBus: Row-Column Bus Topology for Optical Network-on-Chip

Weiwei Fu, Tianzhou Chen
College of Computer Science and Technology, Zhejiang University,
Hangzhou 310027, P. R. China, phone: 86 571 87951793
dapan@zju.edu.cn

Abstract—Future chip multi-processors (CMPs) will require high-performance yet low-energy networks-on-chip to provide scalable communication for the increasing number of cores. CMOS-compatible nanophotonic interconnects have recently emerged as a promising candidate for replacing traditional electrical network-on-chip, by providing high bandwidth chip-wide communication at low latency and power. In this paper, we utilize the emerging nanophotonic technology and propose RCBus, an optical network-on-chip topology in which multiple nanophotonic buses are uniformly distributed in each row and column of a two-dimensional torus network. Each bus employs a dense-wavelength-division-multiplexing optical waveguide to enable high-bandwidth express traversal along each dimension. We use multiple write single-read schemes and virtual channel based token protocol to implement light-speed arbitration and flow control. The optical-electrical interface is also studied in the paper. The proposed optical network-on-chip can exploit the advantages of both photonic bus and direct torus topology. Simulation results under synthetic and real-world application traffic show that RCBus is 2X better in terms of performance compared to typical electrical links.

Thanks to advances in CMOS-compatible photonic elements, optical interconnects (ONoC) have recently emerged as a practical alternative to CMP interconnect infrastructure. ONoC helps to fill the need for many-core performance by providing high bandwidth on-chip communication at low latency and power [15]. ONoC allows fast signal propagation at light speed, high bandwidth transmitting via dense wavelength division multiplexing (DWDM), and dissipates low energy which is insensitive to distance.

A. Building blocks of ONoC

Laser source, optical waveguides, splitters and ring resonators are major components of an ONoC, as shown in Fig. 1. Laser source is often located off-chip, and emits laser into the on-chip waveguides to provide multi-wavelength light source. Silicon-compatible waveguides constrain the laser inside to carry light signals. Optical splitters can divert a fraction of laser power from one waveguide and inject it onto another one to distribute power and data. Ring resonators act as the electrical-optical converter, which modulate electrical signals onto laser, or demodulate laser back into electrical signals. The ring resonators are placed next to the fiber and can be tuned into resonance to couple laser into the ring, or out of resonance to let the light pass by. By changing the resonance states of the ring, bit sequences can be modulated on laser and propagated to the photonic receiver. The receiver simply turns on its germanium-doped ring resonators to detach the modulated light from the waveguide and reestablishes the original electrical signals. By employing DWDM, up to 64 or 128 wavelengths can be carried independently on a single waveguide [1], which promise much higher bandwidth compared to typical electrical links.

B. Network Topologies for ONoC

Several nanophotonic on-chip interconnect architectures...
have been proposed [1], [2], [7], [8], [18], [20]. Two major categories are bus-based and direct-topology based.

Bus-based ONoC topologies treat the optical waveguides as shared buses and allow direct connection between two far apart nodes. Our proposed RCBus topology is based on the HP Corona [1] architecture. Corona uses Multiple Write, Single Read (MWSR) interconnect, in which each node has a dedicated, single-destination, multiple-source communication channel. A MWSR waveguide can only be read by its particular home (i.e. destination) node, but all other nodes can modulate signal onto it. For example, Fig. 2 shows the 4-node MWSR channels. P0, P1 and P2 may write messages on a particular channel destined for P3, and P3 is the home node and has specific detectors for the channel to demodulate messages. Every node has its own home channel.

![Fig. 2. MWSR optical interconnects.](image)

In addition to data channels, extra waveguides are used in Corona for token-based arbitration and flow control to manage access permission to each shared channel for data collision avoidance. A token is an encoded bit sequence that is modulated and propagated in the same way as data. Two kinds of optical arbitration protocols are proposed in [4]: token channel and token slot. The token channel arbitration method is derived from Token Ring LAN standard [5]. [4] proposed fair token channel and fair token slot mechanisms for starvation avoidance. Quality-of-service (QoS) support for MWSR nanophotonic NoC is studied in [6] and [3].

Mesh and torus are very popular topology in conventional electrical NoC. It can be extended to electrical-optical hybrid NoC by leveraging nanophotonic switching elements and optical links along with their electrical counterpart. M. Petracca [10] studied the characteristics of physical optical devices to build a non-blocking all-photonic switch, which is used to setup direct optical path. A hybrid torus optical-electronic NoC was proposed in the study, which consists of two closely related sub-networks. The electrical sub-network is packet-based, which is responsible for routing and optical path setup and transmitting short packets. The other sub-network uses optical switch and links to build a high-bandwidth circuit-switched data network to transmit large amount of data. Another direct topology for ONoC named HOME was proposed in [16], which uses similar architecture, but manages the electrical and optical sub-network in a hierarchical fashion.

II. PROPOSED RCBUS ARCHITECTURE

A. RCBus Topology

Corona architecture implements light-speed point-to-point signal delivering through a large optical crossbar for global arbitration, which however, introduces considerable overhead. Besides, a chip-wide token ring with up to 64 nodes induces high round-trip latency for a token and brings more challenges on arbitration efficiency and fairness. So a local bus with less contending nodes will be preferred. Hybrid optical-electrical NoCs with direct topologies, on the other hand, is easy to scale, however suffer from several limitations: (1) it uses hop-by-hop electrical network to perform arbitration and establish optical paths, which is inefficient and slow; (2) circuit-oriented networks require circuit setup and teardown processes, inducing great latency and hence large performance penalty when an established optical path is seldom reused before disassembly; (3) optical switching introduces a large number of waveguide crossings and the optical signal integrity mainly depends on physical devices.

In this paper, we propose RCBus, a Row-Column Bus architecture that employs optical buses to form a torus topology. In RCBus, the large global MWSR token-ring bus of Corona is partitioned into several small ones, which are uniformly distributed in each row and column of a two-dimensional torus network. Unlike the circuit-switched torus ONoC [10], RCBus avoids waveguide coupling and employs MWSR buses, which are able to provide fast transmission along each dimension. Specifically, for a 8x8 torus network, RCBus uses 16 symmetric optical buses, eight of them (a to h) are row buses and the others (i to p) are called column buses, as shown in Fig. 3(a).

Most packets should experience two-phase delivering through a row bus and a column bus respectively. The process of changing dimension for a packet can be done in electrical or optical domain. Since waveguide crossing and high-bandwidth coupling incur large crosstalk in optical signals, we choose to process it electrically. Thus, an additional pair of O/E and E/O conversion is required. Since each waveguide mounts only 8 nodes, RCBus uses much less ring resonators, which are the main contributors to ONoC power. Besides, as each waveguide only covers one dimension, the round-trip period of both token and data packet can be reduced to only 1 or 2 cycles at 5GHz [1]. Fast token capturing and home collection make the all-optical arbitration fast and effective, as will be discussed later.

Fig. 3(b) illustrates another RCBus topology called RCBus-2, in which the number of waveguides is halved and each waveguide now has two rows (columns) of nodes. Comparison of the two topologies will be made in Section 3.

![Fig. 3. (a) RCBus topology (b) RCBus-2 topology.](image)
B. Optical-electrical Interface Microarchitecture

Optical-electrical interface is critical to the performance of an ONoC. We use an on-chip router to connect local processing elements (PEs) to the optical network. The router is responsible for housing incoming flits (either from the local PEs or from optical demodulators) and dispatching them to different output units, such as the optical modulators or input network interfaces (i.e., ejection ports), as illustrated in Fig. 4.

Buffer structure is an important part of router design. Corona uses virtual output queuing (VOQ) [9] scheme for packets buffering, in which every flit is queued according to their destination. VOQ can totally eliminate head of line (HoL) blocking by allowing sources with requests destined for different destinations to make these requests independently. However, for a 64-node network, each source has to keep 63 VOQs, which is very costly, even if a control table (CAM) is maintained to book-keep information of each queue. Besides, since every node cannot send too many flits simultaneously due to limited optical power supply, the number of nominations (i.e., requests of sending a flit on its optical channel through capturing a token) is supposed to be small in order to mitigate token waste [4]. Hence, the full VOQ structure is neither practical nor necessary.

In this work, we prefer using the conventional worm-hole-based virtual channel (VC) structure [11]. Each VC can only store successive flits from one packet to avoid flit inter-leaving, which is inevitable in VOQ. Additional output links is added to the router, which corresponds to the E/O modulator. An extra dedicated buffer named optical launching buffer, is placed at the end of each link to house flits requesting the optical channels. The output port for the modulator is indistinguishable from the other ports to router operation, for the router only sees an additional physical channel. Since only minimal modification is made to the state-of-the-art router architecture, our router design is compatible with the electrical network and NIs, making it a good candidate for interfacing both electrical and optical networks.

The optical launching buffer is also organized in VCs, as shown in Fig. 4. The number of VCs is equal to the number of nominations.

![Fig. 4. Optical-electrical interface microarchitecture.](image)

When a VC is allocated to a packet, its optical destination is recorded and at the same time, the particular microring detector should be turned on, in order to grab the token and modulate on the waveguide. Since the number of nominations and VCs is small, its overhead is significantly reduced compared to the full VOQ or control-table based organization.

The low-radix electrical crossbar’s role lies in multiplexing buffered flits onto its output links. Dimension switching of signal transmission can be done in this crossbar, by demodulating (i.e., O/E converting) the signal, switching it electrically to the other dimension’s launching buffer, and converting it to optical signal again.

C. VC-based Token Arbitration and Flow Control

We make some modifications to the token channel and token slot arbitration scheme [4] for RCBus to support virtual channel (VC) flow control, which wins popularity in electrical router and NI architectures.

We use a single DWDM waveguide for token-based arbitration in each data channel. Each token occupies $n$ wavelengths, where $n$ is the number of VCs in the optical receiver. Thus an 8-node optical channel with 8 VCs each will require a 64-wavelength waveguide for arbitration, which is exactly the same as data channels.

For token channel protocol, credit-based flow control can be realized via encoding the VC information on its dedicated wavelength channel, including the credit number (i.e., free buffer space) of the VC and its VC state: idle or active. A head flit of a packet must first perform the VC allocation (VA) stage, in which an arbitrary idle VC is picked up and allocated to the packet. Upon successfully allocated an output VC, the source may send all the flits of the packet available in buffer, and then reinjects the token, with its VC credits consumed and the state being active. When the home node absorbs its multi-wavelength token, the information of all the VCs is updated and reinjects it if there’s free buffer space in any VC.

As each home node may emit multiple tokens in its token waveguide and the source is unable to update the VC information in token slot protocol, the flow control mechanism must be implemented at the home node. We use 3-state VC encodings in all the tokens which are released at every cycle. State “i” stands for idle VC, state “p” stands for partial occupancy, which means that the VC is allocated but there’s buffer space available, and state “f” indicates that the VC is full. Each token is labeled according to the VC state, except for the following limitations. At cycle i, a token in which only VC-0 is admitted to be idle, is released to the network, and at cycle i+1 only VC-1 can claim to be idle, and so on. A bubble is added between adjacent periods. During the bubble, every node cannot send tokens with idle VC signal to avoid VA collision, even if the corresponding VC is free. The duration of the bubble is equal to the delay of modulation. Our VA mechanism guarantees that an idle VC cannot be granted to two requesters at the same time, and no starvation occurs during VA process.

D. Hardware Overhead Comparison

Table I illustrates the hardware overheads of the baseline electrical NoC, Corona, RCBus and RCBus-2. All the network topologies contain 64 nodes, and 64-wavelength DWDM is used for optical topologies. The electrical NoC
router has more input and output ports, which requires large arbiters and buffer space, and is not area- and energy-efficient. Corona uses many ring resonators (including arbitration rings), which consume significant power for ring trimming. Besides, long waveguides should acquire more power from laser source to cope with optical power loss when the light passes a great number of rings. RCBus adopts one fourth the number of rings and shorter waveguides compared to Corona, and moderate router size compared to an electrical NoC router. However, 16 waveguides are required to cover all rows and columns. RCBus-2 has 8 waveguides but require more rings to support each 16-node optical bus.

**TABLE 1. OVERHEAD COMPARISON OF DIFFERENT TOPOLOGIES**

<table>
<thead>
<tr>
<th>Overhead</th>
<th>Electrical NoC</th>
<th>Corona</th>
<th>RCBus</th>
<th>RCBus-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Router ports</td>
<td>5X5</td>
<td>2X2</td>
<td>3X3</td>
<td>3X3</td>
</tr>
<tr>
<td>Electrical links</td>
<td>352</td>
<td>192</td>
<td>256</td>
<td>256</td>
</tr>
<tr>
<td>Waveguide bundles</td>
<td>-</td>
<td>1</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>Ring resonators</td>
<td>266240</td>
<td>73728</td>
<td>139264</td>
<td></td>
</tr>
<tr>
<td>Waveguide length</td>
<td>4X</td>
<td>1X</td>
<td>1X</td>
<td></td>
</tr>
</tbody>
</table>

III. EVALUATION

We evaluate our RCBus and RCBus-2 ONOC on a cycle-accurate simulator and compare it with the traditional electrical mesh NoC and Corona. All router designs use 3-stage pipelining and have 8 input VCs. The VCs are 4 flits in depth, which is the same as the length of a packet. For NoC, both token channel and token slot arbitration are modeled and evaluated. Each node can only send 3 flits simultaneously, due to the limited transient power supply. The optical launching buffer contains 16 VCs, and thus the number of maximum concurrent nominations is limited to 16.

In order to test the proposed optical arbiters under a variety of traffic conditions, we use the synthetic uniform and bit-complement traffic patterns [11], and also real-world application traces. Four benchmark Traces (rtview, swaptions, blackscholes, ferret) selected from PARSEC [12] is collected from Gem5 full-system simulator [14] and simulated by our ONOC simulator.

For synthetic traffic tests, we randomly generate traffic according to the injection rate (number of generated flits per cycle per node). All simulations execute for 30,000 cycles. However, we ignore the first 3,000 cycles to eliminate warm-up transients. The results are presented in average packet latency versus offered load in terms of packet injection rate.

For trace-based evaluation, we assume a 64-tile CMP. Each tile contains a processing core, a private L1 cache, and a shared L2 cache bank. Each node has a network interface which directly connects to a concentrated router. Four miss status holding registers (MSHRs) are maintained for each L1 cache to allow multiple memory transactions to be launched simultaneously. We use MESI directory-based coherence protocol and collect all the memory access and cache coherence packets as the input traffic of our simulator. We run the simulator until 1,000,000 memory requests are launched and present the normalized IPC (instructions per cycle) and power results, which will be shown in the following sections.

A. Performance evaluation

Our simulation starts with a latency comparison of RCBus and other networks under synthetic traffic. Fig. 5 plots the average message (packet) latency (in cycles/packet) as the function of injection rate (in flits/cycle) under uniform random and bit-complement traffic. We can see that RCBus performs better than the electrical NoC in zero-load latency by 1.7X for token channel arbitration, and 1.9X for token slot on average. It is because RCBus leverages nanophotonic buses to transmit signals at the speed of light. The baseline NoC however, is required to deliver packets hop by hop, and each packet must experience 3 pipeline stages of routers along its way, so the average latency is relative high even under zero loads. However, RCBus has longer zero-load latency than Corona by 8%, because most packets have to be modulated onto the optical channels twice and thus the whole pipeline stages of the crossing router is added to its latency.

When network load is high, RCBus gains on average 24% throughput improvement than Corona on average for all traffic patterns, because the long token round-trip time and high collision possibility affect the efficiency and fairness of arbitration in Corona. Besides, RCBus achieves higher bandwidth and twice the fan-out of each node.

The latency curve of RCBus-2 is also shown in Fig. 5. RCBus-2 uses waveguides with more nodes compared to RCBus, but only one half channels are provided. The features and parameters, as well as the performance results, fall in between RCBus and Corona.

The token slot arbitration mechanism has much better performance than token channel, since token slot achieves up to 100% bandwidth utilization. Our RCBus architecture is compatible with both protocols and has similar throughput improvement than Corona.

![Fig. 5. Latency-injection rate curves for different traffic patterns.](image-url)
Traces from some PARSEC benchmarks are used to compare the performance of the various NoC architectures. We use IPC as the comparison metric in our comparison. The IPC results are normalized to the baseline electrical NoC, as shown in Fig. 6. The normalized IPC results show the benefit of our RCBus design, since RCBus under token channel and token slot arbitration schemes, achieves speedup of nearly 1.9X and 2.2X respectively, than the baseline NoC, and performs 4.5% better than Corona on average.

Fig. 6. Normalized IPC for 4 benchmarks from PARSEC.

B. Power estimation

We use ORION [13] to estimate energy of on-chip electrical routers and links. For optical networks, laser and ring resonators are the main contributors to power consumption. The laser power should overcome losses due to modulation inefficiencies, transmission losses in the waveguide and insertion of off-resonance rings. Besides, ring modulators and detectors must be trimmed to compensate for fabrication error, since their functions are sensitive to temperature [17]. Thus, external heating for micro-rings is required and is another static power budget contributor. We estimate laser source according to results in [2] and [4]. Dynamic power of optical components mainly comes from photonic modulation and demodulation. We assume the energy consumption to modulate and demodulate a 128-bit flit to be 23pJ [17].

We present static power estimation in Fig. 7 for different networks. The electrical mesh NoC has large routers, which continuously consume much leakage energy because large buffers and high-radix crossbars are thirsty for power. Energy dissipation of the electrical part of the ONoCs is much less compared to the electrical NoC. RCBus uses a bit more energy than Corona, since an extra port is added to the router from dimension switching. The optical static power of RCBus and RCBus-2 is comparable to Corona. Though there’re more waveguides consuming significant laser power, the power for ring heating is much less, since RCBus and RCBus-2 reduces the amount of micro-rings on the data path by a factor of 4 and 2 respectively, compared to Corona.

Next we present results of dynamic power (in mW) versus offered load under bit-complement traffic in Fig. 8. It can be seen that RCBus consumes only 16% dynamic power on average compared to conventional electrical NoC. The activities of buffer reading and writing, VC allocation, switch allocation, crossbar and link traversal introduce significant power dissipation. In contrast, optical transmission is very power-efficient, because modulation and demodulation of signals are fast (about 75ps [19]) while consuming little energy. Corona achieves least power dissipation because its electrical components are simple. Besides, each flit is modulated and demodulated only once, which saves much energy. Under higher load, the curves for all the topologies become flat due to the limited throughput of NoC (i.e. the capability of transmitting messages). We can see that for bit-complement traffic, RCBus and RCBus-2 can achieve higher throughput than Corona and electrical NoC at moderate power consumption.

Finally we show the normalized energy-delay product (EDP) of different ONoC topologies for PARSEC benchmarks traffic in Fig. 9. EDP is measured as the product of power consumption and the total completion time of all the memory transactions in the benchmark. It is shown that RCBus has similar EDP performance to Corona. For token slot protocol, RCBus achieves 7.4% less EDP but has 4% more EDP than Corona for token channel on average. RCBus has better performance than Corona at the cost of more static and dynamic energy, so their overall energy-efficiency is very similar.
IV. CONCLUSIONS

Technology scaling requires the provision of high-performance and low-power on-chip interconnects for many-core applications. Nanophotonic technology is an emerging solution for future on-chip interconnects and provides several significant advantages over metallic interconnects. Yet, current architectural proposals of optical network-on-chip are limited either in scalability or require inefficient electrical control networks. In this paper, we present a novel ONoC architecture named RCBus. RCBus is based on torus topology and has optical token-ring buses at every row and column. The router microarchitecture for RCBus is proposed in the paper, which acts as the optical-electrical interface. We make further modifications to token channel and token slot arbitration scheme to support virtual channel flow control. Simulation results under synthetic traffic and real-world traces show that our proposed RCBus topology achieves best overall performance and power ratio, especially under high network load. For future work, we plan to investigate other bus-based topologies and bus read-write modes. Besides we’ll explore alternatives to the arbitration method in both electrical and optical domains.

REFERENCES


