

LETTER

An Optimized Low-Power Optical Memory Access Network for Kilocore Systems

Tao LIU[†], *Nonmember*, Huaxi GU^{†a)}, *Member*, Yue WANG[†], and Wei ZOU[†], *Nonmembers*

SUMMARY An optimized low-power optical memory access network is proposed to alleviate the cost of microring resonators (MRs) in kilocore systems, such as the pass-by loss and integration difficulty. Compared with traditional electronic bus interconnect, the proposed network reduces power consumption and latency by 80% to 89% and 21% to 24%. Moreover, the new network decreases the number of MRs by 90.6% without an increase in power consumption and latency when making a comparison with Optical Ring Network-on-Chip (ORNoC).

key words: optical network-on-chip, memory access

1. Introduction

The number of cores on a single chip and corresponding memory capacity are increasing dramatically to meet the requirement for memory access network of future exascale computer. In order to respond to this change, three-dimensional integration technology as well as optical interconnect technology is proposed and applied. Some classic networks, such as crossbar, λ -router, have strong scalability but they are constricted by huge power consumption due to many waveguide crossings when network scales.

Ring architecture is put forward to alleviate waveguide crossings. D. Vantrease *et al.* [1] propose Corona, whose 64 clusters communicate through an optical broadcast ring. Le Beux *et al.* [2] bring forward ORNoC using multiple waveguides and diverse directions of rotation to provide sufficient communication optical links considering the limited available wavelengths and waveguides. K. Wang *et al.* [3] design a ring-based waveguide bundle to contain the cores and ranks. K. Wen *et al.* [4] put forward a reconfigurable optical memory interconnect based on ring-based NoC to connect cores and ranks. However, the longest optical link has to go across half of the MRs in ring architecture, which dramatically increases the pass-by loss and cannot be integrated easily. Therefore, in views of the unsolved problem about heavy use of MRs in ring architecture, the trade-off between the requirement for high performance and the cost of MRs, including pass-by loss and integration difficulty, becomes a very urgent problem.

In this letter, we propose a low-power optical memory access network for kilocore systems. Compared with classic optical network, such as ORNoC, the new network

greatly decreases the cost of MRs without an increase in power consumption and latency, which will be more evident as network scales.

2. Architecture

2.1 Overview of the Architecture

The proposed low-power optical memory access network takes full advantage of BMRs to alleviate the scaling bottlenecks such as the cost of MRs. In order to alleviate integration difficulty, we adopt a 2.5D layout of N -core 4-memory ($M + 2$)-layer memory access architecture in Fig. 1 (a). M layers each place the same N/M cores and the other two layers place optical subnetworks and memory respectively. Each of core and memory layers is connected to the optical subnetworks by through-silicon vias (TSVs). Because the architecture is expanded as four-core groups each layer in all designs, the relationship between N , M , and the number of groups k can be defined as $N = 4Mk$.

On the optical layer, there are $N/2M$ optical subnetworks in total. Half of them are responsible for the core-to-memory communication, while others are applied into the memory-to-core communication.

Take the 1024-core 4-memory 6-layer architecture for example. Each core layer contains 64 groups of cores with the same color in Fig. 1 (a) and each group has four cores. The communication from core to memory uses switch units of 64 optical subnetworks, while the remaining 64 optical subnetworks are used to make the communication from memory to core possible.

2.2 Optical Subnetwork

To better explain the architecture, Fig. 1 (b), taking M as 4, gives a close-up to two optical subnetworks in Fig. 1 (a). A subnetwork concludes optical transmitters/receivers and optical switch units. Laser sources supply optical power for several subnetworks. Optical transmitters use BMRs to modulate transactions, while optical receivers use BMRs to demodulate core-to-memory transactions and use narrow-band microring resonators (NMRs) to demodulate memory-to-core transactions since cores sharing the same optical subnetwork are located in different layers.

The crucial components of the optical subnetwork are the new optical switch units. Each subnetwork has 6 BMRs. The realization of BMRs is determined and limited by two

Manuscript received November 12, 2018.

Manuscript publicized February 4, 2019.

[†]The authors are with the State Key Laboratory of Integrated Service Networks, Xidian University, 710071 Xi'an, China.

a) E-mail: hxgu@xidian.edu.cn

DOI: 10.1587/transinf.2018EDL8234

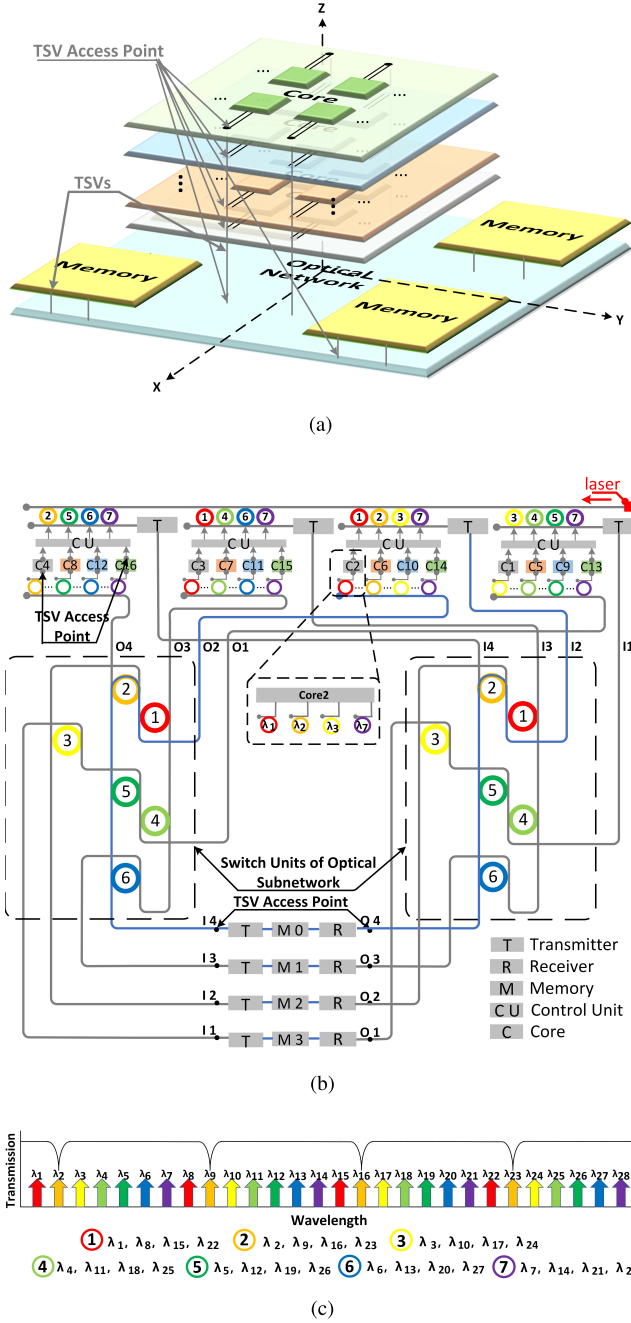


Fig. 1 (a) 2.5D layout of N -core 4-memory ($M+2$)-layer memory access architecture. (b) Detailed layout of two optical subnetworks. (c) The relationship between the BMRs and wavelengths.

major factors: the free spectral range (FSR) of the MRs, which represents the maximum available bandwidth for the wavelength-parallel system, and the wavelength channel spacing utilized by the optical transmission system [5]. The FSR of a MR is approximated by:

$$FSR \approx \frac{\lambda_0^2}{n_{Group} \times L} = \frac{\lambda_0^2}{n_{Group} \times 2\pi r} \quad (1)$$

Here, n_{Group} is the group index of the waveguide that forms the MR, L is the optical path length for one round trip

through the MR, and r is the radius of the MR. The relationship between the serial number of BMRs and all the resonant wavelengths $\lambda_1, \lambda_2, \dots, \lambda_{7M}$ can be defined as follows.

$$BMR_i = \{\lambda_j \mid j = i + 7 \times a\}, 0 < a < M, a \in N^+ \quad (2)$$

When M equals 4, the relationship between them can be depicted in Fig. 1(c). Compared with NMRs where one wavelength is allowed to couple at a time, BMRs with proper radius in Fig. 1(c) can couple four different wavelengths. Instead of arranging independent optical subnetwork for each core layer, the new network enables cores to share the optical subnetwork with those with different z coordinate but the same x and y coordinates, which decreases the number of MRs drastically in kilocore systems.

3. Communication Method

3.1 Overview of Communication Method

By using the technology of Wavelength Division Multiplexing (WDM), the proposed architecture can send multiple core-to-memory and memory-to-core transactions in parallel. When core-to-memory transactions are ready to send, the transmitters will convert electronic signals into optical signals and the control units will switch on specific BMR to couple optical signals based on the routing table. After being detected and converted by optical receivers, transactions will be sent to memory stacks and be processed in order. Specially, when multiple core-to-memory transactions from different cores are sent simultaneously, memory controllers rotate the scheduling priority in a round-robin fashion from bank to bank [13] to alleviate overhead caused by pending transactions.

When it comes to memory-to-core transactions, the case is similar except for the way of demodulating. Because cores sharing the same optical subnetwork are located in different layers, we use NMRs to demodulate these transactions to different cores.

3.2 Case Study

In order to elaborate the routing table, a N -core 4-memory 6-layer system is taken as an example. N -core system is divided into $N/16$ identical parts and each part has two independent optical subnetworks. According to Fig. 1(a)(b), we construct Table 1 to show the details about adopted wavelength, and corresponding BMR and optical links in the cases of specific core-to-memory (optical link1) and memory-to-core (optical link2) transactions of one part. For example, when core 2, 6, 10, 14 send transactions to memory 0 simultaneously, the control units will switch on the BMR_2 to couple optical signals with wavelength $\lambda_2, \lambda_9, \lambda_{16}, \lambda_{23}$ and these signals from I2 can transmit in parallel through optical waveguides. After receiving electronic signals converted by receivers at O4, the memory will record which subnetwork and core they come from and adopt round-robin,

Table 1 Routing table.

Core	Optical link1	Optical link2	BMR_i	Wavelength
1, 5, 9, 13	I1-O1	I1-O1	7	$\lambda_7, \lambda_{14}, \lambda_{21}, \lambda_{28}$
1, 5, 9, 13	I1-O2	I2-O1	3	$\lambda_3, \lambda_{10}, \lambda_{17}, \lambda_{24}$
1, 5, 9, 13	I1-O3	I3-O1	4	$\lambda_4, \lambda_{11}, \lambda_{18}, \lambda_{25}$
1, 5, 9, 13	I1-O4	I4-O1	5	$\lambda_5, \lambda_{12}, \lambda_{19}, \lambda_{26}$
2, 6, 10, 14	I2-O1	I1-O2	3	$\lambda_3, \lambda_{10}, \lambda_{17}, \lambda_{24}$
2, 6, 10, 14	I2-O2	I2-O2	7	$\lambda_7, \lambda_{14}, \lambda_{21}, \lambda_{28}$
2, 6, 10, 14	I2-O3	I3-O2	1	$\lambda_1, \lambda_8, \lambda_{15}, \lambda_{22}$
2, 6, 10, 14	I2-O4	I4-O2	2	$\lambda_2, \lambda_9, \lambda_{16}, \lambda_{23}$
3, 7, 11, 15	I3-O1	I1-O3	4	$\lambda_4, \lambda_{11}, \lambda_{18}, \lambda_{25}$
3, 7, 11, 15	I3-O2	I2-O3	1	$\lambda_1, \lambda_8, \lambda_{15}, \lambda_{22}$
3, 7, 11, 15	I3-O3	I3-O3	7	$\lambda_7, \lambda_{14}, \lambda_{21}, \lambda_{28}$
3, 7, 11, 15	I3-O4	I4-O3	6	$\lambda_6, \lambda_{13}, \lambda_{20}, \lambda_{27}$
4, 8, 12, 16	I4-O1	I1-O4	5	$\lambda_5, \lambda_{12}, \lambda_{19}, \lambda_{26}$
4, 8, 12, 16	I4-O2	I2-O4	2	$\lambda_2, \lambda_9, \lambda_{16}, \lambda_{23}$
4, 8, 12, 16	I4-O3	I3-O4	6	$\lambda_6, \lambda_{13}, \lambda_{20}, \lambda_{27}$
4, 8, 12, 16	I4-O4	I4-O4	7	$\lambda_7, \lambda_{14}, \lambda_{21}, \lambda_{28}$

bank-rotation scheduling to process transactions.

The case is similar when memory 0 sends transactions to one of core 2, 6, 10, 14, except for switching on one of $NMR_2, NMR_9, NMR_{16}, NMR_{23}$ at O2 to couple optical signals from I4 with one of wavelengths $\lambda_2, \lambda_9, \lambda_{16}, \lambda_{23}$ respectively. The other NMRs NMR_1, NMR_3, NMR_7 in detail enlargement of Fig. 1 (b) are used when memory 1, 3, 2 send transactions to core 2 respectively.

4. Analysis and Simulation

DRAMSim2 is employed to simulate the transactions between cores and ranks [6] and analyze the performance improvement of the proposed network. The 100,000 transactions of three applications are extracted from Netrace with PARSEC benchmark in chronological order and the results of latency, power consumption, and the number of MRs are shown and analyzed. In this simulation, the total memory storage is set up to 4GB, which means the size of each rank is 1GB.

4.1 Latency

Figure 2 (a) shows simulation results of average latency and the worst case latency, in which the electronic bus, ORNoC, and the new network running three applications from Netrace. The x axis represents the type of the applications, while the y axis is on behalf of the corresponding cycle interval to accomplish 100,000 traces. Since we mainly focus on latency difference caused by different interconnect links, we adopt different parameters of bandwidth (electronic bus 16Gbps per lane based on PCIe 4.0 standard, optical network 25Gbps per lane [12]), but we derive the same parameters of memory related operations from initialization file “DDR3_micron.64M_8B_x4_sg15.ini” in DRAMSim2. The results indicate that the new network and ORNoC are similar when it comes to latency, while the new network reduces average delay and worst case delay by 21% to 44% and 25% to 41% when compared with the electronic bus interconnect.

Table 2 Insertion loss and optical devices parameters.

Parameter	Value [7]–[11]
Waveguide crossing $L_{crossing}$	0.16 dB
Waveguide bend L_{bend}	0.005 dB/90°
Drop into a MR L_{drop}	0.6 dB
Pass by a MR $L_{through}$	0.005 dB
Transmitter through port $L_{transmitter}$	1.2dB
Receiver drop port $L_{receiver-drop}$	0.5 dB
Receiver through port $L_{receiver-through}$	0.05 dB
Receiver sensitivity S	−22 dBm
Laser efficiency η	8.2%
P_{tuning}	250 fJ/bit×bandwidth
$P_{transmitter} + P_{receiver}$	2 pJ/bit×bandwidth
P_{memory}	500 fJ/bit×bandwidth

4.2 Power Consumption

The total power consumption of optical network can be calculated by following equation sets.

$$\begin{cases} P = P_{laser} + P_{tuning} + P_{transmitter} + P_{receiver} + P_{memory} \\ P_{laser} = 10^{(S+L_{total}+10 \times \lg n)/10} / \eta \\ L_{total} = L_{insertion} + L_{transmitter} + L_{receiver} + L_{distribution} \\ L_{insertion} = L_{drop} + L_{crossing} + L_{bend} + L_{through} \end{cases} \quad (3)$$

where $P_{laser}, P_{tuning}, P_{transmitter}, P_{receiver}, P_{memory}$ respectively refer to the power consumed by all lasers, by tuning MRs, by transmitters, by receivers, by memory stacks and memory controllers. Except for P_{laser} , other four kinds of power consumption are relative to network bandwidth, influenced by specific applications. Table 2 shows specific parameters used in calculation.

Figure 2 (b) shows parts and sum of the worst case power consumption of the electronic bus, ORNoC, and the new network. Due to the huge difference in power consumption between the optical network and the electronic bus, we use the primary vertical axis to represent the power consumption of the optical network, while the power consumption of the electronic bus is represented by the secondary vertical axis. The results demonstrate that the power consumption of the new network is similar to that of ORNoC but less than that of the electronic bus interconnect by 80% to 89%.

4.3 The Number of MRs

Figure 2 (c) shows the number of MRs in 64-/256-/1024-core networks. As a whole, the number of MRs of the new network is 90.6% smaller than that of ORNoC since the application of BMRs enables cores in different vertical layers to share the same optical subnetwork. As network scales, this advantage will be more evident in kilocore systems.

5. Conclusion

In this letter, we propose a low-power optical memory access network for kilocore systems. The unmatched feature

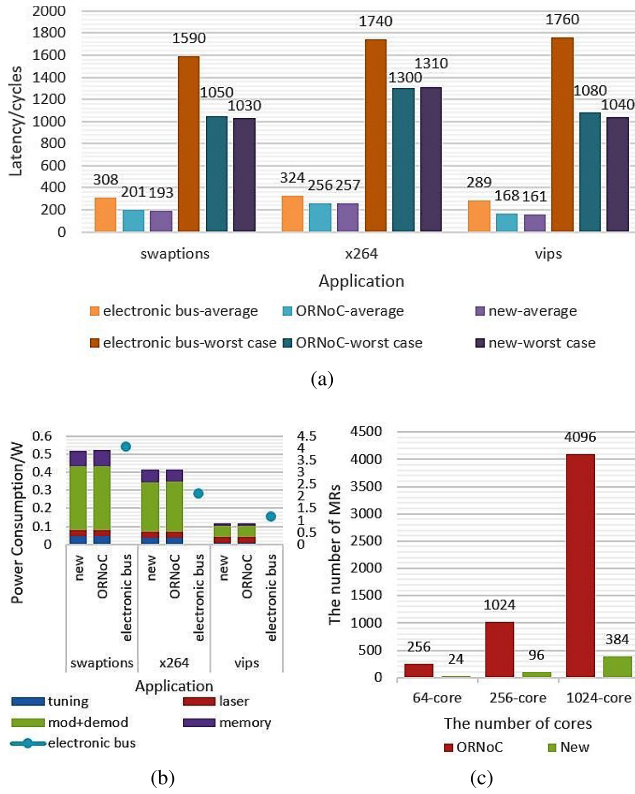


Fig. 2 (a) Comparisons of latency between the electronic bus, ORNoC and the new network. (b) Comparisons of power consumption between the electronic bus, ORNoC, and the new network. (c) The number of MRs of 64-/256-/1024-core network.

of the new network is that BMRs are fully used to decrease the cost of MRs, including pass-by loss and integration difficulty. The proposed communication method implements wavelength routing. The simulation results show that the new network decreases power consumption and latency by 80% to 89% and 21% to 24% when compared with traditional electronic bus interconnect. Moreover, the new network reduces the number of MRs by 90.6% without sacrificing the performance of power consumption and latency when making a comparison with ORNoC.

Acknowledgments

This work was supported by the National Science Foundation of China under Grants 61634004, and Grant 61472300, the Fundamental Research Funds for the

Central Universities Grant No.JB180309 and No.JB170107, the key research and development plan of Shaanxi province No.2017ZDCXL-GY-05-01, and the General Armament Department and Ministry of Education United Fund under Grant No.6141A0224-003.

References

- [1] D. Vantrease, R. Schreiber, M. Monchiero, M. McLaren, N.P. Jouppi, M. Fiorentino, A. Davis, N. Binkert, R.G. Beausoleil, and J.H. Ahn, "Corona: System implications of emerging nanophotonic technology," 2008 International Symposium on Computer Architecture, pp.153–164, June 2008.
- [2] S. Le Beux, J. Trajkovic, I. O'Connor, G. Nicolescu, G. Bois, and P. Paulin, "Optical ring network-on-chip (ORNoC): Architecture and design methodology," 2011 Design, Automation & Test in Europe, pp.1–6, March 2011.
- [3] K. Wang, H. Gu, Y. Yang, and K. Wang, "Optical interconnection network for parallel access to multi-rank memory in future computing systems," *Opt. Express*, vol.23, no.16, pp.20480–20494, 2015.
- [4] K. Wen, H. Guan, D.M. Calhoun, D. Donofrio, and J. Shalf, "Silicon photonic memory interconnect for many-core architectures," 2016 IEEE High Performance Extreme Computing Conference (HPEC), pp.1–7, Sept. 2016.
- [5] K. Bergman, L.P. Carloni, A. Biberman, J. Chan, and G. Hendry, "Photonic network-on-chip design," *Integr. Circuits Syst.*, vol.68, 2013.
- [6] K. Wang, H. Gu, Y. Yang, K. Chen, and K. Wang, "On the design of a 3D optical interconnected memory system," *IEICE Electron Express*, vol.11, no.19, 20140664, 2014.
- [7] W. Bogaerts, P. Dumon, D. Van Thourhout, and R. Baets, "Low-loss, low-cross-talk crossings for silicon-on-insulator nanophotonic waveguides," *Opt. Lett.*, vol.32, no.19, pp.2801–2803, 2007.
- [8] F. Xia, L. Sekaric, and Y. Vlasov, "Ultracompact optical buffers on a silicon chip," *Nat. Photonics*, vol.1, no.1, pp.65–71, 2007.
- [9] X. Zheng, S. Lin, Y. Luo, J. Yao, G. Li, S.S. Djordjevic, J.-H. Lee, H.D. Thacker, I. Shubin, K. Raj, J.E. Cunningham, and A.V. Krishnamoorthy, "Efficient WDM laser sources towards terabyte/s silicon photonic interconnects," *J. Lightwave Technol.*, vol.31, no.24, pp.4142–4154, Dec. 2013.
- [10] Y. Audzevich, P.M. Watts, A. West, A. Mujumdar, S.W. Moore, and A.W. Moore, "Power optimized transceivers for future switched networks," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol.22, no.10, pp.2081–2092, Oct. 2014.
- [11] M. Kennedy and A.K. Kodi, "Laser pooling: Static and dynamic laser power allocation for on-chip optical interconnects," *J. Lightwave Technol.*, vol.35, no.15, pp.3159–3167, Aug. 2017.
- [12] K. Schmidtke, "Designing 100G optical connections," n.p., March 2017, <https://code.facebook.com/posts/1633153936991442/designing-100g-optical-connections/>
- [13] B. Jacob, S.W. Ng, and D.T. Wang, *Memory systems: Cache, DRAM, disk*, Morgan Kaufmann, 2008.