A Silicon Photonic Chip-Scale AWGR Switch for High Performance Computing Systems

Runxiang Yu¹, Stanley Cheung¹, Yuliang Li¹, Katsuari Okamoto², Roberto Proietti³ and S. J. B. Yoo¹
¹Department of Electrical Engineering, University of California, Davis, One shields avenue, Davis, CA, 95616, USA
²AiDi Corporation, 2-2-4 Takezono, Tsukuba, Ibaraki, 305-0032 Japan
³sbyoo@ucdavis.edu

Abstract: This paper demonstrates a silicon-photonic AWGR-based optical switch for HPC systems. Simulations show high throughput and low latency even at high input load. A fabricated silicon-photonic AWGR switch with 32 Tx/Rx pairs showed error-free performance.

OCIS codes: (200.4650) Optical interconnects (250.3140) Integrated optoelectronic circuits

1. Introduction

Low latency, and high-throughput interconnection is essential for future high performance computing (HPC) applications. Interconnect networks based on electronic multistage topologies result in large latencies and high power consumption [1]. On the other hand, integrated optics may enable the continued scaling of capacity. Recent advances in silicon photonic components are paving the way for device integration for large optical systems. AWGR based interconnect architectures for HPC and datacenter, for example, LIONS (formerly DOS) [2] and Petabit [3] have drawn strong attention due to its wavelength routing capability and dense interconnectivity. They also benefit from the low latency and high throughput offered by optical wavelength parallelism. In particular, LIONS for datacenters uses single transmitter per node with tunable wavelength converters (TWCs) in front of the AWGR. The 1-by-k DEMUXs and k parallel receivers accommodate up to k concurrent packets. The contented packets enter an electrical loopback buffer and re-enter the AWGR [2]. This paper presents a new silicon photonic AWGR switch for on-chip communication in HPC systems [4]. Thanks to the short distance between nodes and the switch, the nodes communicate with the controller in the electrical domain, and the packets stored in the input queues are transmitted only upon receiving the grants. Therefore, it does not require electrical loopback buffer and TWCs as in the previous LIONS. In particular, multiple TX/RX pairs per node reduce contention and increase throughput. Simulation results show that, even with only two TX/RX pairs, the end-to-end latency and throughput significantly improves compared to its electrical counterpart, especially at high input load. Finally, we demonstrate a silicon photonic AWGR switch prototype with four TX/RX pairs on each node. The device footprint is approximately 1.2 mm by 2.4 mm.

2. Switch architecture and performance study

Fig. 1(a) The proposed interconnect architecture for chip-scale high performance computing systems and performance study on (b) end-to-end latency and (c) throughput of the proposed architecture with 8 nodes under uniform random traffic.

Fig. 1(a) shows the proposed chip-scale interconnect architecture. AWGR with fixed k_t transmitters and k_r receivers at each node provides us with a scalable solution. k_t ring modulators at each node generate the data packets. An off-chip comb generator provides the N wavelengths required by the cyclic frequency AWGR for wavelength routing. The low-speed electrodes on the input rings are for aligning the ring resonances with the AWGR passbands while the high-speed pads are for data modulation. Only low speed pads are required for the output rings. The receiver reads the information on the de-multiplexed channel after Optical-to-Electrical (O/E) conversion. Due to the presence of multiple TX/RX ring pairs, one node can communicate with multiple other nodes simultaneously. We use an architecture-level simulator to analyze the performance of the proposed switch based on an 8×8 AWGR. We assume uniform random traffic with a packet size of 1024B. The inter-arrival time between two packets follows Bernoulli distribution. The line rate is set at 10 Gb/s. We define the maximum offered load as N × line rate, which
does not change with $k_t$ and $k_r$. There is a 16- KB input buffer for each transmitter. There is a 10-ns guard time between any consecutive packets. Configuration with $k_t = 1$ and $k_r = 1$ represents the electrical switch based on input queuing (IQ) crossbar topology. Fig. 1(b) and (c) shows that multiple TX/RX pairs provide significant boost in performance due to the increased statistical multiplexing and the enhanced instantaneous rate at each AWGR inputs and outputs. We observe zero packet loss and 100% throughput for all the cases where $k_t \geq 2$ and $k_r \geq 2$.

3. Design and characterization of an 8×8 switch prototype with $k_t = 4$ and $k_r = 4$

Fig. 2(a) shows the device fabricated at the IME foundry using OPSIS-IME library and our AWGR design. Unlike the configuration in Fig. 1(a), the ring in Fig. 2(b) did not have the additional waveguide. Instead, for simplicity, the fabricated chip utilized edge couplers to couple external light sources. However, this deviation from the original design (Fig. 2(e)) degrades the on-off extinction ratio of the ring under reverse bias, so we adopted carrier injection method instead for data modulation at 0.3 Gb/s without pre-emphasis [5]. We injected external light at 1064 nm to generate carriers for resonance tuning. OpSIS-IME now offers designs with electrical tunability larger than one FSR. Fig. 3(a) illustrates the experimental setup for the 1-by-4 routing demonstration. A Bias Tee combines a 2^3-1 PRBS sequence at 0.3-Gb/s from a Pulse Pattern Generator (PPG) with the DC bias to drive the input ring. With proper alignment, eventually the modulated light enters the corresponding photo-detector. A second Bias Tee extracts the O/E converted high-speed signal for eye diagram and BER measurements. Fig. 3(b) shows the measured transmission spectra of the 8×8 200-GHz AWGR. Fig. 3(c) shows measured BER for the ring pairs in Fig. 3(a) at 0.3 Gb/s. We observe < 1-dB penalty between different pairs. Achieving 1E-10 BER requires ~4-dBm power, which can be reduced by improving the fiber coupling loss (~3dB), AWGR loss (~9dB) and waveguide loss (3 dB/cm).

4. Conclusion

We propose a scalable chip-scale optical switch for HPC systems by leveraging wavelength routing characteristics of the AWGR and silicon photonic integration. We prove its feasibility through simulation results and experiment.

5. References