Microsecond Optical Switching Network of Processor SoCs with Optical I/O

S. Moazeni¹, J. Henriksson¹, T. J. Seok¹, M. T. Wade², C. Sun³, M. C. Wu³, and V. Stojanović¹

¹Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720, USA
²Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea
³Ayar Labs, Inc., San Francisco, CA 94107, USA
Email: smoazeni@berkeley.edu

Abstract: We demonstrate an OCS network with microsecond switching time between processors featuring monolithic ring-resonator-based WDM-compatible optical I/Os. This solution can solve EPS scalability challenge and enable novel architectures for emerging disaggregated and heterogeneous data-center and HPC.


1. Introduction

Exponential growth of Internet traffic and the emergence of cloud-based computationally intensive applications such as Machine Learning require high-bandwidth and low-latency internal networks in data-center and high performance computers (HPC). Emergence of fast nonvolatile memories like FastFlash and 3D Xpoint [1] with microseconds of access latency further increases the need for fabric solutions with microsecond latencies. However, conventional electrical packet switching (EPS) networks cannot be leveraged for this aim in a cost and energy efficient way [2]. Hybrid electro-optical switching networks such as Helios [3] have been proposed to tackle this challenge, but their usage is currently limited at the core of the network due to their large switching time (10s of ms) and cost of 3D mirror-based MEMS switches. Wavelength selective switching (WSS) 2D mirror-based MEMS switches have been deployed to reduce the switching time to 10s of microseconds [4], however, the number of ports for this type of devices are limited and not suitable for high-radix networks. In this work, we are demonstrating a microsecond optical circuit-switching (OCS) network (Fig. 1a) between processor system-on-chips (SoC) with optical I/O [5] for the first time using MEMS-actuated adiabatic coupler switches [6] in silicon photonic technologies to solve the EPS scalability and energy-efficiency issues for the next generations of data-centers and HPCs. Silicon photonics is a promising technology to realize compact, low-cost, and energy efficient optical switches and transceivers in CMOS-compatible processes. Here, we are using ring-resonator based wavelength division multiplexed (WDM) optical transceivers implemented in an unmodified 45nm SOI CMOS process along with dual-core processors and other electrical blocks on a single die as the network nodes.

Achieving microsecond-scale switching in conjunction with high-bandwidth and energy efficient WDM photonic SoCs also enables many possibilities for imminent disaggregated and heterogeneous data-center and HPC architectures [7]. Data movement and memory access dominates today’s computing latency and energy, which can be significantly improved by utilizing monolithically integrated optical transceivers in computing SoCs and interconnecting them via high-radix and fast silicon photonic switches at both intra- and inter-cluster levels.

Fig. 1: (a) Block-diagram of optical switching network demonstration, (b) Silicon photonics MEMS switch chip, (c) Electro-optically packaged processor SoC.
2. Microsecond Optical Switching

A 4×4 version of the 64×64 MEMS silicon photonics switch reported in [6] was redesigned for O-band and fabricated at a commercial foundry with 150nm lithography process. The switch is based on highly scalable matrix architecture with two orthogonal sets of bus waveguides and low-loss multimode interference (MMI)-based crossings at the intersections. The switching operation is performed via two movable orthogonal adiabatic couplers and a bended waveguide located above each intersection. The adiabatic couplers and the MEMS structures are made out of polysilicon, while the bus waveguides are implemented on 220nm thick silicon-on-oxide substrate. In the ON state, MEMS actuators bring down the couplers so that light is first coupled from one bus waveguide to the coupler and then, after a 90-degree bend, is coupled down to the output bus waveguide via second coupler. In the OFF state the adiabatic couplers are located far above from the bus waveguides so that light in input bus waveguides propagate uninterrupted.

The switch chip is electrically wire-bond packaged on a PCB (Fig. 1b). On-chip vertical grating couplers (VGC) have been used to couple the light in/out of a linear fiber. Measured fiber-to-fiber optical loss is -8.8dB at ON state, of which -7.6dB is due to input/output VGCs. Because of the vertical gap-closing design of the MEMS actuators, the switch can do digital switching and exhibits a bi-stable region between 26V and 36V (Fig. 2a). Measured extinction ratio (ER) limited by the measurement and it can be larger than 60dB [6]. The switching drive waveform in Fig. 2b is used to pre-bias the switch at switching thresholds during transmission to minimize the switching time (1.2μs).

![Switching waveform](image)

Fig. 2: (a) Measured transfer characteristics, (b) Measured transient response of a switch unit cell.

3. Processor SoC with WDM Optical I/O

The processor SoC used in this work contains over 70 million transistors and 850 photonic components that can work together to provide logic, memory, and interconnect functions [5]. Electro-photonic integration is implemented monolithically in a 45nm SOI CMOS process without any modifications to the native process (‘zero-change’ approach) [8]. We used WDM-compatible transceiver rows on these chips to establish a 5Gb/s (per channel) link in the network at 1280nm wavelength. Each row has 11 ring-resonator modulator/detectors providing ~3.2THz free spectral range (FSR) suitable for dense WDM (DWDM). Modulator drivers are running at 1.2V nominal voltage of this process so that the ring-modulator sees ~0.5 to -0.7 voltages. Receiver is also implemented using SiGe resonant photo-detectors (PD) followed by a TIA and double data-rate (DDR) samplers on the electronic side. Fig. 1c shows fully packaged SoC, where it is electrically flip-chip packaged on a PCB and optically attached to fiber arrays. Thermal tuning and locking of the ring-resonances have been addressed through the integration of optical power inside the ring-cavity and tuning an embedded micro-heater via a closed-loop feedback [9]. This tuning is done initially in the network and rings can be kept locked to corresponding wavelengths while switch is off by integrating the small optical leakage power of switch output. Transmitter, receiver, and thermal tuning circuitry energy-efficiencies are 170fJ/b, 560fJ/b, and 144fJ/b, respectively at 5Gb/s.

4. System Demonstrations

Fig. 1a shows the block-diagram of the optical switching network for a single channel operation composed of 2 transceiver chips and a 4×4 switch. Transceiver chips are configured via control FPGAs and an off-chip clock source. A tunable laser source at 1280nm wavelength with 7dBm optical power is shared between two transmitters. Modulated lights are coupled into switch inputs and one of switch outputs is amplified via semiconductor optical amplifier (SOA) with ~23dB gain and fed into the receiver. All ring-resonances at transmitters and receiver side were tuned to this wavelength and this is done using embedded microheaters inside the rings with 3.8μW/GHz thermal efficiency. Light is coupled in and out of the transceiver chips using VGC with 5dB loss. Switch chip fiber-
to-fiber loss is ~9dB. One switch is kept ON, while the other switch element is controlled by the signal shown in Fig. 3a generated by an arbitrary waveform generator and an amplifier to provide voltages required for actuation. Measured switching delay is ~1.2µs as expected from switch characteristics. Optical output of the switch is monitored via a 10/90 splitter and a commercial TIA before coupling into the receiver and is shown in two states with one or both transmitters operating at 5Gb/s in the pattern mode (Fig. 3b, c). Modulator’s ER is 3dB with 4dB insertion loss (IL). ER/IL can be further enhanced by critically coupling the ring-modulator for this wavelength since the current design was optimized for 1180nm. Receiver’s bit error rate (BER) for both of its samplers is measured and shown in Fig. 3d at 2.75dBm average input optical power. Achieving BERs of below 1E-10 in this link eliminates the need of forward error correction (FEC) and consequently improves the latency.

Fig. 3: (a) Switch control signal and optical output, (b, c) transmitters and (d) receiver performances at 5Gb/s.

5. Conclusion

We have demonstrated the first microsecond-switching OCS network between two optical transmitters and a receiver. The number of switch chip ports can be scaled up to 64 and beyond with negligible optical loss penalties [6] unlike other microsecond switching silicon photonic mechanisms (such as ring-resonator or MZI-based switching) and also sub-μs switching times are achievable by optimizing post-processing and MEMS actuator design further. Moreover, data-rate per channel and optical energy-efficiency of SoCs can be drastically improved using advanced modulation techniques [10] and optimized VGC (sub-2dB loss) [11] and PD designs. This energy efficient and low-cost silicon photonic solution can be utilized to meet the latency and bandwidth demands of future large-scale data-centers and HPCs networks.

This work was supported by DARPA POEM program award HR0011-11-C-0100 and HR0011-11-2-0021.

6. References