Broadband Silicon Photonic Packet-Switching Node for Large-Scale Computing Systems

Wenjia Zhang, Lin Xu, Qi Li, Hugo L. R. Lira, *Student Member, IEEE*, Michal Lipson, and Keren Bergman, *Fellow, IEEE*

***Abstract*— We present a broadband packet-switching node that utilizes silicon photonic technology. The node design uses a silicon microring for switching functionality, leverages in-flight header processing for arbitration, and has a tunable driving circuit for thermal-effect mitigation. Moreover, these integrated microring switches are capable of scaling to tremendously high port counts in a compact area, which are attractive for data-center networks. We experimentally characterize the extinction ratio of the switch for varying packet durations, interarrival times, and driving voltages and demonstrate an error-free routing of 10-Gb/s wavelength-striped packets with lengths of up to 1536 ns. We further study the resonance thermal drifting for long-hold-time packet switching through carrier injection and show thermal-effect mitigation using a pre-emphasized gating signal.**

***Index Terms*— Electrooptic devices, interconnections, microresonators, packet switching.**

I. Introduction

THE emerging proliferation of cloud computing applications has created the need for more powerful data centers (DCs). These DCs are projected to interconnect 100,000 or more ports with approximately one petabit-per-second aggregate bandwidth [1]. Current DCs are interconnected by electronic networks based on commodity Ethernet switches and routers. However, existing networks have not been able to satisfy the bandwidth requirement without significantly increasing power consumption and cost [2]. Recent research advances in photonic devices have enabled the design of advanced photonic network architectures, which are attractive solutions for alleviating both the communication bottlenecks and power consumption in DC networks [3, 4].

The switching granularity and network flexibility of the DC node have a significant influence on scalability and system performance. Current proposed photonic interconnection architectures commonly prefer optical circuit switching, in which wavelength-parallel optical paths are provisioned and manipulated by a separate electronic control plane [3, 4]. These networks favor large messages so that the circuit setup overhead is sufficiently amortized over the duration of the message transmission. However, for those applications that require divergent and unpredictable communications between servers, optical packet-switching (OPS), which encodes broadband payloads and headers onto one or multiple wavelengths, is preferred and has been recognized as a viable approach [5].

In this letter, we propose the usage of a microring resonator (MRR) for the wavelength-striped packet-switching node design. The MRR is an optical switching device that offers high bandwidth density, low power consumption, nanosecond-scale response, and compatibility with the complementary metal- oxide-semiconductor (CMOS) process [6–9]. In particular, these integrated MRR switches are attractive for DC networks since they can scale to tremendously high port counts in a compact area. To the best of our knowledge, this is the first packet-switching node structure that utilizes silicon photonic technology. This node design could support in-flight header processing and has a thermally-aware driving technology that can stabilize switching performance for long duration packets.

In this letter, we develop an experimental testbed by combining customized digital control logic and a MRR switch. First, we experimentally characterize the extinction ratio of the MRR device for varied packet durations, inter-arrival times, and driving voltages. This characterization is used to determine optimal operation parameters for switching performance. Then, we demonstrate routing of 10-Gb/s wavelength-striped packets with durations of up to 1536 ns, achieving error free operation (10^{-12} bit-error rates). Finally, mitigation of thermal effects using the pre-emphasis technique is also experimentally confirmed for 3.84 μs optical packets.

II. Node Architecture Overview

Wavelength-striped packet-switching architectures transmit header information on several dedicated wavelength channels (one bit per wavelength). This allows for an extremely simple detection scheme and results in minimized latency and network-processing power consumption [5]. Considering the features of MRR switch, our packet-switching node (Fig. 1(a)) filters out single optical header wavelength for electronic processing at each hop and generates the appropriate driving signal to mitigate the thermal effect by using the logic circuit and a digital-to-analog converter (DAC).

Manuscript received November 11, 2011; revised January 27, 2012; accepted January 30, 2012. Date of publication February 3, 2012; date of current version March 28, 2012. This work was supported in part by the National Science Foundation Engineering Research Center for Integrated Access Networks (subaward Y503160).

W. Zhang is with the Department of Electronic Engineering, Columbia University, New York, NY 10027 USA, and also with the Institution of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: wz2207@columbia.edu).

L. Xu, Q. Li, and K. Bergman are with the Department of Electronic Engineering, Columbia University, New York, NY 10027 USA (e-mail: lx2140@columbia.edu; ql2163@columbia.edu; bergman@ee.columbia.edu).

H. L. R. Lira and M. Lipson are with the School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853 USA (e-mail: hll35@cornell.edu; ml292@cornell.edu).

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LPT.2012.2186796

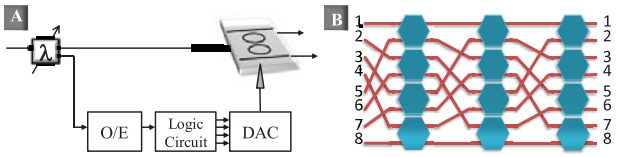

Fig. 1. (a) OPS node consisting of a wavelength filter, receiver, logic circuit, digital–analog converter (DAC), and switch device. (b) Example of Omega network topology.

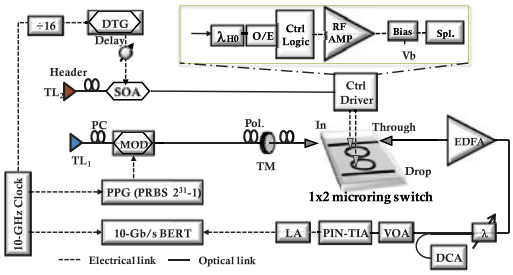

Fig. 2. Experimental setup using the silicon electrooptic MRR switch, header-control logic circuit, and tunable driving circuits.

The switching of light between the through port and drop port of a single MRR is accomplished by injecting electrical carriers through an embedded PIN diode, which results in resonance tuning of the device. This tuning includes a fast blue shift from the free carrier dispersion effect and a slow red shift from the thermal effect [9]. The thermal effect becomes significant when long duration packets are switched, and consequently the switching performance is degraded due to resonance drifting. In order to mitigate the thermal effect and thereby obtain the best switching performance, a control logic circuit parses the header for DAC to generate the driving signal (Fig. 1(a)), which is determined through device characterization. In future integrated iterations, the added processing delay could be compensated by on-chip optical buffers [10].

Architectures utilizing this switching node can be scaled to support large interconnection networks. An 8×8 Omega network (Fig. 1(b)), each stage filtering out one header wavelength, is one example network topology [5]. Future integration of optical functionalities and electronic circuits into a single package is an attractive way to effectively overcome the challenges such as cost, power consumption and footprint.

III. Experimental Setup

Figure 2 shows the experimental setup for our packet-switching node design. The 1×2 microring switch used in this experiment is a cascade double-ring switch. The characterization and switching performance of this device for data rates of up to 40-Gb/s have been previously reported in [7–9]. Two continuous tunable lasers (TL), centered at 1551.05 nm and 1555.75 nm, generate the continuous-wave (CW) light and are subsequently modulated with data and header information, respectively. A pulse pattern generator (PPG) drives the data wavelength with a 10-Gb/s, $2^{31}-1$ pseudorandom bit sequence (PRBS) pattern. The modulated optical data passes through a fiber polarizer, and then is coupled into the silicon waveguide using a tapered fiber. A data timing generator (DTG) gates the header wavelength channel

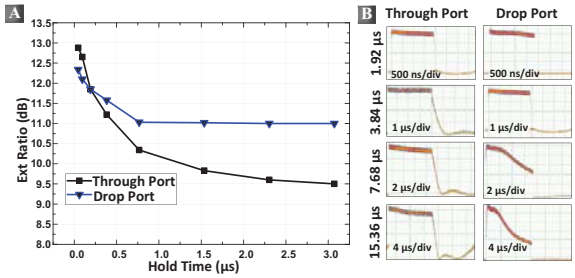

Fig. 3. (a) Output optical packet-extinction ratio from both the through and drop ports for different hold times. (b) Through-port and drop-port waveform distortion arising from long hold times in the MRR switch.

by switching a semiconductor optical amplifier (SOA). The detection of the header wavelength channel is performed by a control driver module, which contains a filter, photo-detector, logic circuit, RF amplifier, bias tee, and splitter. The gated CW signals have an average rise time of 1.66 ns and fall time of 0.85 ns for the through port, defined as the time between 10% and 90% of the signal amplitude. The optical signal egressing from the chip is amplified by an erbium-doped fiber amplifier (EDFA), filtered, and then passes through a variable optical attenuator (VOA). The optical signal is passed into a PIN receiver with transimpedance amplifier (TIA) and limiting amplifier (LA), and examined on a bit-error-rate tester (BERT). A common 10-GHz clock source synchronizes the DTG, PPG, BERT and DCA. A data communications analyzer (DCA) is used to examine the temporal performance of the link.

IV. Results

A. Switching Functionality Characterization

To optimize the active switching performance, we first measure the extinction ratio of output packets egressing from the through and drop ports with different packet sizes, inter-arrival times, and driving voltages. The packet size and arrival times are modeled with a square wave with varying hold times and duty cycles. The data wavelength channel is configured to be on resonance with the MRR in the passive state. This enables the signal to be switched to the through port when the driving voltage signal is set high and to the drop port when the driving voltage signal is set low. We first switch the device actively using optimized driving voltages for varying hold times. As shown in Fig. 3(a), the packet extinction ratio for both the through and drop ports are above 10 dB with packet hold times of approximately 1 μs, confirming the feasibility of packet-switching for message sizes of 1518 bytes (maximum Ethernet frame size) in a 10-Gb/s link. Furthermore, packet waveform distortion for hold times of up to 15 μs are recorded (Fig. 3(b)). The progressive decrease in extinction ratio and increase in waveform distortion indicates significant switching performance degradation, arising from both the free carrier effect and thermal effect inside the MRR switch.

We then pass a CW light through the microring switch while keeping a constant 1536 ns period gating signal but with varied duty cycles from 30% to 70%. The results (Fig. 4) show significant extinction ratio variation when using different driving voltages for each duty cycle, especially for the through port. The proposed tunable driver allows an optimized driving

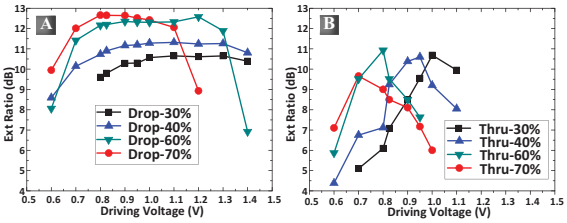

Fig. 4. Extinction ratio of the gating signal with variable driving voltages and duty cycles (30%, 40%, 60%, and 70%) for the (a) drop port and (b) through port. The waveform utilizes a 1536-ns period.

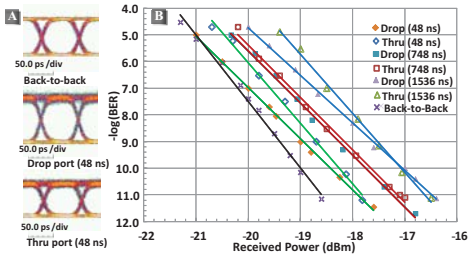

Fig. 5. (a) Eye diagrams for back-to-back and after chip with 48-ns hold time. (b) BER curves for packetized optical signals with 48-, 748-, and 1536-ns hold times, from both outputs of the silicon photonic microring resonator, as well as bypassing the chip in the back-to-back case.

signal for both the through and drop ports even when the traffic pattern varies during operation.

B. Switching Performance Using Optimized Control Driver

In order to validate packet-switching performance with different packet lengths, the 10-Gb/s encoded optical signal is gated by the control driver with 48 ns, 748 ns and 1536 ns packet hold times. Eye diagrams for the chip output with a hold time of 48 ns are shown in Fig. 5(a). Back-to-back measurements are also recorded by replacing the chip with an attenuator tuned to the equivalent insertion loss of 17 dB. BER measurements on the packetized data for both outputs are recorded in Fig. 5(b). We observe error-free operation and approximately 0.5, 1.3 and 2-dB power penalty for 48 ns, 748 ns and 1536 ns packets for both outputs, respectively. This penalty is attributed to the switching instability induced by thermal effects during the relatively long hold time.

C. Thermal Effect Mitigation Using Pre-Emphasis

To mitigate distortions caused by thermal effects as packet hold times exceed 3.84 μs, we emulate the logic circuit and DAC in Fig. 1(b) with an arbitrary waveform generator (AWG). We replace the original square gating signal with a pre-emphasized waveform from the AWG to mitigate the thermal effect, and subsequently to produce improved switching performances. Fig. 6(a) illustrates the spectrum of the resonant response of the device during active switching. As the heat induced by current during the active state causes resonance red shift (t = 0), at the instantaneous moment of switching from active to passive (t = t_1), the resonance continues red shifting and output power from through port at 1551.05 nm will shift from A to B, causing a negative pulse dip shown in the inset of Fig. 6(b). The gradual cooling of the MRR

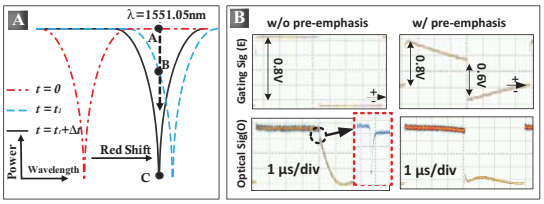

Fig. 6. (a) Spectrum of the resonant response during switching. (b) Thermal-effect mitigation using the pre-emphasis.

will subsequently shift the resonance and thereby the output power returns back to C again. We balance the thermal effect and free carrier dispersion effect by gradually reducing current to shorten the resonance shift when switching from active to passive. The gating waveform we used in this approach is presented in Fig. 6(b) and the improved switching optical signal is shown as well.

V. Conclusion

We have proposed the packet-switching node structure that utilizes silicon photonic technology and experimentally characterized MRR-based switching node operation for varying packet durations, inter-arrival times and driving voltages. We demonstrated error free transmissions of 10-Gb/s optically addressed signals of hold times up to 1536 ns for both the through and drop ports of a MRR switch. The degradation from thermal effects was mitigated using a pre-emphasized driving signal. These experimental results validate the viability of using MRRs as an integrated switching component for optical packet-switching nodes in large-scale computing systems.

Acknowledgment

The authors would like to thank J. Chan for his support in the experiment and writing.

References

[1] A. Vahdat, M. Al-Fares, N. Farrington, R. N. Mysore, G. Porter, and S. Radhakrishnan, "Scale-out networking in the data center," *IEEE Micro*, vol. 30, no. 4, pp. 29–41, Jul./Aug. 2010.

[2] M. Al-Fares, A. Loukissas, and A. Vahdat, "A scalable, commodity data center network architecture," in *Proc. ACM SIGCOMM*, 2008, pp. 1–12.

[3] N. Farrington, *et al.*, "Helios: A hybrid electrical/optical switch architecture for modular data centers," in *Proc. ACM SIGCOMM*, 2010, pp. 1–12.

[4] G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. E. Ng, M. Kozuch, and M. Ryan, "C-through: Part-time optics in data centers," in *Proc. ACM SIGCOMM*, 2010, pp. 1–12.

[5] A. Shacham, B. G. Lee, and K. Bergman, "A wideband, non-blocking, 2×2 switching node for a SPINet network," *IEEE Photon. Technol. Lett.*, vol. 17, no. 12, pp. 2742–2744, Dec. 2005.

[6] B. G. Lee, A. Biberman, J. Chan, and K. Bergman, "High-performance modulators and switches for silicon photonic networks-on-chip," *IEEE J. Sel. Topics Quantum Electron.*, vol. 16, no. 1, pp. 6–22, Jan./Feb. 2010.

[7] A. Biberman, *et al.*, "Broadband silicon photonic electrooptic switch for photonic interconnection networks," *IEEE Photon. Technol. Lett.*, vol. 23, no. 8, pp. 504–506, Apr. 15, 2011.

[8] L. Xu, J. Chan, A. Biberman, H. L. R. Lira, M. Lipson, and K. Bergman, "DPSK transmission through silicon microring switch for photonic interconnection networks," *IEEE Photon. Technol. Lett.*, vol. 23, no. 16, pp. 1103–1105, Aug. 15, 2011.

[9] H. L. R. Lira, S. Manipatruni, and M. Lipson, "Broadband hitless silicon electro-optic switch for on-chip optical networks," *Opt. Express*, vol. 17, no. 25, pp. 22271–22280, Dec. 2009.

[10] F. Xia, L. Sekaric, and Y. Vlasov, "Ultracompact optical buffers on a silicon chip," *Nature Photon.*, vol. 1, no. 1, pp. 65–71, 2007.