

A Low-Latency Optical Switch Architecture using integrated μm SOI-based Contention Resolution and Switching

G. Mourgias-Alexandris^{*a}, M. Moralis-Pegios^a, N. Terzenidis^a, M. Cherchi^b, M. Harjanne^b, T. Aalto^b, K. Vyrsoinos^c, N. Pleros^a

^aSchool of Informatics, Aristotle University of Thessaloniki, 54624, Greece

^bVTT Technical Research Centre of Finland, Espoo Finland

^cSchool of Physics, Aristotle University of Thessaloniki, 54624, Greece

* mourgias@csd.auth.gr

ABSTRACT

The urgent need for high-bandwidth and high-port connectivity in Data Centers has boosted the deployment of optoelectronic packet switches towards bringing high data-rate optics closer to the ASIC, realizing optical transceiver functions directly at the ASIC package for high-rate, low-energy and low-latency interconnects. Even though optics can offer a broad range of low-energy integrated switch fabrics for replacing electronic switches and seamlessly interface with the optical I/Os, the use of energy- and latency-consuming electronic SerDes continues to be a necessity, mainly dictated by the absence of integrated and reliable optical buffering solutions. SerDes undertakes the role of optimally synergizing the lower-speed electronic buffers with the incoming and outgoing optical streams, suggesting that a SerDes-released chip-scale optical switch fabric can be only realized in case all necessary functions including contention resolution and switching can be implemented on a common photonic integration platform. In this paper, we demonstrate experimentally a hybrid Broadcast-and-Select (BS) / wavelength routed optical switch that performs both the optical buffering and switching functions with μm -scale Silicon-integrated building blocks. Optical buffering is carried out in a silicon-integrated variable delay line bank with a record-high on-chip delay/footprint efficiency of $2.6\text{ns}/\text{mm}^2$ and up to 17.2 nsec delay capability, while switching is executed via a BS design and a silicon-integrated echelle grating, assisted by SOA-MZI wavelength conversion stages and controlled by a FPGA header processing module. The switch has been experimentally validated in a 3×3 arrangement with 10Gb/s NRZ optical data packets, demonstrating error-free switching operation with a power penalty of $<5\text{dB}$.

Keywords: Integrated Delay Lines, Echelle Grating, Thick-SOI Platform, On-chip Buffer, All-optical Switching, Feed-forward Buffer, Disaggregated Data Center, Optical Interconnects

1. INTRODUCTION

The on-going effort to address the exponential requirements of Data Centers (DC) networks in terms of bandwidth and energy consumption has led to the successful development of different silicon photonic platforms, already utilized in a variety of optical transceivers. Among the well-established platforms, the μm -Silicon-on-Insulator, also known as “Thick-SOI”, stands out for its relaxed fabrication tolerance and polarization insensitive structures [1], while its commercial employments have led to up to 200 Gb/s optical modules, with an effective line rate of 50 Gb/s [2]. The incompetence of current electronic switches to keep up with these line rates, the increased pressure of disaggregated architectures on the DC network, in terms of latency and throughput, along with the aforementioned technology advances, creates the required momentum to mitigate switching from the electrical to the optical regime [3]. However, for such a transition to take place several critical functionalities, such as buffering and switching, need to be integrated in the same technology platform to allow for high-radix, compact switches.

One of the technology platforms, utilizing the Thick-SOI technology, meeting the required credentials for such an undertaking is VTT’s μm -SOI platform [4]. Exploiting this platform’s advantages, integrated delay lines featuring record high delay-to-footprint values have already been utilized for on-chip buffering [5], while μm -radius Euler bends allow for compact on chip waveguide routing [6]. Moreover, integrated wavelength selective switches, realized by Echelle gratings have already been demonstrated in emerging applications scenarios like, CPU multi-socket routing [7], while

the co-integration of Semiconductor Optical Amplifier (SOA) switches, has validated the wavelength conversion operation allowing nano-second switching operation [8].

In this paper, we experimentally demonstrate the concurrent use of a μm -SOI delay line bank, with a μm -SOI based 3x3 Echelle grating router, as building blocks of an integrated optical switch architecture, that can be expanded to hundreds of ports with higher order Echelle gratings routers. The delay line bank is utilized for buffering, offering variable delays of 6.8ns up to 17.2ns, while the 3x3 Echelle grating, provides wavelength routing assisted by SOA-MZI optical switches. The control plane functionalities are carried out in an FPGA responsible for header processing and packet scheduling. The optical switch architecture employed, features a hybrid Broadcast-and-Select/Wavelength routing scheme, capable of coping with the high throughput and low latency requirements in emerging disaggregated DC applications [9], while its building blocks can in principal be integrated in the same chip. Error free operation is obtained for 10Gb/s data, revealing <5dB power penalty.

2. SILICON BASED OPTICAL SWITCH ARCHITECTURE

Fig. 1(a) and (b) illustrate the two chips, i.e the μm -SOI delay lines and the 3x3 μm -SOI Echelle grating router, that we utilized for our proposed architecture. The integrated delay lines are based on 4 μm thick, 1.875 μm wide strip waveguides, separated by 2 μm gaps, utilizing Euler bends with an effective bending radius of 69 μm , for a 90° bend. The waveguides exhibit a loss of about 0.13dB/cm [10], while a record high induced delay versus footprint efficiency is achieved at 2.6ns/mm², compared to other integrated waveguides featuring higher than 10ns delays [4].

The 3x3 Echelle gratings were fabricated in the same μm -SOI platform, employing a compact grating length of only 150 μm . For this specific implementation, the resulting chip footprint was increased due to packaging limitations that imposed the matching of the output waveguides pitch with the typical pitch of fiber arrays at 250 μm . An illustration of the final packaged device is presented in Fig. 1 (b). The gratings were designed to have a free spectral range of 100nm and channel spacing of 15nm, while the 3-dB bandwidth was measured to be 4.5 \pm 0.3nm, closely monitoring the design parameters. The chip has been characterized in more detail in [7], were it was utilized as a wavelength selective switch for a multi-socket CPU interconnect.

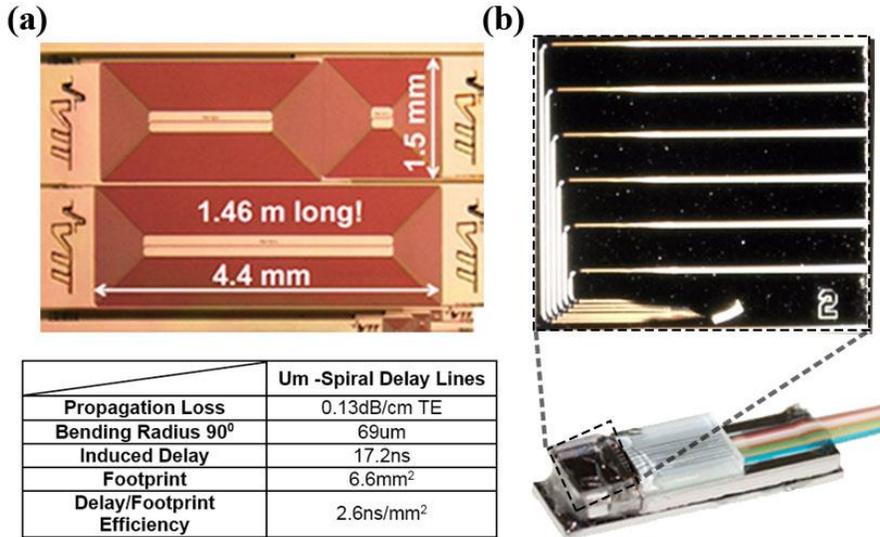


Figure 1. (a) Microscope image of μm SOI delay lines, (b) Microscope image of μm 3x3 Echelle grating along with packaged device.

A version of the optical switch architecture presented in [8], utilizing the aforementioned um-SOI delay lines and 3x3 Echelle grating router, is illustrated in Fig. 2, where a 9x9 configuration is achieved with the input ports organized in 3 Switch planes. Each Plane groups 3 input signals, and employs a hybrid Broadcast and Select (B&S) scheme, where the Broadcast function is realized by optical splitters (Stage A) and the Select function is performed by SOA-MZI based Tunable Wavelength Converters (TWC) controlled by an FPGA controller. At the same time, buffering functionality, utilizing um-SOI integrated delay lines, provides contention resolution when congestion arises. Each buffering block, denoted as Integrated Delay-Line Bank (DLB), offers distributed buffering at an Output Echelle grating basis, by the use of optical combiners and respective multiplexers (MUX) and de-multiplexers (DE-MUX) (Stage B). In Stage C, an extra TWC followed by the 3x3 Echelle grating and a tunable laser (TL) enables routing of the optical packets to any output. It should be noted, that the scale of the switch is determined by the radix of the employed Echelle grating and can be in principle expanded to e.g. a 256x256 switch, by using a 16x16 Echelle grating.

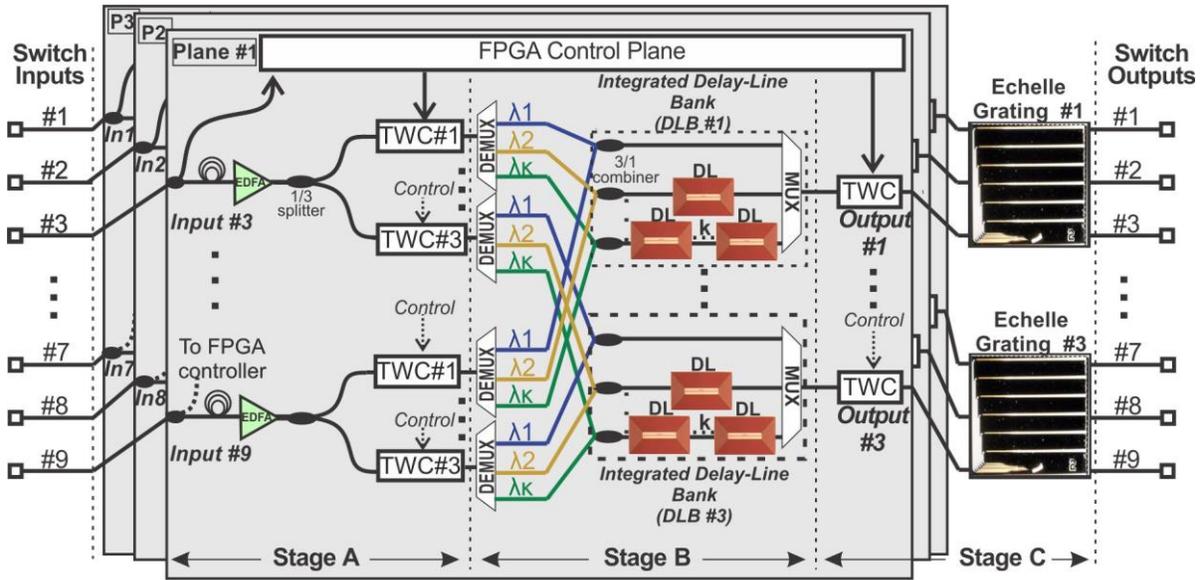


Figure 2. Layout of the proposed architecture utilizing μ m-SOI delay lines and Echelle gratings.

3. EXPERIMENTAL SETUP AND RESULTS

To evaluate the transmission characteristics of the proposed switch architecture, when integrated um-SOI building blocks are utilized, we experimentally demonstrate a single plane of the 9x9 switch, utilizing a 17.2ns inducing integrated um-SOI delay line for packet buffering and a 3x3 Echelle grating. Fig. 3 presents the corresponding experimental setup. Network traffic was emulated via an Altera Stratix V FPGA, utilizing an SFP and XFP transceiver, for generating two input packet streams, while the same board was used to perform the control plane's functionalities, such as header processing and routing control.

The two input packet streams are composed of two 103 bit-long 10.3125Gb/s NRZ data packets and one dummy packet, used to retain synchronization at the receiver side, while an intra-packet guard-band of 35 bits is used. Part of the incoming optical power is fed to the FPGA controller for header processing, via a 70/30 coupler. A 1/3 splitter is then employed to emulate the losses of the 9x9 switch, while a 456ns inducing fiber delay line was used to delay the optical signal, during processing time, and finally an EDFA was employed to compensate for splitting loss. Both XFP and SFP data streams were split into two identical signals before arriving, as control signals, at input ports of In#1WC1 and In#2WC1 respectively. Simultaneously, two laser beams at 1549.6nm and 1550.4nm were fed into a LiNbO₃ modulator, to produce 111bit, 28MHz envelopes, subsequently multiplexed in an AWG and fed as In#1WC1 input signal into port G. The aforementioned procedure was used for generating the input signals of In#2WC1. Resulting optical signals from ports C and B of In#1WC1 and In#2WC1 respectively, were then demultiplexed in separate AWGs and combined again

using 1/3 combiners, based on the delay path they are traversing. The delay path incorporating the 17.2ns integrated um-SOI delay line, utilizes an additional SOA, biased with a holding beam to reduce its noise figure. The resulting signals are then multiplexed again in an AWG, and are injected as control signals of Out1WC after properly amplified in an EDFA and filtered in a 5nm Band Pass Filter. Finally, a Tunable Laser Source (TLS), is injected into a modulator to produce 105bit, 28 MHz envelope signals and fed at Out1WC port D as input signal, at the appropriate wavelength (1535nm, 1551.5nm, 1565.7nm). Finally, an integrated 3x3 Echelle grating is used at the In#2WC1 output port G, allowing routing to any of the Echelle output ports. An EDFA amplifier, along with a bandpass filter, is employed at the Echelle's Output, to compensate for the chip's optical losses. All the WCs utilized the differentially biased scheme, allowing up to 40Gb/s operation [11], while the required electrical signals were generated by the FPGA.

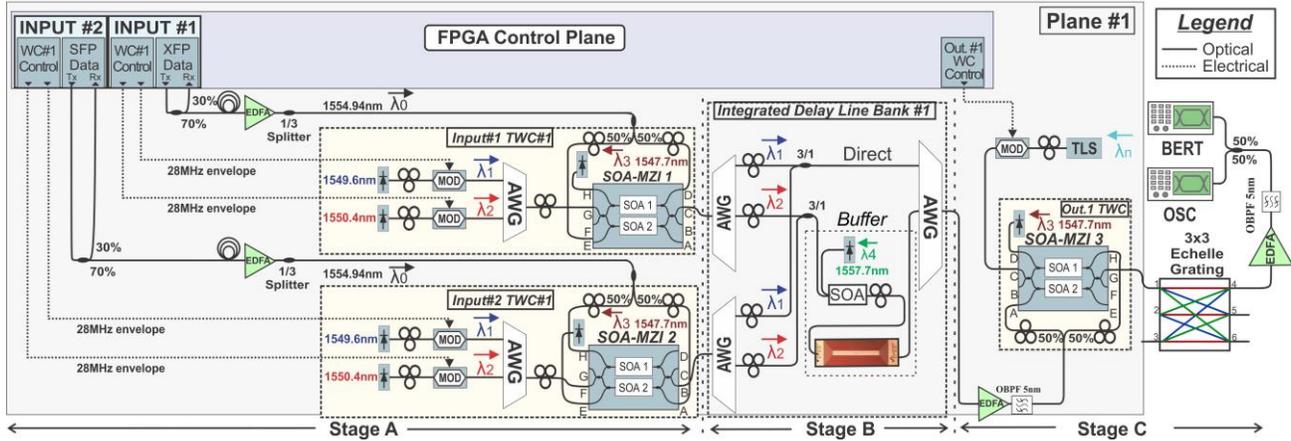


Figure 3. Experimental setup of a 2x3 plane utilizing one optical buffer and a 3x3 Echelle grating Router

Figure 4 illustrates the experimental results, obtained when the two data streams are fed into the switch. Each stream comprises two packets (A, B) and (C, D), while the corresponding traces and eye diagrams are depicted in Fig. 4 (a) and (b) respectively. For demonstration purposes, we emulated a congestion scenario, where incoming packets from both input ports contend for the same output destination. The FPGA enabled control plane, generates the required control signals, and at the DLB output the packets are re-arranged with new resulting order (A-C-B), depicted in Fig. 4 (c), while packet D is dropped to be re-transmitted in the next timeslot. Exploiting the wavelength routing characteristics provided by the 3x3 Echelle grating, packets (A) and (B) were switched to Output port 1 and 2 of 3x3 Echelle grating respectively. The control signals of Out1WC, are depicted in Fig. 4 (d) while the captured traces, along with the respective eyes, at the Echelle's outputs are illustrated in Fig. 4 (e) and (g). Optical spectra at Echelle's Output 1 and 2 are depicted at Fig. 4 (f) and (h) respectively. Finally, BER measurements were performed for each packet at the Echelle's output, revealing error-free operation with a power penalty ranging from 3dB to 5dB. This variation in power penalty is due to the different noise characteristics of the two input MZIs and the variation in the effectiveness of the wavelength conversion for different input wavelengths.

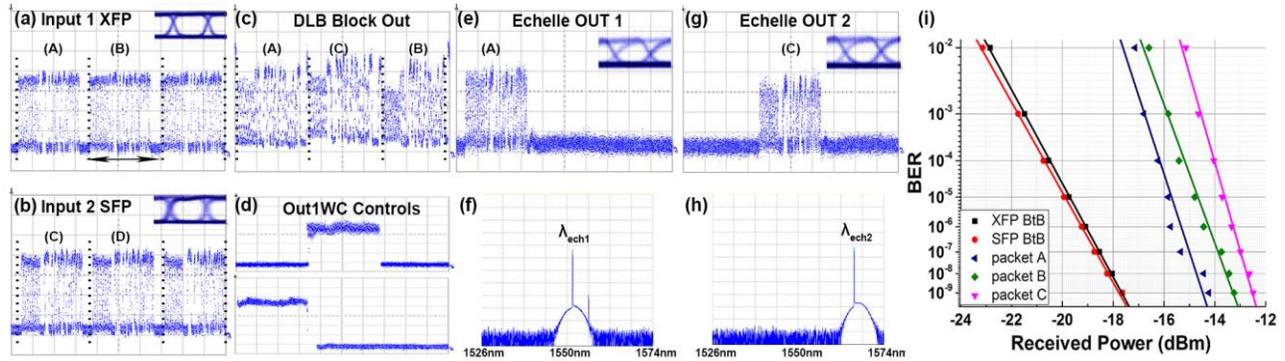


Figure 4. (a) Incoming XFP and (b) SFP data streams traces and eyes (c) Trace at Delay Line Bank Output (e),(f) Trace ,eye and spectra at Echelle Output 1 (g),(h) Trace ,eye and spectra at Echelle Output 2 (i) BER measurements.

4. ACKNOWLEDGEMENTS

This work was supported by the European Commission through projects H2020 projects ICT-STREAMS (Contract 688172) and L3MATRIX (Contract 688544).

REFERENCES

- [1] Feng, Dazeng, *et al.*, "High-speed GeSi electroabsorption modulator on the SOI waveguide platform," in *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 6, pp. 64-73, Nov.-Dec. 2013.
- [2] <http://ir.mellanox.com/releasedetail.cfm?ReleaseID=1018004>.
- [3] Q. Chen, *et al.*, "Hardware Programmable Network Function Service Chain on Optical Rack-Scale Data Centers," in *Proceedings of OFC*, Th2A.35, San Diego (2017).
- [4] T. Aalto, *et al.*, "Silicon photonics for optical connectivity: Small footprint with large dimensions," in *IEEE Proceedings of Optical Interconnects*, pp. 40-41, San Diego, 2015
- [5] M. Moralis-Pegios, *et al.*, "On-Chip SOI Delay Line Bank for Optical Buffers and Time Slot Interchangers," in *IEEE Photonics Technology Letters*, vol. 30, no. 1, pp. 31-34, Jan.1, 1 2018
- [6] M. Cherchi, *et al.*, "Dramatic size reduction of waveguide bends on a micron-scale silicon photonic platform," *OpEX* 21, 17814-17823 (2013)
- [7] G. Dabos, *et al.*, "Thick-SOI Echelle grating for any-to-any wavelength routing interconnection in multi-socket computing environments," in *Proceedings of SPIE* Vol. Vol. 10109. (2017)
- [8] D. Fitsios, *et al.*, "Dual SOA-MZI Wavelength Converters Based on III-V Hybrid Integration on a μm -scale Si Platform," in *IEEE Photonics Technology Letters*, vol. 26, no. 6, pp. 560-563, March 15, (2014)
- [9] M. Moralis-Pegios, *et al.*, "A low-latency high-port count optical switch with optical delay line buffering for disaggregated data centers," p. 10538-4, accepted at SPIE (2018)
- [10] M. Cherchi, *et al.*, "Low-loss spiral waveguides with ultra-small footprint on a micron scale SOI platform," in *Proceedings of SPIE*, 899005, March 2014
- [11] D. Apostolopoulos, *et al.*, "A SOA-MZI NRZ Wavelength Conversion Scheme With Enhanced 2R Regeneration Characteristics," in *IEEE Photonics Technology Letters*, vol. 21, no. 19, pp. 1363-1365, Oct.1, 2009.