

Multi-wavelength selective crossbar switch

AKHILESH S. P. KHOPE,^{1,*} MITRA SAEIDI,¹ RAYMOND YU,² XINRU WU,³ ANDREW M. NETHERTON,¹ YUAN LIU,¹ ZEYU ZHANG,¹ YUJIE XIA,¹ GAREY FLEEMAN,¹ ALEXANDER SPOTT,¹ SERGIO PINNA,¹ CLINT SCHOW,¹ ROGER HELKEY,¹ LUKE THEOGARAJAN,¹ ROD C. ALFERNESS,¹ ADEL A. M. SALEH,¹ AND JOHN E. BOWERS¹

¹Electrical and Computer Engineering Department, University of California, Santa Barbara (UCSB), CA 93106, USA

²The Institute of Optics, University of Rochester, 480 Intercampus Drive, River Campus, Rochester, NY 14627, USA

³Department of Electronic Engineering, The Chinese University of Hong Kong, Hong Kong, China

*akhileshk@ece.ucsb.edu

Abstract: Here we demonstrate an 8x4 multi-wavelength selective ring resonator based crossbar switch matrix implemented in a 220-nm silicon photonics foundry for interconnecting electronic packet switches in scalable data centers. This switch design can dynamically assign up to two wavelength channels for any port-port connection, providing almost full connectivity with significant reduction in latency, cost and complexity. The switch unit cell insertion loss was measured at 0.8 dB, with an out-of-band rejection of 32 dB at 400 GHz channel separation. All the ring resonator heaters were thermally tuned, with heaters controlled by a custom 64-channel DAC driver. Detailed measurements on the whole switch showed standard deviation of 2 dB in losses across different paths, standard deviation of 0.33 nm in resonant wavelength and standard deviation of 0.01 nm/mW in ring heater tuning efficiency. Data transmission experiments at 40 Gbps showed negligible penalty due to crosstalk paths through the switch.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Optical interconnection networks used for intra- and inter-chip communications in future exascale systems can achieve improved scalability, i.e., higher throughput, lower latency and increased energy efficiency, compared to their electrical counterparts [1]. Optical switches with application in datacenters can be divided into two categories, fat pipe switches and wavelength selective switches [2–16]. Fat pipe switches switch the entire Wavelength Division Multiplexed (WDM) signal between input/output (I/O) ports. Multiple variants based on broadband Mach-Zehnder interferometers have been reported in a Silicon Photonic platform [17], Global Foundries 90-nm platform [18] and Silica Platform in [19] and a ring resonator based fat pipe switch is reported in [20]. We demonstrate thermally-tuned switching, which is compact and provides μs switching, which matches datacenter requirements. Nanosecond switching can be achieved with electro-optic tuning, at the cost of larger switch footprint and higher optical loss. Wavelength selective switches with single wavelength connectivity based on Arrayed Waveguide Grating Router (AWGR) [21], cascaded microring crossbar switch [22], modular switch and select architecture [23], switch with multi-casting functionality [24] have been reported. A high radix on chip architecture is proposed in [25].

Multi-wavelength selective switches promise greater flexibility in connection patterns between I/O ports as compared to both single wavelength connectivity and fat pipe switches. In such switches, it is possible to setup connections using any combination of the input wavelengths. We proposed a multi-wavelength selective switch in [26, 27]. Here, we demonstrate and analyze an optical interconnect switch architecture operating in the wavelength-time domain, and having interconnection capability ranging from static all-to-all wavelength connectivity (which is suitable

for systems exchanging short messages) to on-demand microsecond-scale dynamically allocated multiple-wavelength connectivity (which is suitable for systems exchanging longer message sizes). For large enough number of wavelengths per port, the switch can simultaneously enable both connectivity modes, thus resulting in efficient, low-latency operations for systems with both short and long messages. This makes the switch suitable for a diverse set of applications ranging from interconnecting multiple processing nodes for High Performance Computing (HPC), to interconnecting individual servers or racks of servers in data centers.

Multiple N input, N output and M wavelengths/port ($N \times N \times M$) switching approaches have been demonstrated [28–30], but they either require too many ring resonators for similar connectivity or have a higher insertion loss. The main difference in our $N \times N \times M$, with a 2 wavelength connection approach as compared to the $N \times N \times M$ crossconnect approach proposed in [28] is that the signal passes through multiple off-resonant ring resonators instead of through drop ports of multiple rings in the $N \times N \times M$ approach. We measured an off resonance loss of 0.17 dB which is much smaller than the drop loss of 1.9 dB. Also the total number of rings in our approach, are much smaller than $N \times N \times M$, as in [28]. In [29], $4 \times 4 \times 4$ switch was reported, which could be used as a building block in a bigger Benes [31] or Clos network [32]. The $N \times N \times M$ Benes switch requires $M(N \log_2(N) - N/2)$ rings for $N \times N \times M$ Benes switch as compared to $2N^2$ in our approach. An $8 \times 8 \times 8$ MEMS wavelength selective switch is demonstrated [30], by using 8 single wavelength switches in conjunction with 1×8 demultiplexers and 8×1 multiplexers. In the above approaches, the number of ring resonators and hence power consumption and footprint is much higher than our approach, but at the same time latency is marginally lower as shown in a later section on contention resolution. Hence, the above systems are over designed on the basis of latency, footprint, and power consumption.

The present crossbar switch configuration makes it possible to optimize the interconnection network dynamically (down to ≤ 100 microsecond reconfiguration time), to respond to slow or fast variations in the workloads and the message-exchange rates in these systems. This switching time optimization requires a centralized control mechanism based on tracking the traffic demand in real time. This is not a unique requirement for our proposal; rather, earlier pioneering works on introducing slow [33] or fast [34] optical time-domain circuit switching in data centers also require centralized control to optimize the interconnection network. In fact, centralized control is currently in general use in many mega data centers [35].

Another important application of our switch in data centers is that it can enable the creation of flexible, scaled-out, low-latency, leaf-spine data center realization that does not require an increase in the overall size or the number of fiber ports of the associated electronic packet switches [36]. The advantage of this approach is that it reduces cabling complexity as the datacenter is scaled to a higher radix and also requires fewer electronic switches which results in a lower overall power consumption.

The main impact of this work is as follows:

1. For a $N \times N$ crossbar switch with M wavelengths per input port, M rings per crosspoint are not required when buffers are present at transmitters. Near full connectivity can be obtained with just two rings per crosspoint. Diminishing reduction in latency is observed as the number of ring resonators are increased beyond 2 per crosspoint. A huge saving in the number of switching elements, tuning power consumption and path loss is also observed for the case of two ring resonators as compared to M ring resonators.
2. We demonstrate an 8×4 switch fabricated in a 220 nm Silicon Photonic platform and designed a driver PCB to characterize the switch.
3. Loss, resonant wavelength and power efficiency are measured for every ring resonator in the switch. This corresponds to measurements on 64 second order resonators and total 256

measurements to extract the tuning parameters. We measured a mean on-chip path loss of 9.94 dB with standard deviation of 2.08 dB for the fabricated 8x4 switch.

- We measured impairments due to crosstalk sources by injecting 40 Gbps data through multiple input ports of the switch and measured the Bit Error Rate (BER) of the signal.

2. Architecture

2.1. System design

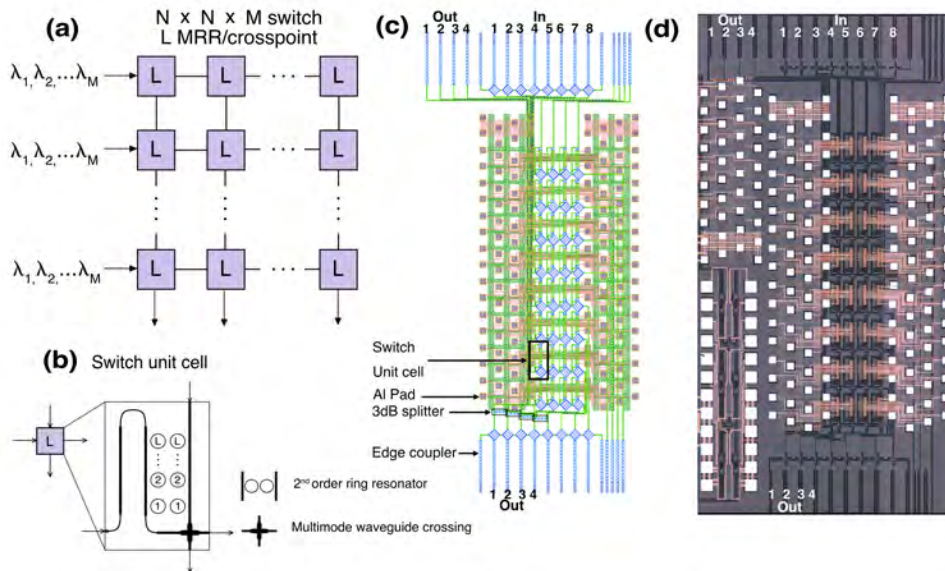


Fig. 1. (a) Schematic of $N \times N \times M$ switch with L MRR (Micro Ring Resonators) per crosspoint (b) Switch unit cell (c) Layout of a 8x4, $L=2$ switch (d) Die Shot of the switch with I/O marked

In Fig. 1(a), we present an $N \times N$ crossbar switch with switching blocks at every intersection. The transmitter injects a WDM signal with M wavelengths in each input port. The ' L ' blocks at the crossings in Fig. 1(a) consist of wavelength-selective components that can switch up to L wavelengths from an input port to an output port. The wavelengths that are not switched pass through these blocks ' L '. For proper operation, the switching elements in any two L blocks should not be tuned to the same wavelength for any row or any column. Thus, wavelengths that have been switched by an L block go straight down to the receiver through the intermediate L blocks below the switching L block. In our design, we use 2nd-order micro ring resonators and a multi-mode waveguide crossing to implement L block in Fig. 1(b). The choice of second order ring resonator [37, 38] instead of first order ring resonator [39] is due to its flat top transmission spectra which minimizes crosstalk through alternate switch paths. The ring resonator drop and through transfer spectrum is very sensitive to coupling coefficients and a slight deviation from the coupling coefficients can result in a non flat drop transfer spectrum for higher order ring filters. It is difficult to satisfy the critical coupling condition for higher order filters due to higher number of distinct coupling coefficients. An experimental demonstration of these non ideal filters is reported in [29]. We thus limited our design choice to second order ring resonator. Each transmitter can transmit up to M wavelengths. Integrated quantum dot lasers are suitable for this application due to their better performance at higher temperatures are reported in [40–42].

We use centralized arbitration, which requires a single shared controller with a separate low-speed control network connecting each transmitter to the controller. A traffic buffer is placed at each of the transmitter sites. System operation is assumed to be synchronous, i.e., all transmitters have access to a common clock and for synchronization of transmission and switching, where time is divided into equal timeslots. These low-bandwidth synchronization signals could be sent optically as part of a control plane. Within each timeslot, the transmitters inform the controller with the number of requested connections and their respective destinations and gets ready for transmission at the start of the next timeslot. The controller then runs the wavelength assignment and contention resolution algorithm (described in the next subsection), assigns wavelengths to each request, informs the transmitters about the wavelength assignments of the various connections, and tunes the appropriate microrings placed in the various switching blocks. The requests that cannot be transmitted are placed in the buffer and transmitted when a wavelength becomes available in a subsequent timeslot. If a connectivity configuration within any timeslot satisfies the following constraints:

1. Any input port can transmit a maximum of M wavelengths, and a maximum of M can be received at any output port. This is due to the placement of M modulators at each transmitter and M photodiodes at each receiver. (All the transmitters use the same M wavelengths.)
2. Maximum of L wavelengths need to be switched from any input port to any output port. All wavelengths at an output port must be distinct.

then the switch can be shown to be nonblocking [26]. The switch is re-arrangeably non-blocking, but is effectively equivalent to fully non-blocking operation as it is reconfigured synchronously between data transmission time slots. For any number of ports (N) and for one wavelength connectivity ($L=1$), no ring tuning is required and the system switching capability is similar to an arrayed waveguide grating router. Contention resolution is required in our switch as resources (wavelengths) are shared between multiple input ports, with the method described in the next subsection.

The impact of switching time on the switch is measured by the ratio of switching time and communication time [43]. A higher switching time results in a lower throughput for a given communication time. For shorter message sizes and for the case of thermo optic tuning, switching time is on the order of 10's of μs . These switching times, exhibit satisfactory throughput for average packet sizes > 100 kB [43]. For smaller message sizes, a switching time on the order of few ns or 10's of ps with a different switching mechanism will be required.

Fig. 1(c) shows the layout of an 8×4 switch. Input/Output (I/O) edge couplers are placed on the same side at a standard pitch of $127 \mu m$ packaging with a fiber V-groove array. We placed a 1×2 3 dB splitter at each output so that waveguides could be routed for testing with I/O on opposite side as well as on same side. Input and output from the same side of the chip is used for testing with V-groove fiber array. This switch was fabricated in a 220 nm thick Silicon Photonic (SiP) platform, with a switch footprint of 1.92 mm \times 4.15 mm as shown in the die shot in Fig. 1(d). All devices in the system were implemented using the AIM Photonics Process Design Kit (PDK) [44].

2.2. Contention resolution

In packet switching terminology, this crossbar switch configuration can be described as an input queued switch, and virtual output queues are used for 100% throughput [45]. In our system simulations, connection requests follow a uniform random distribution and are generated within each timeslot for transmission in the subsequent timeslots. To distinguish between request arrival times within a timeslot, we divide the timeslot into k parts. The probability that a request is

generated by each transmitter within each of these k parts is given by $p = \Lambda/k$, where Λ is the mean rate of request generation per timeslot for each transmitter. We choose a large enough k such that the request generation per timeslot effectively follows a Poisson distribution with mean Λ .

The connectivity matrix is defined as the matrix of requests that satisfies the conditions for nonblocking mentioned in the previous subsection. The buffer matrix represents the requests placed in the buffer. In both of these matrices, the rows and columns correspond to the input and output ports, respectively; and each element corresponds to a list of times at which the packets are generated within a timeslot. For lowest latency, we would like to maximize the number of requests transferred from the buffer matrix to the connectivity matrix. The procedure to generate the connectivity matrix such that it satisfies a nonblocking condition is given as follows.

The connectivity matrix is initially empty and all the new requests are placed in the buffer matrix. The oldest M packets from each row and column of the buffer matrix are transferred to the connectivity matrix such that each input-output link supports less than L packets. At this stage, a heuristic algorithm to maximize the number of packets transmitted is implemented. The requests in the connectivity matrix are now transmitted at the start of the next timeslot.

Note that, so far in this algorithm we have not checked the wavelength continuity of the links, i.e., whether or not an end-to-end wavelength can be assigned to each connection without encountering wavelength contention or dropping any requests from the connectivity matrix. If the connectivity matrix satisfies the two constraints given in the previous subsection, then it turns out that there exists an efficient, contentionless wavelength assignment algorithm associated with it and thus wavelength continuity can be automatically satisfied. This is similar to decomposition of a doubly stochastic matrix as described in [46] and is reported in [27].

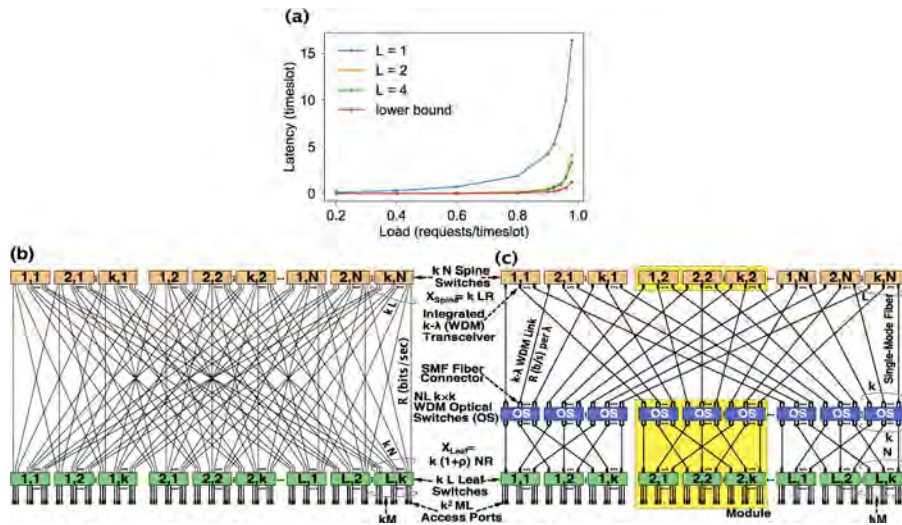


Fig. 2. (a) Latency of the timeslots is plotted against the load in the switching network. Load is defined as average requests/timeslot. (b) Leaf spine architecture without optical switch (c) Leaf spine architecture with WDM optical switch.

A traffic analysis with all to all uniform random traffic shown in Fig. 2(a) shows marginal reduction in latency for $L > 2$, i.e. more than 2 resonators per switch crossing. For $L = 1$, the system reduces to a single link, which corresponds to the classic $M/D/1$ queueing model Eq. (1) [47].

$$Latency(L = 1, M \geq N) = \Lambda / (1 - \Lambda) \tag{1}$$

In general, a lower bound on latency can be obtained by modeling the system as an M/D/m multi-server queue [47], with $m = L = M$, for which blocking occurs only due to the M wavelength limit at each transmitter. Note that the latency for $L = 2$ is only very slightly worse than the lower bound.

A possible application of the WDM switch, which is described in [36], is depicted in Figs. 2(b)–2(c), which shows two designs of a leaf spine data center with the same number of servers. Figure 2(b) shows the architecture with high-radix electronic packet switch and single-wavelength fibers. Figure 2(c) shows an architecture utilizing lower-radix electronic switches, as the individual fibers with single wavelength Fig. 2(b) are replaced with WDM links with integrated transceivers in Fig. 2(c), which shows a marked reduction in the number of fibers. In addition, the optical switches enable dynamical reconfiguration of the data center to respond to varying workloads.

3. Devices and packaging

The photonic switching chip was fabricated on 220 nm Si photonics platform at AIM Photonics foundry at SUNY Albany [44]. The waveguides used for routing were deep etched 400 nm with an expected loss of ≤ 2 dB/cm, although subsequent process optimization has improved this waveguide loss significantly. Edge couplers were used for optical I/O. Diced dies had a $100 \mu\text{m}$ trench etch for fiber coupling.

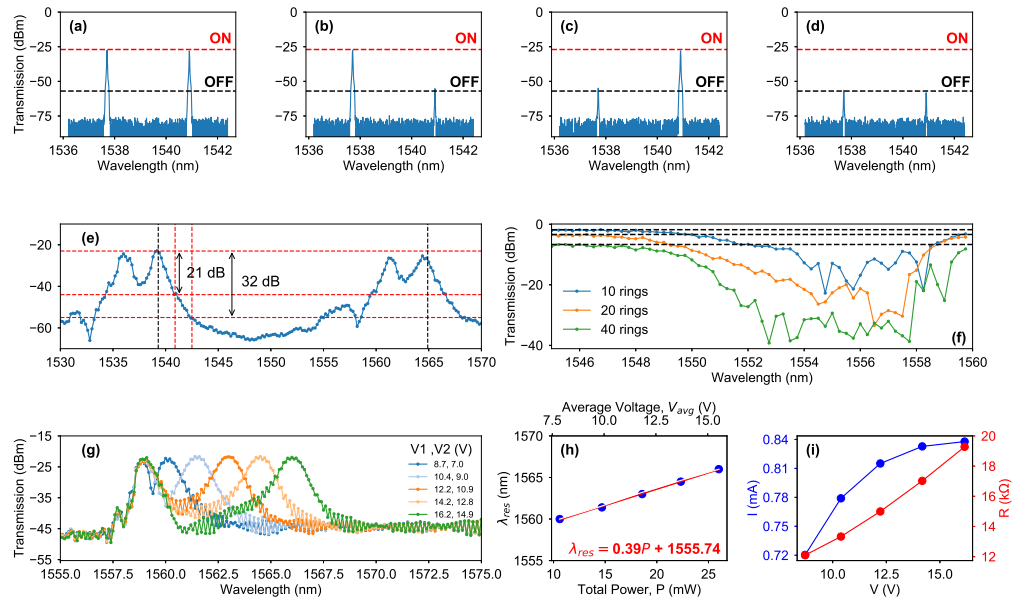


Fig. 3. Two λ switching is demonstrated in (a) λ_1 ON and λ_2 ON, (b) λ_1 ON and λ_2 OFF, (c) λ_1 OFF and λ_2 ON, (d) λ_1 OFF and λ_2 OFF, (e) Transmission spectra of a typical unit cell, the black lines correspond to one FSR = 25.6 nm and typical out of band rejection is 32 dB for 400 GHz spacing (f) Off resonance loss of 10, 20, and 40 ring resonators in series, (g) Tuning curve of a micro-ring resonator in a unit cell, (h) Tuning efficiency of 0.39 nm/mW is calculated from the tuning curve. Here we plot λ_{res} vs Total Power, P (mW) and Average voltage V_{avg} (V) applied to the second order ring resonator, (i) Heater I-V and R-V is plotted and shows that heater resistance changes monotonically with bias voltage.

Figures 3(a)–3(d) shows Optical Spectrum Analyzer (OSA) traces demonstrating two wavelength switching for a unit cell in the switch (Fig. 1(b)). For this experiment two wavelength channels at 400 GHz/3.2 nm spacing were injected with a fused silica splitter into the input port

1 and light is collected from output port 1. Here we observe a 32 dB extinction between 'ON' and 'OFF' signal. In the 'ON' state, voltages on the two heaters of a second order ring resonator are optimized for peak transmission at the output port, a few examples are given in Fig. 3(g). In the 'OFF' state, no voltage is applied to the ring resonators. In this measurement, ring resonators have a resonance wavelength of 1534.5 nm at zero bias voltage. This is the guard band where all the ring resonators filters are 'parked' when no connection is assigned to them. Two channels at 400 GHz spacing at 1537.7 nm and 1540.9 nm are used in the measurement. These channels correspond to a power consumption of 8.2 and 16.4 mW at a measured tuning efficiency of 0.39 nm/mW of power supplied for the second order ring resonator.

Figure 3(e) shows the filter spectrum at the output port 3 with input at port 3. The figure shows transmission spectra of a unit cell with channels spaced at 3.2 nm (400 GHz) channel spacing. We measure a 100 GHz optical 3 dB bandwidth of the filter and an out-of-band rejection of 21 dB/32 dB at 200 and 400 GHz channel spacing respectively. The ring resonators have a measured Free Spectral Range (FSR) of 25.6 nm. The PDK second order ring resonators are capable of both electro-optic tuning and thermal tuning. For electro optic tuning, the resonance wavelength was blue-shifted and attenuated with a loss of 3.6 dB/nm. Hence only thermal tuning was used for filter tuning. In Fig. 3(f), a through port measurement is done on 10, 20 and 40 second order ring resonators, only one resonator in each arrangement is connected to electrical pads to save space on the layout. The measured through loss was 0.17 dB per ring resonator and 1.9 dB for the drop port. This measurement was done to calculate the off resonance loss of a ring resonator filter, which is a crucial factor for scalability of switches. Figure 3(f) shows that many ring resonators in the series arrangement are not aligned to the same resonant wavelength. Standard deviation of ring resonance frequency as well as of free spectral range on a different process are well documented in [48].

Transmission spectra for the unit cell of the switch at different voltages are shown in Fig. 3(f), and in Fig. 3(h) we extract the tuning curve for the switch ring resonators. The tuning efficiency of the second order rings is 0.39 nm/mW with a measured resonance wavelength at zero bias of 1554.74 nm. The ring resonator tuning range was maximum 11 nm. I-V sweep of ring resonator heaters and resistance vs voltage is shown in Fig. 3(i). The figure shows that resistance changes from 12 k Ω to 20 k Ω as voltage is changed from 8 V to 16 V. The same plot also shows the change in current supplied. As a higher voltage is applied, resistance increases due to self heating and a detailed analysis of thermal effects on ring resonators is reported in [49]. A smaller drive voltage $V < 10$ V is preferred for CMOS and HBT processes. Thus a smaller ring resistance than the one measured is preferred for electronic photonic integration.

The drive voltage of the microring resonators could be reduced by reducing the resistance of the heaters on the microring resonators. This increases the power supplied to the heaters as $Power = V^2/R$. A side-effect of this approach is increased drive current as $I = V/R$. This might complicate electronic driver ASIC (Application Specific Integrated Circuit) design. Another way to reduce the drive voltage and thus power consumption is to use the substrate undercut technique [50]. This approach reduces the power consumption but at the same time increases the switching time. In our foundry, we did not have the option for substrate undercut.

The calculated power required to tune the microring by one FSR is 65 mW. With a better process and more vias per ring resonator we expect to increase the tuning range to one FSR. In this case, on an average all micro-rings in the 8x4 switch require 32.5 mW and thus the total average power consumption for the 64 micro-rings is 2.08 W. All paths (path is defined as a connection on a given channel) have a power consumption of 32.5 mW on an average and per bit power consumption per channel is 0.8125 pJ/bit at 40 Gbps Non Return to Zero (NRZ) and 0.325 pJ/bit for 50 Gbaud-PAM4. This is calculated as the ratio of average power consumption to tune a microring/the line rate.

The PCB was designed in Altium Designer and fabricated by Advanced Circuits. The PCB was

designed to have a PIC of up to 6 mm x 9 mm mounted on its surface with one edge overhanging the edge of the board by 250 μm in order to easily accommodate edge coupling to a V-groove fiber array. Standard design rules for the widths and separations of signal traces were maintained, but a more advanced mechanical drill diameter of 300 μm was selected to reduce wire bond pad size. A maximum of 72 signal and 20 ground wirebondable pads on the PCB were due to a constraint on the maximum wirebond length of 8.5 mm set by the vendor.

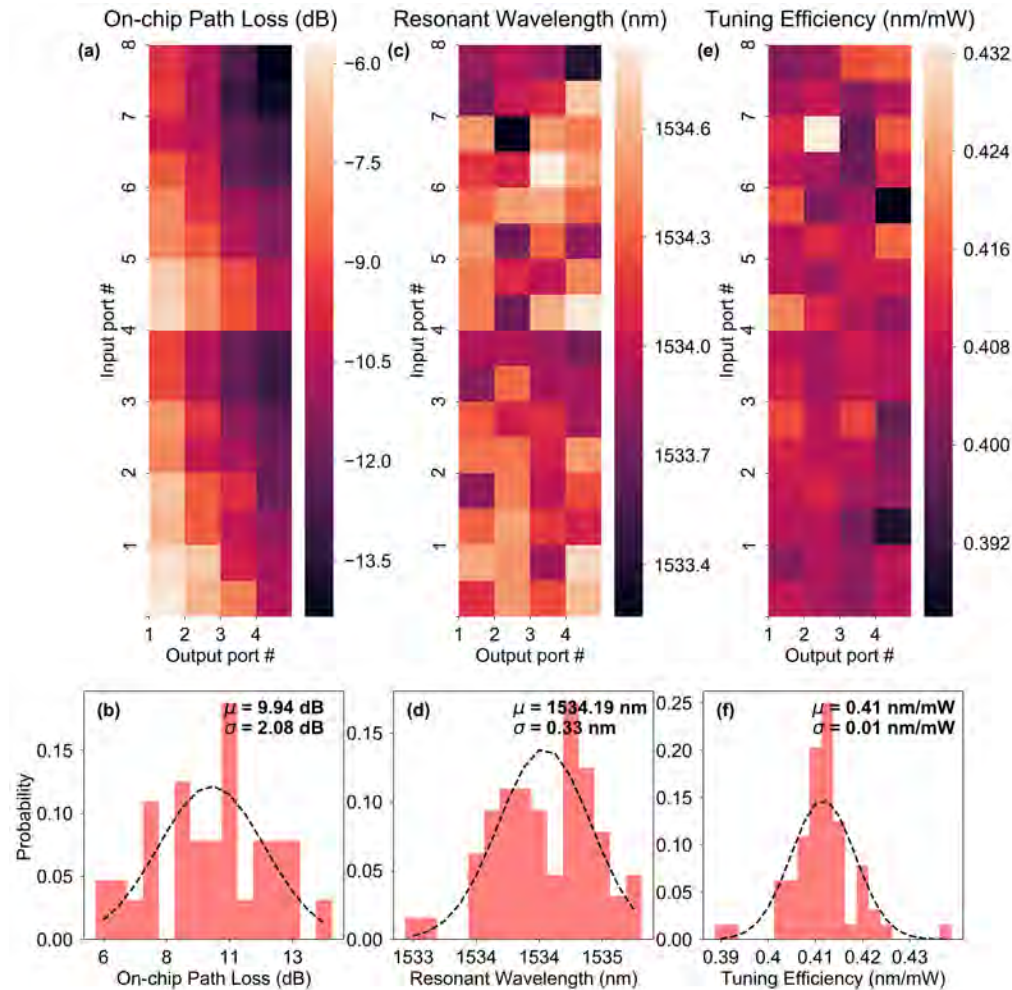


Fig. 4. Heatmaps of (a) On-chip Path Loss, (b) Resonant Wavelength and (c) Tuning Efficiency at different locations in the 8×4 switch, Corresponding histograms of (d) On-chip Path Loss, (e) Resonant Wavelength and (f) Tuning Efficiency.

In Figs. 4(a)–4(f) we characterize a single switch with 64 2nd-order ring resonators. In Figs. 4(a) and 4(b) we measure an on-chip path loss across different paths in the 8×4 with mean 9.94 dB and standard deviation of 2.08 dB. The minimum and max on-chip path loss in the switch path were 6 dB and 14 dB. The test setup loss including chip to fiber coupling loss was 13 dB. These values correspond to input 1 - output 1 and input 8 - output 4. This results in a waveguide crossing loss of 0.3 dB, through port loss of 0.3 dB and drop loss of 4.7 dB, measured in the C-band. In Figs. 4(c)–4(d) we measure a resonant wavelength with mean 1534.19 nm and standard deviation of 0.33 nm. In Figs. 4(e) and 4(f) we measure tuning efficiency with mean

and standard deviation of 0.41 nm and 0.01 nm.

We performed 256 total measurements to generate Figs. 4(a)–4(f), 4 measurements on each of the 64 ring resonators of the 8×4 switch. For each measurement the two voltages on the micro rings are optimized to get the best peak drop port transmission [51]. As reported in [52], the peak power is unimodal function of the heater voltages, i.e. has only one peak and the optimization is automated with Python. In such cases, it is faster to optimize the voltages through a genetic algorithm rather than a grid search on the two heater voltages. All the measurements in this paper were conducted with Random Mutation Hill Climbing (SHC) [53]. For fast convergence an approximate tuning curve is generated, and we assign random guess voltages to the ring heaters depending on the tuning curve and the wavelength of interest. As ring resonator resonant wavelengths follow a probability distribution given by Fig. 4(d), we apply noise sampled from standard normal distribution to the voltages and select the best pair of voltages to maximize the peak optical transmission. Noise can be scaled by a scalar, which can control the search space around the random guess voltage.

The heater driver was designed to control 64 heaters configured as four banks, each bank of 16 was driven by an off-the-shelf 16-channel DAC. A USB to Serial-Parallel-Interface (SPI) converter IC(FT4222) was used for controlling the DACs. DAC outputs were connected to photonic IC via a 64-wide cable. Individual control of the DACs was achieved by configuration words sent to the DACs via the SPI. Since 4 separate ICs were used the FT4222 served as the master while the DACs were configured as slaves. A Graphical-User-Interface (GUI) was designed in Visual Basic to allow programmability and ease of use.

Since the ring heaters require a 10V drive and a current drive of 2mA, the LTC2668 4-channel 12-bit DAC was chosen. The 12-bit DAC yields a 2.44mV LSB step, which is more than adequate for the wavelength control required by the PIC. An additional consideration for the DAC was the bandwidth. Accurate switching characterization of the heater necessitates the DAC's bandwidth be at least an order of magnitude faster than the time constant of the heater. The DAC settling time was 8 μs , yielding a rise time of about 1.5 μs , adequate for measuring the PIC heater which has a 15.92 μs rise time. Careful design of the PCB, including separating the analog and digital signals and ground planes, ensured low coupling noise due to the electronic switching.

Heater time constant was measured by using a photodetector (xpdv2120R) while stepping the heater driver. In the case of the double ring switching element, one ring is fixed at 10V or 0V (rise versus fall times) while the other ring was stepped via the heater driver. This procedure was repeated for different voltage differences and the switching time was recorded.

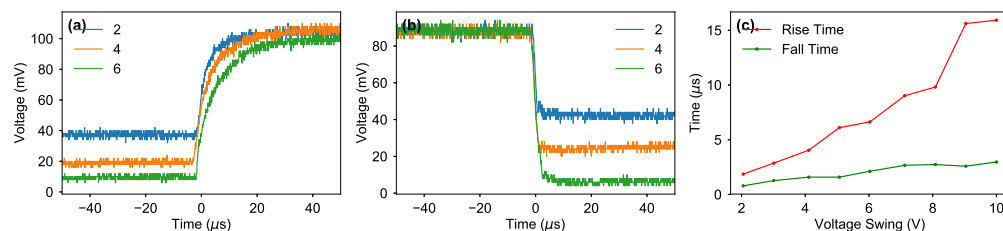


Fig. 5. (a) rise time for three different voltage swing, (b) fall time for three different voltage swing and (c) rise time, and fall time vs voltage swing.

Figure 5(a) shows the rise time for three different voltage swings. In this experiment, peak transmission of ring resonator was optimized at voltage of 10 V and voltage on one of the rings out of double ring resonator was changed by the voltage swing to 10 V. Figure 5(b) repeats the same experiment in reverse for the fall time. Figure 5(c) shows that Rise time and fall time change monotonically increase with the voltage difference swing applied to the ring resonators. From the figure we can see that the worst case switching time is 15.92 μs . A rise time from

10% to 90% is measured and shown for three different voltage differences. As we tune from one channel to another, higher wavelength separation results in higher switching time. This large signal switching is very different from the small-signal switching time reported in papers on ring resonators [54] where switching time $\tau < 5 \mu\text{s}$ was measured by applying a small electrical signal around the 3 dB point of the filter. This small-signal switching time does not apply to wavelength selective switches as tuning across all channels is required, hence the motivation for measurements in Fig. 5(c). If all the ring resonator are calibrated, drive voltages stored in a look-up table and all ring resonators are set to the their respective channels simultaneously, then the switching time of the switch is the worst case large signal switching time ($15.92 \mu\text{s}$) of an individual micro-ring resonator.

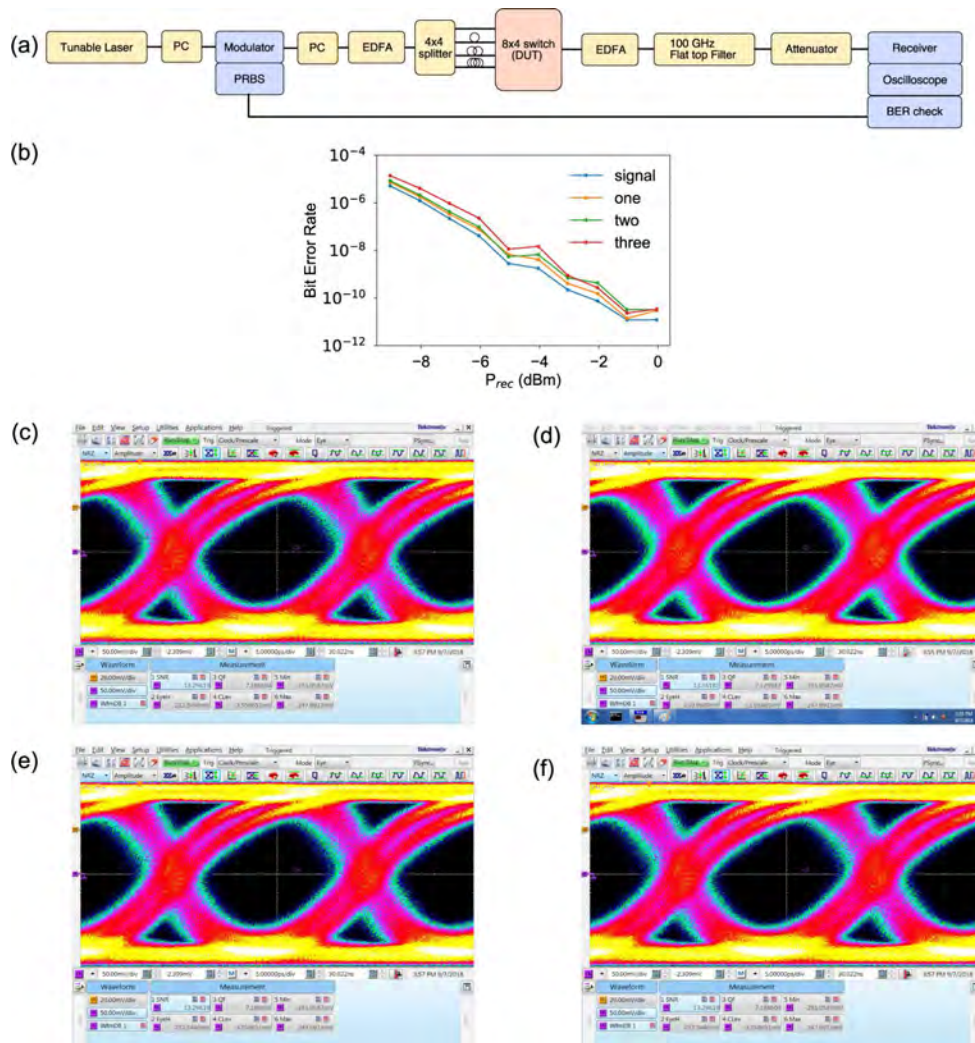


Fig. 6. (a) Test setup from multi channel BERT testing at 40 Gbps (b) Multi-wavelength crosstalk measurement for (signal) 0, 1, 2, and 3 crosstalk sources, Eye diagram for (c) no crosstalk source, (d) one crosstalk sources, (e) two crosstalk sources, and (f) three crosstalk sources

The Bit Error Rate (BER) test setup for multiple input crosstalk impairment testing is shown

in Fig. 6(a). In this experiment, modulated signal at 1537.6 nm is split with a 4×4 splitter and injected into input port 1, 2, 3 and 4. We calculated the coherence length of the Yenista Tunics T100S-HP source as 238 m corresponding to the 400 kHz linewidth. Fiber delays of 500m, 1.4 km and 6.4 km were used to decorrelate the optical phases of different input channels, which would otherwise occur due to the optical source coherence time. We used a 40 Gbps modulator from Sumitomo (T.MXH1.5-4OPD-ADC-LV) with measured insertion loss of 4 dB. An optical preamplifier and post amplifier was required to compensate for interconnect losses including losses from 4×4 splitter and modulator insertion loss. A Finisar high gain 43 Gb/s photoreceiver XPRV2022(A) was used for detection of transmitted data. Figure 6(b) shows a negligible penalty in BER when multiple crosstalk sources are added to the input port. Figs. 6(c)–6(f) correspond to eye diagrams with no crosstalk sources and sources at input 2, 3 and 4 respectively.

4. Discussion

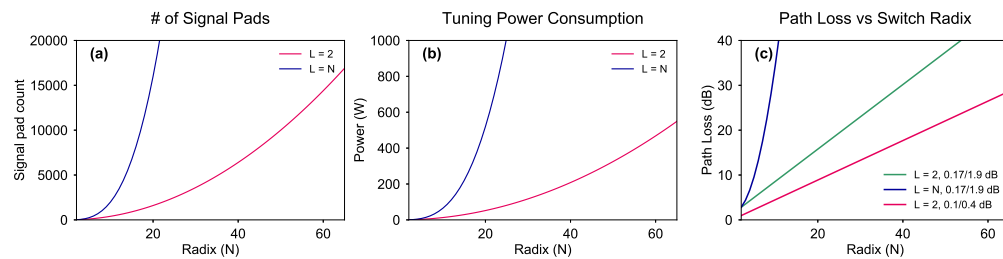


Fig. 7. (a) Number of signal pads which is proportional to complexity (the number of switching elements) (b) tuning power consumption (c) worst case path loss vs switch radix.

Number of signal pads which is proportional to complexity (the number of switching elements is plotted against the Radix (N) in Fig. 7(a). For an arbitrary L and N the *Complexity* = LN^2 . Figure 7(b) shows a plot of tuning power consumption vs Radix(N). Here a 30 nm/FSR tuning is assumed, where on average each ring is tuned by a half FSR. Thus *Tuning power consumption* = $15LN^2$. The scalability simulations of worst case path loss vs switch radix (N) for different L and off loss/drop loss values is shown in Fig. 7(c). In this plot, we observe 10.25 dB reduction in the insertion loss of the longest path between $L=2$, $N=32$ and 0.17/1.9 dB and 0.1/0.4 dB microring resonator losses. For $L = N$ and $N = 32$ we observe a 321.29 dB reduction in insertion loss while at the same time reducing power consumption and complexity and hence footprint of the switch as shown in Figs. 7(a) and 7(b).

The insertion loss of the switch is path dependent. The loss difference between the smallest loss path and the largest loss path needs to be within the receiver dynamic range. A constraint on the receiver dynamic range of 10 dB, sets the difference in minimum and maximum path dependent insertion loss of 10 dB. For an $L = 2$, 32×32 and 64×64 switch this corresponds to a loss of 0.16 dB and 0.08 dB per unit cell where both resonators are tuned off resonance. This corresponds to an off resonance ring loss of 0.065 dB and 0.025 dB assuming ultralow loss silicon photonic waveguide crossings are used [55]. This projected future loss is much lower than the measured off resonance loss of 0.3 dB from histograms in Fig. 4(b).

All the resonators voltages were optimized with an open loop control with photo detectors outside the chip package. A close loop control similar to [52] is possible if on chip wavelength selective taps are available. Another option is to control the switch by calibrating the ring heater vs temperature plot for all the ring resonators in the switch.

5. Conclusion

We analyzed an optical interconnect switch architecture operating in the wavelength-time domain, and having interconnection capability ranging from static all-to-all wavelength connectivity (which is suitable for systems exchanging short messages) to on-demand microsecond-scale dynamically allocated multiple-wavelength connectivity (which is suitable for systems exchanging longer message sizes). This architecture allows the creation of a flexible, scaled-out, low-latency, leaf-spine data center realization that does not require an increase in the overall size or the number of fiber ports of the associated electronic packet switches. This approach reduces cabling complexity and also requires fewer electronic switches which results in a lower overall power consumption. The crossbar switch configuration makes it possible to optimize the interconnection network dynamically (down to ≤ 100 microsecond reconfiguration time), to respond to slow or fast variations in the workloads and the message-exchange rates in these systems.

We demonstrated a 8×4 multi-wavelength selective switch implementation that can switch up to two wavelengths and gives almost full connectivity. A driver PCB with 64 channel DAC was built to characterize the switch. Large signal switching time measurements were also conducted on second order ring resonators and we measured a larger large-switching time as compared to the small-signal switching time reported in literature. Histogram measurements on the whole switch show a low standard deviation in path insertion losses, resonant wavelength without bias and tuning efficiency. 40 Gbps BERT measurement with multiple crosstalk sources showed negligible penalty due to incoherent crosstalk. We conclude that for multi-wavelength selective switch with port count greater than 32, a lower off resonance ring loss is required.

Funding

AIM Photonics

Acknowledgments

The authors thank AIM Photonics for funding this work and providing foundry services [56, 57]. The authors thank Takako Hirokawa, Alfredo Torres, Nicholas Volet, Tin Komljenovic and Paolo Pintus at UCSB, Ben Moss and Erman Timurdogan at Analog Photonics, and Jeremiah Hedding and Brett Attaway at SunyPoly for insightful discussions.

References

1. K. Bergman, S. Borkar, D. Campbell, W. Carlson, W. Dally, M. Denneau, P. Franzon, W. Harrod, K. Hill, J. Hiller, S. Karp, D. Keckler, Stephenand Klein, R. Lucas, M. Richards, A. Scarpelli, S. Scott, A. Snively, T. Sterling, R. S. Williams, and K. Yelick, "Exascale computing study: Technology challenges in achieving exascale systems," Def. Adv. Res. Proj. Agency Inf. Process. Tech. Off. (DARPA IPTO), Tech. Rep **15** (2008).
2. R. Stabile, A. Albores-Mejia, and K. Williams, "Monolithic active-passive 16×16 optoelectronic switch," *Opt. Lett.* **37**, 4666–4668 (2012).
3. R. Stabile, A. Rohit, and K. Williams, "Monolithically integrated 8×8 space and wavelength selective cross-connect," *J. Light. Technol.* **32**, 201–207 (2014).
4. A. Rohit, J. Bolk, X. J. Leijtens, and K. A. Williams, "Monolithic nanosecond-reconfigurable 4×4 space and wavelength selective cross-connect," *J. Light. Technol.* **30**, 2913–2921 (2012).
5. L. Qiao, W. Tang, and T. Chu, " 32×32 silicon electro-optic switch with built-in monitors and balanced-status units," *Sci. Reports* **7**, 42306 (2017).
6. L. Qiao, W. Tang, and T. Chu, " 16×16 non-blocking silicon electro-optic switch based on mach-zehnder interferometers," in *Optical Fiber Communication Conference*, (Optical Society of America, 2016), pp. Th1C–2.
7. D. Celo, D. J. Goodwill, J. Jiang, P. Dumais, C. Zhang, F. Zhao, X. Tu, C. Zhang, S. Yan, J. He, M. Li, W. Liu, Y. Wei, D. Geng, H. Mehrvar, and E. Bernier, " 32×32 silicon photonic switch," in *OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS), 2016 21st*, (IEEE, 2016), pp. 1–3.
8. L. Lu, S. Zhao, L. Zhou, D. Li, Z. Li, M. Wang, X. Li, and J. Chen, " 16×16 non-blocking silicon optical switch based on electro-optic mach-zehnder interferometers," *Opt. Express* **24**, 9295–9307 (2016).
9. Z. Lu, D. Celo, H. Mehrvar, E. Bernier, and L. Chrostowski, "High-performance silicon photonic tri-state switch based on balanced nested mach-zehnder interferometer," *Sci. Reports* **7**, 12244 (2017).

10. N. Dupuis, B. G. Lee, A. V. Rylyakov, D. M. Kuchta, C. W. Baks, J. S. Orcutt, D. M. Gill, W. M. Green, and C. L. Schow, "Design and fabrication of low-insertion-loss and low-crosstalk broadband 2×2 mach-zehnder silicon photonic switches," *J. Light. Technol.* **33**, 3597–3606 (2015).
11. R. Konoike, K. Suzuki, T. Inoue, T. Matsumoto, T. Kurahashi, A. Uetake, K. Takabayashi, S. Akiyama, S. Sekiguchi, K. Ikeda, S. Namiki, and H. Kawashima, "Lossless operation of soa-integrated silicon photonics switch for 8×32 -gbaud 16-qam wdm signals," in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, (IEEE, 2018), pp. 1–3.
12. B. G. Lee, A. Biberman, P. Dong, M. Lipson, and K. Bergman, "All-optical comb switch for multiwavelength message routing in silicon photonic networks," *IEEE Photonics Technol. Lett.* **20**, 767–769 (2008).
13. Y. Vlasov, W. M. Green, and F. Xia, "High-throughput silicon nanophotonic wavelength-insensitive switch for on-chip optical networks," *Nat. Photonics* **2**, 242 (2008).
14. K. Suzuki, R. Konoike, J. Hasegawa, S. Suda, H. Matsuura, K. Ikeda, S. Namiki, and H. Kawashima, "Low insertion loss and power efficient 32×32 silicon photonics switch with extremely-high- δ plc connector," in *Optical Fiber Communication Conference*, (Optical Society of America, 2018), pp. Th4B–5.
15. Q. Cheng, A. Wonfor, R. V. Penty, and I. H. White, "Scalable, low-energy hybrid photonic space switch," *J. Light. Technol.* **31**, 3077–3084 (2013).
16. Q. Cheng, S. Rumley, M. Bahadori, and K. Bergman, "Photonic switching in high performance datacenters," *Opt. Express* **26**, 16022–16043 (2018).
17. K. Tanizawa, K. Suzuki, M. Toyama, M. Ohtsuka, N. Yokoyama, K. Matsumaro, M. Seki, K. Koshino, T. Sugaya, S. Suda, G. Cong, T. Kimura, K. Ikeda, S. Namiki, and K. Hitoshi, "Ultra-compact 32×32 strictly-non-blocking si-wire optical switch with fan-out lga interposer," *Opt. Express* **23**, 17599–17606 (2015).
18. N. Dupuis, A. V. Rylyakov, C. L. Schow, D. M. Kuchta, C. W. Baks, J. S. Orcutt, D. M. Gill, W. M. Green, and B. G. Lee, "Nanosecond-scale mach-zehnder-based cmos photonic switch fabrics," *J. Light. Technol.* **35**, 615–623 (2017).
19. T. Goh, A. Himeno, M. Okuno, H. Takahashi, and K. Hattori, "High-extinction ratio and low-loss silica-based 88 strictly nonblocking thermo-optic matrix switch," *J. Light. Technol.* **17**, 1192 (1999).
20. P. DasMahapatra, R. Stabile, A. Rohit, and K. A. Williams, "Optical crosspoint matrix using broadband resonant switches," *IEEE J. Sel. Top. Quantum Electron.* **20**, 1–10 (2014).
21. R. Yu, S. Cheung, Y. Li, K. Okamoto, R. Proietti, Y. Yin, and S. Yoo, "A scalable silicon photonic chip-scale optical switch for high performance computing systems," *Opt. Express* **21**, 32655–32667 (2013).
22. A. W. Poon, X. Luo, F. Xu, and H. Chen, "Cascaded microresonator-based matrix switch for silicon on-chip optical interconnection," *Proc. IEEE* **97**, 1216–1238 (2009).
23. D. Nikolova, D. M. Calhoun, Y. Liu, S. Rumley, A. Novack, T. Baehr-Jones, M. Hochberg, and K. Bergman, "Modular architecture for fully non-blocking silicon photonic switch fabric," *Microsyst. Nanoeng.* **3**, 16071 (2017).
24. Z. Su, E. Timurdogan, M. Moresco, G. Leake, D. Coolbaugh, and M. Watts, "Wavelength routing and multicasting network in ring-based integrated photonics," in *Integrated Photonics Research, Silicon and Nanophotonics*, (Optical Society of America, 2015), pp. IT4A–3.
25. Z. Wang, J. Xu, P. Yang, Z. Wang, L. H. K. Duong, and X. Chen, "High-radix nonblocking integrated optical switching fabric for data center," *J. Light. Technol.* **35**, 4268–4281 (2017).
26. A. S. Khope, A. A. Saleh, J. E. Bowers, and R. C. Alferness, "Elastic wdm crossbar switch for data centers," *IEEE Opt. Interconnects (OI) Conf.* pp. 48–49 (2016).
27. A. A. M. Saleh, A. S. P. Khope, J. E. Bowers, and R. C. Alferness, "Elastic wdm switching for scalable data center and hpc interconnect networks," *OptoElectronics Commun. Conf. (OECC) Photonics Switch. (PS)* (2016).
28. R. A. Soref and B. E. Little, "Proposed n-wavelength m-fiber wdm crossconnect switch using active microring resonators," *IEEE Photonics Technol. Lett.* **10**, 1121–1123 (1998).
29. Y. Goebuchi, M. Hisada, T. Kato, and Y. Kokubun, "Optical cross-connect circuit using hitless wavelength selective switch," *Opt. express* **16**, 535–548 (2008).
30. T. J. Seok, J. Luo, Z. Huang, K. Kwon, J. Henriksson, J. Jacobs, L. Ochikubo, R. S. Muller, and M. C. Wu, "Mems-actuated 8×8 silicon photonic wavelength-selective switches with 8 wavelength channels," in *CLEO: Science and Innovations*, (Optical Society of America, 2018), pp. STu4B–1.
31. V. E. Beneš, *Mathematical theory of connecting networks and telephone traffic*, vol. 17 (Academic press, 1965).
32. C. Clos, "A study of non-blocking switching networks," *Bell Syst. Tech. J.* **32**, 406–424 (1953).
33. N. Farrington, G. Porter, S. Radhakrishnan, H. H. Bazzaz, V. Subramanya, Y. Fainman, G. Papen, and A. Vahdat, "Helios: a hybrid electrical/optical switch architecture for modular data centers," *ACM SIGCOMM Comput. Commun. Rev.* **41**, 339–350 (2011).
34. G. Porter, R. Strong, N. Farrington, A. Forencich, P. Chen-Sun, T. Rosing, Y. Fainman, G. Papen, and A. Vahdat, "Integrating microsecond circuit switching into the data center," *SIGCOMM Comput. Commun. Rev.* **43**, 447–458 (2013).
35. A. Singh, J. Ong, A. Agarwal, G. Anderson, A. Armistead, R. Bannon, S. Boving, G. Desai, B. Felderman, P. Germano, A. Kanagala, J. Provost, J. Simmons, E. Tanda, J. Wanderer, U. HÄfzlze, S. Stuart, and A. Vahdat, "Jupiter rising: A decade of clos topologies and centralized control in google's datacenter network," *ACM SIGCOMM Comput. Commun. Rev.* **45**, 183–197 (2015).
36. A. A. Saleh, "Scaling-out data centers using photonics technologies," in *Photonics in Switching*, (Optical Society of America, 2014), pp. JM4B–5.

37. B. E. Little, S. T. Chu, H. A. Haus, J. Foresi, and J.-P. Laine, "Microring resonator channel dropping filters," *J. Light. Technol.* **15**, 998–1005 (1997).
38. J. K. Poon, J. Scheuer, S. Mookherjea, G. T. Paloczi, Y. Huang, and A. Yariv, "Matrix analysis of microring coupled-resonator optical waveguides," *Opt. Express* **12**, 90–103 (2004).
39. W. Bogaerts, P. De Heyn, T. Van Vaerenbergh, K. De Vos, S. Kumar Selvaraja, T. Claes, P. Dumon, P. Bienstman, D. Van Thourhout, and R. Baets, "Silicon microring resonators," *Laser Photon. Rev.* **6**, 47–73 (2012).
40. G. Kurczveil, C. Zhang, A. Descos, D. Liang, M. Fiorentino, and R. G. Beausoleil, "On-chip hybrid silicon quantum dot comb laser with 14 error-free channels," ISLC 2018 (2018).
41. J. C. Norman, Z. Zhang, D. Jung, Y. Wan, M. Kennedy, A. Torres, R. W. Herrick, A. C. Gossard, and J. E. Bowers, "High performance quantum dot lasers epitaxially integrated on si," in *Quantum Communications and Quantum Imaging XVI*, vol. 10771 (International Society for Optics and Photonics, 2018), p. 107710D.
42. D. Jung, Z. Zhang, J. Norman, R. Herrick, M. Kennedy, P. Patel, K. Turnlund, C. Jan, Y. Wan, A. C. Gossard, and J. E. Bowers, "Highly reliable low-threshold inas quantum dot lasers on on-axis (001) si with 87% injection efficiency," *ACS Photonics* **5**, 1094–1100 (2017).
43. S. Rumley, M. Bahadori, R. Polster, S. D. Hammond, D. M. Calhoun, K. Wen, A. Rodrigues, and K. Bergman, "Optical interconnects for extreme scale computing systems," *Parallel Comput.* **64**, 65–80 (2017).
44. A. Photonics, "American Institute of Manufacturing Integrated Photonics," <http://www.aimphotonics.com/pdk/> (2016).
45. N. McKeown, A. Mekkittikul, V. Anantharam, and J. Walrand, "Achieving 100% throughput in an input-queued switch," *IEEE Transactions on Commun.* **47**, 1260–1267 (1999).
46. G. Birkhoff, "Tres observaciones sobre el algebra lineal," *Univ. Nac. Tucuman, Ser. A* **5**, 147–154 (1946).
47. D. P. Bertsekas, R. G. Gallager, and P. Humblet, *Data Networks*, vol. 2 (Prentice-Hall International New Jersey, 1992).
48. A. V. Krishnamoorthy, X. Zheng, G. Li, J. Yao, T. Pinguet, A. Mekis, H. Thacker, I. Shubin, Y. Luo, K. Raj, and J. E. Cunningham, "Exploiting cmos manufacturing to reduce tuning requirements for resonant optical devices," *IEEE Photon. J.* **3**, 567–579 (2011).
49. M. Bahadori, A. Gazman, N. Janosik, S. Rumley, Z. Zhu, R. Polster, Q. Cheng, and K. Bergman, "Thermal rectification of integrated microheaters for microring resonators in silicon photonics platform," *J. Light. Technol.* **36**, 773–788 (2018).
50. P. Dong, W. Qian, H. Liang, R. Shafiha, D. Feng, G. Li, J. E. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "Thermally tunable silicon racetrack resonators with ultralow tuning power," *Opt. Express* **18**, 20298–20304 (2010).
51. C. Manganelli, P. Pintus, F. Gambini, D. Fowler, M. Fournier, S. Faralli, C. Kopp, and C. Oton, "Large-fsr thermally tunable double-ring filters for wdm applications in silicon photonics," *IEEE Photonics J.* **9**, 1–10 (2017).
52. A. S. Khope, T. Hirokawa, A. M. Netherton, M. Saeidi, Y. Xia, N. Volet, C. Schow, R. Helkey, L. Theogarajan, A. A. Saleh, J. E. Bowers, and R. C. Alferness, "On-chip wavelength locking for photonic switches," *Opt. Lett.* **42**, 4934–4937 (2017).
53. M. Mitchell, J. Holland, and S. Forrest, "Relative building-block fitness and the building block hypothesis," *D. Whitley, Foundations Genet. Algorithms* **2**, 109–126 (2014).
54. M. R. Watts, W. A. Zortman, D. C. Trotter, G. N. Nielson, D. L. Luck, and R. W. Young, "Adiabatic resonant microrings (arms) with directly integrated thermal microphotonics," in *2009 Conference on Lasers and Electro-Optics and 2009 Conference on Quantum electronics and Laser Science Conference*, (IEEE, 2009), pp. 1–2.
55. Y. Ma, Y. Zhang, S. Yang, A. Novack, R. Ding, A. E.-J. Lim, G.-Q. Lo, T. Baehr-Jones, and M. Hochberg, "Ultralow loss single layer submicron silicon waveguide crossing for soi optical interconnect," *Opt. Express* **21**, 29374–29382 (2013).
56. A. Photonics, "American Institute of Manufacturing Integrated Photonics," <http://www.aimphotonics.com/> (2016).
57. A. Photonics, "American Institute of Manufacturing Integrated Photonics, west coast hub," <https://aim.ucsb.edu/> (2016).