A Scalable, Partially Configurable Optical Switch for Data Center Networks

William Maxwell Mellette, Glenn M. Schuster, George Porter, George Papen, and Joseph E. Ford

***Abstract*—Optical circuit switching may be instrumental in meeting the cost, energy, and aggregate bandwidth requirements of future data center networks. However, conventional MEMS beam-steering cross-connects cannot provide submillisecond switching with the port count necessary for data centers. Here, we investigate a novel noncrossbar *selector switch* architecture and pupil-division switching layout to improve optical switching performance by relaxing the requirement of arbitrary switch configurability. This architecture and switch design enable MEMS beam-steering micromirrors to scale to microsecond response speeds while supporting high port count and low loss switching, and can realize a number of useful interconnection topologies. We present the design, fabrication, and experimental characterization of a proof-of-principle prototype using a single comb-driven MEMS mirror to achieve 150 μs switching of 61 ports between four preprogrammed interconnection mappings. We demonstrate the scalability of this switch architecture with a detailed optical design of a 2048-port selector switch with 20 μs switching time.**

***Index Terms*—Microelectromechanical devices, optical communication, optical fiber switches, optical interconnections.**

I. Introduction

OPTICAL circuit switching may augment or replace electronic switching and meet the size and bandwidth demands of future data center networks [1]–[4]. For optical switching to be adopted in the data center, however, it must provide energy-efficient switching that reduces the net capital and operational cost of the overall network without degrading overall performance. At the physical layer, three aspects of an optical switch impact its practical feasibility: insertion loss, port count, and switching speed.

Optical transceivers account for a large fraction of total data center network cost [5]. Replacing electronic switches with transparent optical switches reduces the number of transceivers required and can reduce total network cost. However, the optical switch insertion loss requires a larger link power budget, which will become increasingly expensive at higher data rates. In fact, datacom manufacturers are already developing 100 Gb/s transceivers [6] with lower power budget and cost than standard long reach (10 km) transceivers with link margins of approximately 6 dB. This technology trend correlates the insertion loss of an optical switch with the network cost, and makes minimizing signal attenuation, crosstalk, distortion, and polarization sensitivity key aspects of optical switch design.

Today's data center networks use a multi-stage folded Clos topology and electronic packet routers to interconnect servers [5], [7]. Each additional stage in the network requires a set of switches and optical interconnection links to the preceding and subsequent stages, increasing cost and cabling complexity. Optical circuit switches have the potential to scale to higher port count and higher per-port bandwidth than electronic switches, reducing cost and cabling complexity by flattening the network. Given that data center networks in production today connect 100,000 servers [5], providing direct connectivity between servers with a monolithic switch is impractical. Instead, transparent optical switches may be used to connect electronically-aggregated groups of servers (e.g. racks, pods, or clusters). Smaller aggregation groups require fewer stages of electronic switching, leading to flatter and less expensive networks, but require optical switches with more ports to interconnect the groups. For example, 2,000 port switches would be required to connect the racks of a 100,000 server network assuming 50 servers per rack.

Optical switches based on microelectromechanical systems (MEMS) actuated tilt-mirrors have been fabricated with over 1,000 ports and less than 4 dB worst-case insertion loss [8], essentially meeting the port count and transmission requirements for deployment in data center networks. However, beam-steering cross-connects have response times on the order of 10 to 100 milliseconds, limiting their role to provisioning point-to-point bandwidth on second-long timescales [1]. While useful for latency-insensitive data migration, many data center applications exhibit short-lived communication patterns between many end-points [9], and cannot effectively utilize slow switching.

Previous work has exposed opportunities for optical switching with microsecond-scale reconfiguration times in data center networks [2]–[4]. Unfortunately, there is a fundamental tradeoff between switching speed, insertion loss, and port count in conventional beam-steering switches: to scale to large port count while maintaining low loss requires micromirrors with larger apertures (and inertia) and/or larger tilting ranges (requiring softer torsion springs), both of which reduce the response speed of the switch [10]. High fill-factor tip-tilt-piston micromirror arrays [11] can allow each optical signal to be steered using a number of faster micromirrors, but this approach may restrict total tilt range if piston and tilt actuators are combined, requires

Manuscript received July 14, 2016; revised December 1, 2016; accepted December 2, 2016. Date of publication December 4, 2016; date of current version December 22, 2016. This work was supported by the National Science Foundation under Grant #1314921.

The authors are with the University of California San Diego, La Jolla, CA 92093-0407 USA (e-mail: wmellett@ucsd.edu; gschuster@eng.ucsd.edu; gmporter@cs.ucsd.edu; gpapen@ucsd.edu; jeford@ucsd.edu).

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

Digital Object Identifier 10.1109/JLT.2016.2636025

more control channels per switched port, and the phase matching requirement restricts the transmission spectral bandwidth for each channel, especially at larger tilt angles.

As an alternative to freespace beam-steering switches, guided-wave switches fabricated with Silicon or III-V materials are being investigated by a number of research groups. Nanosecond to microsecond reconfiguration speeds are possible with optical switches based on electro-optic modulation, semiconductor optical amplification, or thermo-optic modulation. However, the accumulated loss and crosstalk induced by their multistage architectures have limited these switches to small port counts ($\leq$8) [12]–[14] or high loss ($>$15 dB) [15]. Alternatively, MEMS-actuated silicon waveguide switching structures with microsecond response times have recently been reported and integrated into a 64-port cross-point matrix with 4 dB on chip loss [16]. However, because the size and complexity of planar cross-point architectures scale as the square of the port count, scaling these switches to hundreds of ports presents chip-area, loss, and yield challenges. Further, significant fiber-chip coupling losses preclude multi-chip topologies. Today, practical packaging approaches use polarization-splitting coupling structures to interface with standard transceivers and fiber, introducing a total insertion loss of over 6 dB [17].

Here, we investigate a novel and potentially useful optical *selector switch* architecture which forgoes non-blocking crossbar configurability, instead enabling rapid selection between a relatively small set of pre-programmed interconnection mappings. This concept can be implemented in multiple switching technologies, and previous work demonstrated a similar concept using wavelength switching for fast selection of interconnection patterns recorded as volume holograms [18]. However, the change to a selector switch architecture allows MEMS beam-steering mirrors to scale to microsecond response speeds while supporting a large number of ports and low-loss switching between the broadband single- or even multi-mode transceivers used in data center networks.

The paper is organized as follows. We discuss the switch architecture, applications, and pupil-division switching in Section II. In Section III we show the design of a 61-port proof-of-principle prototype switch based on commercial-off-the-shelf (COTS) optical components, which we fabricate and characterize in Section IV. In Section V we present the design of a low-loss 2,048-port switch with a 20 μs response time which uses a custom MEMS device and micro-optic port mapping structures.

II. Selector Switch Architecture

The proposed *selector switch* differs from conventional optical cross-connects in both its architecture and its basic optical switching principle.

A. Partial Configurability

Instead of implementing all $N!$ possible port-mappings of an $N \times N$-port crossbar, the selector switch selects between a small subset $k << N!$ of these mappings. In this sense, the selector switch can be thought of as *partially configurable*. While many interconnection networks have been designed to leverage the arbitrary configurability of crossbars, partially configurable switches can be used to implement a number of useful network topologies. For instance, a network with full connectivity can be constructed from partially configurable circuit switches. The set of shuffle-equivalent network topologies (e.g. Banyan, Perfect Shuffle, Crossover), are typically implemented in space as multistage interconnection networks with $\log_2 N$ stages [19]. These networks provide full connectivity between N ports using a minimum number of connections. The network diameter (number of *hops* data makes as it traverses the network) is $\log_2 N$ in these networks. A partially configurable switch can realize these network topologies by multiplexing in time (rather than space), cycling through a set of $k = \log_2 N$ port mappings. Other network topologies with $k = O(N)$ port mappings [20] can also be constructed from partially configurable switches. Any partially configurable network design will have tradeoffs between the number of switches, number of port mappings, and the network diameter. Optimization of these factors is possible given a particular application, but is outside the scope of this paper and will be considered in a separate publication.

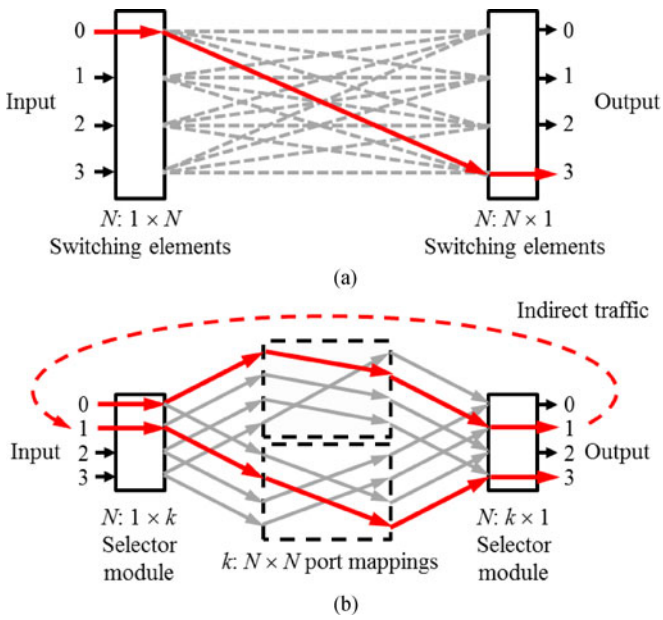

Fig. 1. A crossbar connects any two ports in a single hop through the switch using $1 \times N$ switching elements. A selector switch uses $1 \times k$ switching elements (here $k = \log_2 N$) to select amongst k port mappings. With $k = \log_2 N$ port mappings, data passes through the switch up to $\log_2 N$ times. For example, two hops are required to send from node 0 to 3, with data passing through intermediate node 1.

To ground our discussion and establish a starting point in the design space of partially configurable switches, we focus on selector switches with $k = \log_2 N$ port mappings. This configuration provides a balance between the hardware complexity ($\log_2 N$ physical port mappings) and network diameter (at most $\log_2 N$ hops). There are a number of equivalent sets of port mappings with logarithmic network diameter [19]; a simple example is to form sets of mappings from port p to ports $p + 2^0,\ 2^1, \ldots,\ 2^{\log_2 N - 1}$ modulo N, where p is indexed from 0 (shown for $N = 4$ in Fig. 1(b)) [21]. With only $\log_2 N$ mappings, data will generally traverse the switch multiple times, but

will be electronically forwarded at intermediate terminal nodes between each optical hop.

Figure 1 shows the optical crossbar and selector switch architectures, each requiring two stages of $1\times N$ and $1\times\log_2 N$ switching elements, respectively. When implemented with MEMS micromirror switching elements, each micromirror in the selector switch needs to resolve $k=\log_2 N$ optical states, as opposed to $k=N$ optical states for the crossbar. This reduction in the number of optical states significantly reduces the aperture and tilt requirements of the micromirror, allowing it to be redesigned for higher speed operation. Previous work quantified the theoretical limits of switching speed for single- and dual-axis MEMS micromirror actuators based on the number of resolvable optical states [10], indicating a significant reduction in switching speed is possible by reducing the number of resolvable optical states by a logarithmic factor. Sections IV and V of this paper quantify the achievable switching speeds in selector switch designs based on previously demonstrated beam-steering micromirror devices, showing between two and three orders of magnitude improvement over conventional optical cross-connects depending on the specific micromirror design.

In this work, we consider selector switches with fixed port mappings which are implemented with either low-loss micro-optic or fiber optic interconnections. Alternatively, the switch could be designed to quickly select between k slowly-reconfigurable interconnection patterns by replacing the hard-wired port mappings in Fig. 1(b) with crossbar switches. This would enable arbitrary configurability while reducing the loss-of-light time during switching, but would increase hardware cost and insertion loss.

The primary component of a selector switch is the *selector module* (see Fig. 1(b)). The module can be designed with N $1\times k$ *individually-switched* elements, allowing signals from each port to be routed through an independently selected port mapping, or as a monolithic $1\times k$ *gang-switched* element which simultaneously selects one mapping for all ports. An individually-switched selector module provides greater flexibility because it can select from and combine different port mappings to expand the effective set of selectable mappings, but requires a micromirror array with individually-controllable mirrors. A gang-switched selector module, on the other hand, is less flexible in that it can only select from the $\log_2 N$ hard-wired mappings, but is simpler to control and less expensive to implement because it requires only a single micromirror and control signal. A prototype gang-switched selector module is designed in Section III and characterized in Section IV. Section V discusses a design compatible with individual switching.

B. Pupil-Division Switching

Conventional cross-connects typically employ a non-imaging layout using two micromirror arrays to steer essentially collimated beams through a freespace volume where beams may intersect (Fig. 2(a)). Arbitrary bijective port mappings are possible because the path of each beam is defined by a pair of dedicated micromirrors. An alternate design uses a Fourier lens between micromirror arrays to achieve $\sqrt{2}$ smaller beam diameter at the micromirrors, but operates under the same principle as above [8].

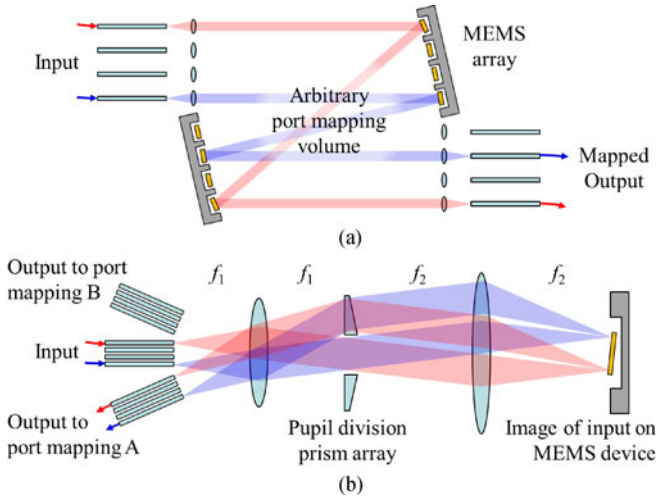

Fig. 2. (a) Schematic of conventional optical cross-connect. (b) Schematic cross section of fiber-coupled selector module. The MEMS device tilts to select prism apertures which refract the arrayed image to couple into different fiber arrays. Port mappings are implemented externally in (b).

We designed the selector module based on a radically different optical configuration, incorporating relay imaging and pupil division instead of the collimated beam steering used in current MEMS cross-connects. Fig. 2(b) shows a schematic cross section of a fiber-coupled selector module, with a 4-f imaging relay and prism array located near the intermediate pupil of the relay. Light from a two dimensional (2-D) array of input fibers is imaged onto a MEMS tilt-mirror device, which tilts in 2-D to direct reflected light through discrete prism apertures, each of which refracts the output image position to couple into a different output fiber array. Each output fiber array interfaces to an external fiber optic port mapping. The telecentricity of the 4-f relay minimizes fiber coupling loss by ensuring that each optical beam couples at normal incidence into its corresponding output fiber core. Because switching occurs by aperture selection in the pupil plane, as opposed to spatially scanning across the fiber array in a conventional cross-connect, the system is more tolerant to angular misalignments of the micromirror. This relaxes the required drive electronics precision and sensitivity to underdamped mirror ringing. The 4-f image relay makes the design compatible with both multimode fiber and space-division multiplexed signals, but we use single mode fiber in the switch designs considered here.

The fundamental switch functionality is similar to that of each channel of a wavelength add-drop switch [22], which has also been adapted for space-division multiplexing with multi-core fiber I/O via the introduction of multi-core fiber breakout into the individual collimators of a single-mode wavelength selective switch [23]. Unlike those designs, the selector switch operates directly on the 2-D arrayed multi-core fiber end face, and allows switching access to each spatial channel in the image plane of the fiber array. Also, because it does not incorporate wavelength selectivity, the selector switch architecture makes use of the additional spatial dimension to support more spatial output ports through two-axis switching.

III. Fiber-Interconnected Selector Module Design

The most straightforward implementation of a selector switch uses fiber optics to realize the desired port mappings between two selector modules (see Figs. 1(b) and 2(b)). In this section, we describe *selector module* designs based on current commercially available fiber arrays and MEMS beam-steering micromirror.

A. Prototype Design Using Commercial Optics

Maximizing the spatial density of the arrayed input signals maximizes the port count of the selector module. We based our design on a 61-core pitch reducing fiber array [24] commercially available from Chiral Photonics. The array maintains the mode-field diameter and numerical aperture of single mode fiber, but is tapered to position the fiber cores in a 2-D hexagonal array with a 37 μm core pitch. The distance r_0 from the center to corner channel in the array was 148 μm. We designed for the C band, with a nominal center-band wavelength of $\lambda = 1560$ nm and a mode waist of $w_0 = 5.2$ μm at the fiber. We used a MEMS device with a single micromirror instead of a micromirror array. This allows the micromirror to be surrounded by a large-area (and correspondingly fast) actuator structure, but also means the relayed image of all 61 signals had to be encircled by the micromirror radius r_m. We allowed the magnification $|M| = f_2 / f_1$ (see Fig. 2(b)) of the 4-f relay to vary to accommodate different micromirror radii. To first order, the mirror radius must be at least

$$r_m = |M| \cdot r_0. \tag{1}$$

The prism apertures were chosen to be a factor of $\xi = 1.3$ larger than the Gaussian beam mode width at the Fourier plane, yielding 97% power transmission through each prism aperture. The micromirror needs to tilt over a mechanical angular range $\pm\theta_m$ sufficient to select between prism apertures. With 61 ports, we need $\log_2 61 = 5.931 \rightarrow$ at most 6 port mappings to implement the logarithmic mappings discussed in Section II. Hexagonally tiling the 6 prism apertures minimized the required mechanical tilt range of the micromirror:

$$\theta_m \approx \frac{1}{|M|} \cdot \frac{\lambda\xi}{\pi w_0}. \tag{2}$$

For a given drive voltage and actuator structure, the switching speed of the micromirror is primarily a function of its radius and tilt range, which respectively determine its mass and torsional stiffness [10]. While equations (1) and (2) could be used to optimize a micromirror for our system, instead we used them to guide our search for a commercially available micromirror with the fastest response which would meet the system requirements. We chose a USB-powered 2-axis electrostatic comb driven micromirror [25] (Mirrorcle Technologies part #A7M8.1) with $r_m = 400$ μm, $\theta_m = \pm 4^\circ$, and sub-millisecond response speed (exact settling time depends on the drive signal, and is measured in Section IV).

The final design steps were to choose relay lenses and the refraction angle of the prisms. The micromirror parameters allowed relay magnifications between 1.8 and 2.7. We targeted a magnification closer to 2.7 because although it increased system length, it maximized the F/# (focal length divided by clear aperture) and minimized the aberrations of Lens 2. We chose commercially available doublet lenses with $f_1 = 25$ mm and $f_2 = 60$ mm, for a relay magnification of $M = 2.4$. To avoid mechanical interference, the fiber arrays needed to be separated laterally by at least the end-face diameter of 600 μm. Minimizing the fiber array separation would also minimize the field-of-view requirement of Lens 1 as well as the refraction angle and chromatic dispersion induced by the prisms, which ultimately limits the spectral bandwidth of the switch. We chose commercially available fused silica prisms with a 5° wedge, which required a fiber array spacing of 1 mm and a field of view of $\pm$ 2.3° at Lens 1. Finally, in order to minimize the off-axis aberrations in Lens 1, we positioned the prism array near Lens 2 instead of at the Fourier plane so the returning beams would enter the lens closer to the optical axis.

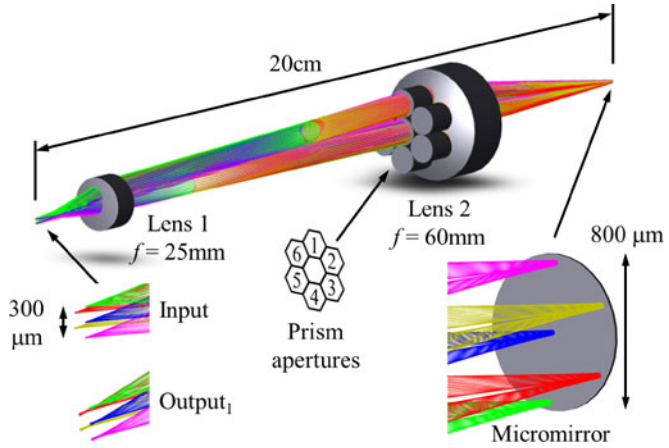

Fig. 3. 3-D Zemax model of prototype 61-port 1×6 selector module, using commercial doublet lenses and fused silica prisms. Ray colors correspond to the center and edge channels of the 61-core fiber array.

Figure 3 shows the prototype selector module with $N = 61$ input ports and six 61-port output arrays modeled in Zemax optical design software. We modeled the spectral transmission of the selector module, including single-mode fiber coupling loss, with the Zemax physical optics propagation tool. The modeled transmission for the center and edge channels in the array is shown in Fig. 4, along with the modeled transmission for the custom optical design described in the following section. The bandwidth is primarily limited by chromatic dispersion from the fused silica prisms. The fiber arrays were aligned to maximize transmission at 1560 nm (corresponding to gain peak of the erbium light source used to characterize the prototype in Section IV), but the transmission peak can be tuned to any waveband by refocusing the lenses and translating the fiber arrays. Peak transmission is limited by lens aberrations and reflection losses from the uncoated fiber arrays.

B. Achromatized Prototype Design with Custom Optics

The modeled insertion loss and bandwidth limits from the COTS lens and prims designs (above) would be unacceptable in a practical data center, especially since the transmission of the full selector switch is half that of the selector module, as light

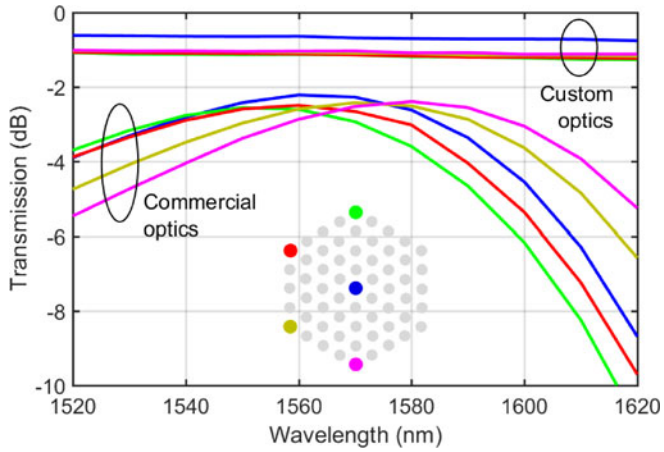

Fig. 4. Modeled transmission spectrum of 61-port 1×6 selector module with (1) commercial optics and (2) custom optics. Color inset shows location of channels in 61-core fiber array.

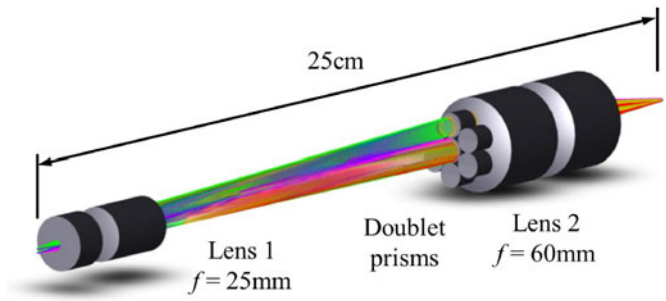

Fig. 5. Zemax model of achromatized 61-port 1×6 selector module with custom triplet lenses and doublet prisms.

must pass through two selector modules in the switch architecture shown in Fig. 1(b). These losses can be significantly reduced by a minor redesign and customized lenses and coatings.

The single-glass prisms were the dominant source of chromatic dispersion in the commercial optics design. We designed a doublet prism using standard crown and flint glasses (Calcium Fluoride and N-BASF64) to provide the same refraction angle but with negligible dispersion over the C-band. The COTS doublet lenses also contributed insertion loss due to aberrations. We designed a pair of triplet lenses in Zemax using standard glasses with diffraction-limited performance over the C-band and the field of view required by the system. We assumed all refractive surfaces (including the fiber arrays) were coated with commercially-available antireflection coatings and the MEMS mirror was coated with gold (97% reflective). The Zemax model of the achromatized selector module is shown in Fig. 5 and the transmission is shown in Fig. 4. The modeled transmission is above -1 dB and substantially flat over >100 nm, which would yield an excellent overall switch transmission of greater than -2 dB.

IV. Prototype Fabrication and Characterization

A. Optomechanical Assembly

We fabricated the 61-port selector module designed in Section III (A) based on COTS optics, and the assembled prototype selector module is shown in Fig. 6(a). The main fabrication challenge was ensuring precise mechanical alignment of the fiber arrays to minimize insertion loss. Positional alignment is the primary driver of fiber coupling efficiency, requiring micrometer accuracy, while angular misalignment (tip/tilt) of up to a few degrees has little impact on coupling [26]. The end faces of the fiber arrays needed to be brought to within one millimeter of contact while avoiding mechanical interference (Figs. 6(b) and 6(c)). To accomplish this, we machined custom aluminum mounts. The central "input" fiber array was attached to a mount fixed to the optical breadboard. The "output" fiber arrays were mounted to goniometers attached to 3-axis roller bearing translation stages. This provided the one rotational and three linear degrees of freedom required to align the output arrays to the fixed input array. Adjustment for tip and tilt of the arrays was not needed.

We fabricated a custom prism array for the switch (Fig. 6(d)) by dicing sections from commercially available antireflection coated fused silica wedges via a diamond saw to form the prism facets. These facets were arranged and bonded to a fused silica flat with optical epoxy to create the hexagonally-tiled array. As described in Section III(A), we used commercially available fiber arrays, lenses, and MEMS micromirror (Fig. 6(e)).

In addition to the input array, we populated four of the six available outputs with fiber arrays. The loss in the interconnection mappings can be minimized by fusion-splicing the output fibers to the required patterns, but for convenience our prototype used mating sleeve connectors for both input and output paths. The left wall of the enclosure provides the 61-fiber patch panel interface to the input array, and the right wall holds the four 61-fiber interfaces to the output arrays. Only four arrays were populated because this provided enough port mappings for integration into a 16 server network testbed. The four output arrays were sufficient to characterize all important aspects of the prototype.

B. Characterization

An erbium amplified spontaneous emission (ASE) light source was used to measure the transmission spectrum of the selector module. The fiber arrays were aligned to maximize transmission at the 1560 nm gain peak of the source. Figure 7 shows the measured transmission spectra for all 61 channels for each of the four output fiber arrays. The passband can be shifted arbitrarily by a simple realignment of the system. The measurements are generally in good agreement with the modeled transmission shown in Fig. 4 (also aligned for maximum transmission at 1560 nm). Array 2 has the most uniform transmission across all channels, while there is the most variance between the channels of array 4. The transmission variance between arrays and between channels within each array was attributed to irregularities in the as-fabricated fiber core positions in the arrays. Modeling showed that core pitch irregularities of $\pm$ 2 μm can reduce coupling by up to 2 dB and/or shift the transmission spectrum by 35 nm, accounting for the observed variability in transmission. Using a laser diode source with a mechanical polarization rotator we confirmed that the selector

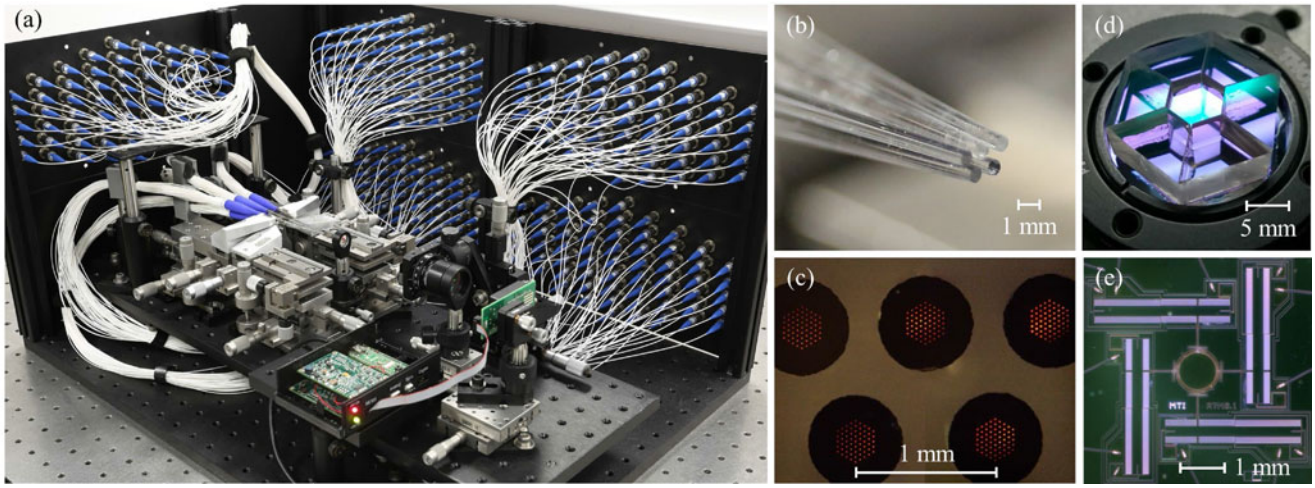

Fig. 6. (a) Prototype selector module. (b) Close-up of fiber arrays. (c) Microscope image of fiber array end faces. (d) Custom antireflection-coated fused silica prism array. (e) Micromirror with large-area comb drive actuators.

module is polarization insensitive, with a loss variation of less than 0.01 dB.

Next, using the same ASE light source, the intra- and inter-array crosstalk of the selector module was measured. The best and worst case nearest neighbor crosstalk was −40 dB at the center of the array, and −30 dB for channels at the edge of the array. The fiber array manufacturer quotes −35 dB crosstalk between adjacent fiber cores, and given that light propagates through two arrays in the selector module, at least −32 dB crosstalk is expected due to the fiber arrays themselves. We measured a worst case inter-array crosstalk of −50 dB.

The response times of the switch measured for each switch transition are shown in Fig. 8. The voltage waveforms were digitally filtered with the inverse transfer function of the micromirror to provide fast switching while suppressing mechanical overshoot. The longest switching time of 151 μs (90% – 90% optical power) occurred while moving the micromirror over its longest angular travel between arrays 1 and 4. This is two orders of magnitude faster than commercial MEMS optical cross-connects with comparable port count. The uncoated fiber/air interface accounts for the return loss during steady state, and can be suppressed with antireflection coatings on the fiber array surface. While the mirror is in motion, it can return a significant fraction of the signal power to the transmitter for certain switch transitions (e.g. 1→4). In our use of the switch in a testbed with commercial datacom transceivers, we found this had no effect on data transmission because no data was communicated during switching, and also because any spurious oscillations in the laser cavity subsided in less than a single 100 ps bit interval.

Finally, we measured the optomechanical stability of the prototype. Figure 9 shows the switch transmission and ambient laboratory temperature over a 15 hour period with no active adjustment. Based on the observed correlation, the selector module transmission varies by 0.05 dB/°C, or 0.1 dB/°C for the full switch. Even lower temperature dependence could be achieved by considering the coefficients of thermal expansion in design of the optomechanical package.

V. Freespace-Interconnected Selector Switch Design

The proof-of-principle prototype described in the previous section established feasibility of the fast pupil-switched optics, but the fiber-based interconnects would present a significant cabling cost for larger switches because the number of interconnection fibers scales as the product of the port count and number of port mappings in the switch. For example, using logarithmic port mappings in a 2,048 port switch would require 2,048 fibers for each of the 11 port mappings, totaling 22,528 interconnection fibers. To circumvent this issue, we describe a design using freespace micro-optics to integrate the port mappings into a single compact optical assembly which combines the entire selector switch diagrammed in Fig. 1(b).

A. Monolithic Switch Assembly

Figure 10 shows a cross sectional illustration of the switch layout with two different light paths through the system. Like the fiber-interconnected switch in Fig. 2(b), this design is also based on 4-f imaging relays and pupil-division switching. However, a number of differences allow this design to incorporate both selector module stages and port mappings into a single assembly. First, this design shares a single fiber array with half the fibers as inputs and half as outputs. Instead of a reduced pitch array, it uses a 2-D fiber array with corresponding microlenses, similar to those employed by cross connect switches [8]. 2-D arrays with core pitches as low as 170 μm and up to 4,096 elements are available from commercial suppliers [27]. The microlenses are designed so they: 1) form a larger beam waist at their output (plane A) to increase signal density, and 2) have a focal point at the fiber face to map any positional misalignment of output beams into angular misalignment at the fiber cores to increase tolerances elsewhere in the system. The beam waists are relayed with demagnification from A onto a micromirror array at B. Each micromirror tilts to direct its beam into one of several 4-f relays, which are defined by lens apertures at C. The beam waists are relayed from plane B to D, where the signals enter the micro-optic interconnection assembly. The interconnection

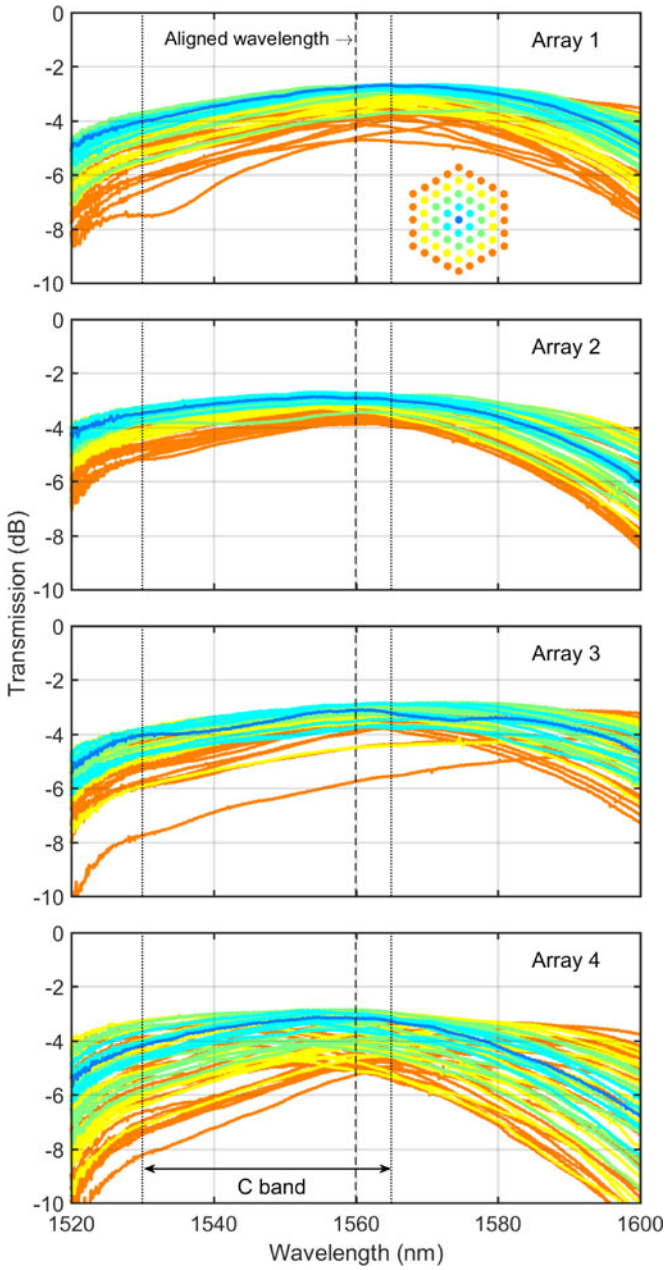

Fig. 7. Measured transmission spectrum through all ports of prototype selector module. Arrays were aligned for maximum transmission at 1560 nm. C band is shown for bandwidth reference. Color inset shows location of channels in 61-core fiber array.

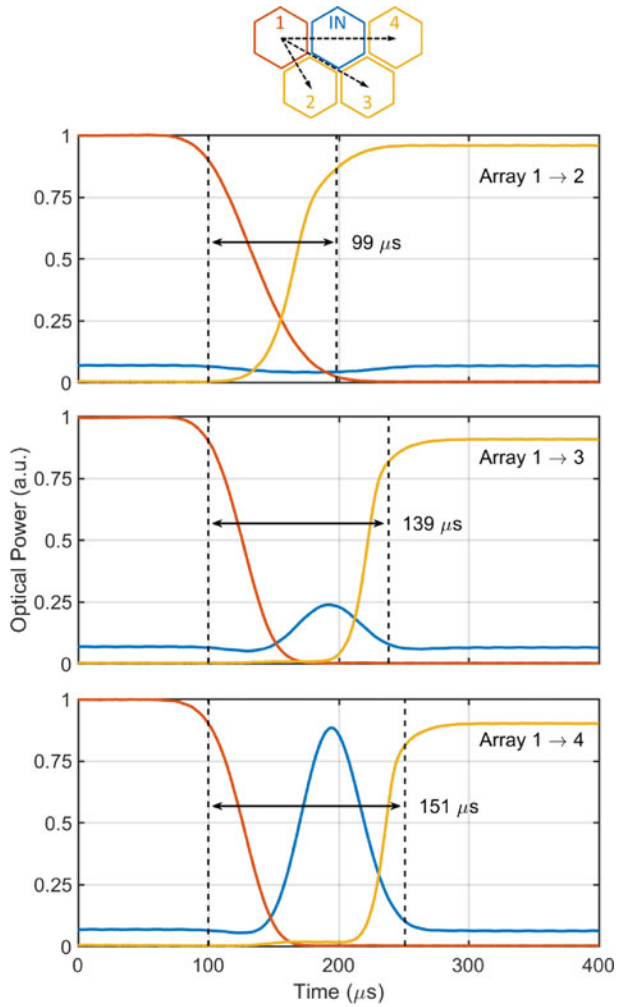

Fig. 8. (Top) illustration of prism apertures with beam trajectories for three representative switching transitions, and (lower) temporal response of switch, with times measured from 90% to 90% power.

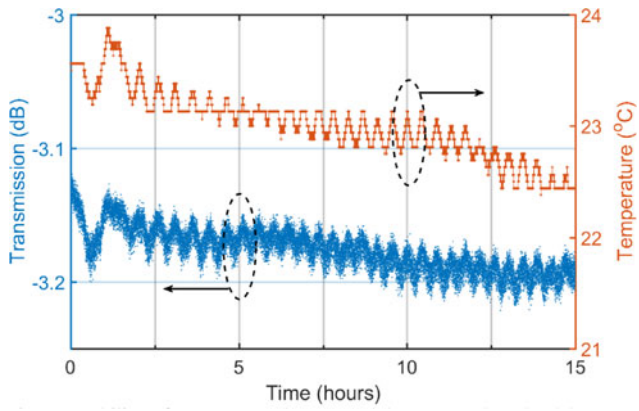

Fig. 9. Stability of prototype selector module measured under laboratory conditions. Loss varies by $0.05\,\mathrm{dB}\,/\,^{\circ}\mathrm{C}$, or $0.1\,\mathrm{dB}\,/\,^{\circ}\mathrm{C}$ for the full switch.

assembly is a set of stacked micro-patterned substrates which spatially rearranges signals through refraction and reflection and sends them back through the switch to couple into output fibers. The different regions of the assembly are uniquely patterned to implement different port mappings.

B. Arbitrary Port Mapping Subassembly

Figure 11(a) illustrates a port mapping subassembly capable of arbitrary mappings and which can be integrated into the switch layout shown in Fig. 10. It uses a single substrate with patterned prism facets on the front side to route signals and curved mirrors on the back side to reflect and refocus the optical beams. Each set of prism facets that constitute a port mapping shares a common curved mirror. To maintain the beam shape across all ports, the curvature of the mirror must match the wavefront curvature of the beam and the maximum differential

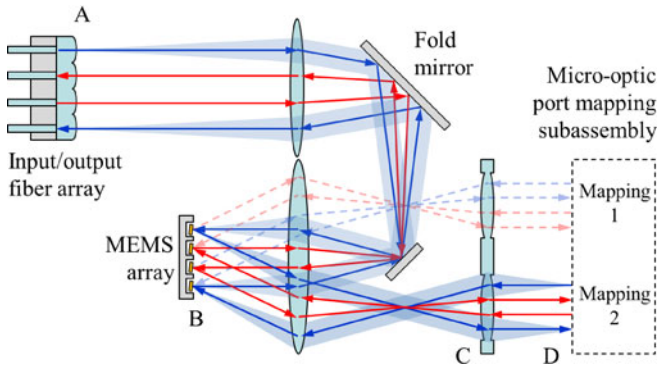

Fig. 10. Schematic cross section of freespace-interconnected switch. Signals are relayed onto a micromirror array, steered, and relayed again onto a micro-optic port mapping assembly. After rearrangement, signals traverse the switch again to couple to output fibers in a shared input/output array. Solid and dashed chief rays indicate light paths through different mappings.

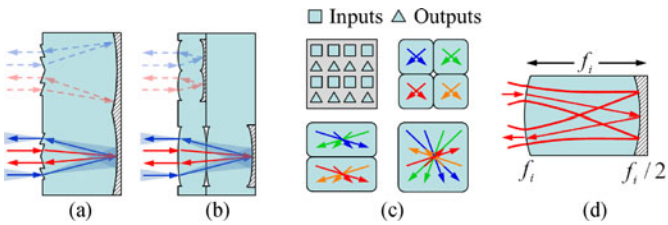

Fig. 11. (a) Arbitrary port mapping subassembly using prism arrays. Solid and dashed lines show light paths through different port mappings. (b) Logarithmic port mapping subassembly using microlens arrays. (c) Head-on view of input/output port arrangement and microlens patterns to implement Crossover port mappings. Ray color indicates port groupings. (d) Side view of a section of the logarithmic port mapping imaging relay.

path length of the beams must be small. A fabrication challenge of this design is the large sag of the curved mirrors, possibly necessitating assembly from bulk optics.

C. Logarithmic Port Mapping Subassembly

While arbitrary port mappings may be required in some situations, logarithmic port mappings support a number of useful interconnection architectures and can be realized in a microlens-based geometry which is more amenable to lithographic fabrication. Here, we examine an implementation of the Crossover network topology, which is isomorphic to the set of logarithmic shuffle-equivalent topologies [28].

Figure 11(b) illustrates the port mapping subassembly, which uses refractive and reflective microlenses in a stack of substrates to define a set of port mapping imaging relays. Figure 11(c) shows a head-on view of how signals are mapped in the Crossover topology in an 8-port example using $\log_2 8 = 3$ port mappings. Figure 11(d) shows a side view of a section of the port mapping imaging relay with a Gaussian beam passing through. The refractive microlens acts as a field lens to route signals through the center of the reflective microlens. The curvature of the reflector is chosen so the beam waists are relayed to the surface of the refractive microlens as the beams exit the relay. The numerical aperture (NA) of the largest microlens increases as more ports and port mappings are added to a single port

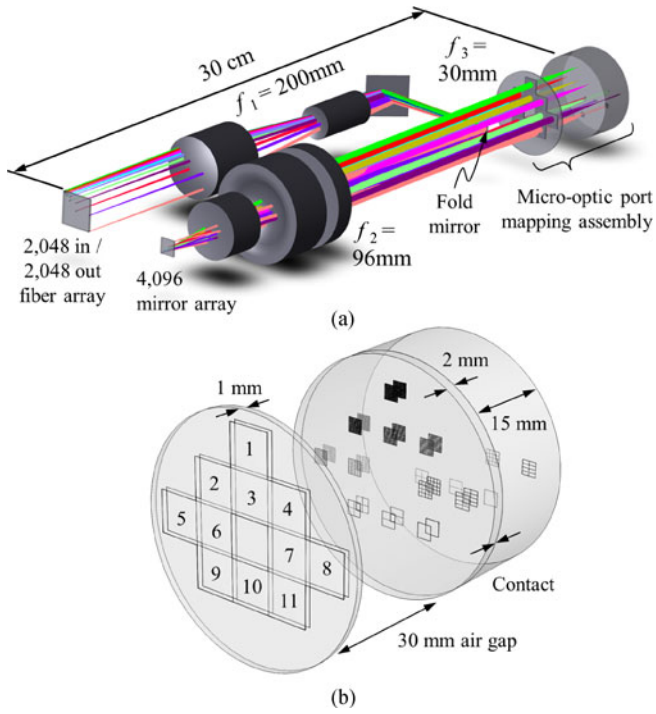

Fig. 12. (a) Zemax model of 2,048 port switch. (b) Close up of port mapping assembly, indicating the 11 port mapping apertures.

mapping substrate, contributing aberrations, path length difference, and lens sag. To lower the NA requirements, multiple substrates can be stacked to accommodate a larger range of microlens focal lengths. Similar stacked assemblies of micro optics have been previously demonstrated [29]. As shown in Fig. 11(b), microlenses can be placed at the intermediate refractive surfaces in the stack to split the optical power required by each surface (and thus reduce the maximum surface sag).

D. Logarithmically-interconnected 2,048-port Switch

Based on components discussed above, we used Zemax to design a 2,048-port switch and model its transmission. Figure 12(a) shows the switch layout. We based the design on a previously demonstrated 4,096-element micromirror array with 20 μs response, $\pm$ 4.3° mechanical tilt, and 120 μm diameter micromirrors [11]. We designed for a 64 $\times$ 64 single mode fiber array with 250 μm core spacing, with 2,048 cores acting as inputs and 2,048 as outputs. A BK7 microlens array attached to the fibers was designed to create an 83.3 μm beam waist at the microlens face for a pitch-to-waist ratio of 3 (for 99% power confinement in the aperture). A 200 mm focal length 2-glass telephoto lens, 96 mm 3-glass lens, and pair of fold mirrors are used to relay the beam waists onto the micromirror array. The lenses were designed with standard glasses and optimized for diffraction-limited performance over the C-band. Beam waists of 40 μm are formed on the 120 μm diameter high-fill-factor micromirrors (again a pitch-to-waist ratio of 3). The focal length ratio of the relay can be modified to accommodate different fiber or micromirror pitches.

Figure 12(b) shows an enlarged view of the port mapping assembly. An array of eleven 30 mm focal length silicon lenses

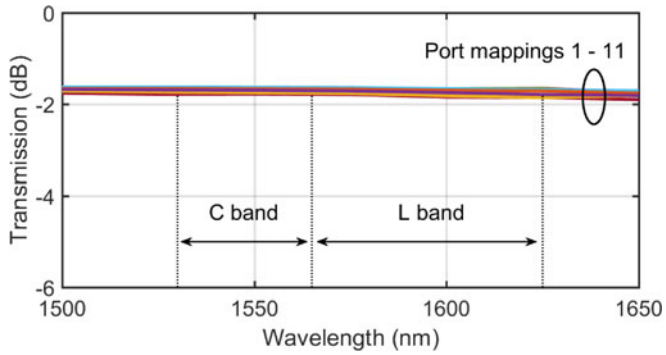

Fig. 13. Modeled worst-case transmission of 2,048 port switch through all 11 port mappings. C and L bands are shown for bandwidth reference.

defines the locations of the $\log_2 2{,}048 = 11$ port mappings. Two micro-patterned silicon substrates, 2 and 15 mm thick, are required for the port mapping relay subassembly in order to keep the maximum microlens sag under 50 μm (sag previously demonstrated in silicon [29]).

Figure 13 shows the modeled transmission of the switch, including fiber coupling, which is > -2 dB over both the C and L bands. All refractive surfaces were assumed to be antireflection coated for 99.75% transmission. The fold mirrors were modeled with enhanced refection coatings (99.5% reflective) and the micromirrors and reflective microlenses were assumed to be gold coated (97% reflective). Tolerancing analysis indicated that a fiber-collimator misalignment of $+$ 2 μm at the input and – 2 μm at the output introduced 3 dB excess loss. A misalignment of 5 μm between any of the substrates in the port mapping assembly introduced 3 dB excess loss.

VI. Conclusion

We have proposed a novel partially configurable optical switch architecture which is highly scalable in speed and port count without compromising transmission performance, potentially meeting the needs of data center networks. We designed and fabricated a 61-port prototype selector module with 150 μs switching time using commercial off the shelf components and measured a center-band overall switch loss of less than 10 dB. Detailed optical designs indicate the loss of the prototype could be reduced to 2 dB with custom optics. The switch can scale to 2,048 ports and a 20 μs response with 2 dB loss using micro-optic port mappings and a previously demonstrated micromirror array. Future work will study the architecture and control aspects of selector switch-based networks.

References

[1] N. Farrington *et al.*, "Helios: A hybrid electrical/optical switch architecture for modular data centers," in *Proc. ACM SIGCOMM Conf.*, 2010, pp. 339–350.
[2] G. Porter *et al.*, "Integrating microsecond circuit switching into the data center," in *Proc. ACM SIGCOMM Conf.*, 2013, pp. 447–458.
[3] A. Vahdat, H. Liu, X. Zhao, and C. Johnson, "The emerging optical data center," in *Proc. Opt. Fiber Commun. Conf.*, 2011, p. OTuH2.
[4] N. Farrington, Y. Fainman, H. Liu, G. Papen, and A. Vahdat, "Hardware requirements for optical circuit switched data center networks," in *Proc. Opt. Fiber Commun. Conf.*, 2011, p. OTuH2.
[5] A. Singh *et al.*, "Jupiter rising: A decade of clos topologies and centralized control in google's datacenter network," in *Proc. ACM SIGCOMM Conf.*, 2015, pp. 183–197.
[6] 100G CWDM4 MSA Technical Specifications, (2015). [Online] Available: http://www.cwdm4-msa.org/wp-content/uploads/2015/12/CWDM4-MSA-Technical-Spec-1p1-1.pdf.
[7] C. Clos, "A study of non-blocking switching networks," *Bell Syst. Tech. J.*, vol. 32, no. 2, pp. 406–424, 1953.
[8] J. Kim *et al.*, "1100 x 1100 Port MEMS-based optical crossconnect with 4-dB maximum loss," *Photon. Technol. Lett.*, vol. 15, no. 11, pp. 1537–1539, 2003.
[9] A. Roy, H. Zeng, J. Bagga, G. Porter, and A. Snoeren, "Inside the social network's (datacenter) network," in *Proc. ACM SIGCOMM Conf.*, 2015, pp. 123–137.
[10] W. Mellette and J. Ford, "Scaling limits of MEMS beam-steering switches for data center networks," *J. Lightw. Technol.*, vol. 33, no. 15, pp. 3308–3318, Aug. 2015.
[11] F. Pardo *et al.*, "Flexible fabrication of large pixel count piston-tip-tilt mirror arrays for fast spatial light modulators," *Microelectron. Eng.*, vol. 84, pp. 1157–1161, 2007.
[12] L. Lu, L. Zhou, Z. Li, X. Li, and J. Chen, "Broadband 4×4 Nonblocking Silicon Electrooptic Switches Based on Mach-Zehnder Interferometers," *Photon. J.*, vol. 7, no. 1, pp. 1–8, 2015.
[13] Q. Cheng, A. Wonfor, J. Wei, R. Penty, and I. White, "Low-energy, high-performance lossless 8×8 SOA switch," in *Proc. Opt. Fiber Conf.*, pp. Th4E-6, 2015.
[14] K. Suzuki *et al.*, "Ultra-compact 8 × 8 strictly-non-blocking Si-wire PILOSS switch," *Opt. Express*, vol. 22, no. 4, pp. 3887–3894, 2014.
[15] K. Tanizawa *et al.*, "Ultra-compact 32 × 32 strictly-non-blocking Si-wire optical switch with fan-out LGA interposer," *Opt. Express*, vol. 23, no. 13, pp. 17599–17606, 2015.
[16] T. Seok, N. Quack, S. Han, R. Muller, and M. Wu, "Large-scale broadband digital silicon photonic switches with vertical adiabatic couplers," *Optica*, vol. 3, no. 1, pp. 64–70, 2016.
[17] A. Mekis *et al.*, "A grating-coupler-enabled CMOS photonics platform," *IEEE J. Sel. Topics Quantum Electron.*, vol. 17, no. 3, pp. 597–608, May/Jun. 2011.
[18] J. Ford, Y. Fainman, and S. Lee, "Reconfigurable array interconnection by photorefractive correlation," *Appl. Opt.*, vol. 33, no. 23, pp. 5363–5377, 1994.
[19] C. Wu and T. Feng, "On a class of multistage interconnection networks," *Trans. Comput.*, vol. C-29, no. 8, pp. 694–702, 1980.
[20] T. Beth and V. Hatz, "A restricted crossbar implementation and its applications," *ACM SIGARCH Comput. Archit. News*, vol. 19, no. 6, pp. 12–16, 1991.
[21] I. Stoica, R. Morris, D. Karger, M. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proc. ACM SIGCOMM Conf.*, 2001, pp. 149–160.
[22] J. Ford, V. Aksyuk, D. Bishop, and J. Walker, "Wavelength add-drop switching using tilting micromirrors," *J. Lightw. Technol.*, vol. 17, no. 5, pp. 904–911, May 1999.
[23] L. Nielson *et al.*, "Spatial superchannel routing in a two-span ROADM system for space division multiplexing," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 783–789, Feb. 2014.
[24] V. Kopp, J. Park, M. Wlodawski, J. Singer, D. Neugroschl, and A. Genack, "Chiral fibers: Microformed optical waveguides for polarization control, sensing, coupling, amplification, and switching," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 605–613, Feb. 2014.
[25] Mirrorcle Technologies, Richmond, CA, USA. (2015). [Online] Available: http://mirrorcletech.com/.
[26] R. Wagner and W. Tomlinson, "Coupling efficiency of optics in single-mode fiber components," *Appl. Opt.*, vol. 21, no. 15, pp. 2671–2688, 1982.
[27] FiberGuide Spec. Sheet. (2012). [Online] Available: http://www.fiberguide.com/wp-content/uploads/2012/08/V-Grooves_Arrays_FINAL.pdf
[28] J. Janhs and M. Murdocca, "Crossover networks and their optical implementation," *Appl. Opt.*, vol. 27, no. 15, pp. 3155–3160, 1998.
[29] T. Suleski and R. Te Kolste, "Fabrication trends for free-space microoptics," *J. Lightw. Technol.*, vol. 23, no. 2, pp. 633–646, Feb. 2005.

Authors' biographies not available at the time of publication.