



## Recent advances in optical technologies for data centers: a review

QIXIANG CHENG,\*  MEISAM BAHADORI, MADELEINE GLICK, SÉBASTIEN RUMLEY, AND KEREN BERGMAN

Lightwave Research Laboratory, Department of Electrical Engineering, Columbia University, New York, New York 10027, USA

\*Corresponding author: qc2228@columbia.edu

Received 27 November 2017; revised 18 September 2018; accepted 1 October 2018 (Doc. ID 314404); published 24 October 2018

**Modern data centers increasingly rely on interconnects for delivering critical communications connectivity among numerous servers, memory, and computation resources. Data center interconnects turned to optical communications almost a decade ago, and the recent acceleration in data center requirements is expected to further drive photonic interconnect technologies deeper into the systems architecture. This review paper analyzes optical technologies that will enable next-generation data center optical interconnects. Recent progress addressing the challenges of terabit/s links and networks at the laser, modulator, photodiode, and switch levels is reported and summarized.** ©2018 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](https://doi.org/10.1364/OPTICA.5.001354)

<https://doi.org/10.1364/OPTICA.5.001354>

### 1. INTRODUCTION

The explosive growth of Internet Protocol (IP) traffic is driving data centers to the so-called “Zettabyte Era,” as predicted by the Cisco Report [1,2] that expects annual global IP traffic to reach over 2.2 zettabytes/year by 2020. Much of this increase in traffic is dominated by video services and associated machine-learning applications. Predictions indicate that video traffic will be over 80% of all IP traffic by 2020—essentially that every second, one million minutes of video will cross the network.

The dramatic growth of cloud computing is further underscoring the need for vast access to data, compute, and storage resources. To accommodate these demands, the trend has been toward mega-data centers with hundreds of thousands of servers that benefit from economies of scale [3].

Despite the aforementioned tremendous growth of traffic into and out of the data center, the nature of the applications requires that at least three quarters of the traffic stay internal to the data center. This internal traffic is often referred to as east–west traffic (as opposed to north–south traffic that enters and exits the data center). For many data centers, the relative percentage of east–west traffic remains approximately the same as the traffic increases [2]. For some applications, such as those in Facebook, the internal traffic may be several orders of magnitude greater [4]. In addition to the internal traffic required to build web pages and search indices, relatively recent machine-learning applications are driving increasing amounts of both computation and traffic on the data center interconnection network. This increase can be seen as a result of the availability of large data sets (“big data”), increased computational capabilities, and advances in machine-learning algorithms [5]. Applications and network architectures drive the traffic patterns in the computer network. In both Google and Facebook, much of the computation does not fit on a single server, partly due to the large sizes of the data sets [4,6].

Data center power consumption is also a matter of significant importance. Not only is delivering more than 100 MW challenging for data center operators (in particular in terms of grid accessibility and reliability), operators must also be responsive to the increased public concerns about climate change and environmental issues, and the ultimate ecological footprint of data center activities [7]. From 1.5% of the total energy consumed in the U.S. at a cost of \$4.5B, the energy consumption of data centers has been predicted to triple by 2020 [8–10]. With this concern in mind, many large companies have made extensive efforts to reduce energy consumption. Future data center interconnects (DCIs) will thus be expected to carry more data while consuming less—the energy dissipated while transmitting a single bit over a link will have to be reduced to  $\sim 1$  pJ from several tens of pJ today [11]. This requires better provisioning of the available communication bandwidth within a data center network [12].

The emergence of these massive-scale data centers has given rise to important engineering requirements, including the need to keep the servers up and running with minimal human intervention, prevent irrecoverable data losses, and adequately exhaust the heat generated by hundreds of thousands of servers. These warehouse data centers require tailored high-bandwidth DCIs that can ensure adequate connectivity of the servers to each other and enable improved resource utilization [3,13,14,15]. At these scales, even apparently small improvements in performance or utilization can have a major impact on the overall network [6].

Here, we provide a review of optical technologies capable of meeting the requirements of the new generation of warehouse-scale intra-data-center interconnects. We start in Section 2 with a review of current trends for improving data center performance that have an impact on interconnects. In Section 3, we examine the optical transceiver technologies that make up the fabrication of a point-to-point optical interconnect. Special emphasis is

focused on technologies that permit the embedding of transceivers close to or within the chips that produce the data to be transmitted. The embedding of transceivers is widely considered to be a route to improved performance through higher bandwidth and energy efficiency. We follow with Section 4, focusing on optical switching for interconnection networks. Optical switches have been proposed to complement or replace conventional buffered packet routers [16,17]. We review the technologies involved in the design, fabrication, and operation of optical switch fabrics that are required. Finally, we summarize our conclusions in Section 5.

## 2. TRENDS IN DATA CENTERS

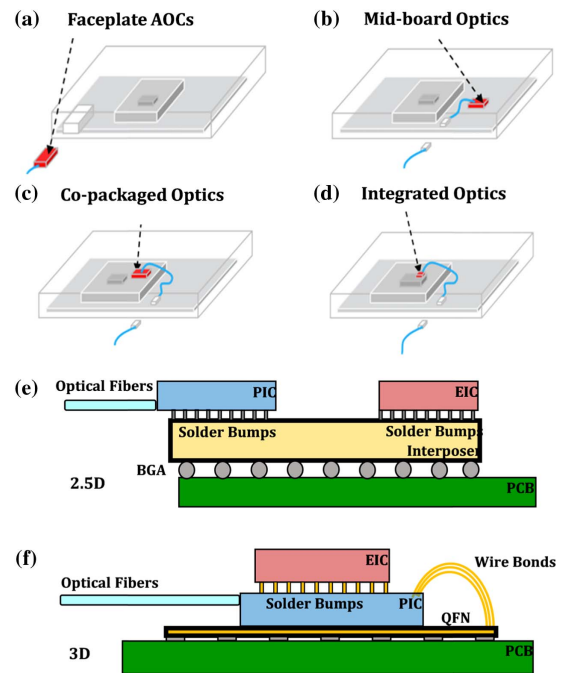
With the growing traffic, there are increasing stresses on the network and hardware. Some machine-learning applications use hundreds terabytes of data and are bounded by available resources. There is considerable effort in industry and academia to enhance performance through improvements at all levels of the architecture, software, and hardware. Proposals for implementations of new solutions, especially those utilizing new hardware, must adhere to data center metrics to maintain low cost. Despite apparent increased hardware cost, Facebook, Google, and Microsoft have found it economically justified to move towards multi-wavelength links, starting with coarse wavelength division multiplexing (CWDM) [18–20].

New architectures have been proposed to improve data center performance, many taking advantage of the high bandwidth density of optics and using optical switches [16,21]. The evaluation of the data center network depends on several metrics beyond those of cost and power consumption of the hardware and interconnect; data throughput and job completion time are also prime metrics. At the system level, these depend on many other factors including, for instance, scheduling packet transmission and congestion control. The description and comparison of these architectures are beyond the scope of this paper. We focus on the performance of the interconnect-level hardware, which is a basic building block of the entire system that can enhance or be a bottleneck to improved performance.

Two current trends for improving data center performance where the use of photonics is enabling are 1) high-bandwidth-density communication links and 2) improved resource utilization through disaggregation.

### A. High Bandwidth Links

There have been considerable advances in high-bandwidth pluggable optical interconnects for the data center. Large-scale data centers adopted optical transmission technologies during the transition from the 1 to 10 Gb/s link data rate between 2007 and 2010. In 2007, Google introduced optical communication in its data centers in the form of 10 Gb/s vertical-cavity surface-emitting laser (VCSEL) and multimode-fiber-based small form-factor pluggable (SFP) transceivers [22] for reaches up to 200 m [Fig. 1(a)]. As the intensity of traffic generated by servers doubles every year [14], transitions from 10 to 40 Gb/s, 40 to 100 Gb/s, and 100 Gb/s to even higher rates were predicted early on [23]. In 2017, 40 Gb/s Ethernet-based DCIs have been deployed in production data centers [14]. 100 Gb/s links have been commercially available since 2014, and are currently installed in production data centers. 400 Gb/s equipment is expected to emerge in the near future [22]. 400G transceivers are being standardized by the efforts of the IEEE 802.3 bs 400 Gb/s Task Force



**Fig. 1.** (a) Optical interface for active optical cables (AOCs) and pluggable transceivers. (b) Optical interface for board-mounted assembly. (c) Co-packaged optics with electronics (2.5D integration on an interposer). (d) Monolithic integration of optics and electronics. (e) Schematic of a 2.5D MCM co-integrating electronics and photonics via an interposer. (f) Schematic of a 3D integrated module. PIC, photonic integrated circuit; EIC, electronic integrated circuit; BGA, ball grid array; PCB, printed circuit boards; QFN, quad-flat no-leads.

on standardizing short-range (500 m – 10 km) intra-data-center interconnects over standard single mode fiber [24,25]. Data center servers will require even higher bitrates to connect their compute capabilities with their peers and the external world [11], especially with the adoption of machine-learning and neuromorphic-based algorithms. Application examples of these algorithms include voice assistants such as Apple Siri, Google Voice Search, and Amazon Alexa, as well as facial recognition applications. The recently introduced DGX-1 station from Nvidia, optimized for machine learning, utilizes 400 Gb/s of network bandwidth to ensure that its compute resources are adequately utilized [18]. In addition to expanded bandwidths, optical equipment with improved energy efficiency and compactness [7] is also expected.

At the link level, it is widely accepted that to achieve the required bandwidth density for the data center, the trend is towards onboard silicon photonics with 2.5D integration on a multichip module (MCM) [Fig. 1(e)] or with more advanced 3D integration using through silicon vias (TSVs) [Fig. 1(f)] for higher bandwidth and considerable energy savings compared to pluggable optics [26,27]. This is partly due to physical area limitations of the front panel of large data center switches [26], channel impairments over the path, and often the need for additional mid-board re-timers. QSFP56 based on 50 Gb/s signaling is expected to increase the front panel bandwidth to 7.2 Tb/s [28].

Although the concept of onboard optical transceivers is not new, the nearer term data center requirements have provoked vendors to push the technology forward to reduce cost by establishing the Consortium for On-Board Optics (COBO). COBO, led by Microsoft, is defining the standard for optical modules that

can be mounted or socketed on a network switch or adapter motherboard. Their initial focus has been on high-density 400 GbE applications [29], with large cloud providers as the early adopters. COBO does not define the optical physical layer but certain requirements that must be met to be considered COBO compliant.

The Integrated Photonics System Roadmap-International (IPSR-I) is developing a common roadmap for low-cost, high-volume manufacturing of photonics for data and telecommunications systems [30]. Given the requirement for high bandwidth density at low cost and power consumption, it is not surprising that silicon photonics, fabricated in high-volume CMOS-compatible foundries [31,32], is a prime candidate for the interconnection network. Among the key findings are that by 2020, there should be early deployment of “2.5-D” integrated photonic technologies, packages in which chips are placed side by side and interconnected through an interposer or substrate. By 2025, there should be pervasive deployment of wavelength-division-multiplexed (WDM) interconnects and the beginnings of commercial chip-to-chip intra-package photonic interconnects. In agreement with interconnect industry trends, the roadmap shows requirements of links to 1 Tbps on boards and 1–4 Tbps within a module by 2020. For the short  $\sim$  cm distance links on the modules, the energy target is on the order of 0.1 pJ/bit.

In Section 3, we review the key optical transceiver technologies that can meet the high bandwidth, low cost, and high energy efficiency requirements. We focus on solutions leveraging the economies of scale of silicon photonics.

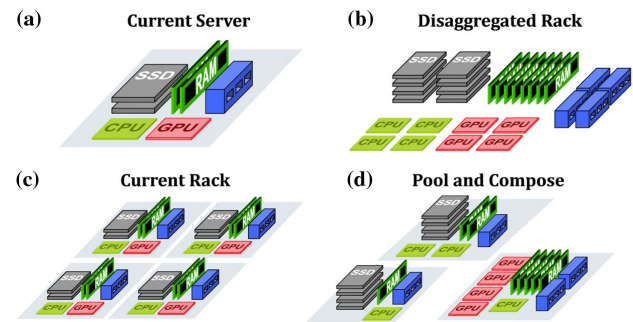
## B. Resource Utilization

The traditional data center is built around servers as building blocks. Each server is composed of tightly coupled resources: CPU, memory, one or more network interfaces, specialized hardware such as GPUs, and possibly some storage systems (hard or solid-state disks). This design has been facing several challenges. The various server elements follow different trends of cost and performance. As updated components become available, upgrading of the CPU or memory in the traditional design would require an entirely new server with new motherboard design [33]. Traditional data centers also suffer from resource fragmentation in cases where resources [CPU, memory, storage Input/Output (IO), network IO] are mismatched with workload requirements. For example, compute-intensive tasks that do not use the full memory capacity or communication-intensive tasks that do not fully use the CPU data gathered from data centers show that server memory for the former could be unused by as much as 50% or higher [34,35]. These challenges become motivations for disaggregation of the server.

### 1. Disaggregation

Disaggregation is a concept in which similar resources are pooled, with the possibility of the different resources being independently upgraded and the system adaptively configured for optimized performance. The network can be disaggregated at different levels, for example at the rack or server scale [34,36] (Fig. 2).

The disaggregated data center requires an interconnection fabric that must carry the additional traffic engendered by the disaggregation, and be high bandwidth and low latency in order to not only maintain, but also improve performance. The network requires a switching fabric to adaptively provision the



**Fig. 2.** Disaggregated rack places resources of different types (a–c) in different parts of the data center compared to traditional servers and uses networking to pool and compose needed resources together. In (d), a logical node can be constructed from distant resources. SSD, solid state drive; GPU, graphics processing unit; CPU, central processing unit; RAM, random-access memory.

computing resources. Although packet-switched networks remain electrical, optical circuit switches are prime candidates for reconfiguration of resources in the disaggregated network. In order to improve both high bandwidth performance and resource utilization, data center architectures with optical switch fabrics have been proposed [35,37,38].

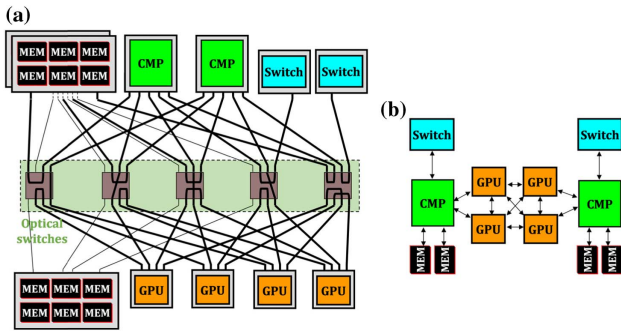
Attention must be paid to any added latency in the interconnect that might lead to performance degradation. Typical latency to memory, in a traditional server where the memory is close to the CPU, is on the order of tens of nanoseconds. The cost of the added interconnect compared to resource savings through improved utilization must also be balanced. Several groups have developed metrics or guidelines to achieve these goals [34,36].

Reference [34] explores a cost/performance analysis, including cost of latency and bandwidth, to determine when a data center disaggregated memory system would be cost competitive to a conventional direct attached memory system. The authors find that from a cost perspective, the current cost of an optically switched interconnect should be reduced by approximately a factor of 10 to be an economically viable solution.

### 2. Bandwidth Steering

We would expect best performance from an architecture that matches the traffic pattern of the application. The traffic pattern has the information regarding the communication between all possible pairs of sources and destinations in the network. Knowledge of the traffic patterns is therefore critical for optimizing the performance of the architecture. However, traffic patterns are often proprietary. In addition, there may be more or less variation depending on the specific network and its applications [37,39,40]. Therefore, there may be insufficient information on the current and future traffic patterns the architecture should support. A solution is the development of flexible, reactive networks that can utilize network resources efficiently and at low cost while meeting bandwidth and latency requirements. We proposed a Flexfly network that uses low to medium radix switches to rewire the interconnect as required by the application, rather than overprovisioning transceivers to achieve a high-bandwidth, low-energy interconnection network bandwidth steering. Bandwidth steering ensures full connectivity by adding optical connections as needed and routing traffic through intermediate





**Fig. 3.** (a) Example of bandwidth steering. Photonic switches may be used to assemble optimized nodes as (b) by configuration of the switches (within the dashed box). MEM, memory; GPU, graphics processing unit; CMP, chip multi-processor.

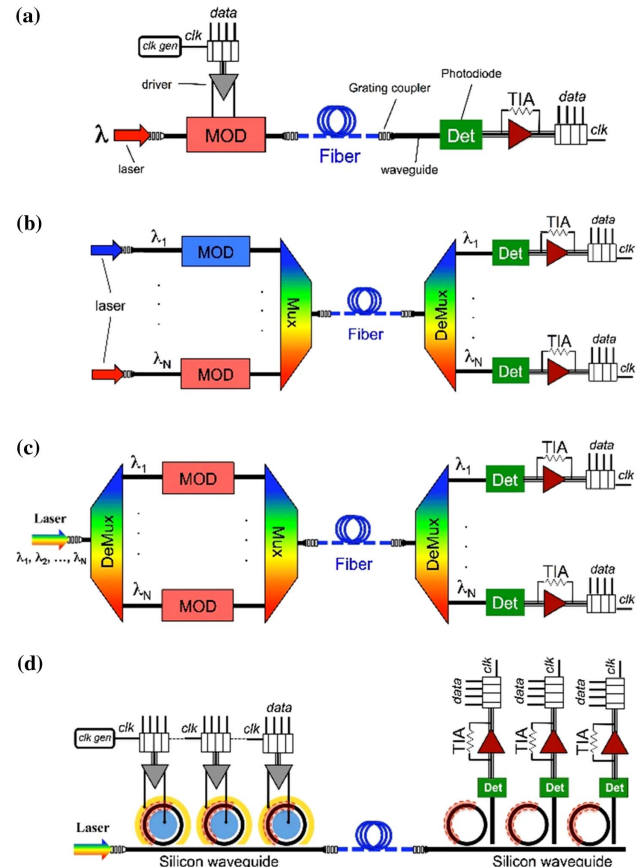
nodes, changing the configuration dynamically to match the application as shown schematically in Fig. 3 [41].

We review optical switching technologies in Section 4 and explore the performance metrics required for optical switches in data centers.

### 3. OPTICAL TRANSCEIVERS FOR DATA CENTERS

Optical transceivers are the fundamental building blocks of optical interconnects. Although all optical communication links have components fulfilling the same basic functions (light source, modulator, photodetector), the specific devices implemented depend on the application and economic considerations. For reliability and short-term development, it would be advantageous to use mature commercially available technologies. However, for intra-data-center networks, low cost (including energy costs) is a primary criterion. Given the much shorter transmission distances and shorter lifetime required of the devices, traditional telecommunications components are over-engineered and too costly to be adopted in the data center. This new and evolving application requires innovative solutions to meet the commercially required performance/cost trade-offs. Initially, and to this date, VCSELs, active optical cables, and parallel fiber transmission have been used inside the data center, where the trade-offs are advantageous over copper cables. With the massive increase of traffic inside the data center, the bandwidths required are increasing dramatically (similarly to telecommunications) and, although the distances are also increasing somewhat as the data center itself grows, the links are still significantly shorter than those used in telecommunications applications. As the lifetimes and other environmental requirements are still not the same as those for telecommunications, there has been an increased interest, in research and commercially, to explore and ready the next generation of devices that will meet the low-cost, low-energy-consumption requirements inside the data center.

Figure 4 schematically shows the anatomy of options for link architectures. The elements of the transceiver are the laser light source, modulator, (de)multiplexer, and photodetector. Figure 4(a) shows the transceiver design for a single channel link. The laser and modulator may be combined into one device element as in, for example, directly modulated VCSELs. Currently, VCSEL-based transceivers and parallel fibers are the dominant technology in the data center. As discussed previously, the roadmap for



**Fig. 4.** Anatomy of various link architectures: (a) single-wavelength point-to-point photonic link; (b) WDM photonic link based on separate lasers and broadband modulators; (c) photonic link based on a comb laser, parallel broadband modulators, and DeMux/Mux. (d) WDM photonic link based on comb laser, cascaded microring resonators, and cascaded drop filters. MOD, modulator; Det, detector; TIA, trans-impedance amplifier; CLK, clock; Mux, multiplexer; DeMux, demultiplexer.

ultra-high-bandwidth, low energy links requires WDM technology leveraging photonic integrated circuits. Figure 4(b) shows an approach, commonly used in telecommunications, combining modulated colored channels using (de)-multiplexers. Broadband optical modulators such as electro-absorption modulators (EAMs) and Mach-Zehnder modulators (MZMs) are used in combination with colored distributed feedback laser arrays. For the Tbps regime, however, a large number of lasers are required, imposing considerable overhead. The development of comb lasers emitting over 100 individual wavelengths is a promising next step. Figure 4(c) shows one potential architecture equipped with DeMux/Mux stages utilizing a multi-wavelength comb source and broadband modulators. Another promising architecture, illustrated in Fig. 4(d), takes advantage of the wavelength selective microring modulators implemented in a cascaded structure, enabling ultra-high on-chip bandwidth density. In this case, although individual lasers may be multiplexed for the WDM light source, a comb source is a promising solution.

#### A. Lasers

##### 1. Vertical-Cavity Surface-Emitting Lasers

Most of today's commercial short-reach (<300 m) optical interconnects employ GaAs-based VCSELs emitting at 850 nm.

VCSELs commonly apply a direct modulation scheme. Due to the small vertical cavity, very high modulation bandwidths (up to 30 GHz [42]) at low bias currents are possible.

To date, the fastest VCSEL-based link with NRZ modulation format is 71 Gb/s [43]. In order to achieve higher capacity and/or longer reach, single-mode [44,45] VCSELs with reduced spectral width are used. Impressive results have been reported utilizing advanced data-encoding schemes in 850 nm VCSEL-based links. Greater than 100 Gb/s transmission over 100 m multimode fiber (MMF), using both carrierless amplitude/phase [44] and four-level pulse-amplitude modulation (PAM4) [46] formats, has been reported. Discrete multi-tone (DMT) modulation schemes have achieved 161 and 135 Gb/s transmission over 10 and 550 m MMF, respectively [45].

Shortwave wavelength division multiplexing (SWDM) can further boost the transmission capacity in a single MMF. Recent work has shown the feasibility of employing four VCSELs at 40 Gb/s/ $\lambda$  in the 850 nm range (855, 882, 914, and 947 nm) transmitted over a 100 m standard OM4 fiber [47]. Very recently, the feasibility of SWDM PAM4 signaling transmission was demonstrated over 300 m of next-generation wideband MMF [48].

## 2. In-Plane Lasers

The increasing demands for longer distance and higher-capacity transmission within data centers favor dense wavelength division multiplexing technology that utilizes in-plane transmitters. As cost is a fundamental design criterion, attention has turned towards leveraging the economies of scale of silicon photonics for high-volume manufacturing.

The advancement in wafer-bonding techniques has led to the development of a new class of lasers, in which III-V gain layers are bonded on silicon wafers. Novel techniques such as transfer printing have been proposed for cost-effective integration of III-V materials on the silicon platform [49]. These so-called hybrid lasers can generally be separated into two groups. The first type is membrane lasers that apply benzocyclobutene or silica on silicon as cladding layers to provide strong optical confinement. In this way, the optical mode stays confined and gets amplified in the semiconductor core layer [50]. The large optical confinement factor enables high modulation speed at low current density [51]. The second type is facet-free evanescent lasers whose cavity is formed in the silicon layer, and the optical mode is either confined in the silicon waveguide overlapping with the III-V layer [52], or in the quantum wells (QWs) with adiabatic coupler transitioning from/to the silicon layer [53]. This method omits the time-consuming and costly optical alignment between the III-V laser and the silicon chip, driving down the assembly and packaging cost. Uncooled, wavelength-stabilized WDM hybrid laser arrays over the 60°C temperature range, i.e., 20°C to 80°C, have been reported [54].

Another promising approach lies in the epitaxial growth of III-V layers on silicon wafers using intermediate buffer layers [55]. The use of quantum dots (QDs) as the gain medium is less sensitive to defects than bulk or QW structures due to carrier localization, while maintaining low threshold, high power, and high temperature insensitivity. This approach can potentially break the wafer size and cost limit of using III-V substrates. An electrically pumped 1300 nm QD-on-silicon laser has been reported with comparable performance to lasers grown with

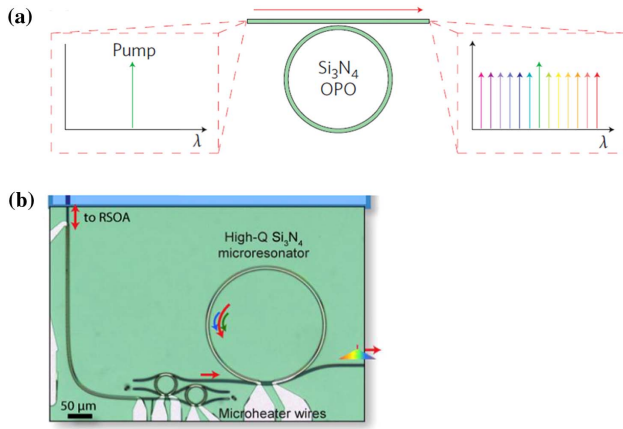
III-V substrates [55]. The thick buffer layers required, however, could be a bottleneck to the monolithic III-V-on-Si integration. The growth of a QD gain layer on silicon and then bonded to a patterned silicon-on-insulator (SOI) wafer, is envisioned as an alternative approach [56]. This solves the wafer size limit for wafer scale III-V to silicon bonding, offering better economies of scale.

## 3. Comb Lasers

Optical frequency combs are an appealing alternative to continuous-wave (CW) laser arrays as sources for DCIs in terms of footprint, cost, and energy consumption. A comb consists of equally spaced lines in the frequency domain which can be used as separate optical carriers for WDM. Since the comb is generated from a single source and has intrinsically equidistant spacing between its lines, it has the potential to eliminate the energy overhead associated with independently tuning many CW lasers to maintain the desired channel locking. Currently, there are two main methods used for generating combs: mode-locking in lasers and nonlinear generation using four-wave mixing in a microcavity. The merits of each are discussed in the following sections.

Comb generation can occur in a laser by inducing a fixed-phase relation between the longitudinal cavity modes in a Fabry–Perot cavity (mode-locking), leading to a stable pulse train in the time domain and therefore a comb with precise spacing in the frequency domain. The mode (channel) spacing can be tuned by changing the cavity length. QD mode-locked semiconductor lasers (QD-MLSLs) are an attractive candidate for DCI sources. The high nonlinear gain saturation of the QD active layer allows one to design a laser with low relative intensity noise. Moreover, intentional inhomogeneous broadening of the gain spectrum can be achieved by the control of size dispersion in QDs, and a 75 nm broad spectrum of emission has been reported [57]. The amplitude and phase noise of QD-MLSLs has been greatly reduced through active rather than passive mode-locking, therefore reducing the optical linewidths of the carriers and increasing the effective bandwidth compatible with coherent systems [58]. QD-MLSL has also been demonstrated by directly growing a passively mode-locked InAs/InGaAs QD laser on silicon [59].

Comb generation has been demonstrated with a silicon nitride ring resonator through the nonlinear process of four-wave mixing (FWM) in an optical parametric oscillator [60]. Numerous equally spaced narrow-linewidth sources can be generated simultaneously using a microresonator with an off-chip CW optical pump, as illustrated in Fig. 5(a) [60]. The pump field undergoes FWM in the resonator and creates signal and idler fields that also satisfy the cavity resonance; these signal and idler fields then seed further FWM, leading to a cascading effect which fills the remaining resonances of the cavity. This yields many equally spaced optical carriers [with spacing depending on the free spectral range (FSR) of the cavity] with a high pump-to-comb conversion efficiency of up to 31.8% when operating in the normal dispersion regime [62]. The device is CMOS compatible ( $\text{Si}_3\text{N}_4$ ) and can be integrated in the current silicon photonics platform. Recently, a chip-scale comb source with an integrated semiconductor laser pumping an ultra-high quality factor ( $Q$ )  $\text{Si}_3\text{N}_4$  ring resonator was reported [61] [shown in Fig. 5(b)]. This low-power-consumption and small-footprint device can run for  $\sim 200$  h from a single AAA battery, making it a strong candidate for future energy efficient DCI sources. [63] also shows a clear direction



**Fig. 5.** (a) On-chip optical comb generator using silicon nitride ring resonator with a single external pump laser [60]. (b) Chip-integrated, ultra low-power comb generator using an electrically pumped RSOA and a high-quality-factor silicon nitride ring resonator [61]. OPO, optical parametric oscillator; RSOA, reflective semiconductor optical amplifier.

towards a fully integrated transmitter which includes the comb source and modulators on a single chip.

Although a very promising direction of research, for comb sources to be adopted in DCIs, they must demonstrate advantages over CW laser arrays in terms of energy efficiency, cost, and footprint. To achieve this, combs should have a relatively flat profile with similar optical power per channel and with each optical channel power greater than the link budget. Furthermore, most of the comb lines must be utilized to ensure that optical power is not wasted. In the case of microresonator comb sources, most demonstrations have low optical power per comb line and require amplification to overcome the link power budget. Additionally, they have poor conversion efficiencies when in the anomalous dispersion regime ( $\sim 2\%$  pump-to-comb conversion efficiency), which poses a major challenge for the power consumption of the comb source when accounting for the wall-plug efficiency including the pump laser.

## B. Modulators

### 1. Electro-Absorption Modulators

Electro-absorption modulators (EAMs) are based on electric field-dependent absorption to alter the intensity of light, taking advantage of the Franz-Keldysh effect (FKE) in bulk semiconductors and the quantum-confined Stark effect (QCSE) in QW structures. FKE- and QCSE-based EAMs have been studied for many years in III-V materials [64–66] due to their advantages in small size, low chirp, and low driving voltage. Recent work demonstrated 100 Gb/s/λ PAM4 transmissions using lumped EAMs [64,65]. By transferring a III-V epitaxy stack to an SOI wafer, a hybrid silicon traveling-wave EAM has achieved a modulation bandwidth of 67 GHz [66].

It has been shown that FKE [67] and QCSE [68] are also effective in germanium (Ge). By epitaxial growth of Ge on Si, the tensile strain makes the band structure of Ge close to that of a direct bandgap material, enhancing the FKE [67]. The bulk Ge with tensile strain, however, normally exhibits the optimal electro-absorption contrast at wavelengths  $>1600$  nm [67]. Later work proposed a material composition of Ge-rich GeSi

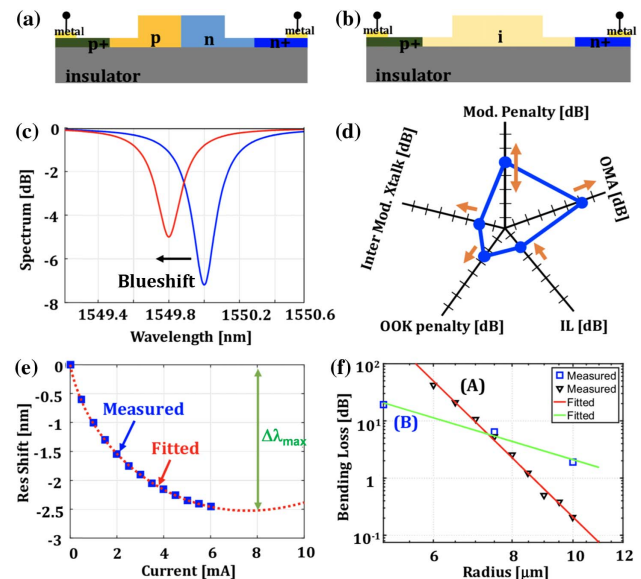
to shift the operating wavelength to the C-band [69]. Both Ge [70] and GeSi [71] based bulk EAMs have achieved 56 Gb/s modulation speed.

### 2. Electro-Refractive Modulators

Electro-refraction modulators (ERMs) operate by changing the index of material in the form of an interferometric or resonant structure. Combined amplitude/phase modulation can be realized with appropriate designs.

III-V ERMs are typically implemented in a Mach-Zehnder interferometer structure. Presently, InP-based phase modulators are usually implemented in deep-etched PIN epitaxial layers, and the core region is mostly a multi-quantum well stack, exploiting the strong nonlinear effect of QCSE. Modulation extinction as high as 14 dB operating at 40 Gb/s with segmented traveling-wave (STW) electrodes has been demonstrated owing to the reduced linear capacitance [72]. Higher-order amplitude/phase modulations of up to 256-QAM operating at 32 GBaud and energy consumption of 6.4 pJ/bit has been realized [73]. Hybrid integrated MZMs of III-V phase shifters on Si waveguide have demonstrated operation up to 40 Gb/s with 11.4 dB extinction ratio [74].

Silicon-based high-speed ERMs rely on the plasma dispersion effect (PDE). A forward-biased PIN diode or reverse-biased PN junction is used to inject or deplete carriers, respectively [75] [see Figs. 6(a) and 6(b)]. Carrier-injection modulators (PIN) exhibit a higher electro-optic efficiency and better modulation extinction but their speed is limited due to the carrier dynamics (recombination lifetime) [79]. MZMs with STW electrodes operating at 56 Gb/s have been reported [80]. Higher-order amplitude and phase modulation formats such as PAM4 at 128 Gb/s [81], binary phase shift keying at 48 Gb/s [82], differential phase shift



**Fig. 6.** (a) Cross-section of a PN-based modulator. (b) Cross-section of a PIN-based modulator. (c) Example of spectral response of a PIN-based microring modulator. (d) Power penalty space of microring modulators based on the spectral shift. (e) Spectral shift of a PIN-based ring modulator as a function of injected current [76]. (f) Measured bending loss of ring resonators as a function of radius reported in [77] and [78] (both horizontal and vertical axes are in log scale). OMA, optical modulation amplitude; OOK, on-off keying; IL, insertion loss.



keying at 10 Gb/s, quadrature phase shift keying (QPSK) at 50 Gb/s [83], and polarization multiplexed QPSK at 112 Gb/s [84] have also been realized with silicon MZMs.

In 2005, it was first demonstrated that a compact footprint ( $\sim 10$   $\mu\text{m}$  diameter) silicon micro-ring resonator (MRR) could provide adequate phase shift by taking advantage of the PDE, leading to more than 10 dB of extinction ratio with a low insertion loss [85]. In 2013, the first 50 Gb/s modulation at 1.96 V peak-to-peak drive voltage for a silicon racetrack MRR was demonstrated [86]. In order to enhance the modulation efficiency and reduce the effective junction capacitance, MRR modulators with ZigZag [87] and interdigitated [88] junctions have been proposed. MRRs have also been used for higher-order amplitude modulation formats such as four-level pulse amplitude modulation (PAM4) at 128 Gb/s [89] and PAM8 at 45 Gb/s [90], with an energy consumption of as low as 1 fJ/bit [90].

### 3. Cascaded Ring Modulators

Small footprint in conjunction with wavelength-selective nature (hence compatible with comb lasers [91]) make MRR modulators highly promising for realizing high-throughput optical interconnects. This is mainly realized by cascading MRRs along a single bus waveguide, as demonstrated by Xu *et al.* [92], Brunina *et al.* [93], and Li *et al.* [94].

In order to quantify the performance of an MRR modulator in a cascaded architecture, the power penalty metric associated with the bit error rate is typically used (a study of the power penalties of an array of cascaded ring modulators is presented in [95]). The modulated light has a certain optical modulation amplitude (OMA) based on the spectral shift of the resonator. Figure 6(c) shows an example of the spectral response of a PIN-based ring modulator in which the spectrum of the MRR experiences a blueshift with the addition of some excess cavity loss. Based on the driving voltage/current of the modulator, the change in the cavity phase shift and roundtrip loss inside the ring and 3 dB bandwidth can be estimated [76,96].

Figure 6(d) shows the dimensions of modulator penalty space in a cascaded WDM arrangement. We have shown that intermodulation crosstalk [95,97] can impact the overall power penalty of modulators in such cascaded arrangement. The tradeoff is between the spacing between the channels and the shift of resonance. A larger shift of resonance results in an improved OMA and lower modulator insertion loss but leads to a higher average loss of optical power due to on-off keying functionality (OOK penalty) and higher intermodulation crosstalk. Starting from a low OMA, the power penalty of the modulator is very large, and gradually increasing the OMA will improve it initially until the intermodulation crosstalk dominates the power penalty and starts to deteriorate. This is depicted as the double-sided arrow on the power penalty axis in Fig. 6(d), where the chosen point qualitatively shows the sweet spot of the power penalty.

In [95], we have shown that the sweet spot for the shift of resonance is close to half of the spacing between channels. However, PIN-based modulators suffer from Ohmic heating due to the injection of current inside the waveguide. Figure 6(e) shows an example of measured results reported in [76], indicating that the Ohmic heating limits the blueshift of the spectrum to about 2.5 nm. This situation is even worse for PN-based ring modulators due to their relatively low electro-optic modulation efficiency [98]. PN-based modulators therefore exhibit higher

optical penalty compared to their PIN-based counterparts, but benefit from operating at higher speeds [99]. The choice of PN or PIN design for ring modulators therefore is based on the desired optical penalty and the operation speed.

A key step to establishing the design space exploration of microring modulators is to relate the spectral parameters ( $Q$  factor, roundtrip loss) to the geometrical parameters (radius, coupling gaps) [100], in which case the bending loss of silicon ring resonators stands out as a critical factor. Figure 6(f) shows two sets of measurements for the bending loss of silicon ring resonators (in dB/cm units) as a function of the radius, reported in [77] (case A) and [78] (case B). A large FSR favors the support of more optical channels in the cascaded WDM configuration but requires a small radius, leading to high bending loss. MRR modulators with free spectral range as large as 22 nm (diameter  $< 10$   $\mu\text{m}$ ) have been demonstrated for dense WDM systems, capable of operating at 15 Gb/s [101].

## C. (De)-Multiplexers

Optical (de)-multiplexers are generally based on thin films [102], diffraction gratings [103], arrayed waveguide gratings (AWGs) [104–109], or microring-based filters. With the advance of photonic integration technology, the latter two have drawn the most attention.

### 1. Arrayed Waveguide Gratings

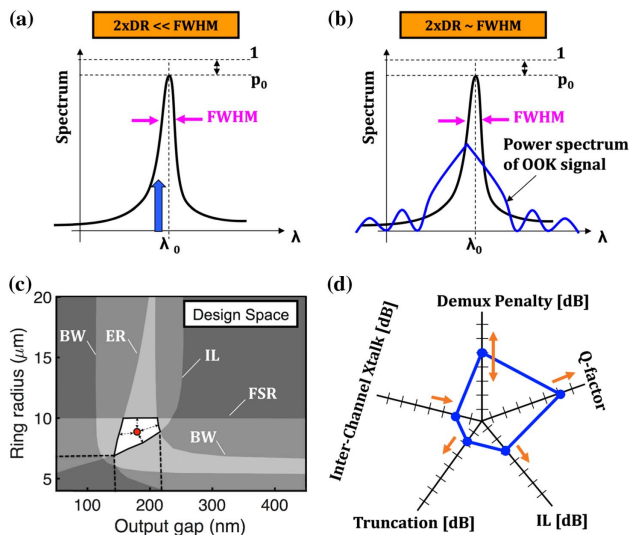
AWGs have been demonstrated in silica-on-silicon [104–106], InP [107,108], and SOI platforms [109,110]. Demonstrations of silica AWGs have shown 400 channels at 25 GHz spacing [104], ultra-low loss of 0.75 dB [105], and ultra-low crosstalk of 82 dB [106]. InP-based AWGs have been integrated with WDM transmitters [107] and receivers [108].

### 2. Microring-Based Drop Filters

With the advent of silicon photonics, MRRs in the form of add-drop structures capable of performing wavelength demultiplexing due to their wavelength selective spectral response have been developed. Based on the desired passband and the rejection ratio of the filter, first-order [111,112] or higher-order [113] add-drop filters are used. Higher-order filters provide a better rejection ratio but suffer from a higher loss in their passbands.

The power penalty of ring filters is typically estimated based on the Lorentzian spectral shape of the filter. As shown in Fig. 7(a), if the data rate of the OOK channel is much smaller than the 3 dB bandwidth of the ring, then the power penalty is simply given by the spectral attenuation of the MRR. However, in [95], we showed that when the data rate is comparable to the bandwidth of the filter [Fig. 7(b)], this interpretation loses its accuracy. We proposed a first-order correction that includes the impact of data rate on the power penalty of the MRR filter and the crosstalk effects in cascaded arrangement [111]. Furthermore, in [78], we explored the design space of silicon-based add-drop filters based on radius and coupling gap spacing. The metrics of the design space are the insertion loss (IL) at the resonance set to  $< 1$  dB, optical bandwidth (BW) set to 10–50 GHz range, FSR of the filter set to  $> 10$  nm, the extinction of resonance set to  $> 30$  dB, and the assumption that the ring is at critical coupling. By overlaying the contours of these metrics, the design space is depicted in Fig. 7(c).

For each individual channel, we proposed an optimization of the add-drop MRR in the cascaded arrangement such that the



**Fig. 7.** (a) Impact of the spectral filtering of a demux ring when data rate is much smaller than the optical bandwidth. (b) Impact of the spectral filtering of a demux ring when data rate is comparable to the optical bandwidth. (c) Design space of a critically coupled demux add-drop ring. (d) Power penalty space of microring demux based on the  $Q$  factor. DR, data rate; FWHM, full width at half maximum; OOK, on-off keying; BW, bandwidth; ER, extinction ratio; IL, insertion loss; FSR, free spectral range.

power penalty associated with the whole demultiplexer array is minimized [95]. This optimization depends on the parameters of the ring, as well as data rate, number of channels, and channel spacing. The power penalty imposed on each channel consists of three parts: (a) Insertion loss of the ring—independent of the number of channels and their data rate. (b) Truncation effect—only dependent on the data rate. Strong truncation arises when the 3 dB bandwidth of the MRR is small compared to the signal bandwidth. (c) Optical crosstalk due to the imperfect suppression of adjacent channels—a function of number of channels and channel spacing. As shown in Fig. 7(d), the  $Q$  factor of the MRRs is the determining factor in the power penalty space [95]. Increasing the  $Q$  will increase the IL of the ring and truncation of OOK signal, but results in suppression of optical crosstalk. Therefore, a sweet spot exists for the minimal penalty.

In order to fully take advantage of MRRs' potentials for optical interconnects, some challenges must be overcome. The main challenges in using silicon-based MRRs as optical modulators or wavelength selective filters can be categorized as follows:

**Thermal Sensitivity:** Thermal effects significantly impact the optical response of silicon-based resonant devices due to the strong thermo-optic coefficient of silicon material. The resonance of a typical silicon MRR is shifted by  $\sim 9$  GHz for each degree Kelvin change in the temperature [114]. We have shown that such thermal drift in high- $Q$  MRRs can impose significant spectral distortion on high-speed OOK signals (more than 1 dB of penalty [95]).

**Self-heating:** The enhancement of optical power inside the MRR is proportional to the finesse (or  $Q$ -factor). In the presence of high optical power inside a high- $Q$  MRR, even a slight internal absorption can lead to a noticeable thermal drift of resonance. A recent transceiver design has proposed a thermal tuning algorithm

based on the statistics of the data stream to counteract this effect [115].

**Fabrication variation:** The spectral parameters of MRRs such as resonance wavelength, FSR, and the 3 dB optical bandwidth largely depend on their geometrical parameters. It is known that current silicon photonic fabrication imposes variations on the dimensions of the waveguides [116]. This results in deviations of the resonance wavelength from the original design [117] and requires thermal tuning, hence degrading the energy efficiency of the link. Various wavelength-locking schemes based on analog and digital feedback [118], bit statistics [115], as well as pulse width modulation and thermal rectification [114] have been proposed and implemented.

**Backscattering:** In applications where narrow optical linewidths (i.e.,  $Q > 10000$ ) are required, even a slight roughness on the sidewalls of the MRRs will cause backreflections inside the ring [119]. The effect of backscattering in MRRs is typically observed in the form of a splitting of the resonance in the spectral response [120]. We have shown that such spectral distortion adds extra complexity to the design of optical link and further narrows down the design space of MRRs [121].

## D. Photodetectors

Cost is the main metric followed by requirements of bandwidth, wavelength, and energy depending on the system level design. For these reasons, to date mostly integrated PINs have been used, although there is research in other types of PDs including uni-traveling-carrier (UTC) and avalanche photodetectors (APD). PIN photodiodes have an intrinsic, i.e., undoped, region between the  $n$ - and  $p$ -doped regions within which photons are absorbed, generating carriers of electrons and holes, and thus photocurrent. UTC PDs are comprised of an un-depleted  $p$ -type photo-absorption layer and a wide-gap depleted carrier-collection layer. In the  $p$ -doped absorbing layer, photo-generated majority holes recombine within the dielectric relaxation time, making electrons the only carrier traveling in the depletion layer [122]. Higher speed can be achieved as electrons having higher velocity than holes, and this also overcomes the space charge effect, enabling high output power. APDs consist of two separate sections: absorption and multiplication regions. A charge layer is inserted in between the two regions to control the electric field distribution, forming a typical separation-of-absorption-charge-multiplication structure [123,124]. PINs are most widely used due to their ultra-compact footprint and low dark current. They are also relatively simple to fabricate, and have shown the best bandwidth-efficiency product. UTC devices are used for applications that require both high bandwidth and high output power. The UTC PD can potentially omit the electrical post-amplification [125], which is beneficial in terms of system simplicity and cost. APDs can largely enhance the receiver sensitivity and thus are preferable for situations that require high link budget [124].

Waveguide-integrated PDs are used to improve the bandwidth-efficiency product, where the bandwidth is primarily limited by the RC-time constant. To couple the light from waveguide to the photodiode, butt coupling [126–128] or evanescent coupling [129–131] have been used. InP-based PIN photodiodes have achieved a responsivity of 0.5 A/W and bandwidths up to 120 GHz [129]. A hybrid PIN PD with a 3 dB bandwidth of 65 GHz and 0.4 A/W responsivity has been shown [130].



An InP-based UTC photodiode with 105 GHz bandwidth and 1.3 dBm RF output power has been reported [132].

Ge/GeSi PDs have used both lateral and vertical geometries, and both butt- and evanescent-coupled schemes have been demonstrated. Evanescent-coupled devices are more tolerant to fabrication errors [133]. A butt-coupled lateral Ge-on-Si PIN photodiode was reported with more than 100 GHz bandwidth, responsivity of 0.8 A/W, and zero-bias 40 Gb/s operation [126]. An evanescent-coupled vertical Ge-on-Si PIN PD with a 3 dB bandwidth of 45 GHz and responsivity of 0.8 A/W has been shown [131]. A vertical Si/Ge UTC PD was demonstrated with a responsivity of 0.5 A/W and 40 GHz bandwidth [134]. Ge-Si-based APDs have shown better sensitivity (-22.5 dBm at 25 Gb/s) than their III-V counterparts due to the superior impact ionization property of silicon [135].

## 4. OPTICAL SWITCHING TECHNOLOGIES FOR DATA CENTERS

### A. Free-Space Optical Switches

Free-space optical switches have been realized by a number of competing technologies, including micro-electromechanical systems (MEMS), beam-steering from Polaris, and liquid crystal on silicon (LCOS). All of the three have been commercialized. Among them, MEMS-based optical switches are the most common and mature free-space switching devices. An electrostatic driver is commonly utilized due to its low power consumption and ease of control. The typical voltage required, however, is relatively high, and can reach up to 100–150 V [16,136]. MEMS spatial switches are realized in both two-dimensional (2D) and three-dimensional (3D) configurations. The 2D implementation is arranged in the crossbar topology and operates digitally as the mirror position is bi-stable. 3D MEMS switches were proposed to support very large-scale optical cross-connect devices [137]. This type of device is assembled by using 2D input and output fiber arrays with collimators. Two stages of independent 2D micro-mirror arrays are used to steer the optical beams in three dimensions. This design requires micro-mirrors implemented with a two-axis tilting structure [137].

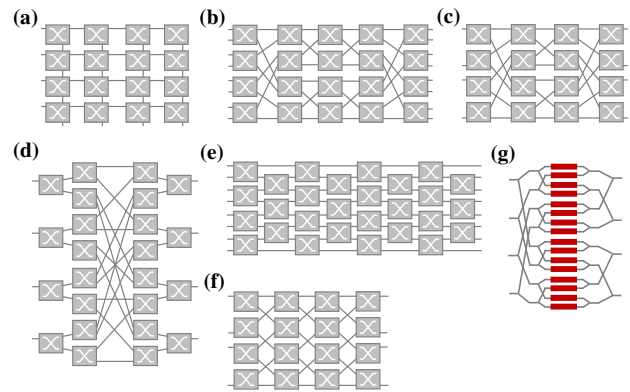
MEMS switch systems support connectivity of hundreds of ports [137]; however, the installation and calibration with surface-normal micro-optics introduces considerable complexity that is ultimately reflected in the cost per port. This is a challenge to the broad introduction of MEMS switches in data centers.

### B. Photonic Integrated Switches

To ensure low cost per port and eventual data center adoption, optical switching technologies must demonstrate a path towards high-volume manufacturing. This generally means lithography-based fabrication and high-level integration. In this subsection, we review these switching technologies, focusing on III-V and silicon platforms.

#### 1. Optical Integrated Switching Technologies and Topologies

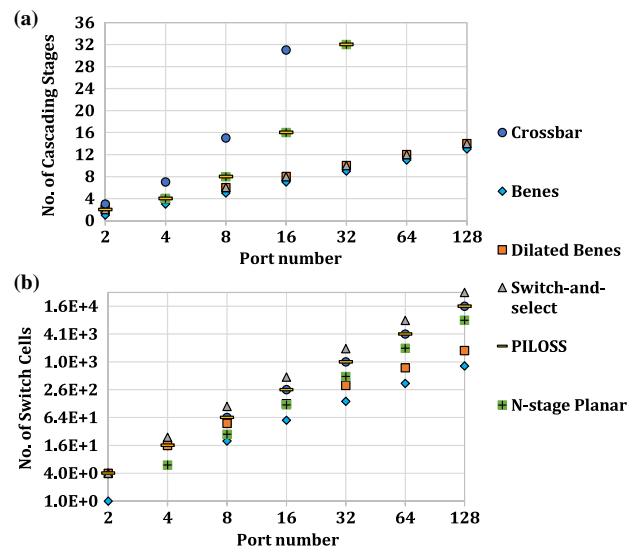
A number of physical mechanisms have been explored to trigger the optical switching process in integrated devices. The most widely applied mechanisms include phase manipulation in interferometric structures, both thermally and electrically actuated, i.e., MZI and MRR, signal amplification/absorption in semiconductor



**Fig. 8.** Schematic of optical switch topologies: (a) Crossbar, (b) Beneš, (c) dilated Beneš, (d) switch-and-select, (e)  $N$ -stage planar, (f) PILOSS, and (g) broadcast-and-select. Red rectangles represent SOA elements.

optical amplifiers (SOAs), and MEMS-actuated coupling between different layers of waveguides.

$N \times N$  optical switch fabrics can be built up by connecting the basic switch cells into switch fabric topologies. The selection of topology depends critically on switching performance. The key figures of merit include blocking characteristics, crosstalk suppression, total number of switch cells, and number of cascading stages. The first factor has significant impact on the interconnect connectivity and scheduling, while the last three largely decide the switch scalability. The commonly utilized topologies include crossbar, Beneš, dilated Beneš, switch-and-select (S&S), path-independent-loss (PILOSS),  $N$ -stage planar, and broadcast-and-select (B&S), shown in Fig. 8, all of which support non-blocking connections [138]. To better reveal the impact of switch topologies, we outline in Fig. 9 the number of cascading stages per path and the total number of switch cells required for each type of topology as a function of port number  $N$  in an  $N \times N$  network, potentially denoting to insertion loss, aggregated crosstalk, footprint, and cost.



**Fig. 9.** (a) Number of cascading stages of switch cells per path and (b) the total number of switch cells required for each type of topology as a function of port number  $N$  in an  $N \times N$  network.

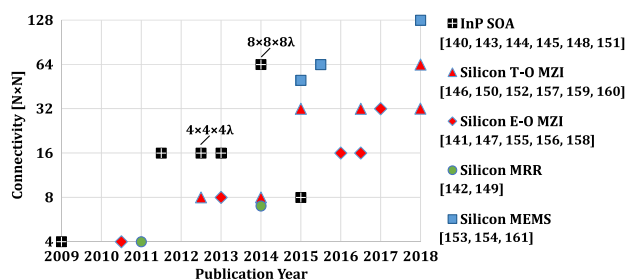
It can be seen that the Beneš architecture requires the smallest number of cascading stages as well as switching cells. It is a recursive network constructed from the Clos architecture with a bank of  $2 \times 2$  switch cells forming each of the input and output stages. The main constraint for Beneš architecture is that each switch cell has two traversing signals, resulting in first-order crosstalk. This imposes stringent requirements on the performance of the elementary switch cells. The dilated Beneš network was proposed to cancel the first-order crosstalk by introducing more cascading stages and switch cells. The  $N$ -stage planar topology was put forward to eliminate waveguide crossings, which could be a drawback to large-scale integration. The S&S topology also blocks first-order crosstalk, but it requires the largest number of switch cells and the central shuffle network becomes complex at high radices. The PILOSS architecture is preferred when uniform performance across all paths is necessary. This is critical to relax the dynamic range requirement on receivers. The PILOSS switch is not immune to first-order crosstalk, though not all switch cells carry two signals at once [139]. The crossbar topology has the highest number of worst-case cascading stages, leading to a large path-dependent loss. However, it is nicely suited to add-drop switching cells such as MRRs and MEMS-actuated couplers, in which case the signal only gets dropped once.

It should be noted that the crossbar, S&S, and PILOSS architectures support either wide-sense non-blocking (WSNB) or strictly non-blocking (SNB) connections that can ease the burden of schedulers, compared to the rearrangeable non-blocking (RNB) networks. RNB networks can establish connections of all permutations of input ports to output ports if rerouting existing connections are allowed. WSNB networks can set up paths between any idle inputs to any idle outputs without interference with existing connections provided that the routing rules are followed, while SNB networks guarantee that any paths can be established without restrictions. Details on switch-blocking characteristics can be found in [138].

## 2. State-of-the-Art

In the last decade, photonic integration technologies have quickly matured to see monolithic integrated circuits of a few thousands of components with increasingly sophisticated functionalities. Figure 10 and Table 1 summarize notable demonstrations of monolithic switch fabrics published during the past 10 years.

InP-based switch fabrics have primarily applied SOA gated elements in the B&S topology [Fig. 8(g)]. B&S networks utilize passive splitters/combiners, and each path is gated by an SOA element. The inherent splitting/combining losses can be compensated by the SOA gain. The B&S architecture provides the prospect of chip-level multicast [162], but the optical loss due



**Fig. 10.** High connectivity optical switch matrix technologies highlighted in terms of input side connectivity.

**Table 1.** Notable Demonstrations of Photonic Integrated Switches

Port Count	Topology	Material Platform	Actuating Element	IL <sup>a</sup> [dB]	CT <sup>a</sup> [dB]	Refs.
4 × 4	B&S	InP	SOA Gate	0	–	[140]
4 × 4	Beneš	SOI	E-O MZI	5.8 <sup>b</sup>	–9	[141]
4 × 4	Hitless	SOI	T-O MRR	–	20 <sup>b</sup>	[142]
16 × 16	Clos-tree	InP	SOA Gate	0 <sup>b</sup>	–	[143]
16 × 16	Clos-tree	InP	SOA Gate	17.2 <sup>b</sup>	50 <sup>c</sup>	[144]
4 × 4 × 4λ	Space-and-λ	InP	SOA Gate	9.4 <sup>b</sup>	40 <sup>c</sup>	[145]
8 × 8	S&S	SOI	T-O MZI	11.8	30	[146]
8 × 8	Custom blocking	SOI	E-O MZI	–	–13	[147]
8 × 8 × 8λ	Space-and-λ	InP	SOA Gate	14.7 <sup>b</sup>	–	[148]
8 × 7	Cross-bar	SOI	T-O MRR	22	20 <sup>c</sup>	[149]
8 × 8	PILOSS	SOI	T-O MZI	6.5 <sup>b</sup>	–23.1	[150]
8 × 8	Clos-tree	InP	SOA Gate	0 <sup>b</sup>	50 <sup>c</sup>	[151]
32 × 32	PILOSS	SOI	T-O MZI	16.8 <sup>b</sup>	–20	[152]
50 × 50	Cross-bar	SOI	MEMS	9.6 <sup>b</sup>	50 <sup>c</sup>	[153]
64 × 64	Cross-bar	SOI	MEMS	5.1 <sup>b</sup>	65 <sup>c</sup>	[154]
16 × 16	Beneš	SOI	E-O MZI	–	–17.9	[155]
16 × 16	Beneš	SOI	E-O MZI	15 <sup>b</sup>	–18	[156]
32 × 32	Custom	SOI	T-O MZI	23 <sup>b</sup>	–35	[157]
32 × 32	Beneš	SOI	E-O MZI	18.5 <sup>b</sup>	–15.1	[158]
32 × 32	PILOSS	SOI	T-O MZI	13.2	–20	[159]
64 × 64	Beneš	SOI	T-O MZI	12 <sup>b</sup>	–	[160]
128 × 128	Crossbar	SOI	MEMS	22.7 <sup>b</sup>	–	[161]

<sup>a</sup>Worst-case reported.

<sup>b</sup>On-chip loss excludes coupling.

<sup>c</sup>Extinction ratio.

to the various signal splits and recombinations discourage scaling beyond  $4 \times 4$  connectivity. Multistage architectures involving cascaded switching elements have been proposed as a trade-off [163].  $16 \times 16$  port count SOA-based switches have been demonstrated using both all-active [143] and passive-active [144] integration schemes. Higher on-chip connectivity was achieved by combining spatial ports with wavelength channels using co-integrated AWGs, scaling up to  $64 \times 64$  connections [148]. Improved performance and/or further scale-up would require a large reduction in component-level excess loss, a more careful design of balancing the summed loss and gain per stage, and a close examination of SOA designs for linear operation [151,154,164,165].

The highly advanced CMOS industry, with mature fabrication and manufacturing infrastructures as well as advances in silicon photonic manufacturing capabilities, has stimulated development in silicon photonic switches. The current record for monolithic photonic switch radix is  $64 \times 64$  and  $128 \times 128$  for a thermo-optic MZI-based Beneš switch [160] and MEMS-actuated crossbar switch [161], respectively. Other notable results include a  $32 \times 32$  thermally actuated PILOSS MZI switch with  $<13.2$  dB insertion loss [159] and  $32 \times 32$  electro-optic MZI-based Beneš switch [158]. Detailed insertion loss (IL) and crosstalk (CT) results are listed in Table 1.

## 3. Challenges

The switch circuits just described show the feasibility of a high-level photonic integration; however, their current performance is far from practical for real applications in data centers. *The intrinsic limitations are loss and crosstalk.* Studies have been carried out to improve the performance of single switch cells. The nested MZI switch fabrics [166,167] with improved crosstalk suppression are

a notable example. The approach from the architectural point of view is discussed in the following section. Insertion loss, however, is more challenging to manage for large-scale switch fabrics. The SOA is a natural solution to provide on-chip gain. A hybrid MZI-SOA design approach was proposed for InP integrated switches, in which the distributed short SOA elements provide additional gain and absorption [168,169]. For the silicon platform, the recent report of a lossless SOA-integrated  $4 \times 4$  PILOSS switch, leveraging the flip-chip bonding technique, was a significant demonstration [170]. In addition to device-level optimization, we have proposed a fabric-wide advanced routing method at the control domain, providing routing strategies that opt for optimal solutions [171,172].

With the scale-up of switch port count, the requirement of efficient calibration/testing methods for such complex integrated circuits is severe and urgent. Power taps are usually utilized with either couplers [158] or built-in photodiodes [157]. The additional components could substantially increase the device insertion loss, complexity, and package cost, leading to a reduction of yield. We proposed an algorithm-based method for device calibration and characterization to eliminate the built-in power checkpoints, and automated implementation was demonstrated to allow cost-effective switch device testing [173–175].

Another challenge would be the packaging solution for large-scale switch circuits. We will review our recent progress on switch packaging in the Section 4.D.

### C. Optical Switching Metrics for Data Centers

In order to benefit from advantages of optical switching in the data center, challenges presented by the traditional architecture must be overcome. The absence of a commercially viable solution for optical random access memory and optical buffering makes it unlikely that true optical packet switching (that is, optical switches that can operate on a packet-by-packet basis) will emerge in the near term. Optical switches thus cannot be considered as a one-to-one replacement for electronic packet switches. The network architecture will likely employ optical switches in some combination with conventional electronic switches for improved performance. Prior research includes C-through [176,177], Helios [21], and Mordia [178]. In all these examples, optical switches are used to adapt the network to specific traffic patterns. In other words, pairs of racks exchanging higher levels of traffic can be awarded more bandwidth using the optical network [12]. A detailed discussion can be found in [16].

Reconfiguration of an optical switch breaks prior links and thus requires phase locking, handshaking, and modification of the routing tables for the new links. A state-of-the-art 25 Gb/s burst-mode receiver requires 31 ns locking time even with interlocking search algorithms [179], and updating routing tables in OpenFlow routers requires on the order of milliseconds [180]. For switch reconfiguration times in the millisecond range, greater performance gains can be obtained if optical reconfigurations are programmed to match task granularity [12].

Optical switches must fit within the optical link power budget. In addition, switch power penalties have an impact on the launch laser power. Adding optical amplifiers as discrete components to each link in a data center would increase the cost and energy consumption per link, and is generally considered an unacceptable solution. The development of lasers and/or amplifiers with

improved wall-plug efficiency would ameliorate the impact of switch power penalty on the end-to-end link power consumption.

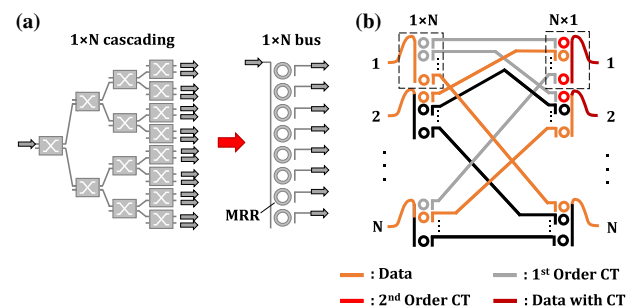
The presence of optical switching in a network can improve the utilization of the limited resources available. Improving utilization of the resources by reconfiguration of disaggregated elements could enable the reduction of components or energy consumption by putting underused components in a sleep mode. It is, however, also possible to mitigate network congestion resulting from intense communications between servers or rack pairs by overprovisioning the network. This reasoning leads us to posit that for optical switching to be competitive, the cost associated with making a network reconfigurable must be considerably smaller than the cost of adding additional links, resulting in similar network performance.

Switches offering the highest number of ports would be preferred over low-port-count switches, as the former will enable more flexibility. However, as shown in [12], the benefits of being able to reorganize a network can be achieved even with modest radix switches, with a break-even point at 8 to 16 ports. This leads us to conclude that for applications in large-scale data center interconnects, optical switches require a minimum number of 16 to 32 ports.

### D. Towards Large-Scale, High-Performance Photonic Integrated Switches for Data Centers

As reviewed in Section 3.B.1, signal degradation resulting from accumulated crosstalk and loss exacerbates as radix scales up with increased cascading stages. To support the scaling of integrated photonic switches, the switch architecture should be re-examined in terms of crosstalk cancellation, number of cascading switch stages, and total number of switch cells.

We therefore proposed an MRR-based modified S&S switching circuit. The concept is illustrated by Fig. 11(a), in which the  $1 \times N$  switch unit is built from MRR add-drop cells assembled in a bus coupled structure. Scaling such a structure requires only adding MRRs to the bus waveguide, which effectively reduces the scaling overhead in loss compared to that of the MZI cascading scheme. The layout of a generic  $N \times N$  S&S MRR-based switch is depicted in Fig. 11(b). This configuration has  $N$  input spatial  $1 \times N$  and  $N$  output spatial  $N \times 1$  units, maintaining the number of drop (i.e., resonating) microrings in any path at two. This topology fully blocks the first-order crosstalk, as illustrated in Fig. 11(b). The wavelength-selective nature of the MRR unit does require wavelength alignment across the switching circuit, adding extra overhead. Various schemes for fast and efficient wavelength locking have been demonstrated, as discussed in Section 3.D.2.



**Fig. 11.** (a)  $1 \times N$  MZI-based cascading structure versus  $1 \times N$  MRR-based bus structure. (b) MRR-based switch-and-select topology. MRR, microring resonator; CT, crosstalk.

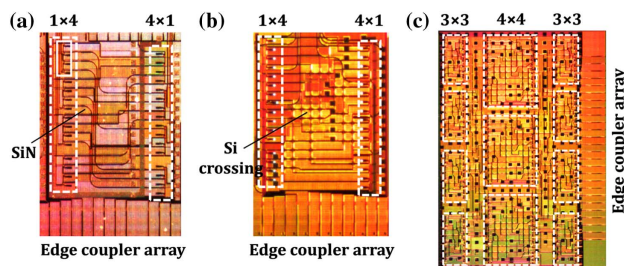


One challenge is that the complexity of the central shuffle network scales exponentially with the port count  $N$ , and it has been recognized that the key limiting factor in loss, crosstalk, and footprint is the waveguide crossings. We proposed leveraging the Si/SiN multi-layer structure to achieve crossing-free designs [181,182], which can substantially improve the performance and scalability. The other challenge lies in the total number of MRRs, which scales poorly at  $2N^2$ . For large port count numbers, we performed further studies on combining the scalable three-stage Clos network with populated S&S stages [183]. The proposed design offers balance that keeps the number of stages to a modest value while largely reducing the required number of switching elements. The scalability is predicted to be  $128 \times 128$  [183].

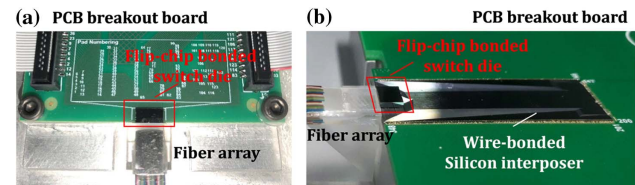
Prototyped devices have been fabricated through the foundry of the American Institute of Manufacturing (AIM photonics). All designs used the pre-defined elements in the process design kit library to ensure high yield and low cost, and followed the design rules of standard packaging houses, i.e., Tyndall National Institute, to be compatible with low-cost packaging solutions. Figure 12 shows the microscope photo of a  $4 \times 4$  Si/SiN dual-layer S&S switch, a  $4 \times 4$  silicon S&S switch, and a  $12 \times 12$  silicon Clos switch, all utilizing thermo-optic MRRs.

All devices have been fully packaged. Small radix switch photonic integrated circuits (PICs) were directly flip-chip bonded onto a PCB breakout board with a UV-cured fiber array, as shown in Fig. 13(a). For densely integrated Clos switches, we developed a packaging platform using the silicon interposer as an electrical re-distribution layer that keeps the PICs ultra-compact for reduced insertion loss. Figure 13(b) presents the packaged  $12 \times 12$  Clos switch, which was first flip-chip bonded onto a silicon interposer and then wire-bonded to a PCB breakout board. A ball grid array (BGA)-based multi-layer silicon interposer with TSVs is currently being taped out, which dramatically shrinks the footprint of the interposers. Excellent testing results were achieved for the fully packaged  $4 \times 4$  S&S switch, with on-chip loss and crosstalk ratio as low as 1.8 and  $-50$  dB, respectively [182].

Looking forward, lossless design would be a significant advantage within data centers because it not only avoids any need for additional optical amplification, but also allows the optical transmitters to operate at moderate output power and removes the need for extensive electrical amplification at the receiver side. We envision a new class of III-V/Si heterogeneously integrated optical switches leveraging the advanced bonding techniques to provide compact, energy-efficient, and low-cost switch fabrics satisfying the data center switching metrics as discussed in Section 4.C. The implementation can follow the approach



**Fig. 12.** Microscope photo of (a)  $4 \times 4$  Si/SiN dual-layered MRR-based S&S switch, (b)  $4 \times 4$  Si MRR-based S&S switch, and (c)  $12 \times 12$  Si MRR-Based Clos switch with populated S&S stages.



**Fig. 13.** (a) Packaged switch device by flip-chip bonding on the breakout PCB board. (b) Packaged  $12 \times 12$  Si MRR-Based Clos switch by flip-chip bonding onto a silicon interposer.

demonstrated in the InP MZI-SOA switch fabrics [184,185] or be combined with the switch-and-select topology of MRR add-drop multiplexers. Detailed discussions can be found in [16].

## 5. CONCLUSIONS

Optical interconnects are moving towards Tb/s scale throughput to keep up with the demands of massively increasing traffic in the data center. Many optical technologies for modern warehouse-scale data centers are maturing, in particular technologies that enable optical systems with a high level of integration as well as technologies that offer large-scale fabrication at low cost. In this paper, we have presented an overview of data center trends in interconnection networks which are most likely to be enabled by advancing photonic capabilities. To that end, we have presented a brief summary of the optical transceiver components (lasers, modulators, (de)multiplexers, and photodetectors) that will enable ultra-high (to greater than Tbps) bandwidth links, in addition to an overview of optical switching technologies that will enable improved resource utilization and thereby lower cost and energy consumption. We expect to see expanding integration of optical technologies into data centers within the next decade, enabling advances in new applications such as machine learning and artificial intelligence and providing cost-effective, fast, and reliable services to users worldwide.

**Funding.** Air Force Research Laboratory (AFRL) (FA8650-15-2-5220); Advanced Research Projects Agency—Energy (ARPA-E) (ENLITENED); U.S. Department of Energy (DOE) (DE-AR0000843); Defense Advanced Research Projects Agency (DARPA) (PERFECT); Rockport Networks, Inc.

## REFERENCES

1. Cisco Visual Networking, "The Zettabyte Era—Trends and Analysis," Cisco white paper (2017).
2. Cisco Global Cloud Index, "Index 2013–2018 (2014)," Cisco Systems Inc., San Jose (2016).
3. A. Greenberg, J. R. Hamilton, N. Jain, S. Kandula, C. Kim, P. Lahiri, D. A. Maltz, P. Patel, and S. Sengupta, "VL2: a scalable and flexible data center network," *SIGCOMM Comput. Commun. Rev.* **39**, 51–62 (2009).
4. N. Farrington and A. Andreyev, "Facebook's data center network architecture," in *Optical Interconnects Conference* (IEEE, 2013).
5. K. Hazelwood, S. Bird, D. Brooks, S. Chintala, U. Diril, D. Dzhulgakov, M. Fawzy, B. Jia, Y. Jia, and A. Kalro, "Applied machine learning at Facebook: a datacenter infrastructure perspective," in *IEEE International Symposium on High Performance Computer Architecture (HPCA)* (IEEE, 2018).
6. S. Kanev, J. P. Darago, K. Hazelwood, P. Ranganathan, T. Moseley, G.-Y. Wei, and D. Brooks, "Profiling a warehouse-scale computer," in *ACM/IEEE 42nd Annual International Symposium on Computer Architecture (ISCA)* (IEEE, 2015).

7. G. Cook and J. V. Horn, *How Dirty is Your Data? A Look at the Energy Choices that Power Cloud Computing* (Greenpeace, 2011).
8. K. Bilal, S. U. Khan, and A. Y. Zomaya, "Green data center networks: challenges and opportunities," in *11th International Conference on Frontiers of Information Technology (FIT)* (IEEE, 2013).
9. J. Koomey, *Growth in Data Center Electricity Use 2005 to 2010* (Analytical, 2011), Vol. 9.
10. K. Church, A. G. Greenberg, and J. R. Hamilton, "On delivering embarrassingly distributed cloud services," in *HotNets* (2008).
11. 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).
12. K. Wen, P. Samadi, S. Rumley, C. P. Chen, Y. Shen, M. Bahadori, K. Bergman, and J. Wilke, "Flexfly: enabling a reconfigurable dragonfly through silicon photonics," in *International Conference for High Performance Computing, Networking, Storage and Analysis, SC16* (IEEE, 2016).
13. A. Vahdat, M. Al-Fares, N. Farrington, R. N. Mysore, G. Porter, and S. Radhakrishnan, "Scale-out networking in the data center," *IEEE Micro* **30**, 29–41 (2010).
14. A. Singh, J. Ong, A. Agarwal, G. Anderson, A. Armistead, R. Bannon, S. Boving, G. Desai, B. Felderman, P. Germano, A. Kanagala, H. Liu, J. Provost, J. Simmons, E. Tanda, J. Wanderer, U. Holzle, S. Stuart, and A. Vahdat, "Jupiter rising: a decade of Clos topologies and centralized control in Google's datacenter network," *Commun. ACM* **59**, 88–97 (2016).
15. S. A. Reinemo, T. Skeie, and M. K. Wadekar, "Ethernet for high-performance data centers: on the new IEEE datacenter bridging standards," *IEEE Micro* **30**, 42–51 (2010).
16. Q. Cheng, S. Rumley, M. Bahadori, and K. Bergman, "Photonic switching in high performance datacenters [Invited]," *Opt. Express* **26**, 16022–16043 (2018).
17. K. Bergman and S. Rumley, "Optical switching performance metrics for scalable data centers," in *21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)* (2016).
18. R. Urata, H. Liu, L. Verslegers, and C. Johnson, "Silicon photonics technologies: gaps analysis for datacenter interconnects," in *Silicon Photonics III* (Springer, 2016), p. 473–488.
19. R. Urata, H. Liu, X. Zhou, and A. Vahdat, "Datacenter interconnect and networking: From evolution to holistic revolution," in *Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2017).
20. A. Chakravarty, K. Schmidtko, V. Zeng, S. Giridharan, C. Deal, and R. Niazmand, "100 Gb/s CWDM4 optical interconnect at Facebook data centers for bandwidth enhancement," in *Laser Science* (Optical Society of America, 2017).
21. 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.* **40**, 339–350 (2010).
22. X. Zhou, H. Liu, and R. Urata, "Datacenter optics: requirements, technologies, and trends (Invited Paper)," *Chin. Opt. Lett.* **15**, 120008 (2017).
23. C. F. Lam, "Optical network technologies for datacenter networks (invited paper)," in *Conference on Optical Fiber Communication (OFC/INFOEC), collocated National Fiber Optic Engineers Conference* (2010).
24. R. Sun, V. Nguyen, A. Agarwal, C.-Y. Hong, J. Yasaitis, L. Kimerling, and J. Michel, "High performance asymmetric graded index coupler with integrated lens for high index waveguides," *Appl. Phys. Lett.* **90**, 201116 (2007).
25. T. Rokkas, I. Neokosmidis, B. Shariati, and I. Tomkos, "Techno-economic evaluations of 400G optical interconnect implementations for datacenter networks," in *Optical Fiber Communication Conference* (Optical Society of America, 2018).
26. A. Ghiasi, "Large data centers interconnect bottlenecks," *Opt. Express* **23**, 2085–2090 (2015). <http://onboardoptics.org/>.
27. R. H. Johnson and D. M. Kuchta, "30 Gb/s directly modulated 850 nm datacom VCSELs," in *Conference on Lasers and Electro-Optics/Quantum Electronics and Laser Science Conference and Photonic Applications Systems Technologies*, San Jose, California (2008).
28. M. Filer, B. Booth, and D. Bragg, "The role of standards for cloud-scale data centers," in *Optical Fiber Communication Conference*, San Diego, California (2018).
29. M. Glick, L. C. Kimmerling, and R. C. Pfahl, "A roadmap for integrated photonics," *Opt. Photon. News* **29**(3), 36–41 (2018).
30. D. Thomson, A. Zilkie, J. E. Bowers, T. Komljenovic, G. T. Reed, L. Vivien, D. Marris-Morini, E. Cassan, L. Viroat, and J.-M. Fédéli, "Roadmap on silicon photonics," *J. Opt.* **18**, 073003 (2016).
31. E. R. H. Fuchs, R. E. Kirchain, and S. Liu, "The future of silicon photonics: not so fast? Insights from 100 G ethernet LAN transceivers," *J. Lightwave Technol.* **29**, 2319–2326 (2011).
32. <https://www.intel.com/content/www/us/en/it-management/intel-it-best-practices/disaggregated-server-architecture-drives-data-center-efficiency-paper.html>.
33. B. Abali, R. J. Eickemeyer, H. Franke, C.-S. Li, and M. A. Taubenblatt, "Disaggregated and optically interconnected memory: when will it be cost effective?" arXiv:1503.01416 (2015).
34. G. Zervas, H. Yuan, A. Saljoghei, Q. Chen, and V. Mishra, "Optically disaggregated data centers with minimal remote memory latency: technologies, architectures, and resource allocation," *J. Opt. Commun. Netw.* **10**, A270–A285 (2018).
35. P. X. Gao, A. Narayan, S. Karandikar, J. Carreira, S. Han, R. Agarwal, S. Ratnasamy, and S. Shenker, "Network requirements for resource disaggregation," in *OSDI* (2016).
36. M. Ghobadi, R. Mahajan, A. Phanishayee, N. Devanur, J. Kulkarni, G. Ranade, P.-A. Blanche, H. Rastegarfar, M. Glick, and D. Kilper, "Projector: Agile reconfigurable data center interconnect," in *Proceedings of the 2016 ACM SIGCOMM Conference* (ACM, 2016).
37. W. M. Mellette, R. McGuinness, A. Roy, A. Forencich, G. Papen, A. C. Snoeren, and G. Porter, "RotorNet: a scalable, low-complexity, optical datacenter network," in *Proceedings of the Conference of the ACM Special Interest Group on Data Communication* (ACM, 2017).
38. T. Benson, A. Akella, and D. A. Maltz, "Network traffic characteristics of data centers in the wild," in *Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement* (ACM, 2010).
39. A. Roy, H. Zeng, J. Bagga, G. Porter, and A. C. Snoeren, "Inside the social network's (datacenter) network," in *ACM SIGCOMM Computer Communication Review* (ACM, 2015).
40. Y. Shen, A. Gazman, Z. Zhu, M. Y. Teh, M. Hattink, S. Rumley, P. Samadi, and K. Bergman, "Autonomous dynamic bandwidth steering with silicon photonic-based wavelength and spatial switching for Datacom network," in *Optical Fiber Communication Conference* (Optical Society of America, 2018).
41. E. Haglund, P. Westbergh, J. S. Gustavsson, E. P. Haglund, A. Larsson, M. Geen, and A. Joel, "30 GHz bandwidth 850 nm VCSEL with sub-100 fJ/bit energy dissipation at 25–50 Gbit/s," *Electron. Lett.* **51**, 1096–1098 (2015).
42. D. M. Kuchta, A. V. Rylyakov, F. E. Doany, C. L. Schow, J. E. Proesel, C. W. Baks, P. Westbergh, J. S. Gustavsson, and A. Larsson, "A 71-Gb/s NRZ modulated 850-nm VCSEL-based optical link," *IEEE Photon. Technol. Lett.* **27**, 577–580 (2015).
43. R. Puerta, M. Agustin, L. Chorchos, J. Toński, J. R. Kropp, N. Ledentsov, V. A. Shchukin, N. N. Ledentsov, R. Henker, I. T. Monroy, J. J. V. Olmos, and J. P. Turkiewicz, "107.5 Gb/s 850 nm multi- and single-mode VCSEL transmission over 10 and 100 m of multi-mode fiber," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2016).
44. C. Kottke, C. Caspar, V. Jungnickel, R. Freund, M. Agustin, and N. N. Ledentsov, "High speed 160 Gb/s DMT VCSEL transmission using pre-equalization," in *Optical Fiber Communication Conference*, Los Angeles, California (2017).
45. F. Karinou, N. Stojanovic, C. Prodaniuc, Z. Qiang, and T. Dippon, "112 Gb/s PAM-4 optical signal transmission over 100-m OM4 multimode fiber for high-capacity data-center interconnect," in *42nd European Conference on Optical Communication (ECOC)* (2016).
46. D. M. Kuchta, T. Huynh, F. Doany, A. Rylyakov, C. L. Schow, P. Pepeljugoski, D. Gazula, E. Shaw, and J. Tatum, "A 4- $\lambda$  40 Gb/s/ $\lambda$  bandwidth extension of multimode fiber in the 850 nm rang," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2015).
47. Y. Sun, R. Lingle, R. Shubochkin, A. H. Mccurdy, K. Balemarthy, D. Braganza, J. Kamino, T. Gray, W. Fan, K. Wade, F. Chang, D. Gazula, G. Landry, J. Tatum, and S. Bhoja, "SWDM PAM4 transmission



- over next generation wide-band multimode optical fiber," *J. Lightwave Technol.* **35**, 690–697 (2017).
49. A. De Groote, P. Cardile, A. Z. Subramanian, A. M. Fecioru, C. Bower, D. Delbeke, R. Baets, and G. Roelkens, "Transfer-printing-based integration of single-mode waveguide-coupled III-V-on-silicon broadband light emitters," *Opt. Express* **24**, 13754–13762 (2016).
  50. S. Arai, N. Nishiyama, T. Maruyama, and T. Okumura, "GaInAsP/InP membrane lasers for optical interconnects," *IEEE J. Sel. Top. Quantum Electron.* **17**, 1381–1389 (2011).
  51. S. Matsuo, T. Fujii, K. Hasebe, K. Takeda, T. Sato, and T. Kakitsuka, "40-Gbit/s direct modulation of membrane buried heterostructure DFB laser on SiO<sub>2</sub>/Si substrate," in *International Semiconductor Laser Conference* (2014).
  52. A. W. Fang, H. Park, O. Cohen, R. Jones, M. J. Paniccia, and J. E. Bowers, "Electrically pumped hybrid AlGaInAs-silicon evanescent laser," *Opt. Express* **14**, 9203–9210 (2006).
  53. H. Duprez, A. Descos, T. Ferrotti, C. Sciancalepore, C. Jany, K. Hassan, C. Seassal, S. Menezo, and B. Ben Bakir, "1310 nm hybrid InP/InGaAsP on silicon distributed feedback laser with high side-mode suppression ratio," *Opt. Express* **23**, 8489–8497 (2015).
  54. B. R. Koch, E. J. Norberg, B. Kim, J. Hutchinson, J.-H. Shin, G. Fish, and A. Fang, "Integrated silicon photonic laser sources for telecom and datacom," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference*, Anaheim, California (2013).
  55. S. Chen, W. Li, J. Wu, Q. Jiang, M. Tang, S. Shuttis, S. N. Elliott, A. Sobiesierski, A. J. Seeds, I. Ross, P. M. Smowton, and H. Liu, "Electrically pumped continuous-wave III-V quantum dot lasers on silicon," *Nat. Photonics* **10**, 307–311 (2016).
  56. J. E. Bowers, J. T. Bovington, A. Y. Liu, and A. C. Gossard, "A path to 300 mm hybrid silicon photonic integrated circuit," in *OFC* (2014).
  57. A. Kovsh, I. Krestnikov, D. Livshits, S. Mikhlin, J. Weimert, and A. Zhukov, "Quantum dot laser with 75 nm broad spectrum of emission," *Opt. Lett.* **32**, 793–795 (2007).
  58. V. Panapakkam, A. P. Anthur, V. Vujicic, R. Zhou, Q. Gaimard, K. Merghem, G. Aubin, F. Lelarge, E. A. Viktorov, L. P. Barry, and A. Ramdane, "Amplitude and phase noise of frequency combs generated by single-section InAs/InP quantum-dash-based passively and actively mode-locked lasers," *IEEE J. Quantum Electron.* **52**, 1–7 (2016).
  59. S. Liu, J. C. Norman, D. Jung, M. Kennedy, A. C. Gossard, and J. E. Bowers, "Monolithic 9 GHz passively mode locked quantum dot lasers directly grown on on-axis (001) Si," *Appl. Phys. Lett.* **113**, 041108 (2018).
  60. J. S. Levy, A. Gondarenko, M. A. Foster, A. C. Turner-Foster, A. L. Gaeta, and M. Lipson, "CMOS-compatible multiple-wavelength oscillator for on-chip optical interconnects," *Nat. Photonics* **4**, 37–40 (2010).
  61. B. Stern, X. Ji, Y. Okawachi, A. L. Gaeta, and M. Lipson, "Fully integrated chip platform for electrically pumped frequency comb generation," in *Conference on Lasers and Electro-Optics*, San Jose, California (Optical Society of America, 2018).
  62. X. Xue, P. H. Wang, Y. Xuan, M. Qi, and A. M. Weiner, "High-efficiency WDM sources based on microresonator Kerr frequency comb," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2017).
  63. J. Pfeifle, V. Brasch, M. Laueremann, Y. Yu, D. Wegner, T. Herr, K. Hartinger, P. Schindler, J. Li, D. Hillerkuss, R. Schmogrow, C. Weimann, R. Holzwarth, W. Freude, J. Leuthold, T. J. Kippenberg, and C. Koos, "Coherent terabit communications with microresonator Kerr frequency combs," *Nat. Photonics* **8**, 375–380 (2014).
  64. M. A. Mestre, H. Mardoyan, C. Caillaud, R. Rios-Müller, J. Renaudier, P. Jennevé, F. Blache, F. Pommereau, J. Decobert, F. Jorge, P. Charbonnier, A. Konczykowska, J.-Y. Dupuy, K. Mekhazni, J.-F. Paret, M. Faugeron, F. Mallecot, M. Achouche, and S. Bigo, "Compact InP-based DFB-EAM enabling PAM-4 112 Gb/s transmission over 2 km," *J. Lightwave Technol.* **34**, 1572–1578 (2016).
  65. S. Kanazawa, T. Fujisawa, K. Takahata, T. Ito, Y. Ueda, W. Kobayashi, H. Ishii, and H. Sanjoh, "Flip-chip interconnection lumped-electrode EADFB laser for 100-Gb/s/λ transmitter," *IEEE Photon. Technol. Lett.* **27**, 1699–1701 (2015).
  66. Y. Tang, J. D. Peters, and J. E. Bowers, "Over 67 GHz bandwidth hybrid silicon electroabsorption modulator with asymmetric segmented electrode for 1.3 μm transmission," *Opt. Express* **20**, 11529–11535 (2012).
  67. S. Jongthammanurak, J. Liu, K. Wada, D. D. Cannon, D. T. Danielson, D. Pan, L. C. Kimerling, and J. Michel, "Large electro-optic effect in tensile strained Ge-on-Si films," *Appl. Phys. Lett.* **89**, 161115 (2006).
  68. Y.-H. Kuo, Y. K. Lee, Y. Ge, S. Ren, J. E. Roth, T. I. Kamins, D. A. B. Miller, and J. S. Harris, "Strong quantum-confined Stark effect in germanium quantum-well structures on silicon," *Nature* **437**, 1334–1336 (2005).
  69. J. Liu, D. Pan, S. Jongthammanurak, K. Wada, L. C. Kimerling, and J. Michel, "Design of monolithically integrated GeSi electro-absorption modulators and photodetectors on an SOI platform," *Opt. Express* **15**, 623–628 (2007).
  70. S. A. Srinivasan, M. Pantouvaki, S. Gupta, H. T. Chen, P. Verheyen, G. Lepage, G. Roelkens, K. Saraswat, D. V. Thourhout, P. Absil, and J. V. Campenhout, "56 Gb/s germanium waveguide electro-absorption modulator," *J. Lightwave Technol.* **34**, 419–424 (2016).
  71. P. D. Heyn, V. I. Kopp, S. A. Srinivasan, P. Verheyen, J. Park, M. S. Wlodawski, J. Singer, D. Neugroschl, B. Snyder, S. Balakrishnan, G. Lepage, M. Pantouvaki, P. Absil, and J. V. Campenhout, "Ultra-dense 16 × 56 Gb/s NRZ GeSi EAM-PD arrays coupled to multicore fiber for short-reach 896 Gb/s optical link," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2017).
  72. S. Akiyama, H. Itoh, S. Sekiguchi, S. Hirose, T. Takeuchi, A. Kuramata, and T. Yamamoto, "InP-based Mach-Zehnder modulator with capacitively loaded traveling-wave electrodes," *J. Lightwave Technol.* **26**, 608–615 (2008).
  73. M. Schell, G. Fiol, and A. Aimone, "DAC-free generation of M-QAM signals with InP segmented Mach-Zehnder modulator," in *Optical Fiber Communication Conference*, Los Angeles, California (2017).
  74. H.-W. Chen, J. D. Peters, and J. E. Bowers, "Forty Gb/s hybrid silicon Mach-Zehnder modulator with low chirp," *Opt. Express* **19**, 1455–1460 (2011).
  75. G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," *Nat. Photonics* **4**, 518–526 (2010).
  76. R. Wu, C. H. Chen, J. M. Fedeli, M. Fournier, R. G. Beausoleil, and K. T. Cheng, "Compact modeling and system implications of microring modulators in nanophotonic interconnect," in *ACM/IEEE International Workshop on System Level Interconnect Prediction (SLIP)* (2015).
  77. H. Jayatilaka, K. Murray, M. Caverley, N. A. F. Jaeger, L. Chrostowski, and S. Shekhar, "Crosstalk in SOI microring resonator-based filters," *J. Lightwave Technol.* **34**, 2886–2896 (2016).
  78. M. Bahadori, M. Nikdast, S. Rumley, L. Y. Dai, N. Janosik, T. Van Vaerenbergh, A. Gazman, Q. Cheng, R. Polster, and K. Bergman, "Design space exploration of microring resonators in silicon photonic interconnects: impact of the ring curvature," *J. Lightwave Technol.* **36**, 2767–2782 (2018).
  79. Q. Xu, S. Manipatruni, B. Schmidt, J. Shakya, and M. Lipson, "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Opt. Express* **15**, 430–436 (2007).
  80. A. Mahendra, D. M. Gill, C. Xiong, J. S. Orcutt, B. G. Lee, T. N. Huynh, J. E. Proesel, N. Dupuis, P. H. W. Leong, B. J. Eggleton, and W. M. J. Green, "Monolithically integrated CMOS nanophotonic segmented Mach Zehnder transmitter," in *Conference on Lasers and Electro-Optics (CLEO)* (2017).
  81. A. Samani, D. Patel, M. Chagnon, E. El-Fiky, R. Li, M. Jacques, N. Abadia, V. Veerasubramanian, and D. V. Plant, "Experimental parametric study of 128 Gb/s PAM-4 transmission system using a multi-electrode silicon photonic Mach Zehnder modulator," *Opt. Express* **25**, 13252–13262 (2017).
  82. Q. Li, R. Ding, Y. Liu, T. Baehr-Jones, M. Hochberg, and K. Bergman, "High-speed BPSK modulation in silicon," *IEEE Photon. Technol. Lett.* **27**, 1329–1332 (2015).
  83. P. Dong, L. Chen, C. Xie, L. L. Buhl, and Y.-K. Chen, "50 Gb/s silicon quadrature phase-shift keying modulator," *Opt. Express* **20**, 21181–21186 (2012).
  84. P. Dong, C. Xie, L. Chen, L. L. Buhl, and Y.-K. Chen, "112-Gb/s monolithic PDM-QPSK modulator in silicon," *Opt. Express* **20**, B624–B629 (2012).
  85. Q. Xu, B. Schmidt, S. Pradhan, and M. Lipson, "Micrometre-scale silicon electro-optic modulator," *Nature* **435**, 325–327 (2005).
  86. T. Baba, S. Akiyama, M. Imai, N. Hirayama, H. Takahashi, Y. Noguchi, T. Horikawa, and T. Usuki, "50 Gb/s ring-resonator-based silicon modulator," *Opt. Express* **21**, 11869–11876 (2013).



87. X. Xiao, X. Li, H. Xu, Y. Hu, K. Xiong, Z. Li, T. Chu, J. Yu, and Y. Yu, "44-Gb/s silicon microring modulators based on zigzag PN junctions," *IEEE Photon. Technol. Lett.* **24**, 1712–1714 (2012).
88. M. Pantouvaki, H. Yu, M. Rakowski, P. Christie, P. Verheyen, G. Lepage, N. V. Hoovels, P. Absil, and J. V. Campenhout, "Comparison of silicon ring modulators with interdigitated and lateral p-n junctions," *IEEE J. Sel. Top. Quantum Electron.* **19**, 7900308 (2013).
89. J. Sun, M. Sakib, J. Driscoll, R. Kumar, H. Jayatileka, Y. Chetrit, and H. Rong, "A 128 Gb/s PAM4 silicon microring modulator," in *Optical Fiber Communications Conference and Exposition (OFC)* (2018).
90. R. Dubé-Demers, S. Laroche, and W. Shi, "Ultrafast pulse-amplitude modulation with a femtojoule silicon photonic modulator," *Optica* **3**, 622–627 (2016).
91. C.-H. Chen, M. A. Seyed, M. Fiorentino, D. Livshits, A. Gubenko, S. Mikhlin, V. Mikhlin, and R. G. Beausoleil, "A comb laser-driven DWDM silicon photonic transmitter based on microring modulators," *Opt. Express* **23**, 21541–21548 (2015).
92. Q. Xu, B. Schmidt, J. Shakya, and M. Lipson, "Cascaded silicon microring modulators for WDM optical interconnection," *Opt. Express* **14**, 9431–9436 (2006).
93. D. Brunina, X. Zhu, K. Padmaraju, L. Chen, M. Lipson, and K. Bergman, "10 Gb/s WDM optically-connected memory system using silicon microring modulator," in *European Conference and Exhibition on Optical Communication. Amsterdam* (Optical Society of America, 2012).
94. J. Li, X. Zheng, A. V. Krishnamoorthy, and J. F. Buckwalter, "Scaling trends for picjoule-per-bit WDM photonic interconnects in CMOS SOI and FinFET processes," *J. Lightwave Technol.* **34**, 2730–2742 (2016).
95. M. Bahadori, S. Rumley, D. Nikolova, and K. Bergman, "Comprehensive design space exploration of silicon photonic interconnects," *J. Lightwave Technol.* **34**, 2975–2987 (2016).
96. R. Wu, C.-H. Chen, J.-M. Fedeli, M. Fournier, K.-T. Cheng, and R. G. Beausoleil, "Compact models for carrier-injection silicon microring modulators," *Opt. Express* **23**, 15545–15554 (2015).
97. K. Padmaraju, X. Zhu, L. Chen, M. Lipson, and K. Bergman, "Intermodulation crosstalk characteristics of WDM silicon microring modulators," *IEEE Photon. Technol. Lett.* **26**, 1478–1481 (2014).
98. O. Dubray, A. Abraham, K. Hassan, S. Olivier, D. Marris-Morini, L. Vivien, I. O. Connor, and S. Menezes, "Electro-optical ring modulator: an ultracompact model for the comparison and optimization of p-n, p-i-n, and capacitive junction," *IEEE J. Sel. Top. Quantum Electron.* **22**, 89–98 (2016).
99. J. B. Quélène, J. F. Carpentier, Y. L. Guennec, and P. L. Maître, "Optimization of power coupling coefficient of a carrier depletion silicon ring modulator for WDM optical transmission," in *IEEE Optical Interconnects Conference (OI)* (2016).
100. G. Li, A. V. Krishnamoorthy, I. Shubin, J. Yao, Y. Luo, H. Thacker, X. Zheng, K. Raj, and J. E. Cunningham, "Ring resonator modulators in silicon for interchip photonic links," *IEEE J. Sel. Top. Quantum Electron.* **19**, 95–113 (2013).
101. M. A. Seyed, R. Wu, C.-H. Chen, M. Fiorentino, and R. Beausoleil, "15 Gb/s transmission with wide-FSR carrier injection ring modulator for Tb/s optical link," in *Conference on Lasers and Electro-Optics, San Jose, California* (2016).
102. T. Saeki, S. Sato, M. Kurokawa, A. Moto, M. Suzuki, K. Tanaka, K. Tanaka, N. Ikoma, and Y. Fujimura, "100 Gbit/s compact transmitter module integrated with optical multiplexer," in *IEEE Photonics Conference* (2013).
103. A. Stavdas, P. Bayvel, and J. E. Midwinter, *Design and Performance of Concave Holographic Gratings for Applications as Multiplexers/Demultiplexers for Wavelength Routed Optical Networks* (SPIE, 1996).
104. Y. Hida, Y. Hibino, T. Kitoh, Y. Inoue, M. Itoh, T. Shibata, A. Sugita, and A. Himeno, "400-channel 25-GHz spacing arrayed-waveguide grating covering a full range of C- and L-band," in *Optical Fiber Communication Conference and Exhibit, OFC, Technical Digest Postconference Edition* (IEEE Cat. 01CH37171) (2001).
105. A. Sugita, A. Kaneko, K. Okamoto, M. Itoh, A. Himeno, and Y. Ohmori, "Very low insertion loss arrayed-waveguide grating with vertically tapered waveguides," *IEEE Photon. Technol. Lett.* **12**, 1180–1182 (2000).
106. S. Kamei, M. Ishii, I. Kitagawa, M. Itoh, and Y. Hibino, "64-channel ultra-low crosstalk arrayed-waveguide grating multi/demultiplexer module using cascade connection technique," *Electron. Lett.* **39**, 81–82 (2003).
107. R. Nagarajan, C. H. Joyner, R. P. Schneider, J. S. Bostak, T. Butrie, A. G. Dentali, V. G. Dominic, P. W. Evans, M. Kato, M. Kauffman, D. J. H. Lambert, S. K. Mathis, A. Mathur, R. H. Miles, M. L. Mitchell, M. J. Missey, S. Murthy, A. C. Nilsson, F. H. Peters, S. C. Pennypacker, J. L. Pleumeekers, R. A. Salvatore, R. K. Schlenker, R. B. Taylor, T. Huan-Shang, M. F. V. Leeuwen, J. Webjorn, M. Ziari, D. Perkins, J. Singh, S. G. Grubb, M. S. Reffle, D. G. Mehuys, F. A. Kish, and D. F. Welch, "Large-scale photonic integrated circuits," *IEEE J. Sel. Top. Quantum Electron.* **11**, 50–65 (2005).
108. X. J. M. L. M. Nikoufard, Y. C. Zhu, J. J. M. Kwaspen, E. A. J. M. Bente, and M. K. Smit, "An 8 × 25 GHz Polarization-Independent Integrated Multi-Wavelength Receiver," in *Optical Amplifiers and Their Applications/Integrated Photonics Research, Technical Digest (CD)* (Optical Society of America, 2004), paper IThB2.
109. X. Wang, S. Xiao, W. Zheng, F. Wang, Y. Li, Y. Hao, X. Jiang, M. Wang, and J. Yang, "Athermal silicon arrayed waveguide grating with polymer-filled slot structure," *Opt. Commun.* **282**, 2841–2844 (2009).
110. X. Fu, J. Cheng, Q. Huang, Y. Hu, W. Xie, M. Tassaert, J. Verbiest, K. Ma, J. Zhang, K. Chen, C. Zhang, Y. Shi, J. Bauwelinck, G. Roelkens, L. Liu, and S. He, "5 × 20 Gb/s heterogeneously integrated III-V on silicon electro-absorption modulator array with arrayed waveguide grating multiplexer," *Opt. Express* **23**, 18686–18693 (2015).
111. M. Bahadori, S. Rumley, H. Jayatileka, K. Murray, N. A. F. Jaeger, L. Chrostowski, S. Shekhar, and K. Bergman, "Crosstalk penalty in microring-based silicon photonic interconnect systems," *J. Lightwave Technol.* **34**, 4043–4052 (2016).
112. L. Chen, N. Sherwood-Droz, and M. Lipson, "Compact bandwidth-tunable microring resonators," *Opt. Lett.* **32**, 3361–3363 (2007).
113. C. L. Manganelli, P. Pintus, F. Gambini, D. Fowler, M. Fournier, S. Faralli, C. Kopp, and C. J. Oton, "Large-FSR thermally tunable double-ring filters for WDM applications in silicon photonics," *IEEE Photon. J.* **9**, 1–10 (2017).
114. 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. Lightwave Technol.* **36**, 773–788 (2018).
115. C. Sun, M. Wade, M. Georgas, S. Lin, L. Alloatti, B. Moss, R. Kumar, A. H. Atabaki, F. Pavanello, J. M. Shainline, J. S. Orcutt, R. J. Ram, M. Popović, and V. Stojanović, "A 45 nm CMOS-SOI monolithic photonics platform with bit-statistics-based resonant microring thermal tuning," *IEEE J. Solid-State Circuits* **51**, 893–907 (2016).
116. P. L. Maître, J. F. Carpentier, C. Baudot, N. Vulliet, A. Souhaité, J. B. Quélène, T. Ferrotti, and F. Boëuf, "Impact of process variability of active ring resonators in a 300 mm silicon photonic platform," in *European Conference on Optical Communication (ECOC)* (2015).
117. M. Nikdast, G. Nicolescu, J. Trajkovic, and O. Liboiron-Ladouceur, "Chip-scale silicon photonic interconnects: a formal study on fabrication non-uniformity," *J. Lightwave Technol.* **34**, 3682–3695 (2016).
118. K. Padmaraju, D. F. Logan, T. Shiraishi, J. J. Ackert, A. P. Knights, and K. Bergman, "Wavelength locking and thermally stabilizing microring resonators using dithering signals," *J. Lightwave Technol.* **32**, 505–512 (2014).
119. F. Morichetti, A. Canciamilla, C. Ferrari, M. Torregiani, A. Melloni, and M. Martinelli, "Roughness induced backscattering in optical silicon waveguides," *Phys. Rev. Lett.* **104**, 033902 (2010).
120. B. E. Little, J.-P. Laine, and S. T. Chu, "Surface-roughness-induced contradirectional coupling in ring and disk resonators," *Opt. Lett.* **22**, 4–6 (1997).
121. M. Bahadori, S. Rumley, Q. Cheng, and K. Bergman, "Impact of backscattering on microring-based silicon photonic links," in *IEEE Optical Interconnects Conference (OI)*, Santa Fe, New Mexico, USA (2018).
122. H. Ito, S. Kodama, Y. Muramoto, T. Furuta, T. Nagatsuma, and T. Ishibashi, "High-speed and high-output InP-InGaAs untraveling-carrier photodiodes," *IEEE J. Sel. Top. Quantum Electron.* **10**, 709–727 (2004).
123. F. Nakajima, M. Nada, and T. Yoshimatsu, "High-speed avalanche photodiode and high-sensitivity receiver optical subassembly for 100 Gb/s ethernet," *J. Lightwave Technol.* **34**, 243–248 (2016).

124. Z. Huang, C. Li, D. Liang, K. Yu, C. Santori, M. Fiorentino, W. Sorin, S. Palermo, and R. G. Beausoleil, "25 Gbps low-voltage waveguide SiGe avalanche photodiode," *Optica* **3**, 793–798 (2016).
125. T. Ishibashi, T. Furuta, H. Fushimi, S. Kodama, H. Ito, T. Nagatsuma, N. Shimizu, and Y. Miyamoto, "InP/InGaAs uni-traveling-carrier photodiodes," *IEICE Trans. Electron.* **83**, 938–949 (2000).
126. L. Vivien, A. Polzer, D. Marris-Morini, J. Osmond, J. M. Hartmann, P. Crozat, E. Cassan, C. Kopp, H. Zimmermann, and J. M. Fédéli, "Zero-bias 40 Gbit/s germanium waveguide photodetector on silicon," *Opt. Express* **20**, 1096–1101 (2012).
127. L. Virot, L. Vivien, J.-M. Fédéli, Y. Bogumilowicz, J.-M. Hartmann, F. Bœuf, P. Crozat, D. Marris-Morini, and E. Cassan, "High-performance waveguide-integrated germanium PIN photodiodes for optical communication applications [Invited]," *Photon. Res.* **1**, 140–147 (2013).
128. D. Feng, S. Liao, P. Dong, N.-N. Feng, H. Liang, D. Zheng, C.-C. Kung, J. Fong, R. Shafiqi, J. Cunningham, A. V. Krishnamoorthy, and M. Asghari, "High-speed Ge photodetector monolithically integrated with large cross-section silicon-on-insulator waveguide," *Appl. Phys. Lett.* **95**, 261105 (2009).
129. A. Beling, H. G. Bach, G. G. Mekonnen, R. Kunkel, and D. Schmidt, "Miniaturized waveguide-integrated p-i-n photodetector with 120-GHz bandwidth and high responsivity," *IEEE Photon. Technol. Lett.* **17**, 2152–2154 (2005).
130. J. Hulme, M. J. Kennedy, R.-L. Chao, L. Liang, T. Komljenovic, J.-W. Shi, B. Szafraniec, D. Baney, and J. E. Bowers, "Fully integrated microwave frequency synthesizer on heterogeneous silicon-III/V," *Opt. Express* **25**, 2422–2431 (2017).
131. C. T. Derose, D. C. Trotter, W. A. Zortman, A. L. Starbuck, M. Fisher, M. R. Watts, and P. S. Davids, "Ultra compact 45 GHz CMOS compatible Germanium waveguide photodiode with low dark current," *Opt. Express* **19**, 24897–24904 (2011).
132. Q. Li, K. Sun, K. Li, Q. Yu, P. Runge, W. Ebert, A. Beling, and J. C. Campbell, "High-power evanescently coupled waveguide MUTC photodiode with >105-GHz bandwidth," *J. Lightwave Technol.* **35**, 4752–4757 (2017).
133. L. Vivien and L. Pavesi, *Handbook of Silicon Photonics* (Taylor & Francis, 2016).
134. M. Piels and J. E. Bowers, "40 GHz Si/Ge uni-traveling carrier waveguide photodiode," *J. Lightwave Technol.* **32**, 3502–3508 (2014).
135. M. Huang, P. Cai, S. Li, L. Wang, T. Su, L. Zhao, W. Chen, C. Hong, and D. Pan, "Breakthrough of 25 Gb/s germanium on silicon avalanche photodiode," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2016).
136. M. Yano, F. Yamagishi, and T. Tsuda, "Optical MEMS for photonic switching-compact and stable optical crossconnect switches for simple, fast, and flexible wavelength applications in recent photonic networks," *IEEE J. Sel. Top. Quantum Electron.* **11**, 383–394 (2005).
137. J. Kim, C. J. Nuzman, B. Kumar, D. F. Liewen, J. S. Kraus, A. Weiss, C. P. Lichtenwalner, A. R. Papazian, R. E. Frahm, N. R. Basavanahally, D. A. Ramsey, V. A. Aksyuk, F. Pardo, M. E. Simon, V. Lifton, H. B. Chan, M. Haueis, A. Gasparyan, H. R. Shea, S. Arney, C. A. Bolle, P. R. Kolodner, R. Ryf, D. T. Neilson, and J. V. Gates, "1100 × 1100 port MEMS-based optical crossconnect with 4-dB maximum loss," *IEEE Photon. Technol. Lett.* **15**, 1537–1539 (2003).
138. H. S. Hinton, *An Introduction to Photonic Switching Fabrics* (Springer Science & Business Media, 2013).
139. K. Tanizawa, K. Suzuki, S. Suda, H. Matsuura, K. Ikeda, S. Namiki, and H. Kawashima, "Silicon photonic 32 × 32 strictly-non-blocking blade switch and its full path characterization," in *21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)* (2016).
140. H. Wang, A. Wonfor, K. A. Williams, R. V. Penty, and I. H. White, "Demonstration of a lossless monolithic 16 × 16 QW SOA switch," in *35th European Conference on Optical Communication* (2009).
141. M. Yang, W. M. J. Green, S. Assefa, J. Van Campenhout, B. G. Lee, C. V. Jahnes, F. E. Doany, C. L. Schow, J. A. Kash, and Y. A. Vlasov, "Non-blocking 4 × 4 electro-optic silicon switch for on-chip photonic networks," *Opt. Express* **19**, 47–54 (2011).
142. N. Sherwood-Droz, H. Wang, L. Chen, B. G. Lee, A. Biberman, K. Bergman, and M. Lipson, "Optical 4 × 4 hitless silicon router for optical Networks-on-Chip (NoC)," *Opt. Express* **16**, 15915–15922 (2008).
143. A. Wonfor, H. Wang, R. V. Penty, and I. H. White, "Large port count high-speed optical switch fabric for use within datacenters [Invited]," *J. Opt. Commun. Netw.* **3**, A32–A39 (2011).
144. R. Stabile, A. Albores-Mejia, and K. A. Williams, "Monolithic active-passive 16 × 16 optoelectronic switch," *Opt. Lett.* **37**, 4666–4668 (2012).
145. A. Rohit, J. Bolk, X. J. M. Leijtens, and K. A. Williams, "Monolithic nano-second-reconfigurable 4 × 4 space and wavelength selective cross-connect," *J. Lightwave Technol.* **30**, 2913–2921 (2012).
146. L. Chen and Y.-K. Chen, "Compact, low-loss and low-power 8 × 8 broadband silicon optical switch," *Opt. Express* **20**, 18977–18985 (2012).
147. B. G. Lee, A. V. Ryljakov, W. M. J. Green, S. Assefa, C. W. Baks, R. Rimolo-Donadio, D. M. Kuchta, M. H. Khater, T. Barwicz, C. Reinholm, E. Kiewra, S. M. Shank, C. L. Schow, and Y. A. Vlasov, "Four- and eight-port photonic switches monolithically integrated with digital CMOS logic and driver circuit," in *Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)* (2013).
148. R. Stabile, A. Rohit, and K. A. Williams, "Monolithically integrated 8 × 8 space and wavelength selective cross-connect," *J. Lightwave Technol.* **32**, 201–207 (2014).
149. 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).
150. K. Suzuki, K. Tanizawa, T. Matsukawa, G. Cong, S.-H. Kim, S. Suda, M. Ohno, T. Chiba, H. Tadokoro, M. Yanagihara, Y. Igarashi, M. Masahara, S. Namiki, and H. Kawashima, "Ultra-compact 8 × 8 strictly-non-blocking Si-wire PILOSS switch," *Opt. Express* **22**, 3887–3894 (2014).
151. Q. Cheng, A. Wonfor, J. L. Wei, R. V. Penty, and I. H. White, "Low-energy, high-performance lossless 8 × 8 SOA switch," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2015).
152. 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 H. Kawashima, "Ultra-compact 32 × 32 strictly-non-blocking Si-wire optical switch with fan-out LGA interposer," *Opt. Express* **23**, 17599–17606 (2015).
153. T. J. Seok, N. Quack, S. Han, and M. C. Wu, "50 × 50 digital silicon photonic switches with MEMS-actuated adiabatic coupler," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2015).
154. T. J. Seok, N. Quack, S. Han, W. Zhang, R. S. Muller, and M. C. Wu, "64 × 64 Low-loss and broadband digital silicon photonic MEMS switches," in *European Conference on Optical Communication (ECOC)* (2015).
155. L. Qiao, W. Tang, and T. Chu, "16 × 16 Non-blocking silicon electro-optic switch based on Mach-Zehnder interferometer," in *Optical Fiber Communications Conference and Exhibition (OFC)* (2016).
156. 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).
157. D. Celso, D. J. Goodwill, J. Jia, P. Dumais, Z. Chunshu, Z. Fei, T. Xin, Z. Chunhui, Y. Shengyong, H. Jifang, L. Ming, L. Wanyuan, W. Yuming, G. Dongyu, H. Mehrvar, and E. Bernier, "32 × 32 silicon photonic switch," in *21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)* (2016).
158. L. Qiao, W. Tang, and T. Chu, "32 × 32 silicon electro-optic switch with built-in monitors and balanced-status units," *Sci. Rep.* **7**, 42306 (2017).
159. 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-D PLC connector," in *Optical Fiber Communications Conference and Exposition (OFC)* (2018).
160. T. Chu, L. Qiao, W. Tang, D. Guo, and W. Wu, "Fast, High-radix silicon photonic switches," in *Optical Fiber Communications Conference and Exposition (OFC)* (2018).
161. K. Kwon, T. J. Seok, J. Henriksson, J. Luo, L. Ochikubo, J. Jacobs, R. S. Muller, and M. C. Wu, "128 × 128 silicon photonic MEMS switch with scalable row/column addressing," in *Conference on Lasers and Electro-Optics*, San Jose, California (Optical Society of America, 2018).
162. K. Wang, A. Wonfor, R. V. Penty, and I. H. White, "Active-passive 4 × 4 SOA-based switch with integrated power monitoring," in *Optical Fiber Communication Conference (OFC)* (2012), paper OTh4F.4.

163. I. White, E. T. Aw, K. Williams, H. Wang, A. Wonfor, and R. Penty, "Scalable optical switches for computing applications [Invited]," *J. Opt. Netw.* **8**, 215–224 (2009).
164. Q. Cheng, M. Ding, A. Wonfor, J. Wei, R. V. Penty, and I. H. White, "The feasibility of building a  $64 \times 64$  port count SOA-based optical switch," in *International Conference on Photonics in Switching (PS)*, Florence (2015), pp. 199–201.
165. S. C. Nicholes, M. L. Mašanović, B. Jevremović, E. Lively, L. A. Coldren, and D. J. Blumenthal, "An  $8 \times 8$  InP Monolithic Tunable Optical Router (MOTOR) packet forwarding chip," *J. Lightwave Technol.* **28**, 641–650 (2010).
166. N. Dupuis, A. V. Ryljakov, C. L. Schow, D. M. Kuchta, C. W. Baks, J. S. Orcutt, D. M. Gill, W. M. J. Green, and B. G. Lee, "Ultralow crosstalk nanosecond-scale nested  $2 \times 2$  Mach-Zehnder silicon photonic switch," *Opt. Lett.* **41**, 3002–3005 (2016).
167. 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. Rep.* **7**, 12244 (2017).
168. Q. Cheng, A. Wonfor, R. V. Penty, and I. H. White, "Scalable, low-energy hybrid photonic space switch," *J. Lightwave Technol.* **31**, 3077–3084 (2013).
169. Q. Cheng, A. Wonfor, J. L. Wei, R. V. Penty, and I. H. White, "Demonstration of the feasibility of large-port-count optical switching using a hybrid Mach-Zehnder interferometer-semiconductor optical amplifier switch module in a recirculating loop," *Opt. Lett.* **39**, 5244–5247 (2014).
170. 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  $8 \times 32$ -Gbaud 16-QAM WDM signal," in *Optical Fiber Communications Conference and Exposition (OFC)* (2018).
171. Q. Cheng, M. Bahadori, and K. Bergman, "Advanced path mapping for silicon photonic switch fabrics," in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (online) (Optical Society of America, 2017), paper SW1O.5.
172. Q. Cheng, M. Bahadori, Y. Huang, S. Rumley, and K. Bergman, "Smart routing tables for integrated photonic switch fabrics," in *European Conference on Optical Communication (ECOC)*, Gothenburg (2017).
173. Y. Huang, Q. Cheng, N. C. Abrams, J. Zhou, S. Rumley, and K. Bergman, "Automated calibration and characterization for scalable integrated optical switch fabrics without built-in power monitors," in *European Conference on Optical Communication (ECOC)*, Gothenburg (2017).
174. Y. Huang, Q. Cheng, and K. Bergman, "Crosstalk-aware calibration for fast and automated functionalization of photonic integrated switch fabrics," in *Conference on Lasers and Electro-Optics*, OSA Technical Digest (online) (Optical Society of America, 2018), paper STh3B.6.
175. Y. Huang, Q. Cheng, and K. Bergman, "Automated calibration of balanced control to optimize performance of silicon photonic switch fabrics," in *Optical Fiber Communications Conference and Exposition (OFC)*, OSA Technical Digest (online) (Optical Society of America, 2018), paper Th1G.2.
176. G. Wang, D. G. Andersen, M. Kaminsky, K. Papagiannaki, T. S. E. Ng, M. Kozuch, and M. Ryan, "c-Through: part-time optics in data centers," *SIGCOMM Comput. Commun. Rev.* **40**, 327–338 (2010).
177. M. Glick, D. G. Andersen, M. Kaminsky, and L. Mummert, "Dynamically reconfigurable optical links for high-bandwidth data center network," in *Optical Fiber Communication Conference and National Fiber Optic Engineers Conference*, San Diego, California (Optical Society of America, 2009).
178. 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* (ACM, 2013), Vol. **43**.
179. A. Ryljakov, J. E. Proesel, S. Rylow, B. G. Lee, J. F. Bulzacchelli, A. Ardey, B. Parker, M. Beakes, C. W. Baks, C. L. Schow, and M. Meghelli, "A 25 Gb/s burst-mode receiver for low latency photonic switch networks," *IEEE J. Solid-State Circuits* **50**, 3120–3132 (2015).
180. Y. Shen, M. H. N. Hattink, P. Samadi, Q. Cheng, Z. Hu, A. Gazman, and K. Bergman, "Software-defined networking control plane for seamless integration of multiple silicon photonic switches in Datacom networks," *Opt. Express* **26**, 10914–10929 (2018).
181. Q. Cheng, L. Y. Dai, M. Bahadori, P. Morrissey, R. Polster, S. Rumley, P. O'Brien, and K. Bergman, "Microring-based Si/SiN dual-layer switch fabric," in *Optical Interconnects*, Santa Fe, New Mexico, USA (IEEE, 2018), pp. 29–30.
182. Q. Cheng, L. Y. Dai, M. Bahadori, N. C. Abrams, P. E. Morrissey, M. Glick, P. O'Brien, and K. Bergman, "Si/SiN microring-based optical router in switch-and-select topology," in *European Conference on Optical Communication (ECOC)* (2018), p. We1C.3.
183. Q. Cheng, M. Bahadori, S. Rumley, and K. Bergman, "Highly-scalable, low-crosstalk architecture for ring-based optical space switch fabrics," in *IEEE Optical Interconnects Conference (OI)*, Santa Fe, New Mexico, USA (2017), pp. 41–42.
184. M. Ding, A. Wonfor, Q. Cheng, R. V. Penty, and I. H. White, "Hybrid MZI-SOA InGaAs/InP photonic integrated switches," *IEEE J. Sel. Top. Quantum Electron.* **24**, 1–8 (2018).
185. Q. Cheng, A. Wonfor, J. L. Wei, R. V. Penty, and I. H. White, "Monolithic MZI-SOA hybrid switch for low-power and low-penalty operation," *Opt. Lett.* **39**, 1449–1452 (2014).