External vs. Integrated Light Sources for Intra-Data Center Co-Packaged Optical Interfaces

Brandon Buscaino, *Member, IEEE*, Elizabeth Chen, *Student Member, IEEE*, James W. Stewart, *Member, IEEE*, Thang Pham, and Joseph M. Kahn, *Fellow, IEEE*

***Abstract*—Co-packaging of optics and electronics for data center switches has been proposed to reduce system-level power consumption by minimizing power-hungry electrical interconnects. Co-packaging optical components near high-temperature electronics, however, can diminish their performance and reliability. Moreover, limitations of the switch environment, such as restricted footprint, power, and number of fiber attachments can limit practical co-packaging implementations. Here, we study four architectures for co-packaged optical interfaces using either single- or multi-wavelength light sources that can be either external to or integrated with the optical interfaces. We model the temperature- and current-dependent performance and reliability of the sources and calculate the link budget for switch bandwidths up to 102.4 Tb/s. We compare architectures based on coherent and direct detection and find that all coherent detection architectures support 102.4 Tb/s switching with over 13 dB link budget, while most direct detection architectures scale to 51.2 Tb/s or 102.4 Tb/s switching with link budgets less than 5 dB. In addition, we demonstrate that external-source architectures require wavelength-division multiplexing at lower switch bandwidths and consequently have link budgets 1-5 dB lower than integrated source architectures. Further, we demonstrate that higher-power lasers can scale external architectures to 102.4 Tb/s switching with over 4 dB link budgets. Finally, for 51.2 Tb/s switching, we show that a reduction in integrated source temperature improves link budgets by 2-4 dB and direct detection external-source architectures can achieve greater than 5 dB link budget with fewer than 300 fiber attachments.**

***Index Terms*—Co-packaged optics, data center switching, optical fiber communication.**

I. Introduction

OVER the past five years, data center Internet Protocol traffic has grown at a compound annual growth rate of 27% [1]. Less than 15% of data center traffic directly reaches end users, however, with nearly three-fourths of data center traffic residing inside data centers [2]. Scaling the throughput of high-capacity switches, which form the backbone of typical hyperscale data centers, is crucial to meet data center traffic demands. In particular, switching bandwidths of up to 51.2 Tb/s will be required within the next 4 years [3].

In the past, switch bandwidth scaling has relied on increased electrical signal speeds and modulation format orders [4]. Further scaling of the speed and modulation format requires additional power, however, nearing the power-handling limits of data center network racks [5], [6]. These prohibitive power demands result from long electrical traces, which are often 20 cm or longer, that are required to connect pluggable front-panel transceivers to the high-speed switching application-specific integrated circuits (ASICs). Intra-data center switches that rely on pluggable transceivers also may soon suffer from a front-panel bottleneck [7].

In order to reduce power consumption and address the front-panel bottleneck, co-packaging of optics and electronics on one or more modules near the switching ASIC has been proposed [8]. In the past several years, significant improvements in silicon photonics have enabled compact and efficient integrated transceivers capable of replacing pluggable front-panel transceivers [9], [10]. These integrated transceivers could be placed near the switching ASIC to minimize electrical interconnects and remove power-hungry retimers.

Placing highly integrated optical components near electronics can introduce significant challenges, however. For example, the high-temperature environment near switching ASICs degrades the reliability and performance of sensitive optical components [11]. Moreover, integrated optical devices typically have higher insertion losses than their bulk counterparts. Finally, a move towards co-packaged optics for data center switches has posed significant logistical challenges for commercial deployment [2].

Of particular interest for intra-data center links based on co-packaged optical interfaces is the light source used at the transmitter. It is still unclear whether the light source should be integrated on the same silicon photonics platform, due to considerations of temperature sensitivity, reliability, and footprint. Additionally, recent advances in multi-wavelength source (MWS) technology has opened up new possibilities for light sources in intra-data center optical interfaces [12], [13]. A single MWS may serve as the light source for several wavelength-division-multiplexed (WDM) channels.

In addition to the light source technology, the ideal modulation format for intra-data center links is still uncertain. While

Manuscript received September 3, 2020; revised November 5, 2020 and December 5, 2020; accepted December 7, 2020. Date of publication December 9, 2020; date of current version April 2, 2021. Brandon Buscaino, Elizabeth Chen, and Joseph M. Kahn acknowledge support from Facebook Inc., Maxim Integrated, and National Science Foundation under Grant ECCS-1740291 E2CDA. *(Corresponding author: Brandon Buscaino.)*

Brandon Buscaino, Elizabeth Chen, and Joseph M. Kahn are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: buscaino@alumni.stanford.edu; echen105@stanford.edu; jmk@ee.stanford.edu).

James W. Stewart and Thang Pham are with the Department of Hardware Engineering, Facebook Inc, Menlo Park, CA 94025 USA (e-mail: jwstewart@fb.com; thangtpham@fb.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JLT.2020.3043653.

Digital Object Identifier 10.1109/JLT.2020.3043653

TABLE I
CO-PACKAGED TRANSCEIVER REQUIREMENTS

Requirement	Value
Maximum Power Consumption	1 kW
Maximum Footprint	6900 mm^2
Minimum PIC Shoreline Density	200 Gb/s/mm
Maximum Fiber Attachments	512
Maximum PIC Temperature	70°C
Maximum PIC Temperature Shift	±15°C
Maximum Annual Failure Rate	1%
Maximum Eye-Safe Optical Power	100 mW

data center networks are dominated by direct-detection formats such as non-return-to-zero on-off-keying (NRZ-OOK) or M-ary pulse amplitude modulation (M-PAM), coherent detection-compatible formats, such as quaternary phase shift keying (QPSK) or M-ary quadrature amplitude modulation (M-QAM), provide superior receiver sensitivity, which could offset high insertion losses or performance penalties associated with silicon photonics [14].

In this paper, we propose four architectures for intra-data center co-packaged optical interfaces that rely on both external and integrated light sources as well as single- and multi-wavelength light sources. We compare the performance, reliability, and scalability of these architectures using modulation formats for either direct or coherent detection.

The remainder of the paper is organized as follows. In Section II we discuss requirements for intra-data center switching and their relevance to link architecture and performance. In Section III we propose four architectures for intra-data center co-packaged optical interfaces based on either external or integrated light sources, as well as single- or multi-wavelength sources. In Section IV we present modeling of the link components, including the light sources as well as integrated transmitter and receiver interfaces. In Section V we utilize the models of the previous section to analyze link performance and reliability. In Section VI we present results for variable- and fixed-bandwidth switch scaling. We present conclusions in Section VII.

II. REQUIREMENTS FOR INTRA-DATA CENTER SWITCHING WITH CO-PACKAGED OPTICS

In this section, we describe typical intra-data center switching requirements and performances, which are summarized in Table I. For the remainder of the paper, we assume that links operate in the O band, centered at 1310 nm, to avoid significant dispersion, and thus power-hungry electronic dispersion-compensation. Additionally, we assume that the symbol rate is limited to 53.125 Gbaud, which is typical for current intra-data center links, to avoid increased power consumption penalties as well as to highlight the scaling of the proposed architectures. We will discuss how to interpret these results for systems utilizing higher symbol rates or higher power sources in Section VI-A. Finally, for multi-wavelength architectures, we assume a minimum channel spacing of 75 GHz, though the channel spacing is typically much larger in optimized architectures, as discussed in Section VI-A1.

A. Power Consumption

As discussed in the Introduction, total switching power consumption can be reduced by limiting the length of electrical interconnects. Some estimates indicate that minimizing electrical interconnect length can reduce power consumption by up to 30% [8]. While data center network racks are typically power-limited to 20 kW [15], each network switch typically consumes less power. Here, we assume that the maximum power consumption of the transceiver elements, which excludes the power consumption of the switching ASIC, is limited to 1 kW. In Section VI, we will relax this constraint to determine the most efficient light sources.

B. Footprint and Shoreline

The total footprint available for the passive optical components and the transceiver electronics is limited by the maximum manufacturable substrate size. Organic substrates, which photonic integrated circuits (PICs) are likely to be mounted on, typically have maximum footprints of around $90 \times 90\ mm^2$ [16], lower than that of ceramic substrates. While PIC sizes are significantly smaller, limited by maximum reticle sizes of around $30 \times 30\ mm^2$, we assume that the smaller PICs are placed on the substrate such that the total maximum footprint is not dependent on the PIC size. In addition, we assume that the footprint of the switching ASIC is $30 \times 40\ mm^2$ [17], resulting in an available footprint for transceiver components of 6900 mm^2.

The initial generation of switches using co-packaged optical interfaces is not expected to fully integrate all transceiver components on a single substrate due to issues with complexity, yield, reliability, and manufacturing. For example, recent experiments have demonstrated the feasibility of co-packaging optics and electronics by using eight 1.6 Tb/s PICs around a 12.8 Tb/s switching ASIC [9]. While this represents one possible implementation, we will not discuss PIC modularization further in this paper but rather assume that the footprint of the organic substrate is used efficiently.

Another potential limitation of co-locating optics near switching electronics is the high required shoreline density, defined as the transceiver bandwidth divided by the occupied switch edge length. In order to minimize electrical interconnect lengths, transceiver electronics need to be located close to the switch edge. For the 51.2 Tb/s switching generation, the expected required bandwidth density is approximately 200 Gb/s/mm [14]. By inserting electrical fan-out, the required bandwidth density can be reduced, at the cost of increased electrical interconnect length.

C. Fiber Attachment Complexity

Increasing the shoreline density of transceiver elements places additional constraints on the packaging complexity, such as fabrication and placement of fiber-to-chip couplers. For co-packaged optical interfaces, either vertical couplers, such as grating couplers, or edge couplers, such as spot-size converters with V-grooves may be used [18], [19]. Edge couplers may only be placed along the edge of the PIC, whereas vertical

couplers can be placed in most areas on the surface. The fiber density is likely not limited by the absolute fiber-to-chip density limit, however, but rather by practical packaging and routing considerations.

Here, we assume that 128 fibers may be attached to each side of the PIC, which corresponds to a fiber density lower than typical standardized multi-fiber ferrules [20]. While this number is consistent with edge-coupler fiber density, it is not an absolute limit. In Section VI-B we will discuss system performance when the total number of fiber attachments is varied.

D. Eye Safety

Current data center optical links operate well below standard eye safety limits in the O band in order to reduce laser power consumption [21]. In order to reduce the number of fiber-to-chip attachments, however, it may be feasible to use high-power (200–500 mW) O band lasers shared across several different optical links [22]. While high-power O band lasers have successfully been demonstrated [23], the commercial viability of such lasers is still uncertain. Additionally, recent international standards on eye safety have increased the maximum eye-safe power to 500 mW in the O band based on estimates of retinal damage [24]. Whether this limit is appropriate to avoid corneal damage in the long-wavelength portion of the O band, however, is not clear [25]. Thus, in this study, we conservatively assume a maximum eye-safe fiber power of 100 mW.

Importantly, it may be possible to utilize light sources with output powers higher than 100 mW by enclosing the exposed fibers inside the switch unit so that data center technicians cannot be exposed to the high powers. Due to high losses of integrated links, the optical power of the transmit and receive fibers is likely to be far lower than the eye safety limit. Placement of high-power light sources inside switching units, however, may compromise serviceability. In Section VI-A we discuss further how to interpret our results when using high-power light sources.

E. Temperature Sensitivity

The environment near the switching ASIC can be at temperatures over 85 °C and can cause active optics such as lasers and amplifiers to become less reliable and perform poorly. Moreover, depending on the utilization of the switch, the temperature near the switching ASIC may vary by up to $\pm 15^{\circ}$C. Variations in temperature may not affect only active optical components, such as optical amplifiers that have temperature-dependent gain peaks, but also passive optics, such as (de)multiplexers. While techniques to mitigate temperature sensitivity, such as locating the active optics at the edges of the PIC, can be employed, we assume that integrated optical components may see a maximum temperature of 70 °C, with temperature swings up to $\pm 15^{\circ}$C [14]. Alternatively, locating active optical devices in a separate package and routing them to the transceivers can significantly reduce temperature dependence. In Section VI-B, we further discuss the effects of advanced cooling techniques to reduce temperature sensitivity of integrated optical components.

F. Optical Reliability

Data center networks are currently designed to avoid significant network impact from optical failures, which we loosely define as the failure of an active optical component necessary for transmit or receive functions. Because active optical elements are located in pluggable transceivers, they can be replaced easily. Nevertheless, due to software and hardware errors, intra-data center switches still fail. Typically, hyperscale data centers report switch annual failure rates of several percent [26], [27].

Integrating optical components near switching electronics inevitably couples optical component and network reliability. The transition away from pluggable transceivers will thus require a re-thinking of failure in data center networks. Due to the cost of servicing or replacing highly integrated devices, which may fail at a higher rate due to high temperatures, switches may need to be designed to flexibly handle failed optical links. Another option is to under-provision network switches and internally or externally switch between broken and backup ports. In this study, we will conservatively assume an annual tolerable failure rate of 1% due to optical component failures, less than typical software and hardware reliability requirements.

III. Architectures for Intra-Data Center Co-Packaged Optical Interfaces

As discussed in Section II, the use of external light sources for intra-data center co-packaged optical interfaces may improve switch performance and reliability as well as reduce the PIC footprint by removing active optics from the high-temperature environment near the switching ASIC. The use of external light sources, however, may increase insertion losses and complexity by increasing the number of fiber attachments.

Additionally, studies have demonstrated the use of multi-wavelength sources for high-capacity data center optical links, leveraging recent breakthroughs in fabrication technologies [12], [28]. In particular, an MWS may be used as the source for WDM communications links to address challenges with traditional laser systems such as laser wavelength stability [29]. Moreover, for architectures where the light source can be shared across several independent optical links, the use of an MWS can eliminate the need for on-chip waveguide crossings.

In the following subsections, we compare four distinct architectures for intra-data center co-packaged optical interfaces based on the use of external vs. integrated light sources as well as multi-wavelength sources vs. single-wavelength sources (SWSs). These architectures are shown in Fig. 1.

A. Single-Wavelength Source Architectures

Fig. 1(a) illustrates two proposed architectures based on single-wavelength sources. Lasers emit at a single wavelength in the O band and may be either externally located and coupled onto the PIC via a fiber-to-chip coupler (FC) or they may be integrated on the PIC, requiring no FCs. The light is then split, if necessary, and routed to the transmitter (TX) and receiver (RX) blocks near the switching ASIC edge. Notably, light is only routed to the receiver elements if coherent detection is being

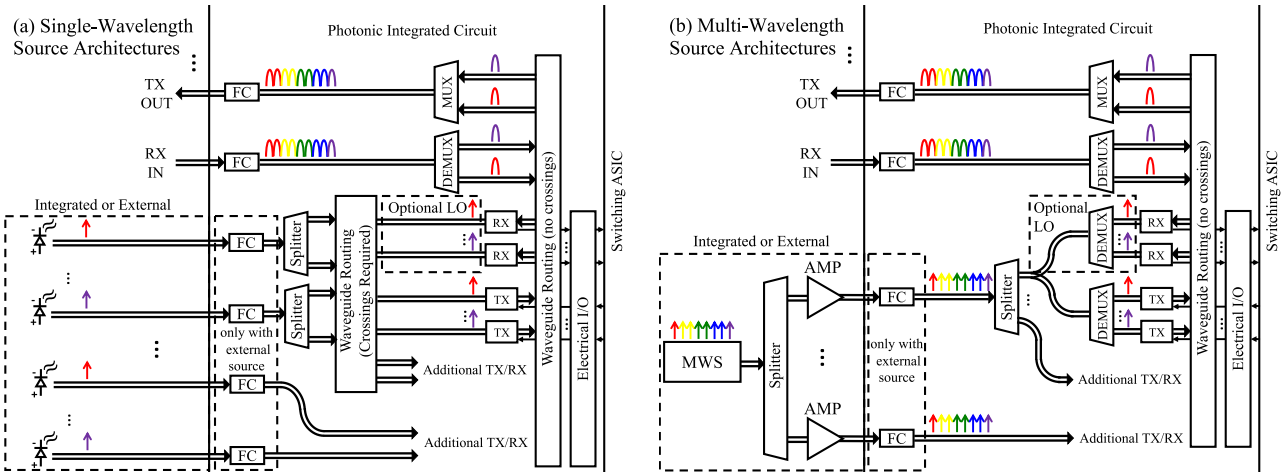

Fig. 1. PIC Architectures for intra-data center co-packaged optical interfaces using (a) single-wavelength sources or (b) multi-wavelength sources. The light sources may be integrated or external, as denoted by a dashed box in the lower-left corner. The light source output is then routed to the transmitter and receiver blocks for use in the optical link. See text for details. Acronyms: fiber-to-chip coupler (FC), local oscillator (LO), transmitter (TX), receiver (RX), multiplexer (MUX).

used, requiring a local oscillator (LO). Additionally, if the optical links use WDM, as shown in Fig. 1(a), the waveguide routing after splitting will require waveguide crossings in order to correctly order the wavelengths. The number of required waveguide crossings depends on the number of sources, wavelengths, and transmitter/receiver pairs.

Once the light source output reaches the TX, it is modulated and routed to a multiplexer (MUX), after which it is output-coupled via a FC to another switch. Fig. 1(a) also demonstrates the receiving functionality of the switch, where an optical signal is input-coupled to the PIC via a FC, demultiplexed (DEMUX), and routed to corresponding receivers. Electrical input/output (I/O) connects the transmitter and receiver blocks to the switching ASIC.

B. Multi-Wavelength Source Architectures

Fig. 1(b) illustrates two proposed PIC architectures based on multi-wavelength sources. Similar to Fig. 1(a), the light source may be externally located or integrated on the PIC. In the former arrangement, the MWS is split and amplified externally before being coupled onto the chip. In the latter, the integrated MWS is split on the PIC and sent to integrated amplifiers (AMPs), which are required to generate sufficient optical power for the optical links. In both cases, the multi-wavelength light is then optionally split again, demultiplexed, and routed to the transmitter and receiver blocks identically to Fig. 1(a).

Notably, the MWS architectures require additional demultiplexers before the transmitter blocks, which will increase the insertion loss of the link. However, MWS architectures do not require any waveguide crossings, as do the SWS architectures in Fig. 1(a). In some circumstances, such as with fewer high-power sources and many wavelengths, the number of required waveguide crossings in SWS architectures can significantly increase footprint, loss, and complexity. MWS architectures, on the other hand, rely on external and integrated amplifiers instead of lasers.

Finally, while the terminology "single-wavelength source" and "multi-wavelength source" is used throughout the rest of the paper, neither of these terms is directly linked to the number of wavelengths in the optical link. For example, the SWS architectures in Fig. 1(a) may use arrays of lasers having the same wavelength or multiple wavelengths. Similarly, in the MWS architectures, the MWS may be, redundantly, a single-wavelength laser or, as shown in Fig. 1(b), a source emitting at several different wavelengths. In Section VI, the scalability of WDM optical links will be discussed.

IV. Component Modeling

A. Active Optics

Due to the variety of laser and amplifier platforms and technologies, determining the optimal optical technology for co-packaged optics, external or integrated, is outside the scope of this paper. Rather, we develop the following generalized models to understand the impact of variations in current and temperature. Variations in the injection current can affect switch power consumption as well as optical reliability. Importantly, the laser current determines the maximum output power and the total number of source fibers. Variations in temperature influence optical reliability, efficiency, and performance.

1) Laser Modeling: We model temperature-dependent effects on laser performance via a self-heating model [30]. Both the threshold current $I_{th} \sim e^{T/T_0}$ and the slope efficiency $\eta \sim e^{-T/T_\eta}$ depend exponentially on temperature, where T_0 is the overall characteristic temperature and T_η is the characteristic temperature for an above-threshold current increment. The characteristic temperatures are determined by several different effects such as Auger recombination, carrier leakage, and temperature dependence of the laser gain and loss coefficients. Following this model, the output laser power P_o can be determined from the following implicit equation [30]:

$$P_o = \eta(T) e^{-(P_d(I)-P_o)Z_T/T_\eta} \left[I - I_{th}(T) e^{(P_d(I)-P_o)Z_T/T_0} \right], \quad (1)$$

TABLE II
LASER PARAMETERS

Parameter	Value	Reference
Slope Efficiency, η	0.3 W/A	[9], [31]
Threshold Current, I_{th}	15 mA	[31], [32]
Series Resistance, R_s	1 Ω	[33], [34]
Voltage Drop, V_d	1.2 V	[31], [32]
Characteristic Temp., T_η	100 K	[31], [35]
Characteristic Temp., T_0	70 K	[31], [35]
Thermal Impedance, Z_T	40°C/W	[35], [36]

where $\eta(T)$ and $I_{th}(T)$ are the slope efficiency and threshold current at a temperature T, $P_d(I) = I^2R_s + IV_d$ is the dissipated electrical power, Z_T is the thermal impedance, I is the injection current, R_s is the series resistance, and V_d is the voltage drop.

Table II shows the laser parameters assumed in this study. Here we model external and integrated laser systems with similar parameters in order to emphasize the temperature-dependent performance of external and integrated lasers. The threshold current and slope efficiency values in Table II are specified for lasers operating at 25 °C. Note that the parameters shown in Table II are used in this study only to calculate the laser optical power and power consumption and necessarily do not take into account the many significant trade-offs associated with laser fabrication.

Furthermore, we assume integrated lasers systems cannot be cooled due to size or complexity constraints and thus operate at the maximum PIC temperature of 70 °C, as specified in Section II-E. Laser sources used for local oscillators in coherent architectures may require temperature control for thermal tuning, though the requirement for thermal tuning may be partially or fully alleviated by including wavelength-selective structures in the sources [37], locating the sources further from the switching ASIC, or using digital frequency offset correction. Additionally, we assume that integrated lasers have a footprint of 0.01×1 mm ^{2} [31] and a center wavelength temperature dependence of 0.08 nm/° C [38]. External laser systems, subject to fewer space constraints, can use thermo-electric coolers (TEC) to decrease junction temperature and boost performance, at the cost of TEC power consumption, P_{TEC}. We assume the TEC coefficient of performance is $\eta_{TEC} = P_d/P_{TEC} = 2$. In Section VI-B, we further study the effect of additional cooling techniques for integrated laser sources and the impact on system performance.

Fig. 2(a) demonstrates the output optical power of the external and integrated lasers with the above parameters. The maximum output power for cooled external lasers (∼125 mW) is significantly higher than that for integrated lasers (∼75 mW), owing to cooling. The wall plug efficiency, indicated on the right axis, shows that external lasers, despite the TEC power consumption, can be more efficient than integrated lasers. Notably, the optical power of external light sources in Fig. 2 does not include laser-to-fiber or fiber-to-PIC coupling losses, though they are included in the link analysis in Section VI.

2) MWS Modeling: The performance of multi-wavelength sources must be quantified to determine the practicality of the MWS architectures in Section III b. Although promising MWS-based optical links have been reported, we model both the external and integrated MWSs as black boxes with 2% wall plug efficiency and maximum optical powers of 50 mW and 30 mW, respectively. This conservative lower bound on performance could simply be achieved by a bank of external lasers that are multiplexed into a single waveguide or one of the other promising technologies described in the above references. Finally, we assume that the total footprint of the MWS is negligible compared to the footprint of the integrated amplifiers.

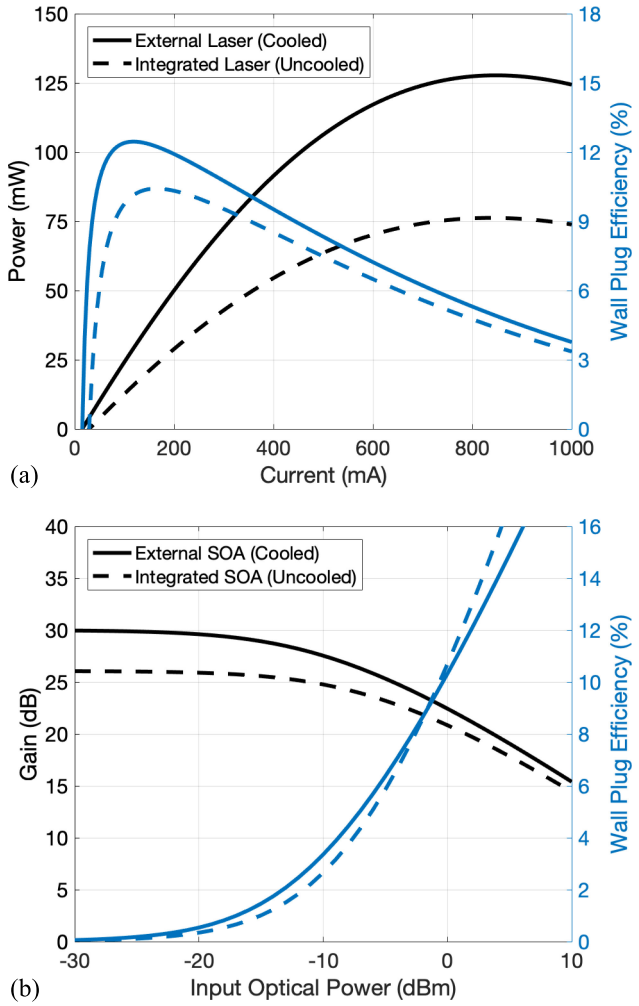

Fig. 2. Performance of external and integrated (a) lasers and (b) semiconductor optical amplifiers (SOAs) for co-packaged optical interfaces based on single- and multi-wavelength sources. External active optical components can achieve higher power due to thermo-electric cooling, at the cost of a decrease in wall-plug efficiency.

3) Amplifier Modeling: Similar to the laser modeling, the amplifier modeling must demonstrate the performance of external and integrated amplifiers for wide ranges of input current and operating temperature values. In the rest of the study, we assume the use of semiconductor optical amplifiers (SOAs) due to their efficiency in the O band and feasibility of integration [33]. The gain G of an SOA can be calculated by the following implicit equation [39]:

$$G = G_0(T)e^{-(G-1)(P_{in}/P_s)}, \tag{2}$$

TABLE III
SOA PARAMETERS

Parameter	Value	Reference
Small Signal Gain, G_0	30 dB	[36], [41]
Saturation Power, P_s	20 dBm	[41], [42]
Forward Voltage Drop, V_d	1.5 V	[41]

TABLE IV
INTEGRATED COMPONENT PARAMETERS

Integrated Component	Insertion Loss	Footprint	Ref.
(De)Multiplexer	3 dB	1 mm^2	[44]
Power Splitter	0.1 dB	30 μm^2	[45]
Waveguide Routing	2 dB	N/A	[46]
Waveguide Crossing	0.05 dB	100 μm^2	[47]
Fiber-to-chip Coupler	1 dB	180 μm^2	[48]
Laser-to-fiber Coupler	1 dB	N/A	[49]
Polarization Multiplexing	1 dB	N/A	[50]

where P_{in} is the input optical power, P_s is the SOA saturation power, and the temperature-dependent unsaturated gain $G_0(T)$ is [39]

$$G_0(T) = e^{(\eta/T)(I - qN_0/\tau_c)}, \tag{3}$$

where I is the injection current, q is the electron charge, N_0 is the carrier density at transparency, τ_c is the carrier lifetime. The efficiency parameter η can be calculated from the unsaturated gain G_0 at a specified current and temperature. Similar to the laser modeling, G_0 is exponentially dependent on the temperature [30].

Table III shows the SOA parameters used for MWS architectures. We assume the SOAs have a transparency current density $qN_0/\tau_c = 50$ mA, a footprint of 0.01×1 mm^2, and negligible wavelength-dependent performance variation [40]. As with the external lasers, external SOAs are cooled with a TEC with a performance coefficient $\eta_{TEC} = 2$.

Fig. 2(b) shows the performance of external and integrated SOAs for MWS architectures for intra-data center co-packaged optical interfaces. External SOAs have higher unsaturated gains due to temperature control while external and integrated SOAs have comparable wall plug efficiency due to TEC power consumption and high-temperature operation, respectively.

B. Passive Optics

Table IV shows insertion losses and footprints of the integrated passive optical components used in this study. The absolute number of optical components is determined by the type of architecture (external vs. integrated and MWS vs. SWS) as well as the number of active light sources used. We assume an athermal (de)multiplexing technology is used to avoid temperature-dependent passband shifts [43]. Section VI-A discusses how the number of active optical components is optimized.

C. Transceiver Characteristics

The integrated transmitter and receivers shown in Fig. 1 are key in determining the overall link performance. Here, we consider transceivers using direct detection modulation formats, non-return-to-zero on-off-keying (NRZ-OOK) and 4-ary pulse amplitude modulation (4-PAM), or coherent detection formats, dual-polarization quaternary phase shift keying (DP-QPSK) and dual-polarization 16-ary quadrature amplitude modulation (DP-16-QAM).

TABLE V
INTEGRATED TRANSCEIVER PARAMETERS

Modulation Format	Energy Efficiency	Areal Density	Linear Density	Ref.
NRZ-OOK	4.2 pJ/bit	38 Gb/s/mm^2	100 Gb/s/mm	[51]
4-PAM	8.5 pJ/bit	65 Gb/s/mm^2	120 Gb/s/mm	[52]
DP-QPSK	30 pJ/bit	8 Gb/s/mm^2	15 Gb/s/mm	[53], [54]
DP-16-QAM	26 pJ/bit	15 Gb/s/mm^2	30 Gb/s/mm	[55], [56]

Table V shows state-of-the-art integrated transceiver parameters for the above direct and coherent modulation formats. As discussed in Section II, we assume a symbol rate of 53.125 Gbaud and state-of-the-art IC process nodes. The transceiver energy efficiency—defined as the total power consumption divided by the bit rate—has been a key parameter governing the evolution and adoption of standards for pluggable transceivers. Likewise, for co-packaged interfaces, energy efficiency influences the choice of modulation/detection method as well as thermal management and component placement. Owing to steady increases in switching capacity per unit area of switching ASICs and the scaling demands of data center switches, energy efficiency requirements have been increasingly driven downward. Because the scaling of IC process nodes will eventually enable lower-power intra-data center links, in the rest of this paper, we will focus on quantifying the power consumption of the light sources, which is impacted by the architecture of the co-packaged optical interfaces, even though it typically represents a small portion of total transceiver power consumption.

While the footprint (areal density) of the transceivers shown in Table V is likely sufficient for future switching generations, outperforming the 7.4 Gb/s/mm^2 required for 51.2 Tb/s switching, the shoreline (linear density) requirement of 200 Gb/s/mm is not satisfied by any of the listed transceivers. The areal density values in Table V are calculated by dividing the total bit rate by the referenced PIC footprint. In calculating the linear density values of Table V, we assume that the maximum electrical interconnect between switch ASIC and transceiver components is 3 mm. As discussed in Section II-B, however, shoreline restrictions can be relaxed by increasing the substrate size or introducing longer electrical fan-outs. Thus, in the rest of this paper, we relax the shoreline density requirements in order to demonstrate the performance of the various architectures and modulation formats.

V. SYSTEM MODELING

A. Link Performance

In this section, we analyze the link performance in the presence of various noise sources for direct and coherent detection

architectures. Reference transmitter and receiver diagrams can be found in [14]. In the following, we assume that the receiver photodetector sensitivity is $R = 1$ A/W, the transimpedance amplifier input-referred noise is $i_N = 15$ pA/$\sqrt{\text{Hz}}$, the receiver bandwidth is $\Delta f_e = 35$ GHz, and that the dark current is negligible. Furthermore, we assume 6.25% coding and Ethernet overhead and a pre-forward error correction (FEC) bit-error ratio $\text{BER}_{req} = 2.4\times10^{-4}$, typical for FEC codes such as RS(544,514).

1) Direct Detection: The modulator insertion loss—defined as the input continuous-wave power divided by the peak output power—depends on the modulator absorptive loss and maximum drive voltage swing. We assume the peak-to-peak drive swing is limited to V_π, so the insertion loss of direct detection modulators is given only by absorptive loss, assumed to be 3 dB [57]. The modulation loss—defined to be the peak output power divided by the average output power—depends on the modulation format and the modulator extinction ratio. We assume an extinction ratio of $10\log_{10}(1/r_{ex}) = 10$ dB, resulting in a modulation loss of 2.6 dB for both NRZ-OOK and 4-PAM.

In the direct detection receivers considered here, employing a single photodetector, thermal noise dominates due to the high insertion loss of integrated links. Amplified spontaneous emission noise from the SOAs in the MWS architectures is negligible, differing from the analysis in [14]. The receiver signal-to-noise ratio (SNR) for NRZ-OOK is $\text{SNR} = (RP_{rx}/\sigma_{th})^2$ where P_{rx} is the received power and the Gaussian-distributed thermal noise has a variance $\sigma_{th}^2 = i_N^2\Delta f_e$. The receiver sensitivity is the received power P_{rx} required such that $Q(\sqrt{\text{SNR}}) = \text{BER}_{req}$. For 4-PAM, we assume equally spaced intensity levels and assume an ideal performance penalty of $10\log_{10}(M-1) = 4.8$ dB with respect to NRZ-OOK. The extinction ratio penalty for both NRZ-OOK and 4-PAM is $\delta_{ex} = (1+r_{ex})/(1-r_{ex}) = 0.9$ dB.

2) Coherent Detection: Coherent modulators driven by a peak-to-peak drive swing of V_π are subject to an additional 3 dB insertion loss compared to direct detection modulators, so the total modulator insertion loss is 6 dB for both DP-QPSK and DP-16-QAM. The modulation loss is 0 dB for DP-QPSK and 2.6 dB for DP-16-QAM.

The coherent receiver requires a local oscillator (LO), as shown in Fig. 1, as well as polarization demultiplexing and 90° optical hybrids. Moreover, because the LO power P_{LO} is typically high, LO-dominated shot noise and thermal noise are both significant in the received signal. The LO-dominated shot noise can be modeled as Gaussian distributed with variance $\sigma_{sh}^2 \approx (1/2)qRP_{LO}\Delta f_e$ where we have assumed that the received power P_{rx} is significantly lower than P_{LO}, valid for the high-loss links discussed here. Thus, the receiver SNR per symbol is

$$\text{SNR} = \frac{R^2P_{rx}P_{LO}}{4\sigma_{sh}^2 + 8\sigma_{th}^2}, \tag{4}$$

where the additional constant factors in the denominator are included because P_{rx} and P_{LO} are referenced to the input of the dual-polarization receiver and thus undergo splitting before arriving at the balanced photodetectors. Because the shot noise is dominated by the local oscillator power, it is reduced by a factor of two relative to a single-polarization receiver.

The receiver sensitivity can be calculated by determining the required receiver power P_{rx} that satisfies the following BER relation for M-QAM:

$$BER \approx \frac{4(1-1/\sqrt{M})}{\log_2(M)}Q\left(\sqrt{\frac{3}{M-1}\text{SNR}}\right), \tag{5}$$

where we note that DP-QPSK is equivalent to DP-4-QAM and the extinction ratio penalty for M-QAM systems is [14]

$$\delta_{ex} = \left(1-\sqrt{r_{ex}\left(1-\frac{1}{(\sqrt{M}-1)^2}\right)}\right)^2. \tag{6}$$

3) Optimized Power Splitting: Light source power is shared between transmitter and receiver blocks in Fig. 1, with a fraction α going to the transmitters and a fraction $1-\alpha$ going to the receiver LOs. We optimize the splitting ratio by differentiating Equation 4 with respect to α, finding the optimal splitting ratio in the presence of LO-dominated shot noise and thermal noise to be

$$\alpha_{opt} = (1+\beta^2) - \beta\sqrt{1+\beta^2}, \tag{7}$$

where β is a dimensionless parameter describing the relative noise power of the LO-dominated shot noise and thermal noise given by

$$\beta^2 = \frac{4i_N^2}{qRe^{-\gamma_{src}}P_{tot}}, \tag{8}$$

where $e^{-\gamma_{src}}$ is the total insertion loss from the light source to the transmitter-LO splitter.

B. Optical Reliability

As discussed in Section II, integrating lasers and SOAs can significantly reduce their reliability. As with the analysis of the performance of the optical devices in Section IV, it is necessary to characterize the reliability as a function of temperature and current.

Here we define the failure of a laser or SOA to be a doubling of the required injection current to maintain constant output power. This model is based on device wear out as we assume that a mature manufacturing process, with infant mortality screening, is used to fabricate the PIC. Moreover, typical random failure rates can be low compared to wear out failures for integrated devices [58]. Thus, we model the failure time t_f of integrated optical components via the following common Arrhenius model:

$$t_f = AI^{-n}e^{E_a/kT}, \tag{9}$$

where A is a constant scale factor, I is the current applied to the device, n is the current scaling factor E_a is the device activation energy, k is the Boltzmann constant, and T is the temperature.

Finally, we assume that the variation in activation energy E_a dominates the distribution in Equation 9 when compared to the current scaling factor, n. Thus, the current- and temperature-dependent failure time can be modeled as a lognormal variable with parameters $\mu = E_a/(kT) + \log(A(I/I_{ref})^n)$ and

TABLE VI
INTEGRATED OPTICAL COMPONENT RELIABILITY

Parameter	Value	Ref.
Median Time to Failure, MTTF	10×10^6 hr	[11], [59]
Activation Energy, E_a	0.7 eV	[60]
Current Scaling Factor, n	1.5	[59]
Normalized Deviation, σ_{ref}	1	[61]

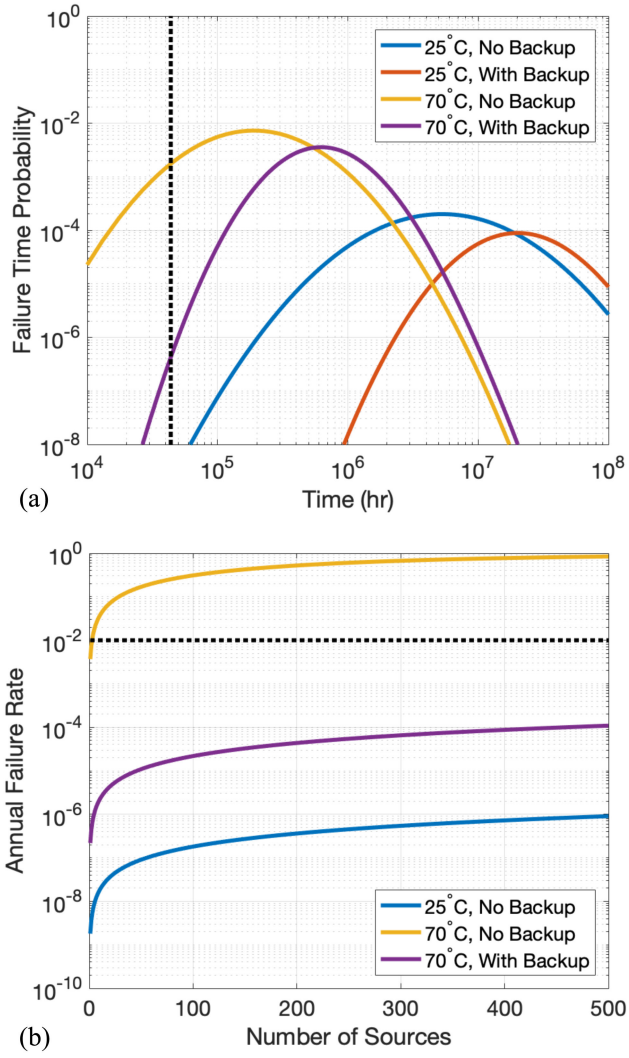

Fig. 3. Reliability of light sources and multi-source systems. The (a) probability density functions of a single light source or a source with backup determine the (b) overall annual failure rate of multi-source systems. High-temperature systems require backup lasers to meet the annual failure rate requirement, shown as a horizontal dashed line in (b). The vertical dashed line in (a) represents a typical data center switch lifetime of five years.

$\sigma = \sigma_{ref}(T_{ref}/T)$, where σ_{ref}, T_{ref}, and I_{ref} are reference parameters which, along with the median time to failure (MTTF), determine the scale factor A. Table VI shows reliability parameters assumed in this study. Owing to the lack of SOA reliability data and the similarities between integrated lasers and SOAs, we assume that the failure distributions of integrated lasers and SOAs are identical.

Fig. 3(a) shows the probability density function (PDF) of the failure time at a reference current of 600 mA. The blue and yellow curves in Fig. 3(a) demonstrate the decrease in reliability associated with operating optical devices at higher temperatures. The typical operating switch lifetime of five years [62] is shown as a dotted vertical line. To offset lower failure times for high-temperature optical components, a backup laser, connected with an optical switch may be included [9]. The red and purple curves in Fig. 3(a) demonstrate the shifted PDFs of low and high temperature devices, respectively.

These failure time distributions can also be applied to systems involving many light sources, such as the architectures of Fig. 1. Fig. 3(b) demonstrates the annual failure rate of multi-source systems, where a failure is defined as the failure of any optical link, as discussed in Section II-F. The maximum allowable failure rate, 1%, is shown as a horizontal dashed line in Fig. 3(b). For high-temperature operation, a backup laser is necessary to ensure sufficient reliability.

VI. RESULTS AND DISCUSSION

A. Scaling Switching Bandwidth

In the following subsections, the previously developed models are used to determine the link budget and minimum light source power consumption of external and integrated source architectures. As discussed in Section IV-C, in order to demonstrate architectural trends, the requirements on transceiver footprint and shoreline density are relaxed. Additionally, we assume that no more than 30% of the total system power consumption may be consumed by the optics, and the fiber loss between switches is 3 dB.

1) Link Budget: Fig. 4 demonstrates the optimized performance of the various architectures as the switch bandwidth is varied. Fig. 4(a) and Fig. 4(d) show the link budget of external and integrated architectures, respectively, versus switch bandwidth. The optimized link budget is calculated by fixing the switch bandwidth and varying the total number of light sources (lasers or SOAs) and the total number of wavelengths, subject to the power constraints. For each combination of the parameters, a link architecture is created and the link budget is estimated. For the remainder of the paper, external architectures are shown as solid lines while integrated architectures are shown as dashed lines. The maximum link budget is shown in Fig. 4(a) and Fig. 4(d).

As expected, coherent modulation formats have much higher link budgets than direct detection formats due to their superior receiver sensitivity. Coherent architectures also scale to 102.4 Tb/s switching with over 13 dB link budget while most direct detection architectures scale to 51.2 Tb/s and 102.4 Tb/s switching generations with link budgets less than 5 dB. For both external and integrated source architectures, Fig. 4(a) and Fig. 4(d) demonstrate that SWS and MWS architectures perform similarly. Integrated-source architectures, however, scale to higher switch bandwidths because of reduced coupling losses, in spite of reduced performance due to increased PIC temperatures.

Fig. 4(b) and Fig. 4(e) show the optimal number of sources for external and integrated architectures, respectively. For the external SWS architecture, plotted as circles in Fig. 4(b), the

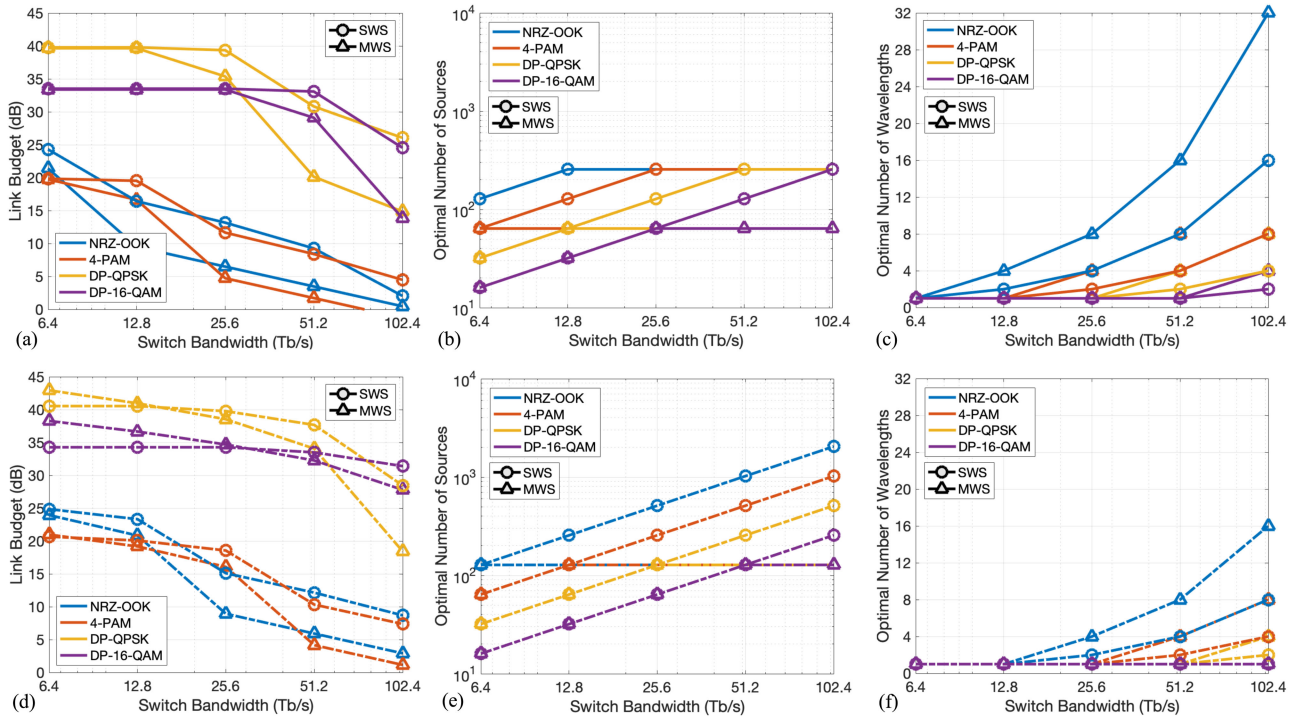

Fig. 4. Performance of optimized designs of (a–c) external and (d–f) integrated source architectures. Coherent modulation formats offer over 20 dB improvement in link budget of (a) external and (d) integrated architectures compared to direct detection formats. The optimal numbers of sources for (b) external and (e) integrated source architectures demonstrate that external architectures are limited by the total number of fiber attachments. The optimal numbers of wavelengths for (c) external and (f) integrated source architectures demonstrate that multi-wavelength links are necessary to scale nearly all optical links to higher switch bandwidths.

total number of sources for each modulation format does not scale beyond 256 due to the maximum fiber attachment requirement. However, for the external MWS architecture, plotted as triangles in Fig. 4(b), the optimal number of sources is lower because the MWS architectures are power-limited. As shown in Fig. 2(a), while the laser power is roughly proportional to input current, the SOA power-current relationship is more complicated. In Fig. 4(b) the optimization algorithm chooses to use fewer SOAs and split the output optical power rather than continue to scale up the number of SOAs. In general, the amount of electrical input power will determine the maximum optimal number of sources for MWS architectures. In Fig. 4(d), the MWS architectures are similarly power-limited, while the number of integrated light sources for SWS architectures scales with switch bandwidth.

Fig. 4(c) and Fig. 4(f) show the optimal number of wavelengths for external and integrated architectures, respectively. External-source architectures are forced to scale to higher numbers of wavelengths at lower switch bandwidths than integrated architectures due to the limited number of fiber attachments. Importantly, all of the architectures studied here have higher link budgets for single-wavelength optical links. This is unsurprising because the addition of wavelength-dependent components such as multiplexers significantly increase the insertion loss. The use of fewer wavelengths also can reduce complexity of integrated components such as lasers, amplifiers, and multiplexers. This effect on insertion loss is demonstrated in Fig. 4(a) and Fig. 4(d) where large drops in link budget are observed when architectures are forced to scale to more than one wavelength.

In order to more easily interpret the link budgets shown in Fig. 4(a) and Fig. 4(d), we show an example link budget breakdown of an optimized external SWS architecture at 51.2 Tb/s in Table VII, which includes all losses from the single-polarization laser output to the integrated receiver. The table compares the link budgets when using 4-PAM and DP-16-QAM, which have link budgets of 8.4 dB and 33.1 dB, respectively. Due to the limited number of fiber attachments, the 4-PAM implementation requires four wavelengths per fiber, while the DP-16-QAM implementation can use a single wavelength per fiber. For the 4-PAM implementation, which has 64 lasers for each of the four wavelengths, the light at each input wavelength must be split in half to support the 128 transceiver modules. In order to order the wavelengths for multiplexing, a maximum of three waveguide crossings are required. Table VII demonstrates that architectures using multiple wavelength incur a penalty of over 6 dB due to (de)multiplexing and required waveguide crossings.

Finally, increasing the output power from the light sources, or scaling the symbol rate beyond 53.125 Gbaud may enable modulation formats to scale differently than shown in Fig. 4 or Table VII. For example, for the 102.4 Tb/s switch generation with external single-wavelength sources, if the maximum optical output power is increased from 100 mW to 200 mW, the link budget would increase by 3 dB, assuming that eye safety requirements and power consumption requirements can be satisfied. Additionally, designs using high-power sources may suffer additional loss due to nonlinear effects in integrated waveguides such as two-photon absorption and four-wave mixing [63]. Similarly, increasing the symbol rate from 53.125 Gbaud to 106.25 Gbaud

TABLE VII
51.2 TB/S EXTERNAL SWS ARCHITECTURE LINK BUDGET

Parameter	4-PAM	DP-16-QAM
Number of Wavelengths	4	1
Number of Sources	256	128
Source Power	17.8 dBm	19.7 dBm
Laser-to-fiber Coupling Loss	1 dB	1 dB
Input Coupling Loss	1 dB	1 dB
Routing Loss	2 dB	2 dB
Splitting Loss (Ratio)	3.1 dB (2)	0 dB (1)
Waveguide Crossing Loss (Number)	0.2 dB (3)	0 dB (0)
LO Splitting Loss	N/A	1.2 dB
Modulator Insertion Loss	3 dB	6 dB
Modulation Loss	2.6 dB	2.6 dB
Tx Multiplexer Loss	3 dB	0 dB
Polarizing Beam Splitter Loss	0 dB	1 dB
Output Coupling Loss	1 dB	1 dB
Fiber Link Loss	3 dB	3 dB
Input Coupling Loss	1 dB	1 dB
Rx Demultiplexer Loss	3 dB	0 dB
Received Optical Power	-6.1 dBm	-0.1 dBm
LO Optical Power	N/A	10.6 dBm
Receiver Sensitivity*	-14.5 dBm	-33.2 dBm
Link Budget	8.4 dB	33.1 dB

*Referenced to input of RX in Fig. 1

between the 51.2 Tb/s and 102.4 Tb/s switch generations, would allow for use of 4-PAM links with four wavelengths per fiber, assuming that the implementation penalty is lower than 8.4 dB, the link budget at the 51.2 Tb/s switch generation, shown in Fig. 4(a).

2) Light Source Efficiency: In the previous section, we assume a maximum power consumption for the light sources. Here this constraint is relaxed and the electrical power is varied to find the minimum power consumption for each architecture. Fig. 5(a) shows the power consumption of the light source for external architectures. Due to the superior receiver sensitivity of coherent modulation formats, lower source power consumption is required, despite power sharing between the LO and transmitter laser. In general, the external MWS architectures consume more power than the SWS architectures because of the fixed electrical power of the external MWS.

Fig. 5(b) shows the power consumption of the light sources for the integrated architectures, which consume less electrical power than the external architectures, despite the reduced performance due to temperature. The lower power consumption of integrated sources, despite being less power-efficient than external sources as shown in Fig. 2, results from two effects. First, integrated source architectures do not require input couplers to route light to the transmitters. Secondly, integrated source architectures can utilize more sources than the fiber attachment-constrained external source architectures, and thus operate in a lower-power, more efficient regime than external sources. This efficient regime of operation allows integrated source architectures to operate with a single wavelength at larger switch bandwidths and thus avoid (de)multiplexer losses. Interestingly, coherent MWS

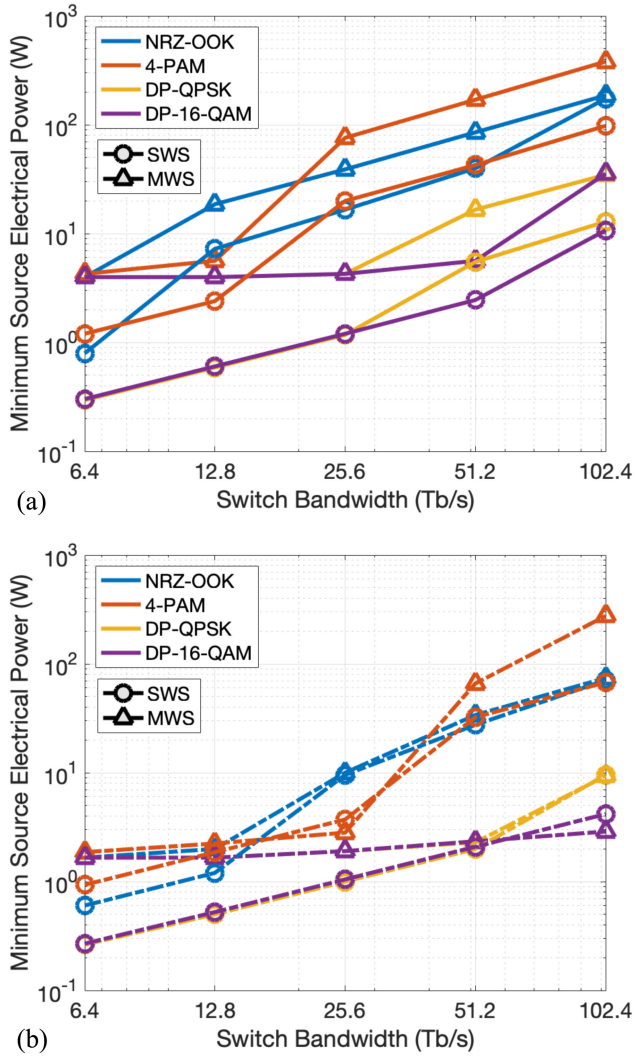

Fig. 5. Minimum electrical power consumption of (a) external and (b) integrated sources. While single-wavelength sources scale more efficiently for external source architectures, multi-wavelength sources are more efficient for integrated architectures and large switching bandwidths.

architectures consume the least power for large switch bandwidths due to the efficiency of SOAs.

B. Fixed-Bandwidth Performance

In the previous subsection, the performance of external and integrated source architectures was demonstrated for various switching bandwidths. Here, we assume a fixed switch bandwidth of 51.2 Tb/s, corresponding to a switch generation still several years away from commercial deployment [3], and vary the number of fiber attachments and PIC temperature.

1) Fiber Attachments: In Section II, a maximum of 512 fiber attachments, including both source fiber attachments and data-carrying fiber attachments, is assumed based on practical packaging limitations as well as edge coupler density. Advances in edge coupler technology, fiber array management, or the use of grating couplers may increase, or further limit, the maximum number of fiber attachments. Fig. 6 demonstrates the performance of external and integrated source architectures for

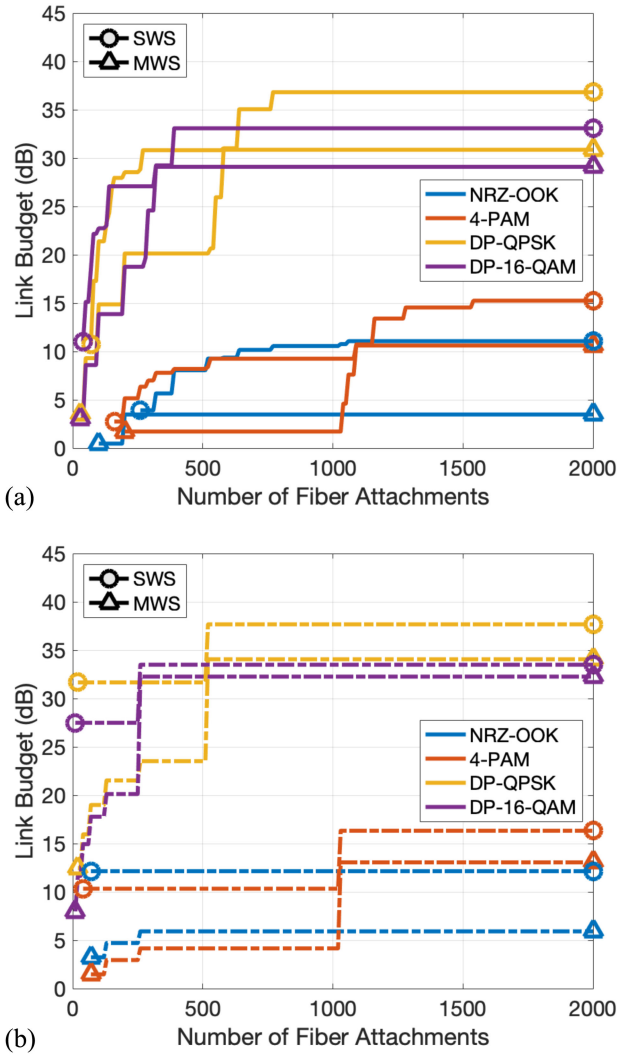

Fig. 6. Link budget of (a) external and (b) integrated source architectures for a 51.2 Tb/s switch versus maximum number of fiber attachments. Because the fiber limit includes data-carrying fibers, reducing the number of fiber attachments often increases the number of required wavelengths. Coherent architectures can be used with fewer fiber attachments because of their superior receiver sensitivity and compatibility with denser WDM links. External source architectures exhibit more variation with number of fiber attachments because fiber attachments must be allocated to connect the external source.

various numbers of fiber attachments connected to a 51.2 Tb/s switch.

Fig. 6(a) shows the performance of external source architectures versus number of fiber attachments. Coherent modulation formats can support architectures with fewer than 100 fiber attachments because they have higher spectral efficiencies and superior receiver sensitivities, enabling the use of many wavelengths per fiber. Additionally, external sources architectures require more fiber attachments than integrated source architectures because fibers must be allocated for the external sources.

Similar to external source architectures, integrated source architectures can use fewer fiber attachments with coherent modulation formats, as shown in Fig. 6(b). While the link budgets of integrated source architectures only change when the number of fibers is doubled, corresponding to a halving of the required bit

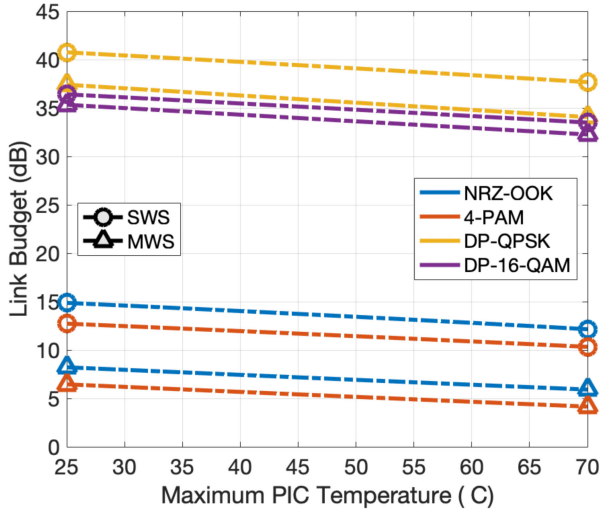

Fig. 7. Link budget of integrated source architectures versus maximum PIC temperature. The variation in link budget over 45 °C is less than 5 dB for all integrated source architectures.

rate per fiber, the link budgets of external source architectures increase at smaller increments of fiber attachments, because additional fiber attachments correspond to the addition of more external light sources, and thus more optical power. For small numbers of fiber attachments, integrated source architectures, in contrast to external source architectures, generally have high link budgets because the number of sources, and thus the total optical power, is unconstrained.

Importantly, decreasing the number of fiber attachments for a fixed switch bandwidth necessitates the use of additional wavelengths to meet the required bit rate per fiber. In some cases, this results in architectures requiring more wavelengths than are practical. For example, the optimized NRZ-OOK, MWS architectures (blue triangles) requiring the fewest fiber attachments in Fig. 6(a) and Fig. 6(b) operate with 64 wavelengths. While this wavelength count may be technically achievable in principle, the complexity of employing this wavelength grid may be impractical for intra-data center systems. Additionally, there may be additional link budget penalties for multiplexing and splitting over large optical bandwidths, although they are not included in this study.

2) Temperature: While a maximum PIC temperature of 70 °C is assumed in Section II, advanced cooling techniques, such as water cooling, can reduce PIC temperature and improve the performance of data center switches [64]. While the complexity of implementation for some of these techniques is high, they may be implemented in future data centers.

Fig. 7 demonstrates the temperature dependent performance of integrated source architectures. Unsurprisingly, the link budget of the integrated architectures decreases with increased PIC temperature. Coherent architectures using integrated sources exhibit higher temperature-dependent variation in performance because the light source is used for both the transmitter and the receiver local oscillator. For coherent and direct detection architectures, the link budget decreases by 2-4 dB when the PIC temperature is increased to 70 °C.

VII. Conclusion

We propose four architectures for intra-data center co-packaged optical interfaces based on either single-wavelength or multi-wavelength light sources that can be either external or integrated. We model the current- and temperature-dependent performance of the light sources in the presence of practical impairments such as limited power consumption, limited footprint, and limited number of fiber attachments. We show that coherent architectures can support 102.4 Tb/s switching with link budgets 13-25 dB higher than comparable direct detection architectures. Nevertheless, we demonstrate that higher-power lasers and higher-symbol-rate electronics can scale direct detection architectures to 102.4 Tb/s switching, assuming eye safety and power consumption is managed appropriately. While integrated sources are less power-efficient than external laser sources and require a backup to achieve reliable performance over typical switch lifetimes, they can be more efficient at a system level because they are not limited by the number of fiber attachments and can thus operate in a more efficient regime. For a fixed switch bandwidth of 51.2 Tb/s, we demonstrate that advanced cooling can increase the performance of integrated architectures by nearly 4 dB. Finally, we show that increasing the number of fiber attachments can improve link budgets by over 10 dB, highlighting the importance of fiber attachment technologies.

References

[1] "Cisco global cloud index: Forecast and methodology, 2016-2021," Feb. 2018, (White Paper).

[2] D. Patterson, I. De Sousa, and L.-M. Achard, "The future of packaging with silicon photonics," *Chip Scale Rev.*, vol. 21, no. 1, 2017, p. 1.

[3] M. Rakowski, "Silicon photonics platform for 50G optical interconnects," Sep. 2017, Cadence Photonic Summit and Workshop 2017.

[4] "100 Gb/s per lane for electrical interfaces and PHYs," Nov. 2017, IEEE 802.3 Consensus Building.

[5] C. Schmidt, C. Kottke, R. Freund, F. Gerfers, and V. Jungnickel, "Digital-to-analog converters for high-speed optical communications using frequency interleaving: Impairments and characteristics," *Opt. Express*, vol. 26, no. 6, pp. 6758–6770, Mar. 2018.

[6] C. Kachris, K. Bergman, and I. Tomkos, *Optical Interconnects for Future Data Center Networks*. Berlin, Germany: Springer Science & Business Media, 2012.

[7] A. Ghiasi, "Large data centers interconnect bottlenecks," *Opt. Express*, vol. 23, no. 3, pp. 2085–2090, Feb. 2015.

[8] A. Bechtolsheim, "Scaling the cloud network," *Open Comput. Summit*, Mar. 20, 2018.

[9] S. Fathololoumi *et al.*, "1.6 Tbps silicon photonics integrated circuit for co-packaged optical-io switch applications," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. T3H.1.

[10] H. Yu *et al.*, "400 Gbps fully integrated DR4 silicon photonics transmitter for data center applications," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. T3H–6.

[11] D. Jung *et al.*, "Impact of threading dislocation density on the lifetime of InAs quantum dot lasers on Si," *Appl. Phys. Lett.*, vol. 112, no. 15, 2018, Art. no. 153507.

[12] Q. Cheng, M. Bahadori, M. Glick, S. Rumley, and K. Bergman, "Recent advances in optical technologies for data centers: A review," *Optica*, vol. 5, no. 11, pp. 1354–1370, 2018.

[13] M. Zhang *et al.*, "Broadband electro-optic frequency comb generation in a lithium niobate microring resonator," *Nature*, vol. 568, no. 7752, pp. 373–377, 2019.

[14] B. Buscaino, B. D. Taylor, and J. M. Kahn, "Multi-Tb/s-per-fiber coherent co-packaged optical interfaces for data center switches," *J. Lightw. Technol.*, vol. 37, no. 13, pp. 3401–3412, 2019.

[15] C. D. Patel, R. Sharma, C. E. Bash, and A. Beitelmal, "Thermal considerations in cooling large scale high compute density data centers," in *Proc. ITherm. 8th Intersoc. Conf. Thermal Thermomech. Phenomena Electron. Syst.*, 2002, pp. 767–776.

[16] E. Timurdogan *et al.*, "400G silicon photonics integrated circuit transceiver chipsets for CPO, OBO, and pluggable modules," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. T3H.2.

[17] P. Maniotis, L. Schares, B. G. Lee, M. A. Taubenblatt, and D. M. Kuchta, "Scaling hpc networks with co-packaged optics," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. T3K.7.

[18] R. Marchetti, C. Lacava, L. Carroll, K. Gradkowski, and P. Minzioni, "Coupling strategies for silicon photonics integrated chips [invited]," *Photon. Res.*, vol. 7, no. 2, pp. 201–239, Feb. 2019.

[19] L. Carroll *et al.*, "Photonic packaging: Transforming silicon photonic integrated circuits into photonic devices," *Appl. Sci.*, vol. 6, no. 12, pp. 3–8, 2016.

[20] K. Suematsu *et al.*, "Super low-loss, super high-density multi-fiber optical connectors," *Furukawa Rev.*, vol. 23, pp. 53–58, 2003.

[21] A. Ghiasi, "400GBase-LR8: A proposal for 10 km objective using 50 Gb/s PAM4 signaling," *IEEE 802.3bs Task Force*, May 2015.

[22] "Co-packaged optics external laser source guidance document," May 2020. [Online]. Available: https://www.facebook.com/CoPackagedOpticsCollaboration/

[23] P. Doussiere, C.-L. Shieh, S. DeMars, and K. Dzurko, "Very high-power 1310 nm InP single mode distributed feed back laser diode with reduced linewidth," in *Novel In-Plane Semiconductor Lasers VI*, C. Mermelstein and D. P. Bour, Eds., vol. 6485. Bellingham, WA, USA: International Society for Optics and Photonics., 2007, pp. 129–136.

[24] "Part 1: Equipment classification and requirements," May 2014, IEC 60825-1:2014 Safety of laser products.

[25] K. Schulmeister, M. Jean, D. Lund, and B. E. Stuck, "Comparison of corneal injury thresholds with laser safety limits," in *Proc. Int. Laser Safety Conf.*, vol. 2019, no. 1, 2019, p. 303, [Online]. Available: https://lia.scitation.org/doi/abs/10.2351/1.5118573.

[26] P. Gill, N. Jain, and N. Nagappan, "Understanding network failures in data centers: measurement, analysis, and implications," in *Proc. ACM SIGCOMM Conf.*, 2011, pp. 350–361.

[27] J. Meza, T. Xu, K. Veeraraghavan, and O. Mutlu, "A large scale study of data center network reliability," in *Proc. Internet Meas. Conf.*, 2018, pp. 393–407.

[28] Y. Li, Y. Zhang, L. Zhang, and A. W. Poon, "Silicon and hybrid silicon photonic devices for intra-datacenter applications: state of the art and perspectives," *Photonics Res.*, vol. 3, no. 5, pp. B10–B27, 2015.

[29] D. J. Blumenthal *et al.*, "Frequency stabilized links for coherent wdm fiber interconnects in the datacenter," *J. Lightw. Technol.*, vol. 38, pp. 3376–3386, 2020.

[30] L. A. Coldren, S. W. Corzine, and M. L. Masanovic, *Diode Lasers and Photonic Integrated Circuits*, 2nd ed. Hoboken, NJ, USA: Wiley & Sons, 2012.

[31] A. Y. Liu *et al.*, "High performance continuous wave 1.3um quantum dot lasers on silicon," *Appl. Phys. Lett.*, vol. 104, no. 4, 2014, Art. no. 041104.

[32] H.-H. Chang *et al.*, "1310 nm silicon evanescent laser," *Opt. Express*, vol. 15, no. 18, pp. 11 466–11 471, 2007.

[33] S. Liu *et al.*, "High-performance o-band quantum-dot semiconductor optical amplifiers directly grown on a cmos compatible silicon substrate," *ACS Photon.*, vol. 6, no. 10, pp. 2523–2529, 2019.

[34] 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*, vol. 14, no. 20, pp. 9203–9210, Oct. 2006.

[35] M. N. Sysak *et al.*, "Hybrid silicon laser technology: A thermal perspective," *IEEE J. Select. Topics Quantum Electron.*, vol. 17, no. 6, pp. 1490–1498, Nov./Dec. 2011.

[36] M. L. Davenport, S. Skendžić, N. Volet, J. C. Hulme, M. J. R. Heck, and J. E. Bowers, "Heterogeneous silicon/III-V semiconductor optical amplifiers," *IEEE J. Select. Topics Quantum Electron.*, vol. 22, no. 6, pp. 78–88, Nov./Dec. 2016.

[37] A. Malik, J. Guo, M. A. Tran, G. Kurczveil, D. Liang, and J. E. Bowers, "Widely tunable, heterogeneously integrated quantum-dot o-band lasers on silicon," *Photon. Res.*, vol. 8, no. 10, pp. 1551–1557, Oct. 2020.

[38] H. Duprez *et al.*, "1310 nm hybrid InP/InGaAsP on silicon distributed feedback laser with high side-mode suppression ratio," *Opt. Express*, vol. 23, no. 7, pp. 8489–8497, Apr. 2015.

[39] G. P. Agrawal, *Fiber-Optic Communication Systems*, 3rd ed. Hoboken, NJ, USA: Wiley, 2002.

[40] D. Jung *et al.*, "Highly reliable low-threshold inas quantum dot lasers on on-axis (001) si with 87% injection efficiency," *ACS photon.*, vol. 5, no. 3, pp. 1094–1100, Dec. 2017.
[41] S. Liu *et al.*, "High efficiency, high gain and high saturation output power quantum dot SOAS grown on Si and applications," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. T4H.3.
[42] K. V. Gasse, R. Wang, and G. Roelkens, "27 db gain iii-v-on-silicon semiconductor optical amplifier with >17 dbm output power," *Opt. Express*, vol. 27, no. 1, pp. 293–302, Jan. 2019.
[43] S. Namnabat *et al.*, "Athermal silicon optical add-drop multiplexers based on thermo-optic coefficient tuning of SOL-gel material," *Opt. Express.*, vol. 25, no. 18, pp. 21 471–21 482, Sep. 2017.
[44] S. Pathak, M. Vanslembrouck, P. Dumon, D. V. Thourhout, and W. Bogaerts, "Compact SOI-based polarization diversity wavelength demultiplexer circuit using two symmetric awgs," *Opt. Express*, vol. 20, no. 26, pp. B493–B500, Dec. 2012.
[45] Z. Lu *et al.*, "Broadband silicon photonic directional coupler using asymmetric-waveguide based phase control," *Opt. Express*, vol. 23, no. 3, pp. 3795–3808, Feb. 2015.
[46] L. Pavesi *et al.*, "Silicon photonics III," *Topics in Applied Physics (Springer Berlin Heidelberg, 2016)*, vol. 122, 2016.
[47] Y. Ma *et al.*, "Ultralow loss single layer submicron silicon waveguide crossing for soi optical interconnect," *Opt. Express*, vol. 21, no. 24, pp. 29 374–29 382, Dec. 2013.
[48] L. Carroll, D. Gerace, I. Cristiani, and L. C. Andreani, "Optimizing polarization-diversity couplers for si-photonics: reaching the -1db coupling efficiency threshold," *Opt. Express*, vol. 22, no. 12, pp. 14 769–14 781, Jun. 2014.
[49] H. Zhou, H. Xu, and J.-a. Duan, "Review of the technology of a single mode fiber coupling to a laser diode," *Opt. Fiber Technol.*, vol. 55, 2020, Art. no. 102097.
[50] C. Li, D. Dai, and J. E. Bowers, "Ultra-broadband and low-loss polarization beam splitter on silicon," in *Proc. Opt. Fiber Commun. Conf.*, 2020, Art. no. Th1A.4.
[51] M. Moralis-Pegios *et al.*, "4-channel 200 gb/s wdm o-band silicon photonic transceiver sub-assembly," *Optics Express*, vol. 28, no. 4, pp. 5706–5714, 2020.
[52] C. Loi *et al.*, "A 400 gb/s transceiver for PAM-4 optical direct-detect application in 16 nm finFET," in *Proc. IEEE Int. Solid-State Circuits Conf.*, 2019, pp. 120–122.
[53] J. K. P. Joseph, M. Kahn and A. Shastri, "Low-complexity coherent interfaces for intra-data center networks," 2020, IEEE Photonics Society Summer Topicals. [Online]. Available: https://ee.stanford.edu/ jmk/pubs/sum.topical.20.WA1.2.pdf
[54] P. Dong, C. Xie, L. Chen, L. L. Buhl, and Y.-K. Chen, "112-gb/s monolithic pdm-qpsk modulator in silicon," *Opt. Expr.*, vol. 20, no. 26, pp. B624–B629, 2012.
[55] B. S. G. Pillai *et al.*, "End-to-end energy modeling and analysis of long-haul coherent transmission systems," *J. Lightw. Technol.*, vol. 32, no. 18, pp. 3093–3111, 2014.
[56] P. Dong, X. Liu, S. Chandrasekhar, L. L. Buhl, R. Aroca, and Y.-K. Chen, "Monolithic silicon photonic integrated circuits for compact 100^{+}Gb/s coherent optical receivers and transmitters," *J. Select. Topics Quantum Electron.*, vol. 20, no. 4, pp. 150–157, Jul./Aug. 2014.
[57] J. Witzens, "High-speed silicon photonics modulators," *Proc. IEEE*, vol. 106, no. 12, pp. 2158–2182, 2018.
[58] J.-S. Huang, "Reliability of optoelectronics," in *Reliability Characterisation of Electrical and Electronic Systems*. Amsterdam, The Netherlands: Elsevier, 2015, pp. 83–114.
[59] S. Srinivasan, N. Julian, J. Peters, D. Liang, and J. E. Bowers, "Reliability of hybrid silicon distributed feedback lasers," *IEEE J. Select. Topics Quantum Electron.*, vol. 19, no. 4, pp. 1 501 305–1 501 305, Jul.–Aug. 2013.
[60] Jia-Sheng Huang, "Temperature and current dependences of reliability degradation of buried heterostructure semiconductor lasers," *IEEE Trans. Device Mater Rel.*, vol. 5, no. 1, pp. 150–154, Mar. 2005.
[61] J. L. Spencer, "Calculating laser diode reliability," in *Proc. IEEE Glob. Telecommun. Conf. Exhib. Commun. Inf. Age*, 1988, vol. 1, pp. 63–69.
[62] A. Shehabi *et al.*, "United states data center energy usage report," Lawrence Berkeley National Lab. (LBNL), Berkeley, CA, USA, Tech. Rep., 2016.
[63] Y. Zhang *et al.*, "Non-degenerate two-photon absorption in silicon waveguides: analytical and experimental study," *Opt. Express*, vol. 23, no. 13, pp. 17 101–17 110, Jun. 2015.
[64] L. Li, W. Zheng, X. Wang, and X. Wang, "Placement optimization of liquid-cooled servers for power minimization in data centers," in *Proc. IEEE Int. Green Comput. Conf.*, 2014, pp. 1–10.

Brandon Buscaino (Member, IEEE) received the B.S. degree in physics, in 2015, and the M.S. and Ph.D. degrees in electrical engineering, in 2016 and 2020 respectively, from Stanford University, Stanford, CA, USA. He currently works with Infinera Corporation as a Coherent Optical Systems Architect. His research interests include coherent optical fiber communications, electro-optic frequency combs, intra-data center link design, and co-packaged optics.

Elizabeth Chen (Student Member, IEEE) received the B.A.Sc. degree in engineering science from the University of Toronto, Toronto, ON, Canada, in 2019. She is working toward the Ph.D. degree in electrical engineering with Stanford University, Stanford, CA, USA. Her current research interests include optical communications and coherent data center links.

James W. Stewart (Member, IEEE) has more than 30 years of experience developing photonic hardware for optical communications. He works with Facebook supporting the Development of Optical Networking Hardware for their hyper-scale data centers. He has been granted more than 20 US patents and has authored or coauthored numerous publications for conferences and journals on optical communication subsystems.

Thang Pham received the M.S. degree in electrical and electronic engineering from Yonsei University, South Korea, in 2009 and the Ph.D. degree in photonics engineering from the Technical University of Denmark, Denmark, in 2013. He was a Visiting Scholar with U.C. Berkeley, California, in 2011. He has been in the role of Optical Engineer with Facebook Inc. since 2019 where he is working on next-generation optical technologies for data-centers. From 2013 to 2019, he was with Finisar Corp. working on development of optical transceivers with advanced modulation formats. He has authored or coauthored more than 50 publications in journals and conferences.

Joseph M. Kahn (Fellow, IEEE) received A.B., M.A., and Ph.D. degrees in physics from the University of California, Berkeley in 1981, 1983, and 1986, respectively. In 1987–1990, Kahn was with AT&T Bell Laboratories. In 1989, he demonstrated the first successful synchronous (i.e., coherent) detection using semiconductor lasers, achieving record receiver sensitivity. In 1990–2003, Kahn was with the Electrical Engineering and Computer Sciences faculty, Berkeley. He demonstrated coherent detection of QPSK in 1992. In 1999, D.-S. Shiu and Kahn published the first work on probabilistic shaping for optical communications. In the 1990s and early 2000s, Kahn and collaborators performed seminal work on indoor and outdoor free-space optical communications and multi-input multi-output wireless communications. In 2000, Kahn and K.-P. Ho founded StrataLight Communications, whose 40 Gb/s-per-wavelength long-haul fiber transmission systems were deployed widely by AT&T, Deutsche Telekom, and other carriers. In 2002, Ho and Kahn applied to patent the first electronic compensation of fiber Kerr nonlinearity. StrataLight was acquired by Opnext in 2009. In 2003, Kahn became a Professor of Electrical Engineering in the E. L. Ginzton Laboratory with Stanford University. Kahn and collaborators have extensively studied rate-adaptive coding and modulation, as well as digital signal processing for mitigating linear and nonlinear impairments in coherent systems. In 2008, E. Ip and Kahn (and G. Li independently) invented simplified digital backpropagation for compensating fiber Kerr nonlinearity and dispersion. Since 2004, Kahn and collaborators have studied propagation, modal statistics, spatial multiplexing and imaging in multi-mode fibers, elucidating principal modes and demonstrating transmission beyond the traditional bandwidth-distance limit in 2005, deriving the statistics of coupled modal group delays and gains in 2011, and deriving resolution limits for imaging in 2013. Kahn's current research addresses optical frequency comb generators, coherent data center links, rate-adaptive access networks, fiber Kerr nonlinearity mitigation, ultra-long-haul submarine links, and optimal free-space transmission through atmospheric turbulence. Kahn received the National Science Foundation Presidential Young Investigator Award in 1991.