Device Scaling Considerations for Nanophotonic CMOS Global Interconnects

Sasikanth Manipatruni, Member, IEEE, Michal Lipson, Fellow, IEEE, and Ian A. Young, Fellow, IEEE

Abstract—We introduce an analytical framework to understand the path for scaling nanophotonic interconnects to meet the energy and footprint requirements of CMOS global interconnects. We derive the device requirements for sub-100 fJ/cm/bit interconnects including tuning power, serialization–deserialization energy, and optical insertion losses. Using CMOS with integrated nanophotonics as an example platform, we derive the energy/bit, linear, and areal bandwidth density of optical interconnects. We also derive the targets for device performance which indicate the need for continued improvements in insertion losses (<8 dB), laser efficiency, operational speeds (>40 Gb/s), tuning power (<100 μW/mm), serialization–deserialization (<10 fJ/bit/Operation), and necessity for spectrally selective devices with wavelength multiplexing (>6 channels).

Index Terms—Coupled resonators, integrated optics devices, integrated optoelectronic circuits, switching.

I. INTRODUCTION

A. A. Framework for Scaling CMOS Nanophotonic Global Interconnects

Increasing computational demands of enterprise and datacom (DC) applications [1], [2] have created a need for scalable interconnect solutions for high-performance computing (HPC). While the present industry focus is on the adoption of interchip optical interconnections [3], [4], the rapid adoption of multicore processors in DC and HPC [5] with high demands on bandwidth density and efficiency [1] may necessitate new interconnect solutions for same-die global interconnects [6]–[9]. Given the rapid progress in CMOS compatible nanophotonics using III–V [10], Germanium [11] as well as Silicon-based [10]–[16] platforms, the on-chip adaptability of optical interconnects for global wires [17] needs to be revisited.

In this paper, we develop a systematic framework for scaling nanophotonic interconnects by using device and system-level arguments. We use CMOS with integrated nanophotonic devices as an example platform but the analytical framework can be applied to other platforms (e.g., [10] and [11]). The device advances in couplers [18], low loss waveguides [19], modulators [20]–[24], switches [25]–[28], multiwavelength devices [29], [30], and detectors [31]–[35] can be put in context with the targets for on-chip integration using this framework.

We derive the total interconnect energy per bit, areal bandwidth density, and linear bandwidth density for a silicon photonic link considering the device parameters. We arrive at a minimal set of features for nanophotonic devices for building a scalable on-chip photonic network. We note that we limit our analysis to how photonic devices can be scaled to meet the on-chip interconnect energy/bit and bandwidth density requirements. We compare the energy/bit/mm, linear bandwidth density of the optical interconnect with generic interconnect targets for CMOS. A direct comparison with a future advanced low swing voltage (LSI) electrical on-chip interconnects is hard to achieve within the scope of this paper since such an analysis has to fundamentally comprehend the variability limits to LSI interconnects [59], [60].

II. FIGURES OF MERIT FOR NANOPHOTONIC INTERCONNECTS

We discuss four critical figures of merit for nanophotonic interconnects based on physical constraints of the optical and electrical properties of a silicon-based material system, namely, 1) energy consumption per bit $E$; 2) interconnect density $\beta$; 3) single-channel bandwidth $f$; and 4) areal bandwidth density $D$.}


S. Manipatruni and I. A. Young are with the Components Research, Intel Corp., Hillsboro, OR 97124 USA (e-mail: sasikanth.manipatruni@intel.com; ian.young@intel.com).

M. Lipson is with the School of Electrical and computer engineering, Cornell University, Ithaca, NY 14850 USA and she is also with the Kavli Institute at Cornell, Ithaca, NY 14853 USA (e-mail: ml292@cornell.edu).

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

Digital Object Identifier 10.1109/JSTQE.2013.2239262

Fig. 1. Minimal nanophotonic link with an optical source, couplers, modulators, waveguide, and a detector. A serializer and deserializer are considered to operate the link at a specific wavelength.
direction for the optical device requirements when the on-chip detector capacitance is low [36], [37]. The total optical interconnect energy per bit can be written (in the absence of tuning power and serialization) as a sum of energy from the source and the electro-optic modulator’s energy as

$$E_{\text{total}} = E_{\text{Source, Detect}} + E_{\text{EO}}$$  \hspace{1cm} (1)

where $E_{\text{Source, Detect}}$ is the energy spent in the source laser and the detector energy; $E_{\text{EO}}$ is the energy spent in electro-optic coding of the electrical information into an optical signal. A lower bound to the interconnect energy can be obtained by assuming that the detector needs to charge a capacitor of capacitance $C_d$ to a voltage $V_c$ corresponding to a specific CMOS node [36]. While this is an aggressive requirement, this assumption lets us derive a minimum bound for energy-per-bit requirements. $E_{\text{Source, Detect}}$ can be written in terms of drive laser parameters and insertion losses as

$$E_{\text{Source, Detect}} > \frac{h\omega}{\eta_L \eta_D \eta_M \eta_C} \frac{V_c C_d}{e} \cdot 10^{4\frac{C_d}{d}}$$  \hspace{1cm} (2)

where $V_c$ is the minimum voltage to which the detector capacitance is to be charged, $\eta_L$ and $\eta_D$ are the quantum efficiencies of the laser and detector normalized to the maximum values, respectively, $\eta_C$ is the laser to waveguide coupling efficiency, $\eta_D$ includes the waveguide to detector coupling efficiency, $\eta_M$ is the modulator insertion loss, $\alpha$ is the insertion loss of the waveguides in dB/cm, and $L$ is the length of the interconnect in centimeter. The previous equation is a reasonable approximation for the following conditions: 1) the detector $R_C$ response is significantly faster than the optical pulsewidth, 2) the received optical power and extinction ratio exceeds the bit-error-rate requirement (see Appendix B) of the link, and 3) the collected optical power at the receiver is always adjusted to allow full voltage at the detector. We also note that an on-chip receiver drives a significantly lower load capacitance (a few transistor gate capacitances on the order of aF-s).

The minimum electro-optic conversion energy per bit $E_{\text{EO}}$ is arrived at using the modal volume of the modulator and the injected charge density for a given transmission change. We assumed a modulator drive voltage $V_m$, electro-optic modal volume $\Theta$, and the optical transmission change $\Delta T$. $dT/dn$ is the spectral sensitivity of the optical device. $dn/dp$ is the change in refractive index $n$ versus carrier concentration $\rho$ in the electro-optic device

$$E_{\text{EO}} > V_m \Theta \Delta T$$  \hspace{1cm} (3)

We show that an idealized nanophotonic interconnect in the absence of tuning power and electrical I/O overheads can achieve sub-100 fJ/bit/cm operation. Modulator switching energy approaching 10 fJ/bit can be expected in the near future in the depletion-based and ultralow modal volume modulators [23], [38]. Fig. 2 shows the energy versus distance scaling of a nanophotonic interconnect with $E_{\text{mod}} = 10$ fJ/bit modulation energy, $C_d = 1$ fF detector capacitance, 1 dB coupling loss, 1 dB modulator insertion loss, $-1$ dB detector efficiency, and 25% efficiency laser source. (See Appendix C.)

### A. Effect of Laser Efficiency on the Energy Per Bit

The power efficiency of the laser has a significant effect on the interconnect energy per bit. In Fig. 3 we show the interconnect energy per bit for varying laser efficiency (defined as optical output power versus electrical power supplied to the laser). The low inefficiency of the laser may arise due to several factors including the requirement for thermoelectric cooling, collection efficiency, and leakage power. At 5% wall plug efficiency, the interconnect energy/bit at 1-cm length can approach 50 fJ/bit/cm, for idealized interconnects with no tuning requirement. The effect of additional insertion loss due to routing and selective devices is described in Appendix E.

### B. Effect of Tuning Nanophotonic Devices to Offset Variability and Temperature Dependence

We show that higher operating speeds of the devices may allow for the averaging of the tuning power required over many bits in order to achieve low energy per bit. Tuning of nanophotonic devices is essential due to the intrinsic temperature dependence of refractive index of solid-state materials, wafer level
variability, with run time operating temperature variability [39].
The total power including the tuning power for modulator and
detector wavelength selective devices can be written as

$$E_{\text{total}} = \frac{h\omega}{\eta_L\eta_D\eta_M\eta_C} \cdot \frac{V_C}{e} \cdot 10^{\frac{C}{10}} + \frac{V_m}{\theta}\frac{\Delta T}{\theta} + \frac{2}{B}P_{\text{tune}} \times \Delta \lambda \quad (4)$$

where we included the tuning power per nanometer of correction
$P_{\text{tune}}$ to correct the operating wavelength of the modulator and
detector by $\Delta \lambda$ and $B$ is the bit rate of the link. In Fig. 4, we
show the effect of the tuning power on the total interconnect
energy. The constant power penalty due to tuning will mandate
operation at higher speeds so that the tuning power can be shared
among more bits per second.

Higher operating speeds of interconnects will be necessary
to achieve an energy/bit below 100 fJ/bit/cm since the tuning
power imposes a significant constraint on the energy efficiency
of nanophotonic interconnects. As shown in Fig. 4, 100 fJ/bit
energy targets can be reached only at 40 Gb/s when a 2 nm
(20 °C) correction is required. The run time temperature control
for the microprocessors is expected to be 20 °C with a spatial
variation of 50 °C in temperature [39]. Hence, significant
advances, in temperature-independent device operation [40] or
highly efficient low overhead tuning schemes remain to be de-
developed [41], [42]. We note that packaging and module level
cooling may significantly change the tuning requirements.

C. Effect of On-Chip Serialize–Deserialize Operations

We show that efficient electrical serialize and deserialize op-
erations are essential to operate the optical links at higher op-
erating speeds. We obtain the optimum operation speeds of
the silicon optical interconnect by including the energy cost of
serialize–deserialize (SerDes) operations and the tuning power.

We modeled the power penalty for SerDes operations as a
constant energy per bit per serialization order. The total energy
of the link can be written as:

$$E_{\text{total}} = \frac{h\omega}{\eta_L\eta_D\eta_M\eta_C} \cdot \frac{V_C}{e} \cdot 10^{\frac{C}{10}} + \frac{V_m}{\theta}\frac{\Delta T}{\theta} + \frac{2}{B}P_{\text{tune}} \times \Delta \lambda + E_{\text{SD}} \frac{B}{2F_{\text{clock}}} \quad (5)$$

where $F_{\text{clock}}$ is the system clock and $E_{\text{SD}}$ is the energy-per-bit
per serialization order $N$. The SerDes are used for scaling the
bit rates beyond twice the system clock. The exact functional
form for the SerDes operations can be different; however, it is
commonly understood that the higher the bit rate and degree
of serialization, the larger the energy for serializing and dese-
erializing. In Fig. 5, we show the effect of SerDes power on the
total energy-per-bit. Some recent examples of optimization for
on-chip serial link SerDes are [52] and [53]. For a large SerDes
energy of 50 fJ/bit per serialization order, we see that the mini-
num of the energy is obtained when no serialization takes place
at 2 * $F_{\text{clock}}$ bit rate. However, for a lower SerDes energy (10
fJ/bit), the penalty due to SerDes is not significant enough to
change the behavior of the interconnect energy. The minimum
energy is then obtained when the interconnect is operated at
the maximum possible drive conditions. (See Appendix D for
SerDes energy scaling with CMOS technology node.)

We also study the effect of the system clock on the behav-
ior of the total interconnect energy per bit considering tuning
power, serialization as well as device insertion losses. The en-
ergy penalty due to serialization can be minimized by operating
at the highest available system clock. We also assumed that a
distributed clock is available throughout the chip. The clock
distribution from the local source to the SerDes is considered
local distribution and is ignored. We see that at a 5-GHz system
clock, with a SerDes power of 10 fJ/bit operation and a tuning
power of 100 μW/nm, 150 fJ/bit operation can be achieved for
all bit rates above 20 Gb/s.

D. Total Interconnect Energy Dependence on Length

We study the total optical interconnect energy as a function of
length including insertion losses, laser, modulator, and detector
efficiency in Fig. 6. Cross-over points of the optical interconnect
energy/bit versus generic interconnects with a fixed energy/unit area are shown in Fig. 6. A high energy/bit interconnect such as a 1 pJ/cm interconnect [43] (e.g., a full swing interconnect with a swing voltage of 0.68 V (ITRS 2011_ORTC-6, $V_{dd}$ for high performance) and capacitance of 140 aF/μm (ITRS Table 2011_INTC2, 2020) will have cross over points as low as a few millimeter. However, an energy-efficient interconnect with 100 fJ/cm [44], [59] will have a longer cross over point. It remains to be seen if the emerging electrical interconnects can meet the on-chip bit error rate and variability requirements [59], [60] given the high aggregated bandwidth of microprocessors [61]. We believe that given the number of interconnects and the aggregated bandwidth in the microprocessor application of interconnects, error correction will be limited due to latency area and power considerations.

IV. LINEAR INTERCONNECT BANDWIDTH DENSITY OF A NANOPHOTONIC INTERCONNECT $\beta$

Linear bandwidth density (LBD) of an interconnect is the bandwidth ($B$ in bits/μm·s) of an interconnect normalized for the width of the interconnect. The interconnect density on a microprocessor scales as the wire pitch of interconnects scale as per ITRS requirements.

The fundamental limit to optical interconnect density is greatly enhanced by the high central carrier frequency and the ability to multiplex a large number of wavelengths [45]. For a nanophotonic waveguide array comprised of waveguides of width $W$, separated in a pitch of $P$, the bandwidth density (per micron) can be written as

$$\beta_{\text{Optical}} = \frac{NB}{P} = NB \left( 0.12 \log_{e} \left( \frac{56.6L}{\pi} \right) \right)^{-1}$$

where $N$ is the number of wavelength division multiplexing (WDM) channels, $B$ is the single-channel bandwidth, $P$ is the waveguide pitch, and $L$ (in microns) is the crosstalk distance in microns. The pitch is the waveguide center to center pitch calculated for 250 nm (height) × 450 nm (width) waveguides such that a 3 dB coupling to the closest waveguide takes place for TE mode over a length of $L$ (in microns) [46]. Novel CAD methods and wavelength allocation methods to separate the waveguides can reduce the effective pitch. Note that unlike the electrical case, the optical bandwidth density is not a strong function of the length of propagation. The dispersion effects enter the analysis as a secondary effect over several meters of propagation [47] enabling 1 Tb/s on a waveguide using WDM [45], thus indicating bandwidth density limits exceeding $10^{12}$ bits/μm·s.

A. Length Dependence of Interconnect Linear Bandwidth Density

Typical ITRS projections for electrical interconnect density at intermediate lengths are in the order of 20–200 Gb/s·μm. Given the scaling trends for the intermediate wires from 76 (2011) to 24 nm (2020), the electrical wires will increasingly be limited in BW density for longer distances (100 to 500 μm, arising from electromagnetic interference, etc). Fig. 8 shows the LBD for optical WDM waveguides plotted with a benchmark 1 μm intermediate wire pitch at 100 Gb/s/μm. For > 150 Gb/s/μm over global/intermediate distances (up to cm), an 8 × 40 Gb/s WDM will be essential.

B. Considerations on Scaling the Number of Channels Using Microresonators

Here, we analyze two critical design considerations for scaling the bandwidth density using WDM: 1) the channel spacing and 2) the total number of channels set by cavity-free spectral range. We use first-order optical microring resonators as example resonators. We note that in general variety of microresonators and higher order designs can be employed. The
wavelength spacing between the resonators can be controlled by considering the effect of waveguide and material dispersion. The functional dependence of resonance position of the rings can be given by

$$\lambda_k = \frac{2\pi(r + \delta r(k)) n_{\text{eff}}(\lambda_k)}{r n_{\text{eff}}(\lambda_0)}$$  (7)

where \(\lambda_k\) is the position of the optical resonance of the \(k\)th microring, \(r\) is the radius of the base microring resonant, and \(\lambda_0\) is the radius perturbation introduced in the \(k\)th ring. We note that for a WDM microring bank spanning several tens of nm \(\delta r(k)\) will be a nonlinear spacing variation obtained by including the variation in \(n_{\text{eff}}(\lambda_0 + \delta \lambda, k)\) due to strong waveguide dispersion of high-index-contrast systems [45], waveguide bending and the material dispersion of the media. The channel spacing is also affected by the amplitude and phase cross talk due to off resonant interaction with the adjacent channels.

A second consideration is the free spectral range of the resonators to enable a large wavelength range for packing the WDM channels. The maximum number of channels that can be packed in a WDM system using microrings of radii \(r + \delta r(k)\) with uniformly spaced channels at spacing \(\delta \lambda\) is given by

$$N = \frac{\Delta}{\delta \lambda} = \frac{\lambda_0}{\delta \lambda} \left(1 - \frac{m n_{\text{eff}}(\lambda_0 + k\delta \lambda)}{n_{\text{eff}}(\lambda_0)}\right)$$  (8)

where \(\Delta\) is the free spectral range in wavelength, and \(m = 2\pi r n_{\text{eff}}(\lambda_0)/\lambda_0\) is the mode order for the base microring. For example a microring resonator of 1.5 \(\mu\)m radius can have an FSR of 62 nm allowing a large number of WDM channels [62]. One can see that a considerable design space is available using microresonators to meet the linear bandwidth density requirement.

V. SINGLE-CHANNEL BANDWIDTH OF A NANOPHOTONIC INTERCONNECT \(F\)

The limit to single-channel bandwidth is decided by the operation speed of the receiver and transmitter. The fundamental limits to the electro-optic device speed are given by free carrier response times [20]–[23], [36] or electro-optic material response time or the driving capacitor time constant [15]. For photodetectors and free carrier dispersion modulators

$$f_{\text{EO/DE}} < \frac{1}{\tau_{\text{min}}} = \frac{v_{\text{sat}}}{n w}$$  (9)

where \(v_{\text{sat}}\) is the saturation velocity of carriers in silicon (set by the optical phonon dispersion), typical values of \(\sim 10^7\) cm/s (for Si, Ge, and III-Vs), \(w = \lambda/15–103\) nm is space rate of decay of the evanescent field of the waveguide [50], and \(n\) is the arbitrary factor chosen such that \(e^{-n}\) gives the factor by which the evanescent field decays. The typical clearance for placing thin film planar doped regions next to nanophotonic waveguides can be estimated to be \(3\lambda/15–310\) nm.

The switching speed of a scaled electrooptic device driven by a scaled single-stage digital logic driver is [54]

$$f_{\text{Drive}} < f_{\text{EO/DE}} = \left(\frac{C_n V_n}{I_n} \left(3 \frac{I_{\text{mod}}}{I_n} + 1.5\right)\right)^{-1}$$  (10)

where \(C_n\), \(V_n\), and \(I_n\) are the capacitance, voltage, and current of a minimum sized transistor at a given technology node, respectively, and \(I_{\text{mod}}\) is the peak current through the modulator. We plot the maximum switching speed of the direct logic drive as a function of the drive current for the modulator in Fig. 9. Gate lengths, voltages, and delays are taken from ITRS HPC PIDS [1]. The voltage and current drive requirements for the EO devices, therefore, should be compatible with scaled CMOS for high-speed operation. The voltage and current drive requirements for the EO devices, therefore, should be compatible with scaled CMOS for the high-speed operation.

VI. AREAL BANDWIDTH DENSITY OF NANOPHOTONIC COMPONENTS \((D\) IN BITS/MM\(^2\))

Areal bandwidth density (ABD) of nanophotonic (transmitters/receivers) is the bandwidth generation/receiving capacity of components divided by the area of the device. The area taken by the wires and waveguides themselves is separately accounted for in the prior, interconnect bandwidth density metric. The role of ABD is to quantify the footprint taken by optical components to provide a certain bandwidth capacity. The modal volume of modulators as well as detectors enhanced by resonance effects is ultimately limited by diffraction limits

$$D_{\text{optical}} \equiv \frac{f}{\text{Area}} < \frac{v_{\text{sat}}}{n w} \left(\frac{\lambda}{2N}\right)^{-2}$$  (11)

where \(N\) is the index refractive index of the guiding medium. The density will have to be adjusted to allow for the driver and receiver circuits (as shown by the driver scaling in Section V). A 1.5-\(\mu\)m radius modulator operating at 10 Gb/s will reach bandwidth density of 1400 Tbit/s.mm\(^2\) [38]. Improved speed, 3-D integration, and ultrasmall modal volumes may be necessary for meeting the CMOS ABD requirements.

VII. DEVICE REQUIREMENTS FOR SCALABLE NANOPHOTONIC INTERCONNECTS

Based on the figures of merit proposed earlier, we present a minimal set of optical device requirements for replacing CMOS global interconnects. However, we note that specific device
requirements derived previously are for a single direct link and not a networked topology [6]–[9]. Four minimal features to enable optical components on-chip are as follows.

1) **High bandwidth, Broadband devices:** Higher speed of operation will allow large interconnect densities and offset the tuning power to reduce the energy/bit. Target speeds are in 10 to 40 Gb/s for modulators with switch bandwidths to allow switching of 40 Gb/s signals.

2) **Compactness:** The dimensions of modulator, detector, switches, and delays directly contribute to the areal density of interconnects and reduce the energy per bit. The target sizes of the modulators and detectors are less than 1 μm². ABD > 500 Tbit/mm²·s and footprint < 10 μm² are essential to meet the requirements of future interconnects.

3) **Multiwavelength:** Multiple wavelength operation is essential for the linear interconnect density scaling. WDM is ideally suited for an on-chip optical interconnect due to complexity, footprint, and optical insertion loss considerations.

4) **CMOS Compatibility:** The modulators, detectors, and switches must operate with available voltage and current requirements of digital CMOS. Compatibility in drive currents and voltages must be ensured so that future technology nodes may allow for direct logic drive operation of the interconnect components.

**VIII. CONCLUSION**

We introduce an analytical framework for scaling nanophotonic interconnects to meet the energy and footprint requirements of CMOS global interconnects. We emphasize that the goal of this paper is to lay out a framework for a scaling path for optical devices and not provide a direct comparison with the several emerging promising technologies such as LSI modulation. The adoption of any of the emerging technologies including photonic interconnects depends not only on the aforementioned figures of merits but on a combination of the HPC computing requirements, activity factors, cost, robustness to variations, and noise margins. The following conclusions can be drawn for the photonic technology scaling requirements for CMOS global interconnects.

1) **Scaling link bandwidth to 40 Gb/s and beyond can enable competitive energy/bit and areal bandwidth density. Improvement in link speed must be accompanied by improvement in SerDes operation.**

2) **Scaling the operational voltages of all electro-optics (<0.6 V) to follow the CMOS voltage scaling is desirable.**

3) **Scaling the number of wavelengths per waveguide is essential to meet the linear bandwidth density of the global interconnects.**

4) **Fundamental limitations to the compactness of the optical devices may mandate 3-D integration. If a viable 3-D
integration scheme does emerge, the photonic device layer may be unconstrained in area.

5) Improvement in thermal stability of the electro-optic detectors and modulators and passive elements beyond 10 μW/K is essential for stable operation of the links. The goal is to provide the performance with no change in the module level thermal management.

6) High conversion efficiency lasers (>25%) and low insertion loss (<8 dB) modulation, wave guiding, and detection schemes are essential for low energy/bit operation.

With the appropriate scaling of device performance, photonic CMOS for on-chip interconnects may emerge as a technology for high-performance computing applications in the CMOS/beyond CMOS era.

APPENDIX A
DERIVING OPTICAL LINK ENERGY

The energy-per-bit of \( E_{\text{Source,Detector}} \) can be derived as follows. At the detector end, the charge through the detector for 1 ON bit (and current) is given by

\[
Q_{\text{injected}} = C_d V_r, \quad i_{\text{detector}} = C_d V_r B. \tag{A1}
\]

The incident optical energy at the detector can be written as

\[
E = \frac{P_{\text{detector}}}{B} = \frac{\hbar \omega}{\eta_d} \frac{C_d V_r}{e}, \tag{A2}
\]

which gives the total electrical energy as

\[
E_{\text{Source,Detector}} > \frac{\hbar \omega}{\eta_L \eta_D \eta_M \eta_C} \frac{V_r C_d}{e} \cdot 10^{\frac{-4}{\eta}}. \tag{A3}
\]

APPENDIX B
BIT ERROR CONSTRAINTS AT THE RECEIVER

For an \( N \) node interconnect network operating at frequency \( f \), the tolerable error rate \( P_{\text{req}} \) for operating with a failure rate of \( R \) over time \( T \) is [60]

\[
P_{\text{req}} < \frac{R}{N f T}. \tag{B1}
\]

For 10 000 on-chip global interconnects operating at 5 GHz with a failure rate of \( 10^{-6} \) over a lifetime of ten years, the required error rate is \( 6.3 \times 10^{-29} \).

Following Beausoli et al. [37], the mean number of photons required in an ON pulse for an error rate of \( P \), for a modulation depth (1 - \( M \)), and for the OFF state of the light pulse is

\[
n_{\text{min}} \approx -2 \ln P \left( 2 - M + 2 \sqrt{1 - M - \frac{M^2}{2 \ln P - \frac{2kTC_d}{e^2}} \right) \tag{B2}
\]

where \( \eta \) is the total quantum efficiency of the detector. For a target error rate of \( 10^{-29} \), this corresponds to 823 collected photons per “ON” pulse at a detector capacitance 1 fF; modulation depth of \( M = 0.9 \) (extinction ratio = \(-10\log_{10}(1-M) = 10 \) dB).

The effect of modulation depth on the required optical power at the receiver (for a 40-Gbit/s signal) is shown in Fig. 10. The minimum number of collected photons required at the given extinction ratio is also shown. Under the assumption of full charging of the detection capacitor (i.e., collected photons = \( C_d V_r/e = 6240 \)), we can see that the tolerable extinction ratio at the receiver is 1.4 dB. Hence, the degradation of the modulated optical signal due to insertion loss should not affect the BER for low insertion losses (<8 dB). For the analysis of this paper, we assumed that the modulators are maintained at optimal modulation depth using a tuning mechanism. We note that the previously received optical power is a lower limit for a receiver less detector. The degradation of SNR due to a TIA has to be accounted for in a receiver-based system [63].

APPENDIX C
EO MODULATOR ENERGY FOR ELECTRO-OPTIC POLYMER MODULATORS

A second common class of modulators compatible with CMOS is electro-optic polymer modulators [15]. The scaling with electro-optic properties for such modulators is as follows:

\[
E_{\text{EO}} > C_m V_m^2 \left( \frac{\chi}{\Delta T} \right)^2 \chi^2 \tag{C1}
\]

where \( C_m \) is the modulator capacitance, \( \Delta T \) is the modulation depth at the modulator, and \( \chi \) is the voltage electro-optic coefficient. The square law dependence with \( \chi \) and \( \Delta T \) are in contrast with carrier injection modulators.

APPENDIX D
SCALING ESTIMATE FOR SERDES ENERGY

We arrived at an energy/bit/\( N \) (\( N \) = order of the SerDes multiplexing) scaling estimate assuming equal time performance at a given node. For equal time response, the ratio of the total channel width of the SerDes circuit is (for 32 nm CMOS versus 11 nm CMOS)

\[
r_w = \frac{\sum W_{11}}{\sum W_{32}} = \frac{C_{g11} V_{11} J_{32}}{C_{g32} V_{32} J_{11}} \approx 0.2696. \tag{D1}
\]

The ratio of the energy/bit/\( N \) can be estimated as

\[
r_E = \frac{C_{g11} V_{11} J_{32}}{C_{g32} V_{32} J_{11}} \sum W_{11} \approx 0.0797. \tag{D2}
\]

Equations D1 and D2 use the values from ITRS 2011, PIDS2 HP CMOS Table III [1].
TABLE III

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>32 nm</th>
<th>111 nm (MG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_g$ (fF/µm)</td>
<td>Ideal Gate Capacitance</td>
<td>$C_{g32} = 0.658$</td>
<td>$C_{g11} = 0.338$</td>
</tr>
<tr>
<td>V (V)</td>
<td>HP Power supply</td>
<td>$V_{32} = 0.87$</td>
<td>$V_{11} = 0.66$</td>
</tr>
<tr>
<td>J (µA/µm)</td>
<td>NMOS drive current</td>
<td>$J_{32} = 1367$</td>
<td>$J_{11} = 1976$</td>
</tr>
</tbody>
</table>

Fig. 11. Effect of insertion loss on the energy/bit.

The estimated SerDes power at 32 nm under a global on-chip synchronous clock without clock recovery is 27 fJ/bit/N [53]. Using the projected scaling ratio of 0.0797, at 11 nm node the estimated energy/bit/order is 2.16 fJ/bit/N (nonideal gate capacitance as predicted by ITRS increases this projected value to 3.35 fJ/bit/N). To study the effect of the SerDes, we have included a wide range of energy estimates of 100 fJ/bit/N to 10 fJ/bit/N in this paper.

APPENDIX E

EFFECT OF INSERTION LOSSES

The energy/bit of the optical interconnect is affected by the insertion losses due to the passive and active optical components. The insertion losses may arise from nonresonant modulator loss, mux, demux filters, and waveguide crossing losses. The change in energy/bit due to total insertion losses is shown in Fig. 11. Insertion losses can also play a major role if the degradation in extinction ratio at the detector reduces the received extinction ratio at the detector below the threshold for high bit error rate. For example, in Appendix D, if the extinction ratio (of the received bits) reduces below 1.4 dB due to insertion loss, the interconnect will be BER limited. (For a modulator ER of 10 dB, this places an 8.6 dB limit on IL.)

REFERENCES

Sasikanth Manipatruni (M’07) received the Graduation degree in electronic engineering from the Indian Institute of Technology, Delhi, India, and the Ph.D. degree in silicon photonics from Cornell University, Ithaca, New York, USA.

He is a Research Scientist in the Exploratory Integrated Devices and Circuits group in Intel Components Research, Hillsboro, OR, USA. He is involved in emerging novel devices to identify beyond CMOS logic technologies. During his Ph.D., he was a co-originator of several silicon photonics components devices including the first 18 GHzs microring modulator, first GHZ polysilicon modulator, EO switches, and silicon nanoporous links. He was a national science fellow Kishore Vaigyanik Protsahan Yojana of the Indian Institute of Sciences (1999–2001); he was at Swiss Federal Institute of Technology, Zurich, Switzerland, in 2004 and Inter University Center for astronomy and astrophysics in 2001. He has 20 patent applications in nanophonics, MRI, spintronics and 50 peer-reviewed journal and conference papers.

Dr. Manipatruni serves as a Peer Reviewer for the OSA, IEEE, and Nature Photonics.

Michal Lipson (SM’07–F’13) received the B.S., M.S., and Ph.D. degrees in physics in the Technion-Israel Institute of Technology, Haifa, Israel.

In December 1998, she joined the Department of Material Science and Engineering, Massachusetts Institute of Technology (MIT) as a Postdoctoral Associate. In 2001, she joined the School of Electrical and Computer Engineering, Cornell University, Ithaca, New York, USA, where she is currently an Associate Professor. She is the inventor of 12 patents regarding novel micron-size photonic structures for light manipulation. She is the author or coauthor of more than 100 papers in the major research journals in physics and optics. Her research interest at Cornell includes novel on-chip nanophotonic devices.

Dr. Lipson is the McArthur Fellow and a Fellow of the Optical Society of America. She received the National Science Foundation CAREER Award, IBM Faculty Award, and Blavatnik award, NY state academy of science.