Addressing Link-Level Design Tradeoffs for Integrated Photonic Interconnects

Michael Georgas, Jonathan Leu, Benjamin Moss, Chen Sun, and Vladimir Stojanović
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139
Email: mgeorgas at mit.edu

***Abstract*—Integrated photonic interconnects have emerged recently as a potential solution for relieving on-chip and chip-to-chip bandwidth bottlenecks for next-generation many-core processors. To help bridge the gap between device and circuit/system designers, and aid in understanding of inherent photonic link tradeoffs, we present a set of link component models for performing interconnect design-space exploration connected to the underlying device and circuit technology. To compensate for process and thermal-induced ring resonator mismatches, we take advantage of device and circuit characteristics to propose an efficient ring tuning solution. Finally, we perform optimization of a wavelength-division-multiplexed link, demonstrating the link-level interactions between components in achieving the optimal degree of parallelism and energy-efficiency.**

I. Introduction

As high performance computing continues to advance towards the many-core regime, it is clear that future processors will require ever-increasing bandwidth from the on-chip interconnect network and off-chip interfaces. These interconnect fabrics already occupy large portions of the chip area and consume a significant fraction of the total power in current generation processors [1]. Furthermore, projections show poor scaling of on-chip wires and I/O bandwidth density with technology [2] [3], highlighting the need for a disruptive interconnect technology to meet the throughput demands and power-efficiency requirements of many-core systems.

Integrated silicon photonics is an emerging technology that demonstrates significant advantages over traditional electrical links for on-chip core-to-core [4] [5] and off-chip core-to-DRAM [6] [3] [7] applications. Of its most salient features, dense wavelength-division-multiplexing (WDM) and distance-insensitive energy-per-bit stand out as key enablers for a faster, denser, and more energy-efficient interconnect fabric.

To realize these potentials, it is essential for device, circuit and system designers to understand the relationships and trade-offs among components in this new technology, as well as the impact of the potential integration scenarios. In this paper, we illustrate these design trade-offs on an example of a complete integrated WDM photonic link, by creating the component models that connect device, process and circuits parameters to link performance and power consumption. Identifying the tuning power as one of the most-significant costs, we propose a robust ring-tuning solution requiring only a fraction of traditional tuning power. We show that careful design-space exploration coupled with new tuning techniques can result in an optimal set of link, circuit, and device parameters. We demonstrate that the choice of per-wavelength data-rate allows us to achieve minimum energy-costs by balancing laser, tuning and circuits backend power. This exploration also highlights the significance of monolithic-integration to minimize the receiver parasitics, reducing the laser power and enabling low-energy, high throughput-density interconnect fabrics.

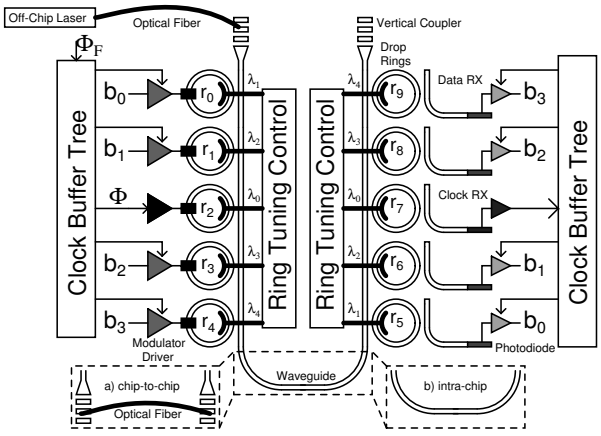


Fig. 1. An integrated WDM photonic link. A continuous-wave (CW) multi-λ laser is coupled onto the chip through a vertical-coupling grating structure. Once on chip, frequency selective ring-resonant modulators encode digital bitstreams onto their resonant wavelengths. Each wavelength propagates along the waveguide (and possibly off-chip) until it is routed through a matching drop ring to an integrated photodiode (PD). An optical receiver forms a bit decision based upon the PD photocurrent. Clock signals are routed both optically along the waveguide and electrically through local H-trees. Ring tuning circuits are used to tune the resonance of the modulator and drop rings.

II. Photonic Link Components

We begin with a discussion concerning the operation of modulators and receivers, the primary data-path elements that form a WDM photonic link (Figure 1). As these two components depend highly on the characteristics of the devices that they use, we motivate a device-technology-driven analysis. We consider two integration scenarios - a monolithic integration of photonic components into the CMOS front-end (polysilicon photonics or thin-BOX SOI photonics), and a hybrid integration scenario with optimized SOI photonic die attached to the CMOS chip via through-silicon vias (TSVs).

978-1-4577-0223-5/11/$26.00 ©2011 IEEE

The monolithic integration will typically have smaller parasitic capacitances between circuits and photonic components, while potentially having higher optical lossess due to fabrication process constraints.

A. Modulator

The optical ring-resonant modulator and driver convert electrical data into the optical domain by on-off keying one CW λ from the multi-λ laser source. Light is modulated by shifting the ring-resonant filter's stopband in and out of the optical channel's wavelength. The stopband is shifted most efficiently using the free-carrier plasma dispersion effect [8] to change the refractive index of the ring material. To avoid the high energy cost of carrier-injection modulators, which have a high on-current due to carrier recombination [9], we focus on reverse-bias driver designs that modulate the depletion width of a vertical P-N junction fabricated in the polysilicon/silicon ring [10].

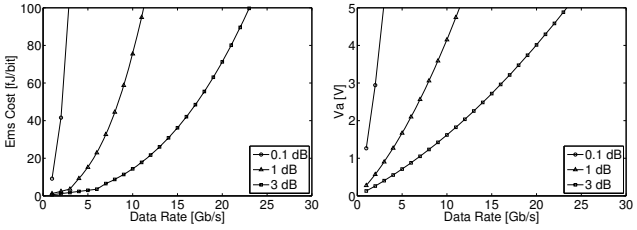
1) Modulator Design: A key system tradeoff exists between the extinction ratio ER of the modulator (a ratio of the on-to-off light intensity), its insertion loss IL, and its total energy cost. A small shift in the ring's Lorentzian frequency response requires a small energy cost, but results in a low ER. The location of the CW laser resonant wavelength with respect to the resonance of the ring is set by the desired IL, ER and data rate DR (bandwidth) requirements, which then set the modulation energy cost. The system designer must balance the modulation energy cost with receiver and laser power, which depend strongly on the extinction ratio, insertion loss and data rate.

2) Device Energy: In this section we calculate the modulation energy required to achieve given specifications. The necessary charge difference between the on- and off-states, ΔQ, is determined in Equation 1 by evaluating the Lorenzian transfer function, where $T_1 = 1/IL$ is the ring's transmissivity in its shifted on-state, $T_0 = 1/(ER{\cdot}IL)$ is its transmissivity in its off-state, and T_n is its transmissivity at the ring's resonant wavelength, Figure 3. Q_0 is the charge difference necessary to shift the ring by its half-width-half-max bandwidth [9], where q is the charge of an electron, $n_g = 4$ is the group index of the ring, and some typical ring modulator parameters are set as $V_{tot} = 2.6 \times 10^{-12}$ cm^3 is the total volume of the ring, $n_f = 3 \times 10^{-21}$ cm^3 is the carrier-induced index change per unit carrier density at $\lambda_0 = 1300$ nm, and $\Gamma = 0.4$ is the overlap of the optical mode with the ring cross-section. The quality factor $\mathbb{Q}_f$ of the ring is set by the required data rate as $\mathbb{Q}_f = \frac{8{\cdot}\pi}{3\lambda_0{\cdot}DR}$, with a maximum value of $1{\times}10^5$, limited by practically achievable optical losses.

$$\Delta Q = Q_0{\cdot}\left(\sqrt{\frac{T_1 - T_n}{1 - T_1}} - \sqrt{\frac{T_0 - T_n}{1 - T_0}}\right), Q_0 = \frac{q \cdot n_g \cdot V_{tot}}{2 \cdot \mathbb{Q}_f \cdot n_f \cdot \Gamma} \quad (1)$$

The modulator diode operates as a varactor in the reverse-bias regime. The required charge difference is integrated onto the nonlinear junction capacitance (Equation 2) to determine the minimum reverse-bias drive voltage. V_a is plotted in Figure 2b as a function of data rate, at fixed ER and various values of IL.

$$\Delta Q = Q(V_a) - Q(0) = \int_0^{V_a} \frac{C_{j0}}{\sqrt{1 + \frac{V}{V_{bi}}}} dV \quad (2)$$



(a) Device energy cost from supply. (b) Reverse bias voltage V_a.

Fig. 2. Device requirements to reach the target $ER_{dB} = 6$dB for various values of IL_{dB}.

If V_a is less than the supply V_{DD}, the energy to charge the junction comes from V_{DD} and the energy-per-bit is $E_{ms} = \Delta Q \cdot V_{DD}/4$ assuming a random data pattern. Otherwise the energy must come from a higher-voltage source, which we assume is generated from the supply with a conversion efficiency of $\eta = 0.5$. With $V_a > V_{DD}$, the energy-per-bit drawn from the supply is $E_{ms} = \Delta Q{\cdot}V_a{\cdot}(1{+}\eta)/4$ (Figure 2a).

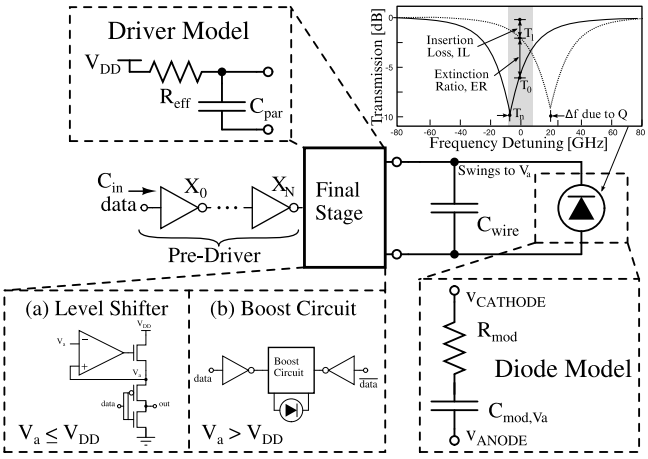
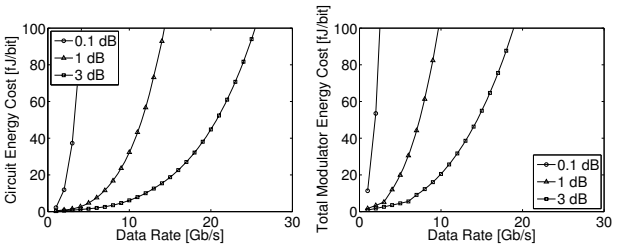


Fig. 3. Electrical model of modulator device and driver.

3) Circuit Energy: The driver model is shown in Figure 3 as an inverter chain pre-driver followed by a final driver stage. The circuit topology of the final drive stage will change based on V_a; If $V_a \leq V_{DD}$, a low-swing topology can be used (Figure 3a); otherwise a voltage-boosting circuit may be necessary (Figure 3b). The final stage is modeled as an effective resistance R_{eff} and a parasitic capacitance C_{par}, connected to wiring capacitance C_{wire}. Logical effort analysis is used to size the driver (W) and pre-driver chains (FO) to meet the data-rate requirements.

$$E_{dr} = \frac{C_{par} + C_{wire}}{4}{\cdot}\max\left(V_a \cdot V_{DD}, V_a^2\right) + \frac{1}{4}{\cdot}\frac{3 \cdot C_g \cdot W}{1 - \frac{1}{FO}}{\cdot}V_{DD}^2 \quad (3)$$

Figure 4 shows the driver energy-per-bit cost to reach $ER =$ 6 dB for various values of IL. For current modulator device



(a) Driver energy cost E_{dr}. (b) Total modulator energy cost $E_{mod,tot} = E_{ms} + E_{dr}$.

Fig. 4. Energy costs to reach the target $ER_{dB} = 6$ dB for various values of IL_{dB}.

technology, to satisfy data rates up to 30 Gb/s, the intrinsic RC time constant of the device allows designers to choose R_{mod} up to 1 kΩ without significant energy penalty, relaxing the modulator optical losses due to contact placement.

B. Optical Data Receiver

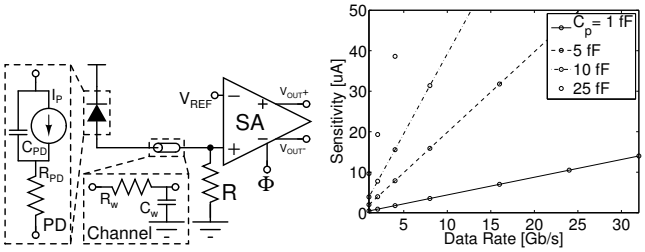
The optical receiver converts optically-modulated data back into the electrical domain by sensing a photocurrent produced by the PD. In contrast to traditional optical receivers which utilize various power-hungry trans-impedance amplifiers (TIA) to combat the large PD parasitic capacitance, monolithic integration offers the opportunity for much-simpler, energy-efficient receiver circuits due to low PD parasitic capacitances. In this section, we illustrate the relationship between sensitivity and power consumption across ranges of data rates and parasitic capacitances, for various receiver topologies, including transimpedance amplifiers (TIA) and integrating receivers.

1) Photodiode: An equivalent model of the PD is shown in Figure 5a, consisting of a capacitance in parallel with a photocurrent-generating source and series resistance. The PD is connected to the front-end through either a Through-Silicon Via (TSV, $C_p \approx 25$ fF) or low-level metal routing ($C_p \approx 5$ fF). All parasitic series resistances are assumed negligible.

2) Sense Amplifier: A regenerative sense-amplifier (SA), known for energy-efficient and scalable operation, creates a full-voltage-swing interface with the digital back-end.

$$d_{min} = v_{sense} + v_{OS,res} + \frac{v_{supply,det}}{CMRR} + v_{margin} + Q^{-1}(BER)\sqrt{v_{noise}^2 + \frac{v_{supply,rand}^2}{CMRR^2}} \quad (4)$$

Equation 4 shows the SA's input swing requirement. The minimum input signal that allows the latch's decision nodes to settle is given by $v_{sense} = V_{DD}e^{-T_{bit}/2\tau}$, with τ measured in simulation. Residue intrinsic offset due to mismatch is compensated by a 5 bit compensation DAC [11], resulting in $V_{OS,res} = 3V_{OS}/2^5$, with V_{OS} estimated from [12]. Deterministic supply noise, $v_{supply,det}$ is assumed as 50 mV, while random noise, $v_{supply,rand}$, is 10 mV [13] [14]. Supply noise is divided by a common-mode rejection ratio of 5. Front-end noise, such as thermal and PD shot noise, is amplified across the front-end and added to the SA input noise in the v_{noise}^2. Finally, v_{margin} accounts for all other un-modeled noise. SA power at 1 Gb/s is approximately 8 μW and is assumed to scale linearly with frequency for the data range up to 32 Gb/s for a given 32 nm technology node.



(a) Resistive optical receiver with PD and channel models shown. (b) Photocurrent Sensitivity.

Fig. 5. Resistive receiver sensitivity.

3) Resistive Receiver: Figure 5a shows a resistive receiver with a SA. Photocurrent is driven across the resistance, R, which is the front-end's gain. The dominant pole is at the input node. Equation 5 (with $R_f = R_{in} = R$) shows that the resistor is penalized for its parasitic capacitance through the parameter k_R, assumed to be 0.4 fF/kΩ [15]. For each data rate and $C_p = C_{PD} + C_w + C_{front-end}$, the maximum R is computed from Equation 5. BER requirements result in a minimum $\Delta I = I_{ON} - I_{OFF}$. The input sensitivity is then $I_{ON} = \Delta I/(1 - 10^{-ER/10})$.

Figure 5b shows that the receiver is able to sense photocurrents of approximately 10 μA for low C_p and data rate. The sensitivity worsens linearly with data rate as R is traded for bandwith. The energy-efficiency remains constant due to the dominance of the SA's digital switching power.

4) TIA: A TIA (Figure 6) breaks the gain-bandwidth limitation of the resistive receiver. Equation 6 shows the receiver's gain and use of feedback to decrease input impedance [16].

$$BW = \frac{1}{2\pi R_{in}\left(C_{PD} + k_R R_f\right)} \quad (5)$$

$$R_{TIA} = \frac{g_m - g_f}{g_f\left(g_m + g_{ds}\right)}, R_{in} = \frac{g_{ds} + g_f}{g_f\left(g_m + g_{ds}\right)} \quad (6)$$

Figure 7a shows sensitivity-optimized designs for different TIA bias powers. In this relatively small C_p environment, large designs are penalized for their increased gate capacitance, requiring a reduction in R_{in} and therefore R_{TIA} and sensitivity. Figure 7b shows sensitivity optimum for various values of v_{margin}.

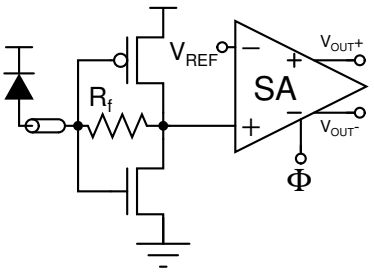
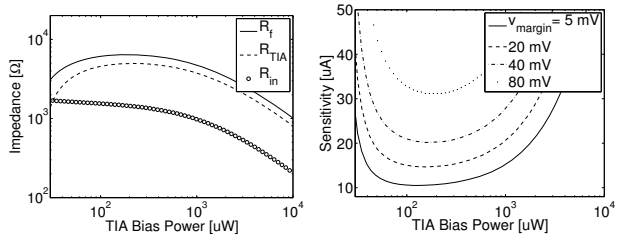


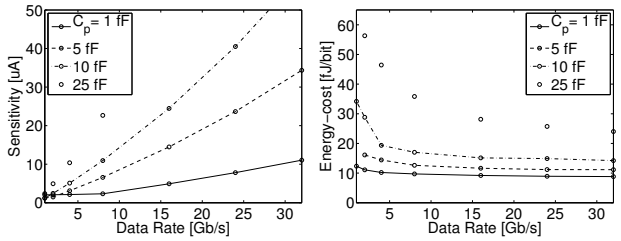
Fig. 6. Transimpedance Amplifier.



(a) v_{margin}=20 mV. (b) Effect of v_{margin}.

Fig. 7. TIA design example at C_p=25 fF, DR=5 Gb/s.

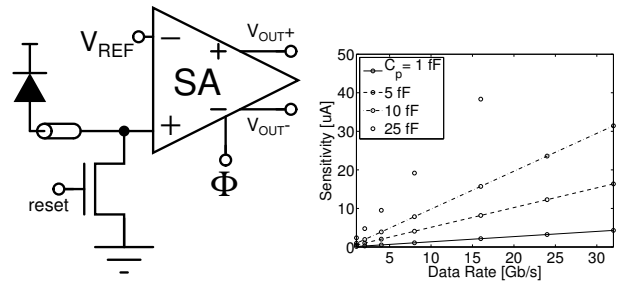
Figure 8 summarizes TIA performance for various values of C_p. Though the TIA achieves sensitivity superior to the resistive receiver, the power consumption is considerably worse.



(a) Photocurrent sensitivity. (b) Energy efficiency.

Fig. 8. TIA performance.

5) Current-Integrating Receiver: The third topology considered is an integrating receiver (Figure 9a), where the photocurrent is converted to a voltage by integrating it onto a capacitor $C_{INT} = C_{PD} + C_w + C_{SA,in}$. The photocurrent is integrated over a fraction (k_{INT}=0.7) of a bit time yielding a front-end gain given by Equation 7.



(a) Integrating receiver. (b) Photocurrent sensitivity.

Fig. 9. Integrating receiver design and performance.

$$R_{INT} = \frac{k_{INT} \cdot T_{bit}}{C_{INT}} \quad (7)$$

Figure 9 shows that the integrating receiver is the best performing of the three receivers considered. The energy-efficiency of the receiver is dominated by the SA as in the resistive receiver. It should be noted that this simple model has several hidden challenges remaining. The voltage on C_{INT} must be reset or at least charge-shared [17], which is partially accounted for through k_{INT}. A small C_{INT} will also suffer from SA kickback, while increasing C_{INT} degrades sensitivity.

C. Optical Clock Receiver

Clock distribution is critical in synchronizing communication channels and functional blocks in high-performance processors. The simplest optical clock receiver considered is a receiverless clocking scheme [18]. By alternately illuminating two PDs stacked in series, a clock signal is generated at the internal node. The only circuitry between the PDs and the clocked node are buffers, minimizing added jitter and reducing circuit power consumption. However, a large optical power is required to create the rail-to-rail voltage swing at the PD. As shown in Section II-B, differential TIAs can be used to amplify the signal at the cost of added noise and circuit power [16] [19].

In addition to bandwidth and sensitivity constraints, a clock receiver output must meet a given jitter specification. While C_p is relatively small, it is still much larger than any circuit loading. The voltage transient slope is then approximated by C_{PD}/I_{ON}. Any transistor or power supply noise is input referred onto this slope, yielding a timing jitter.

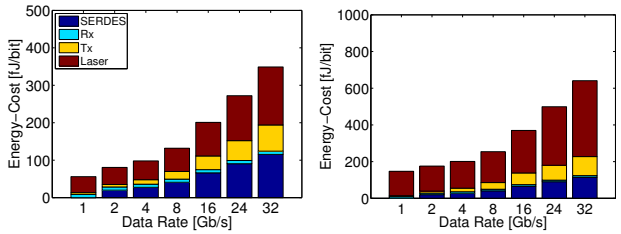
D. Single Channel Link Tradeoffs

To illustrate the interactions between the modulator and receiver and the impact on wall-plug laser power, we perform a power optimization across modulator insertion loss, extinction ratio, and receiver topologies for different link datarates. Figure 10 shows the energy-per-bit breakdowns for four integration scenarios.

TABLE I
LINK EVALUATION PARAMETERS

Parameter	Value
Process Node	32 nm Bulk CMOS
V_{DD}	1.0 V
Device to Circuit Parasitic Cap C_P	5-25 fF
Wavelength Band λ_0	1300 nm
Photodiode Responsivity	1.1 A/W
Wall-plug Laser Efficiency P_{laser}/P_{elec}	0.3
Channel Loss	10-15 dB
Insertion Loss IL_{dB} (Optimized)	0.05-5.0 dB
Extinction Ratio ER_{dB} (Optimized)	0.01-10 dB
Bit Error Rate (BER)	10^{-15}
Core Frequency	1 GHz
SERDES Topology	Mux/Demux Tree

In all plots, the laser power is the dominant energy consumer, increasing quickly with data rate as aggressive modulation rates force a relaxation of modulator insertion loss and extinction ratio. We can see that the laser power is highly sensitive to C_P, as the laser power for C_P=25 fF is roughly 5X that of C_P=5 fF. Though the modulator tries to offset the laser cost by increasing its extinction ratio and decreasing insertion loss, it inevitably reaches a limit on its capabilities. The higher loss simply amplifies the laser power component, resulting in a 3X laser power difference between the 15 dB and 10 dB loss cases. Matching previous analysis, the optimization chose the integrating receiver as the optimal receiver in all scenarios. Though our results present a grim outlook for the C_P=25 fF (optical die with TSV) scenario, we note that lower losses



(a) Loss=10 dB, C_P=5 fF (b) Loss=15 dB, C_P=5 fF

(c) Loss=10 dB, C_P=25 fF (d) Loss=15 dB, C_P=25 fF

Fig. 10. Data rate tradeoffs for a single photonic link for 4 integration scenarios. C_P=5 fF represents monolithic integration, while C_P=25 fF is expected for a TSV connection to an optical die. Channel losses of 10 dB and 15 dB correspond to on-chip and chip-to-chip links, respectively.

may be achievable with a dedicated optical die, allowing TSV integration to remain competitive.

III. Towards a Full WDM Link

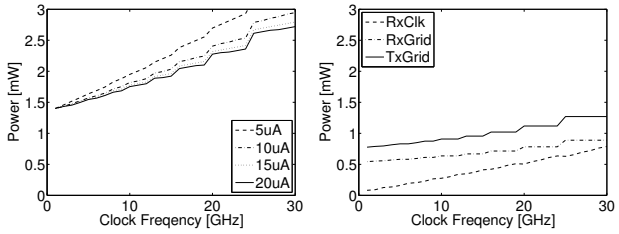
Expanding upon our analysis for a single-channel data link, we explore the additional backend components required in a high-speed multi-channel WDM link. We present a model of optical clock distribution and source-synchronous clocking. Then, we outline techniques for tackling resonance mismatches of optical ring resonators, a key challenge in nanophotonic integration.

A. Optical Clock Distribution

An example point-to-point optical WDM source-forwarded link is shown in Figure 1. Clock transmission occurs on λ_0. At the differential receiver, the signal is regenerated, buffered, and used to clock the other data receivers. Since clock-TX and data-TX share the same clock fabric, relative jitter between the sent clock and data is minimal.

On top of previously discussed benefits of optical links, optical clock signaling does not suffer power from rail injected noise or crosstalk, so no jitter is added in the channel. This obviates the need for an RX PLL/DLL, greatly reducing the power and area overhead. The low latency of optical waveguides also means that all data receivers can operate on the same clock phase.

By accounting for the total clock load at the TX or RX, including wire routing, the power consumption of the clock distribution is modeled for a given clock frequency. The expected timing jitter can be derived from [20]. In Figure 11a we fix C_{PD} at 5 fF, and plot the total clock power across frequency, with the criteria that the timing jitter at the data receivers is less than 3% unit interval (UI). With higher clock frequency, fewer channels are needed for a given total data throughput, decreasing the distribution endpoint capacitance. As we increase the frequency, the jitter requirement (in seconds) tightens, requiring an increase in power. When I_{ON} increases, the jitter performance of the receiver improves proportionally, allowing for less electrical power. Figure 11b shows the power consumption of the TX clock tree, RX clock tree and clock receiver circuit for I_{ON}=10 μA and jitter of 3 % UI.



(a) Power across Ion (b) Power Breakdown at Ion=10uA

Fig. 11. Power vs. datarate per channel, jitter is fixed to be within 3% UI, and C_{PD}=5fF

B. Ring Resonance Mismatch

An integrated WDM link relies heavily upon optical ring resonators to perform channel selection using the ring's resonant frequency. Dependent upon both device geometry and the index of refraction, large ring resonance mismatches can arise from limited process tolerances and temperature changes. For rings built with gate poly-silicon on commercial CMOS bulk processes, process variation can result in resonance mismatches of up to 90 GHz and absolute die-to-die mismatches of 600 GHz or more [21]. Similarly, mismatches with standard deviations in the range of 20-70 GHz for same-die and 150-220 GHz for die-to-die have been observed for rings built with SOI [22]. As local- and systematic-level process mismatches differ greatly in magnitude and tuning implications, we model them using σ_{rL} and σ_{rS}, corresponding to the standard deviations characteristic of local ring-to-ring and global systematic mismatches, respectively.

A strong thermal dependence in the index of refraction of silicon causes ring resonances to drift with temperature. $\frac{\Delta f}{\Delta T}$ in the range of -10 GHz/K have been observed [21] [23], implying that a shift of several hundred GHz can be expected in a hostile thermal environment, such as that of a high-performance processor. Unlike static process variations, however, thermal fluctuations are time-dependent, requiring active tuning to stabilize ring resonances. At the same time, strong temperature dependence allows for simple and effective thermal compensation of process mismatches. Recently, *athermal* ring resonators using polymer-based cladding [24] have also been proposed as a solution to undesired thermal-induced resonance drifts. Their inability to be thermally tuned, however, means that any process-induced mismatch must be compensated by UV trimming on a per ring basis, potentially limiting commercial scalability.

C. Ring Tuning Techniques

To mitigate mismatch introduced by process and temperature, we present several strategies for ring tuning. As an example, we consider the problem of tuning a set of receive-side rings to a set of WDM channel frequencies placed at fixed

frequency intervals (Figure 12). Given that ring resonances repeat (with separation between peaks defined as the ring's free spectral range, or FSR), we require that all channel frequencies fit within one FSR.

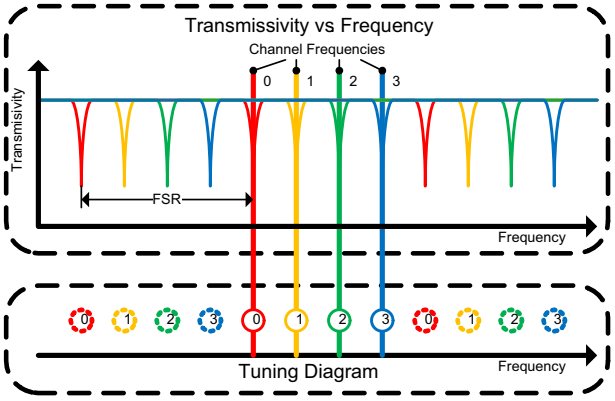


Fig. 12. Mapping between a frequency transmissivity vs frequency to an equivalent tuning diagram for a set of perfectly aligned rings. Different WDM Channels are depicted by color and constitute the 4 vertical lines. There are only 4 rings present in this picture, as each ring's resonance repeats every FSR. In the tuning diagram, these extra resonances are depicted as dashed circles.

To tune ring resonances, resistive heaters are fabricated alongside each ring for thermal control, and are driven by a relatively low-bandwidth, receive-data driven control loop. Though large resonance shifts can be achieved, this straightforward approach comes at a steep power cost – the inability to cool down implies that a large fabrication frequency bias must be applied to the rings such that the resonant frequency of each ring remains greater than the corresponding channel frequency at any operating temperature and process corner (Figure 13). Given high uncertainty in absolute ring resonances due to process variations, this bias may need to be as large as 1 THz, requiring temperature increases of 100 K or more. Though post-process steps such as undercut [21] can increase thermal isolation (increasing heating efficiency), additional problems such as thermal cross-talk amongst rings make such large temperature ranges impractical and power-inefficient.

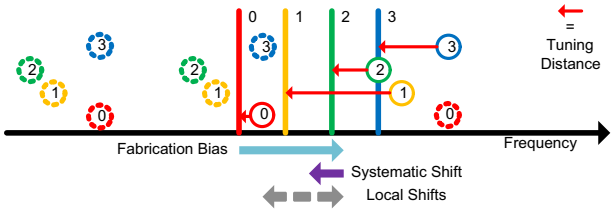


Fig. 13. Tuning ranges of full thermal tuning for a set of rings affected by systematic and local variations. A large fabrication bias is applied to keep ring resonances above channel frequencies.

In light of this, we propose an electrical backend capable of bit re-shuffling (Figure 16). As opposed to tuning to an assigned channel (as in the full-thermal case), rings simply tune to the nearest channel with the electrical backend performing all necessary bit-ordering operations. The barrel-shifter compensates for systematic process and temperature shifts common to all rings in the set by exploiting ring

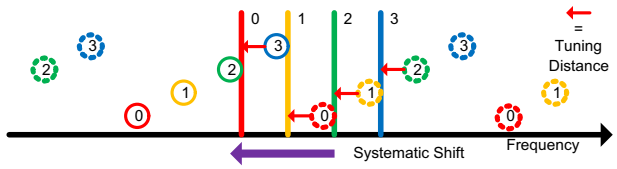


Fig. 14. When a systematic variation shifts all ring resonances by some frequency, we can still allow rings to align themselves with the nearest channel and electrically barrel-shift to reposition the bits using the backend. Note that we have wrapped around the next resonance for rings 0, 1 and 2.

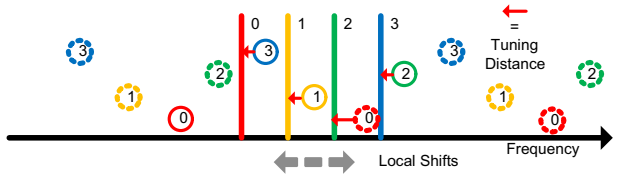


Fig. 15. Bit re-ordering multiplexers allow rings to tune to the nearest channel even when local variations throw ring resonances out of order.

resonance repetition (Figure 14). A set of muxes then perform bit-reordering in the event that local variations shift a ring's resonance beyond that of its neighbors, shown in Figure 15. The necessary tuning distance of a ring becomes proportional to the channel separation, independent of systematic variations. The electrical backend also allows for the option to have more rings/receivers than the number of channels. With all rings spaced evenly across the FSR, extra rings reduce the mean tuning distance and allow for further tuning power reduction. This must be balanced with area costs and a more expensive backend as it requires a larger barrel-shifter and a higher degree of bit-reorder muxing.

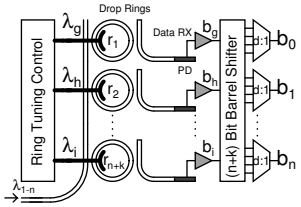


Fig. 16. Electrical bit re-shuffler backend for the receive-side rings of an n-bit channel with k extra rings and receivers. Rings simply tune to the nearest channel, not necessarily maintaining bit positions and relying upon the electrical backend to align bits to the expected positions. Though not shown, the same concept can be applied at the modulation-side.

By drastically reducing the tuning range, the idea of electrical tuning becomes highly promising. Using the same resonance detuning principle used for reverse-biased modulators, electrical ring tuning requires no static power and is able to tune-in and tune-out much more quickly than thermal tuning, albeit with a far inferior tuning range (typically sub-100 GHz). With the bit-shuffler backend, however, the required tuning distances are decreased and can often be covered without engaging the heaters when the channel spacing is small.

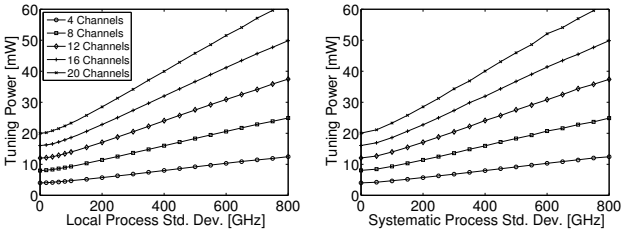
D. Ring Tuning Model

To evaluate each of our tuning strategies, we develop a *Monte Carlo* based tuning model. For each tuning scenario, a set of rings in a ring filter bank with some desired resonances is fabricated. To simulate the effects of local process

variations, we randomize the resonance of each ring using σ_{rL} and apply global systematic variations using σ_{rS}. The model then attempts to tune the set of rings across a range of temperatures. If successful, the tuning power cost is reported. The experiment is performed 1000 times for each parameter combination (fabrication bias, number of extra rings, etc.) to find the optimum tuning strategy for a given yield target.

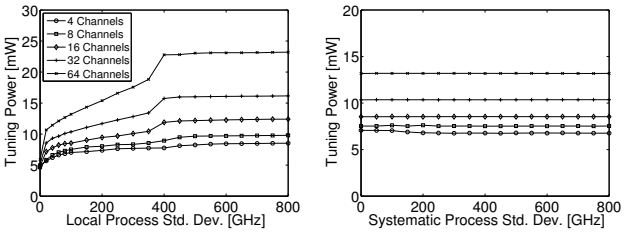
TABLE II
TUNING MODEL EVALUATION PARAMETERS

Parameter	Value
Aggregate Link Throughput	64 Gb/s
Free Spectral Range (FSR) R=3um ring	4 THz
Heating Efficiency	44 K/mW [21]
Tuning Efficiency $\frac{\Delta f}{\Delta T}$	10 GHz/K
Local Process Variation σ_{rL}	varies
Systematic Process Variation σ_{rS}	varies
Temperature Range	300-360 K
Process (for electrical backend)	32 nm Bulk CMOS
Tuner Controller Power	10 uW/Ring
Electrical Tuning Limit	50 GHz
Yield Target	99 %



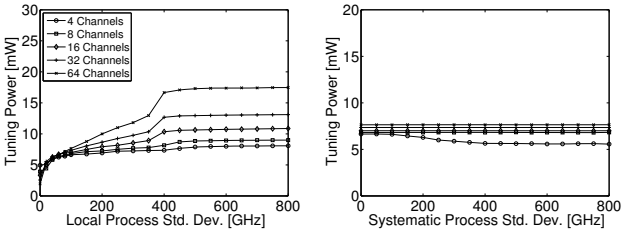
(a) Power vs σ_{rL}, (σ_{rS}=100 GHz) (b) Power vs σ_{rS}, (σ_{rL}=100 GHz)

Fig. 17. Tuning power vs process variation at various channelizations for the full thermal tuning scenario.



(a) Power vs σ_{rL}, (σ_{rS}=100 GHz) (b) Power vs σ_{rS}, (σ_{rL}=100 GHz)

Fig. 18. Tuning power vs process variation at various channelizations with an electrical backend capable of bit reshuffling.



(a) Power vs σ_{rL}, (σ_{rS}=100 GHz) (b) Power vs σ_{rS}, (σ_{rL}=100 GHz)

Fig. 19. Tuning power vs process variation at various channelizations with a bit reshuffling electrical backend and electrical assisted tuning.

The power needed to perform full thermal tuning is shown in Figure 17 across a range of process variations (σ_{rS}, σ_{rL}) and channelizations. Against process variations, the power cost of full thermal tuning is linear with both σ_{rS} and σ_{rL}, stemming from the increase in fabrication bias needed in order to maintain the same yield given higher process variations. The increase in tuning power is also linear with the number of channels, tracking the increase in the number of rings that require tuning.

Using an electrical backend to perform bit reshuffling, we show that tuning power can be successfully decoupled from σ_{rS} (Figure 18). Local variations (σ_{rL}) still affect the tuning power, as a larger σ_{rL} requires a larger degree of bit-reorder multiplexing. The tuning power also scales gracefully with the number of channels, owing to the decrease in channel-to-channel separation (and tuning distance) of each ring. Electrically-assisted tuning with an electrical backend allows for even further reductions in tuning power. As shown in Figure 19, cases with high numbers of channels benefit most as the channel separation is small enough to be covered electrically, without using heaters. Using this backend, we demonstrate a 5-10X tuning power reduction at dense WDM channelizations while maintaining tuning robustness across a range of process variations.

IV. WDM PHOTONIC LINK EVALUATION

In this section, we perform a full link-level optimization and evaluation of a WDM link to quantify energy consumption tradeoffs. In our evaluation, we explore links with 4 different aggregate throughput design points, 64 Gb/s, 256 Gb/s, 512 Gb/s, 1024 Gb/s, corresponding to minimum, medium, high, and maximum bandwidth scenarios.

Figure 20 shows that tuning power dominates at lower data-rates (since there are more channels given fixed throughput) and decreases with data-rate. Modulator, laser, SERDES, and receiver energies increase with data-rate and dominate at high rate-rates. At all throughput scenarios, an optimal energy balance is achieved at around 4-8 Gb/s. An overall energy-optimal point occurs at less than 200 fJ/bit for a link with 256 Gb/s of aggregate throughput and 4 Gb/s data-rate.

At the energy optimal point, we see that the energy consumption is roughly an even 3-way split between tuning, laser, and mod/rx/SERDES. As tuning power is now mostly dominated by the backend electrical components, this energy will scale favorably with technology and can be optimized using custom design. A full electrical tuning backend is also unnecessary on both modulate- and receive-side – barrel-shifts and bit-reordering only need to be performed once – meaning backend power can be cut by another 50 %. Refinement of photodetector responsivity and parasitic capacitances as well as lower-loss optical devices with improved electrical laser efficiencies can bring about further reductions in wall-plug laser power. It can be expected that energy/bit will drop to sub-100 fJ with device development, process scaling and overall link component refinement.

V. CONCLUSION

Integrated photonic interconnects are a promising solution to the throughput demands of future many-core processors.

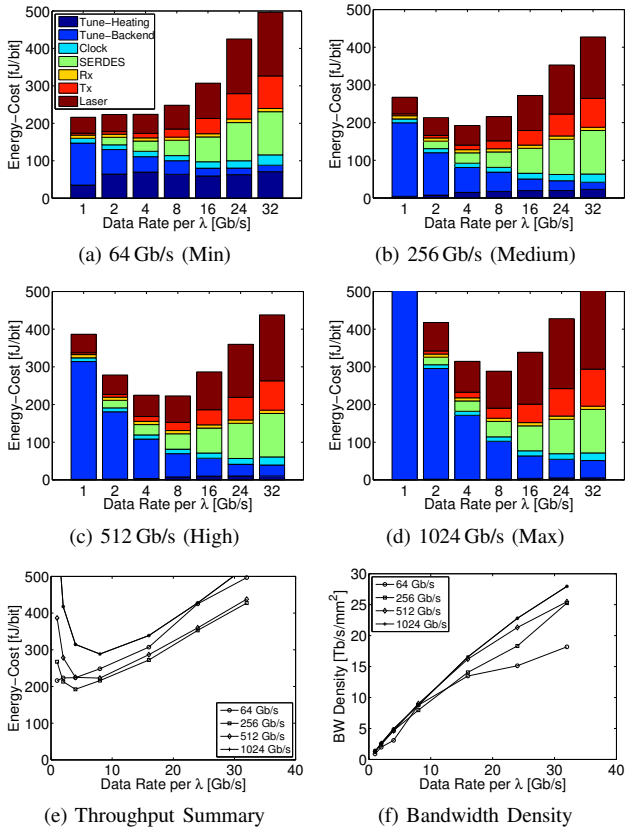


Fig. 20. Optimized power vs. data-rate for different aggregate link throughputs for Loss=10 dB, C_P=5 fF. For tuning, we assume a bit-reshuffler backend and electrically-assisted tuning with local variation σ_{rL}=40 GHz and systematic variation σ_{rS}=200 GHz. Note that the number of WDM channels changes with data rate (Channels = Throughput / Data-Rate).

As an emerging technology, circuit, device and architecture designers require insights concerning the impact of device and circuit parameters on link-level figures of merit.

This work presented a design-space exploration of a WDM integrated photonic link, facilitated through a set of circuit and device models that captured the optical-electrical tradeoffs of each link component. The modulator model showed the relationship between the modulation energy and the laser power, set through the extinction ratio and insertion loss specifications. Similarly, the optical receiver models demonstrated the degradation of sensitivity with data rate, which translated directly into increased laser power requirement. The impact of clock distribution was factored into the link-level analysis. Finally, to reduce ring tuning power, we proposed an electrical bit-shuffling backend, allowing for dense WDM and robustness against process and thermal variations.

Using our models, we performed co-optimization across all link components for a complete WDM integrated photonic link. We found that relatively low (sub 10 Gb/s) data-rates per link yielded optimal energy-efficiency accross a range of system throughputs. We showed that the photonic link is highly sensitive to parasitic capacitances present at the receiver input and that optical integration using TSVs to connect to an optical die could result in significant overhead in the laser power. This study illustrated that monolithic integration of photonic components can offer interconnect solutions with high throughput-density and energy-efficiency.

Acknowledgements

The authors would like to thank the Integrated Photonics teams at both University of Colorado, Boulder and MIT, in particular Milos Popović, Jeff Shainline and Jason Orcutt. This work was supported in part by DARPA, NSF, FCRP, IFC, Trusted Foundry, Intel, APIC, MIT CICS, and NSERC.

References

[1] R. J. Riedlinger *et al.*, "A 32nm 3.1 billion transistor 12-wide-issue Itanium; processor for mission-critical servers," in *ISSCC '11*, feb. 2011, pp. 84 –86.

[2] S. Palermo *et al.*, "A 90 nm CMOS 16 Gb/s transceiver for optical interconnects," *JSSCC*, vol. 43, no. 5, pp. 1235–1246, May 2008.

[3] C. Batten *et al.*, "Building manycore processor-to-dram networks with monolithic silicon photonics," *HOTI '08*, pp. 21–30, Aug. 2008.

[4] G. Kurian *et al.*, "ATAC: a 1000-core cache-coherent processor with on-chip optical network," in *Proceedings of PACT '10*. New York, NY, USA: ACM, 2010, pp. 477–488.

[5] A. Joshi *et al.*, "Silicon-photonic clos networks for global on-chip communication," in *Proceedings of NoCS '09*, may 2009, p. 91.

[6] D. Vantrease *et al.*, "Corona: System implications of emerging nanophotonic technology," in *Proceedings of ISCA '08*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 153–164.

[7] S. Beamer *et al.*, "Re-architecting DRAM memory systems with monolithically integrated silicon photonics," in *Proceedings of ISCA '10*. New York, NY, USA: ACM, 2010, pp. 129–140.

[8] R. Soref *et al.*, "Electrooptical effects in silicon," *IEEE Journal of Quantum Electronics*, vol. 23, no. 1, January 1987.

[9] Q. Xu *et al.*, "12.5 Gbit/s carrier-injection-based silicon micro-ring silicon modulators," *Optical Society of America*, vol. 15, no. 2, 2007.

[10] M. Watts, D. Trotter, R. Young, and A. Lentine, "Maximally confined silicon microphotonic modulators and switches," in *IEEE Lasers and Electro-Optics Society, 2008. LEOS 2008. 21st Annual Meeting of the*, nov. 2008, pp. 457 –458.

[11] K.-L. Wong *et al.*, "Offset compensation in comparators with minimum input-referred supply noise," *JSSCC*, vol. 39, no. 5, pp. 837 – 840, 2004.

[12] S.-H. Woo *et al.*, "Offset voltage estimation model for latch-type sense amplifiers," *Circuits, Devices Systems, IET*, vol. 4, no. 6, pp. 503 –513, 2010.

[13] E. Alon *et al.*, "Circuits and techniques for high-resolution measurement of on-chip power supply noise," *JSSCC*, vol. 40, no. 4, pp. 820–828, 2005.

[14] S. Naffziger *et al.*, "The implementation of a 2-core, multi-threaded itanium family processor," *JSSCC*, vol. 41, no. 1, pp. 197 – 209, 2006.

[15] S. Assefa *et al.*, "Cmos-integrated optical receivers for on-chip interconnects," *Selected Topics in Quantum Electronics, IEEE Journal of*, vol. 16, no. 5, pp. 1376 –1385, sept.-oct. 2010.

[16] C. Kromer *et al.*, "A low-power 20-ghz 52-db-ohm transimpedance amplifier in 80-nm CMOS," *JSSCC*, vol. 39, no. 6, pp. 885–894, 2004.

[17] A. Emami-Neyestanak *et al.*, "A 1.6 gb/s, 3 mw cmos receiver for optical communication," *VLSI '02*, pp. 84–87, 2002.

[18] A. Bhatnagar *et al.*, "Receiverless clocking of a cmos digital circuit using short optical pulses," in *LEOS '02*, vol. 1, 2002, pp. 127 – 128.

[19] F. Tavernier *et al.*, "High-speed optical receivers with integrated photodiode in 130 nm cmos," *JSSCC*, vol. 44, no. 10, pp. 2856 –2867, 2009.

[20] A. Strak *et al.*, "Analysis of timing jitter in inverters induced by power-supply noise," in *DTIS '06*, sept. 2006, pp. 53 –56.

[21] J. S. Orcutt *et al.*, "Nanophotonic integration in state-of-the-art CMOS foundries," *Opt. Express*, vol. 19, no. 3, pp. 2335–2346, Jan 2011.

[22] S. Selvaraja *et al.*, "Fabrication of uniform photonic devices using 193nm optical lithography in silicon-on-insulator," *ECIO '08*, 2008.

[23] M. S. Nawrocka *et al.*, "Tunable silicon microring resonator with wide free spectral range," *Applied Physics Letters*, vol. 89, no. 7, p. 071110, 2006.

[24] V. Raghunathan *et al.*, "Athermal silicon ring resonators," in *Integrated Photonics Research, Silicon and Nanophotonics*. Optical Society of America, 2010, p. IMC5.