High-Speed Serial I/O Design for Channel-Limited and Power-Constrained Systems

Samuel Palermo
Texas A&M University

Introduction

Dramatic increases in processing power, fueled by a combination of integrated circuit scaling and shifts in computer architectures from single-core to future many-core systems, has rapidly scaled on-chip aggregate bandwidths into the Tb/s range [1], necessitating a corresponding increase in the amount of data communicated between chips to not limit overall system performance [2]. Due to the limited I/O pin count in chip packages and printed circuit board (PCB) wiring constraints, high-speed serial link technology is employed for this inter-chip communication.

The two conventional methods to increase inter-chip communication bandwidth include raising both the per-channel data rate and the I/O number, as projected in the current ITRS roadmap (Figure 1). However, packaging limitations prevent a dramatic increase in I/O channel number. This implies that chip-to-chip data rates must increase dramatically in the next decade, presenting quite a challenge considering the limited power budgets in processors. While high-performance I/O circuitry can leverage the technology improvements that enable increased core performance, unfortunately the bandwidth of the electrical channels used for inter-chip communication has not scaled in the same manner. Thus, rather than being technology limited, current high-speed I/O link designs are becoming channel limited. In order to continue scaling data rates, link designers implement sophisticated equalization circuitry to compensate for the frequency dependent loss of the bandlimited channels [3-5]. Although, with this additional complexity comes both power and area costs, necessitating advances in low-power serial I/O design techniques and consideration of alternate I/O technologies, such as optical inter-chip communication links.

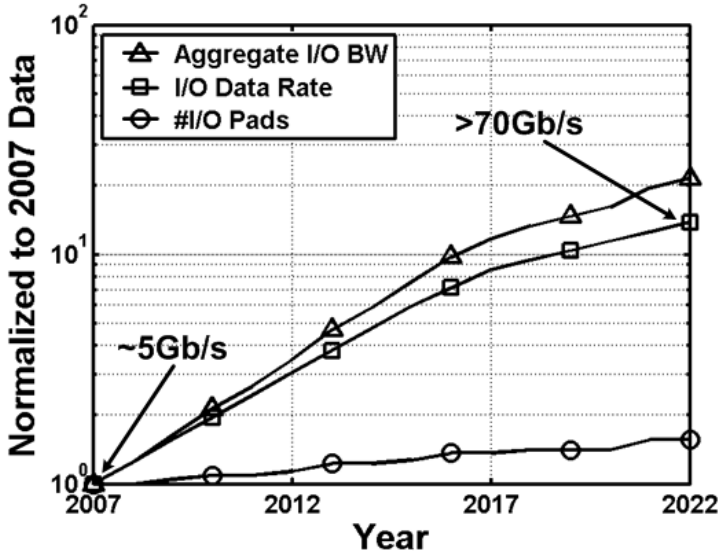


Figure 1 I/O scaling projections [6].

This chapter discusses these challenges associated with scaling serial I/O data rates and current design techniques. The section "High-Speed Link Overview" describes the major high-speed components, channel properties, and performance metrics. A presentation of equalization and advanced modulation techniques follows in section "Channel-Limited Design Techniques". Link architectures and circuit techniques that improve link power efficiency are discussed in section "Low-Power Design Techniques". Section "Optical Interconnects" details how optical inter-chip communication links have the potential to fully leverage increased data rates provided through CMOS technology scaling at suitable power efficiency levels. Finally, the chapter is summarized in the section "Conclusion".

High-Speed Link Overview

High-speed point-to-point electrical link systems employ specialized I/O circuitry that performs incident wave signaling over carefully designed controlled-impedance channels in order to achieve high data rates. As will be described later in this section, the electrical channel's frequency-dependent loss and impedance discontinuities become major limiters in data rate scaling.

This section begins by describing the three major link circuit components, the transmitter, receiver, and timing system. Next, it discusses the electrical channel properties that impact the transmitted signal. The section concludes by providing an overview of key technology and system performance metrics.

Electrical Link Circuits

Figure 2 shows the major components of a typical high-speed electrical link system. Due to the limited number of high-speed I/O pins in chip packages and printed circuit board (PCB) wiring constraints, a high-bandwidth transmitter serializes parallel input data for transmission. Differential low-swing signaling is commonly employed for common-mode noise rejection and reduced crosstalk due to the inherent signal current return path [7]. At the receiver, the incoming signal is sampled, regenerated to CMOS values, and deserialized. The high-frequency clocks which synchronize the data transfer onto the channel are generated by a frequency synthesis phase-locked loop (PLL) at the transmitter, while at the receiver the sampling clocks are aligned to the incoming data stream by a timing recovery system.

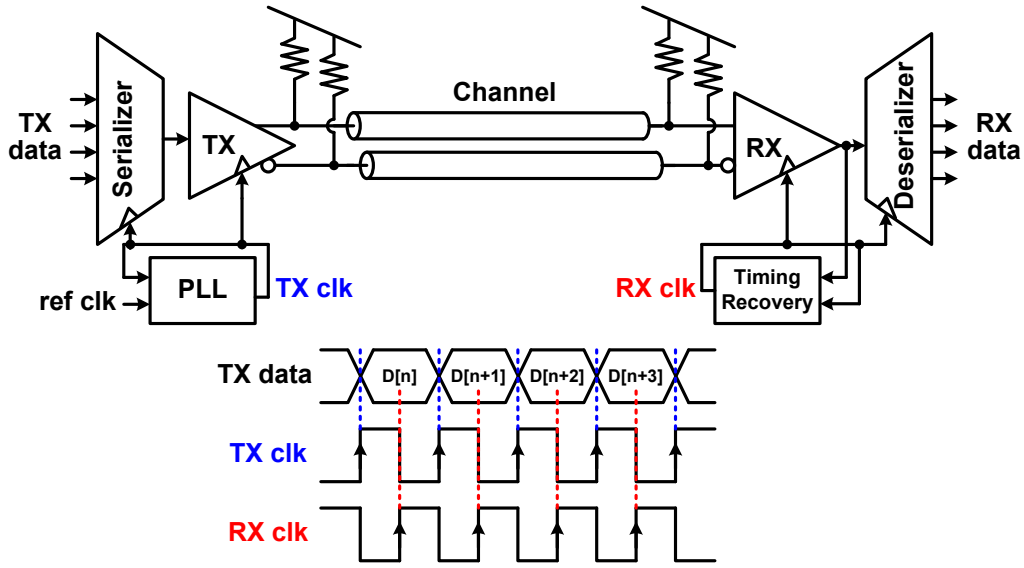


Figure 2 High-speed electrical link system.

Transmitter

The transmitter must generate an accurate voltage swing on the channel while also maintaining proper output impedance in order to attenuate any channel-induced reflections. Either current or voltage-mode drivers, shown in Figure 3, are suitable output stages. Current-mode drivers

typically steer current close to 20mA between the differential channel lines in order to launch a bipolar voltage swing on the order of ±500mV. Driver output impedance is maintained through termination which is in parallel with the high-impedance current switch. While current-mode drivers are most commonly implemented [8], the power associated with the required output voltage for proper transistor output impedance and the "wasted" current in the parallel termination led designers to consider voltage-mode drivers. These drivers use a regulated output stage to supply a fixed output swing on the channel through a series termination which is feedback controlled [9]. While the feedback impedance control is not as simple as parallel termination, the voltage-mode drivers have the potential to supply an equal receiver voltage swing at a quarter [10] of the common 20mA cost of current-mode drivers.

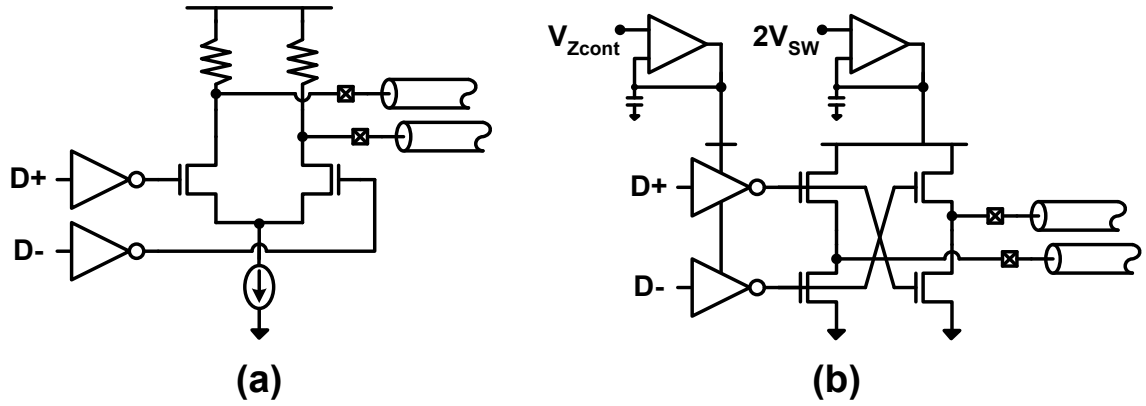


Figure 3 Transmitter output stages: (a) current-mode driver, (b) voltage-mode driver.

Receiver

Figure 4 shows a high-speed receiver which compares the incoming data to a threshold and amplifies the signal to a CMOS value. This highlights a major advantage of binary differential signaling, where this threshold is inherent, whereas single-ended signaling requires careful threshold generation to account for variations in signal amplitude, loss, and noise [11]. The bulk of the signal amplification is often performed with a positive feedback latch [12,13]. These latches are more power-efficient versus cascaded linear amplification stages since they don't dissipate DC current. While regenerative latches are the most power-efficient input amplifiers, link designers have used a small number of linear pre-amplification stages to implement equalization filters that offset channel loss faced by high data rate signals [14,15].

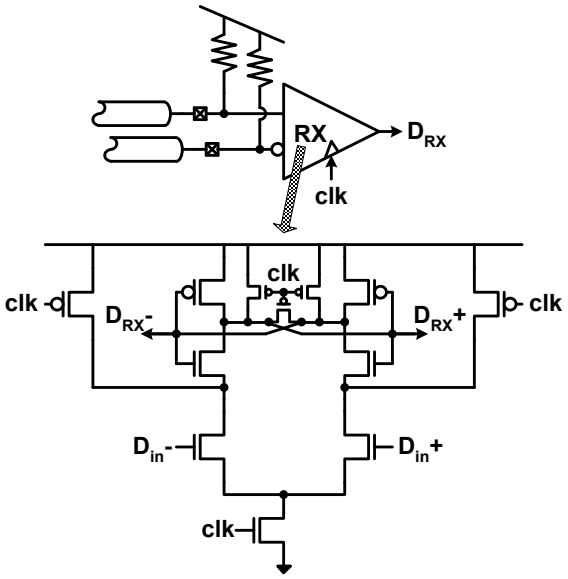


Figure 4 Receiver input stage with regenerative latch [12].

One issue with these latches is that they require time to reset or "pre-charge", and thus to achieve high data rates, often multiple latches are placed in parallel at the input and activated with multiple clock phases spaced a bit period apart in a time-division-demultiplexing manner [16,17], shown in Figure 5. This technique is also applicable at the transmitter, where the maximum serialized data rate is set by the clocks switching the multiplexer. The use of multiple clock phases offset in time by a bit period can overcome the intrinsic gate-speed which limits the maximum clock rate that can be efficiently distributed to a period of 6-8 fanout-of-four (FO4) clock buffer delays [18]].

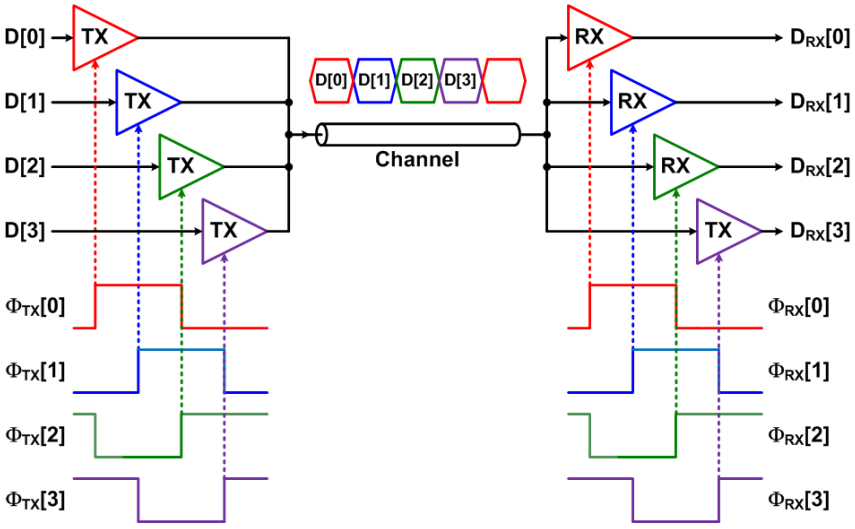


Figure 5 Time-division multiplexing link.

Timing System

The two primary clocking architectures employed in high-speed I/O systems are embedded and source synchronous clocking, as shown in Figure 6. While both architectures typically use a PLL to generate the transmit serialization clocks, they differ in the manner in which receiver timing recovery is performed. In embedded clocking systems, only I/O data channels are routed to the receiver, where a clock-and-data recovery (CDR) system extracts clock information "embedded" in the transmitted data stream to determine the receiver clock frequency and optimum phase position. Source synchronous systems, also called forwarded-clock systems, use an additional dedicated clock channel to forward the high-frequency serialization clock used by multiple transmit channels to the receiver chip, where per-channel de-skewing circuitry is implemented. The circuits and architectural trade-offs of these timing systems are discussed next.

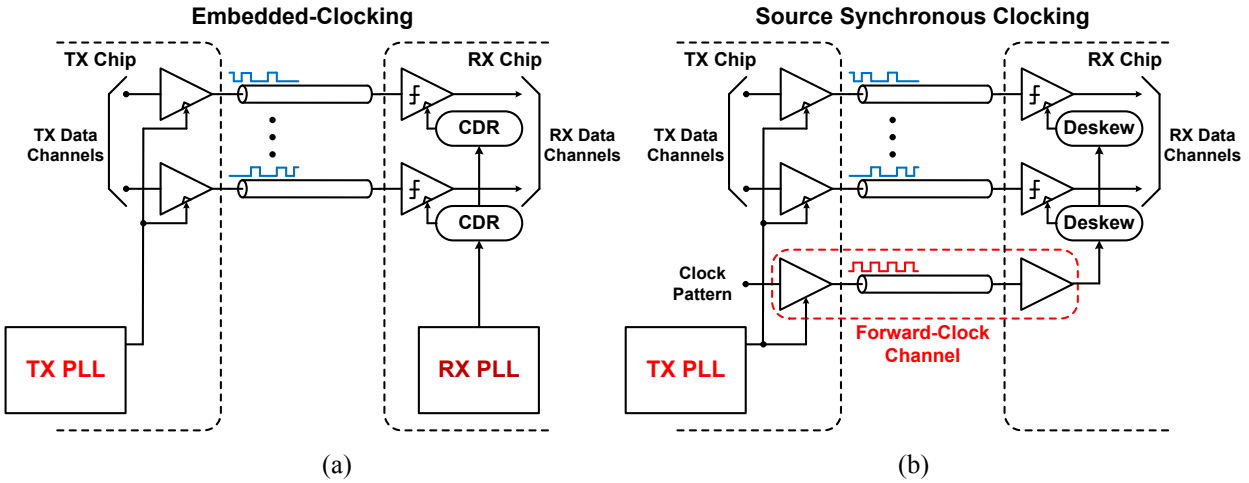


Figure 6 Multi-channel serial link system: (a) embedded-clocking architecture, (b) source synchronous clocking architecture.

Figure 7 shows a PLL, which is often used at the transmitter for clock synthesis in order to serialize reduced-rate parallel input data and also potentially at the receiver for clock recovery. The PLL is a negative feedback loop which works to lock the phase of the feedback clock to an input reference clock. A phase-frequency detector produces an error signal which is proportional to the phase difference between the feedback and reference clocks. This phase error is then filtered to provide a control signal to a voltage-controlled oscillator (VCO) which generates the output clock. The PLL performs frequency synthesis by placing a clock divider in the feedback

path, which forces the loop to lock with the output clock frequency equal to the input reference frequency times the loop division factor.

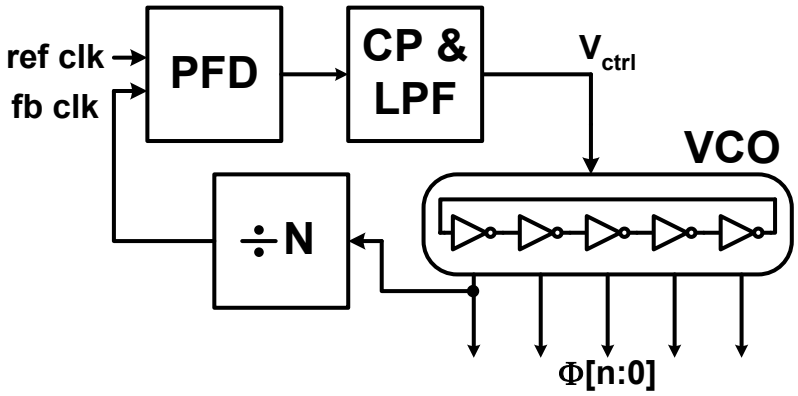


Figure 7 PLL frequency synthesizer.

It is important that the PLL produce clocks with low timing noise, quantified in the timing domain as jitter and in the frequency domain as phase noise. Considering this, the most critical PLL component is the VCO, as its phase noise performance can dominate at the output clock and have a large influence on the overall loop design. LC oscillators typically have the best phase noise performance, but their area is large and tuning range is limited [19]. While ring oscillators display inferior phase noise characteristics, they offer advantages in reduced area, wide frequency range, and ability to easily generate multiple phase clocks for time-division multiplexing applications [16,17].

Also important is the PLL's ability to maintain proper operation over process variances, operating voltage, temperature, and frequency range. To address this, self-biasing techniques were developed by Maneatis [20] and expanded in [21,22] that set constant loop stability and noise filtering parameters over these variances in operating conditions.

At the receiver, timing recovery is required in order to position the data sampling clocks with maximum timing margin. For embedded clocking systems, it is possible to modify a PLL to perform clock recovery with changes in the phase detection circuitry, as shown in Figure 8. Here the phase detector samples the incoming data stream to extract both data and phase information. As shown in Figure 9, the phase detector can either be linear [23], which provides both sign and magnitude information of the phase error, or binary [24], which provides only

phase error sign information. While CDR systems with linear phase detectors are easier to analyze, generally they are harder to implement at high data rates due to the difficulty of generating narrow error pulse widths, resulting in effective dead-zones in the phase detector [25]. Binary, or "bang-bang", phase detectors minimize this problem by providing equal delay for both data and phase information and only resolving the sign of the phase error [26].

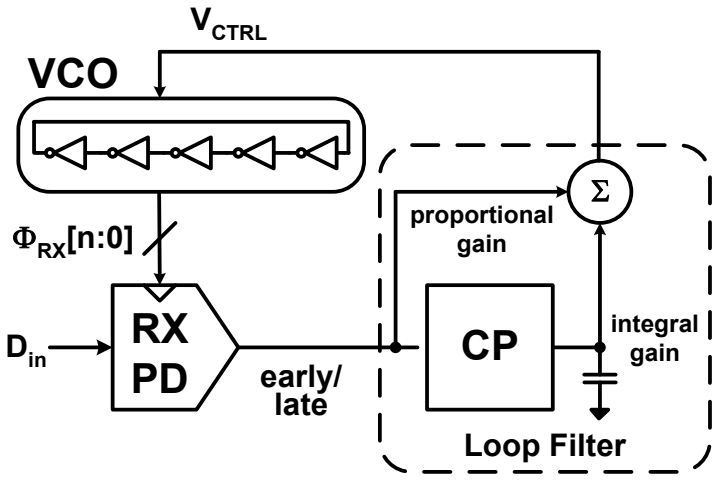


Figure 8 PLL-based CDR system.

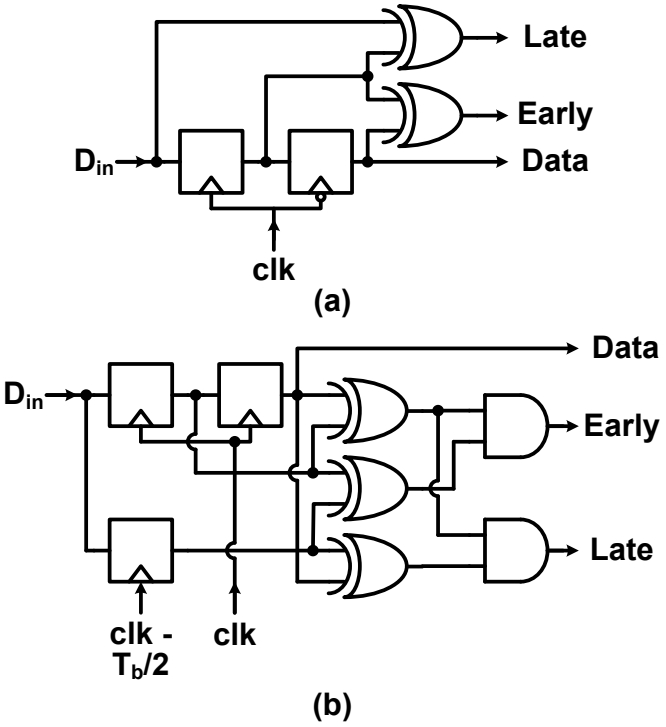


Figure 9 CDR phase detectors: (a) linear [23], (b) binary [24].

While a PLL-based CDR is an effective timing recover system, the power and area cost of having a PLL for each receive channel is prohibitive in some I/O systems. Another issue is that CDR bandwidth is often set by jitter transfer and tolerance specifications, leaving little freedom to optimize PLL bandwidth to filter various sources of clock noise, such as VCO phase noise. The nonlinear behavior of binary phase detectors can also lead to limited locking range, necessitating a parallel frequency acquisition loop [27]. This motivates the use of dual-loop clock recovery [11], which allows several channels to share a global receiver PLL locked to a stable reference clock and also provides two degrees of freedom to independently set the CDR bandwidth for a given specification and optimize the receiver PLL bandwidth for best jitter performance.

An implementation of a dual-loop CDR is shown in Figure 10. This CDR has a core loop which produces several clock phases for a separate phase recovery loop. The core loop can either be a frequency synthesis PLL [28] or, if frequency synthesis is done elsewhere, a lower-complexity delay-locked loop (DLL) [11]. An independent phase recovery loop provides flexibility in the input jitter tracking without any effect on the core loop dynamics and allows for sharing of the clock generation with the transmit channels. The phase loop typically consists of a binary phase detector, a digital loop filter, and a finite state machine (FSM) which updates the phase muxes and interpolators to generate the optimal position for the receiver clocks. Low-swing current-controlled differential buffers [11] or large-swing tri-state digital buffers [29] are use to implement the interpolators.

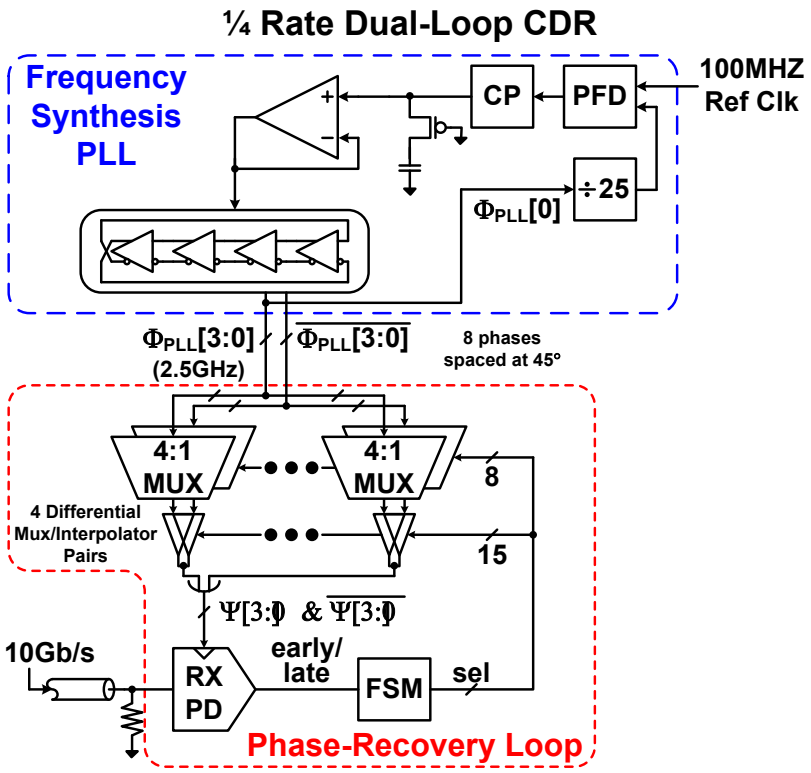


Figure 10 Dual-loop CDR for a 4:1 input demultiplexing receiver.

Phase interpolator linearity, which is both a function of the interpolator design and the input clocks' phase precision, can have a major impact on CDR timing margin and phase error tracking. The interpolator must be designed with a proper ratio between the input phase spacing and the interpolator output time constant in order to insure proper phase mixing [11]. Also, while there is the potential to share the core loop among multiple receiver channels, distributing the multiple clock phases in a manner that preserves precise phase spacing can result in significant power consumption. In order to avoid distributing multiple clock phases, the global frequency synthesis PLL can distribute one clock phase to local DLLs, either placed at each channel or shared by a cluster of receive channels, which generate the multiple clock phases used for interpolation [30,31]. This architecture can be used for source-synchronous systems by replacing the global PLL clock with the incoming forwarded clock [30].

While proper design of these high-speed I/O components requires considerable attention, CMOS scaling allows the basic circuit blocks to achieve data rates that exceed 10Gb/s [14, 32]. However, as data rates scale into the low Gb/s, the frequency dependent loss of the chip-to-chip electrical wires disperses the transmitted signal to the extent that it is undetectable at the receiver

without proper signal processing or channel equalization techniques. Thus, in order to design systems that achieve increased data rates, link designers must comprehend the high-frequency characteristics of the electrical channel, which are outlined next.

Electrical Channels

Electrical inter-chip communication bandwidth is predominantly limited by high-frequency loss of electrical traces, reflections caused from impedance discontinuities, and adjacent signal crosstalk, as shown in Figure 11. The relative magnitudes of these channel characteristics depend on the length and quality of the electrical channel which is a function of the application. Common applications range from processor-to-memory interconnection, which typically have short (<10") top-level microstrip traces with relatively uniform loss slopes [14], to server/router and multi-processor systems, which employ either long (~30") multi-layer backplanes [3] or (~10m) cables [33] which can both possess large impedance discontinuities and loss.

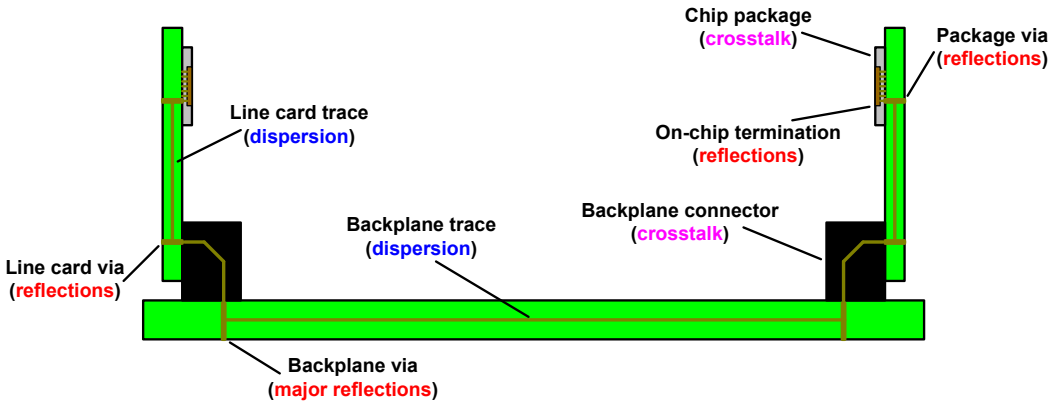


Figure 11 Backplane system cross-section.

PCB traces suffer from high-frequency attenuation caused by wire skin effect and dielectric loss. As a signal propagates down a transmission line, the normalized amplitude at a distance x is equal to

$$\frac{V(x)}{V(0)} = e^{-(\alpha_R + \alpha_D)x}, \tag{1}$$

where α_R and α_D represent resistive and dielectric loss factors [7]. The skin effect, which describes the process of high-frequency signal current crowding near the conductor surface,

impacts the resistive loss term as frequency increases. This results in a resistive loss term which is proportional to the square-root of frequency

$$\alpha_R = \frac{R_{AC}}{2Z_0} = \frac{2.61\times10^{-7}\sqrt{\rho_r}}{\pi D 2 Z_0}\sqrt{f}\,, \tag{2}$$

where D is the trace's diameter (in), ρ_r is the relative resistivity compared to copper, and Z_0 is the trace's characteristic impedance [34]. Dielectric loss describes the process where energy is absorbed from the signal trace and transferred into heat due to the rotation of the board's dielectric atoms in an alternating electric field [34]. This results in the dielectric loss term increasing proportional to the signal frequency

$$\alpha_D = \frac{\pi\sqrt{\varepsilon_r}\,\tan\delta_D}{c} f\,, \tag{3}$$

where ε_r is the relative permittivity, c is the speed of light, and $\tan\delta_D$ is the board material's loss tangent [15].

Figure 12 shows how these frequency dependent loss terms result in low-pass channels where the attenuation increases with distance. The high-frequency content of pulses sent across such these channel is filtered, resulting in an attenuated received pulse whose energy has been spread or dispersed over several bit periods, as shown in Figure 13(a). When transmitting data across the channel, energy from individual bits will now interfere with adjacent bits and make them more difficult to detect. This intersymbol interference (ISI) increases with channel loss and can completely close the received data eye diagram, as shown in Figure 13(b).

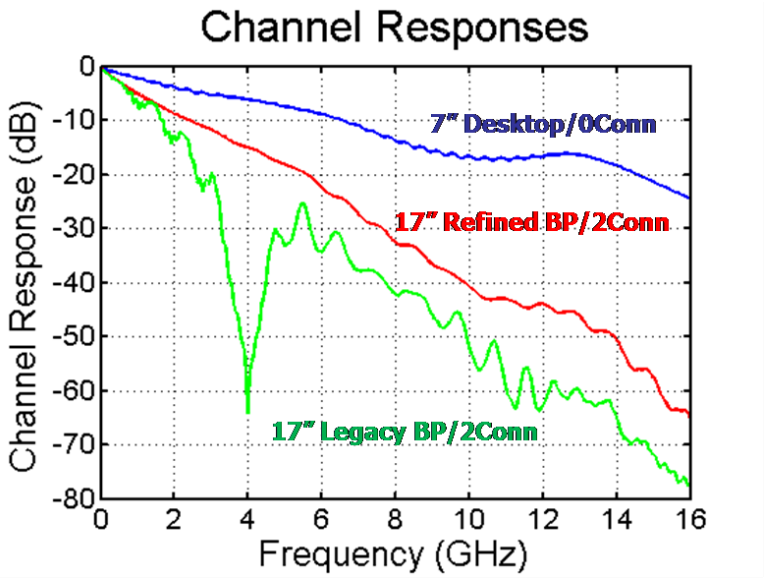
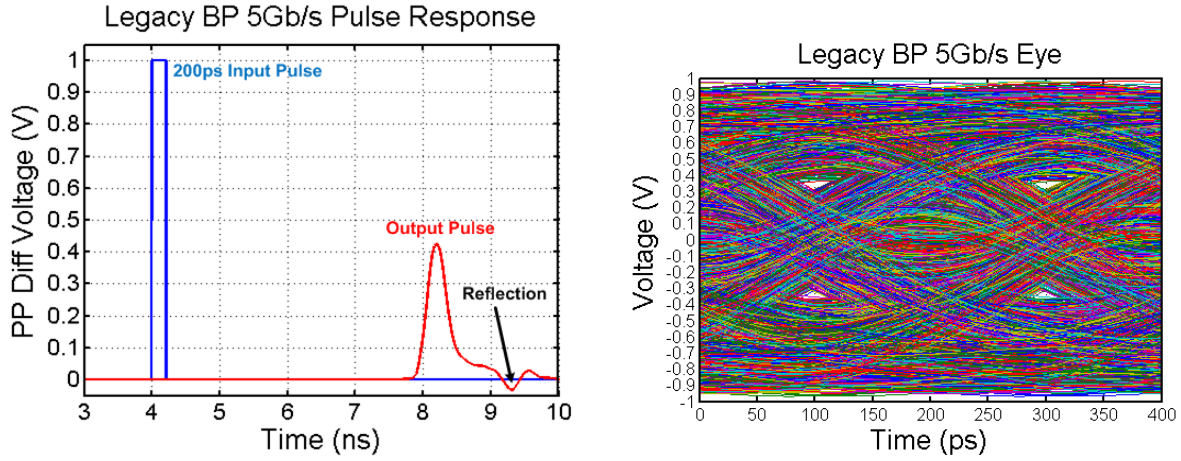


Figure 12 Frequency response of several channels.



(a) (b)

Figure 13 Legacy backplane channel performance at 5Gb/s: (a) pulse response, (b) eye diagram.

Signal interference also results from reflections caused by impedance discontinuities. If a signal propagating across a transmission line experiences a change in impedance Z_r relative to the line's characteristic impedance Z_0, a percentage of that signal equal to

$$\frac{V_r}{V_i} = \frac{Z_r - Z_0}{Z_r + Z_0} \tag{4}$$

will reflect back to the transmitter. This results in an attenuated or, in the case of multiple reflections, a time delayed version of the signal arriving at the receiver. The most common

sources of impedance discontinuities are from on-chip termination mismatches and via stubs that result with signaling over multiple PCB layers. Figure 12 shows that the capacitive discontinuity formed by the thick backplane via stubs can cause severe nulls in the channel frequency response.

Another form of interference comes from crosstalk, which occurs due to both capacitive and inductive coupling between neighboring signal lines. As a signal propagates across the channel, it experiences the most crosstalk in the backplane connectors and chip packages where the signal spacing is smallest compared to the distance to a shield. Crosstalk is classified as near-end (NEXT), where energy from an aggressor (transmitter) couples and is reflected back to the victim (receiver) on the same chip, and far-end (FEXT), where the aggressor energy couples and propagates along the channel to a victim on another chip. NEXT is commonly the most detrimental crosstalk, as energy from a strong transmitter (~$1V_{pp}$) can couple onto a received signal at the same chip which has been attenuated (~$20mV_{pp}$) from propagating on the lossy channel. Crosstalk is potentially a major limiter to high-speed electrical link scaling, as in common backplane channels the crosstalk energy can actually exceed the through channel signal energy at frequencies near 4GHz [3].

Performance Metrics

High-speed link data rate can be limited both by circuit technology and the communication channel. Aggressive CMOS technology scaling has benefited both digital and analog circuit performance, with 45nm processes displaying sub-20ps inverter fanout-of-four (FO4) delays and capable of measured nMOS f_T near 400GHz [35]. As mentioned earlier, one of the key circuit constraints to increasing data rate is the maximum clock frequency that can be effectively distributed, which is a period of 6-8 FO4 clock buffer delays. This implies that a half-rate link architecture can potentially achieve a minimum bit period of 3-4 FO4, or greater than 12Gb/s in a 45nm technology, with relatively simple inverter-based clock distribution. While technology scaling should allow the data rate to increase proportionally with improvements in gate delay, the limited electrical channel bandwidth is actually the major constraint in the majority of systems, with overall channel loss and equalization complexity limiting the maximum data rate to lower than the potential offered by a given technology.

The primary serial link performance metric is the bit-error-rate (BER), with systems requiring BERs ranging from 10^{-12} to 10^{-15}. In order to meet these bit-error-rate (BER) targets, link system design must consider both the channel and all relevant noise sources. Link modeling and analysis tools have matured rapidly as data rates have significantly exceeded electrical channel bandwidth. Current state-of-the-art link analysis tools [36-38] use statistical means to combine deterministic noise sources, such as ISI and bounded supply noise and jitter, random noise sources, such as Gaussian thermal noise and jitter, and receiver aperture in order to estimate link margins at a given BER. The results of a link analysis tool can be visualized with a statistical BER eye (Figure 14) which gives the overall link voltage and timing margins at a given BER.

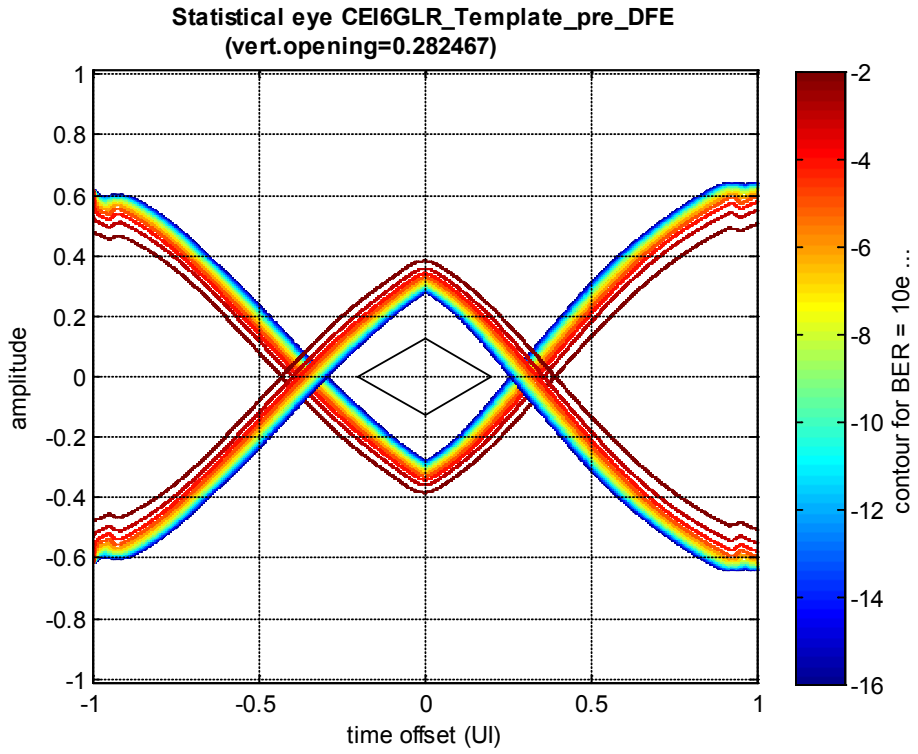
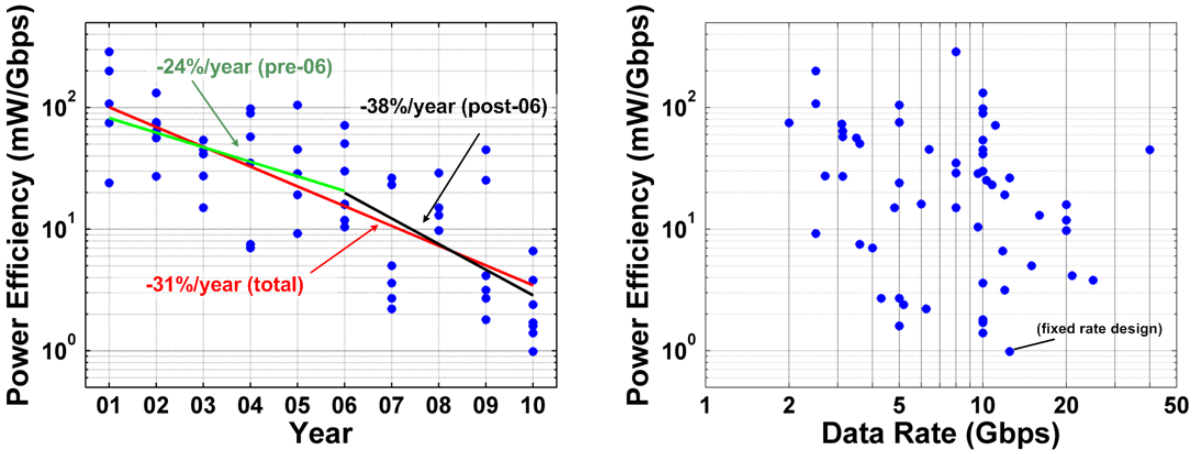


Figure 14 6Gbps statistical BER eye produced with StatEye simulator.

Another metric under increased emphasis is power efficiency, expressed by the link power consumption divided by the data rate in units of mW/Gb/s or pJ/bit. This is a due to systems demanding a rapid increase in I/O bandwidth, while also facing power ceilings. As shown in Figure 15, the majority of link architectures exceed 10mW/Gb/s. This is a consequence of several reasons, including complying with industry standards, supporting demanding channel

loss, and placing emphasis on extending raw data rate. Recent work has emphasized reducing link power [15,30,39], with designs achieving sub-3mW/Gb/s at data rates up to 12.5Gb/s.



(a) (b)

Figure 15 High-speed serial I/O trends: (a) power efficiency vs year (b) power efficiency vs data rate.

Channel-Limited Design Techniques

The previous section discussed interference mechanisms that can severely limit the rate at which data is transmitted across electrical channels. As shown in Figure 13, frequency dependent channel loss can reach magnitudes sufficient to make simple NRZ binary signaling undetectable. Thus, in order to continue scaling electrical link data rates, designers have implemented systems which compensate for frequency dependent loss or equalize the channel response. This section discusses how the equalization circuitry is often implemented in high-speed links, and other approaches for dealing with these issues.

Equalization Systems

In order to extend a given channel's maximum data rate, many communication systems use equalization techniques to cancel intersymbol interference caused by channel distortion. Equalizers are implemented either as linear filters (both discrete and continuous-time) that attempt to flatten the channel frequency response, or as non-linear filters that directly cancel ISI based on the received data sequence. Depending on system data rate requirements relative to

channel bandwidth and the severity of potential noise sources, different combinations of transmit and/or receive equalization are employed.

Transmit equalization, implemented with an FIR filter, is the most common technique used in high-speed links [40]. This TX "pre-emphasis" (or more accurately "de-emphasis") filter, shown in Figure 16, attempts to invert the channel distortion that a data bit experiences by pre-distorting or shaping the pulse over several bit times. While this filtering could also be implemented at the receiver, the main advantage of implementing the equalization at the transmitter is that it is generally easier to build high-speed digital-to-analog converters (DACs) versus receive-side analog-to-digital converters. However, because the transmitter is limited in the amount of peak-power that it can send across the channel due to driver voltage headroom constraints, the net result is that the low-frequency signal content has been attenuated down to the high-frequency level, as shown in Figure 16.

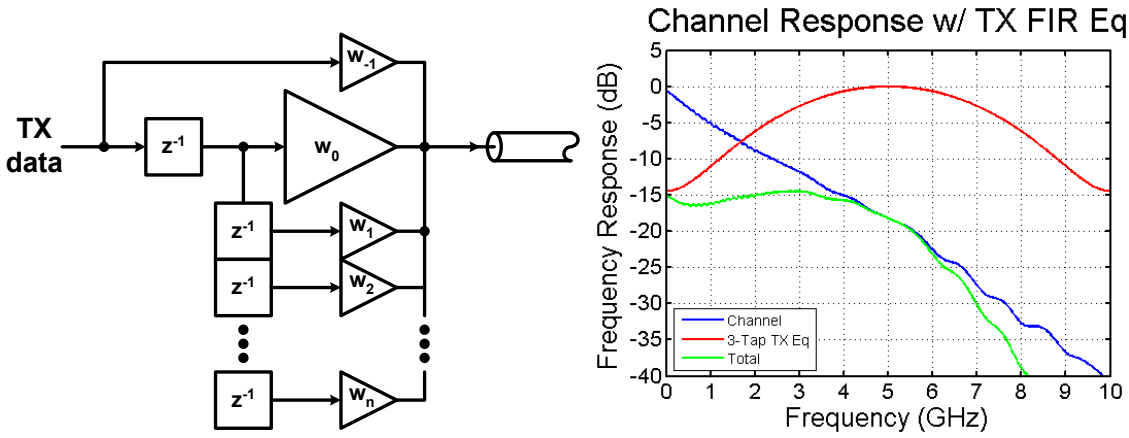


Figure 16 TX equalization with an FIR filter.

Figure 17 shows a block diagram of receiver-side FIR equalization. A common problem faced by linear receive side equalization is that high-frequency noise content and crosstalk are amplified along with the incoming signal. Also challenging is the implementation of the analog delay elements, which are often implemented through time-interleaved sample-and-hold stages [41] or through pure analog delay stages with large area passives [42,43]. Nonetheless, one of the major advantage of receive side equalization is that the filter tap coefficients can be adaptively tuned to the specific channel [41], which is not possible with transmit-side equalization unless a "back-channel" is implemented [44].

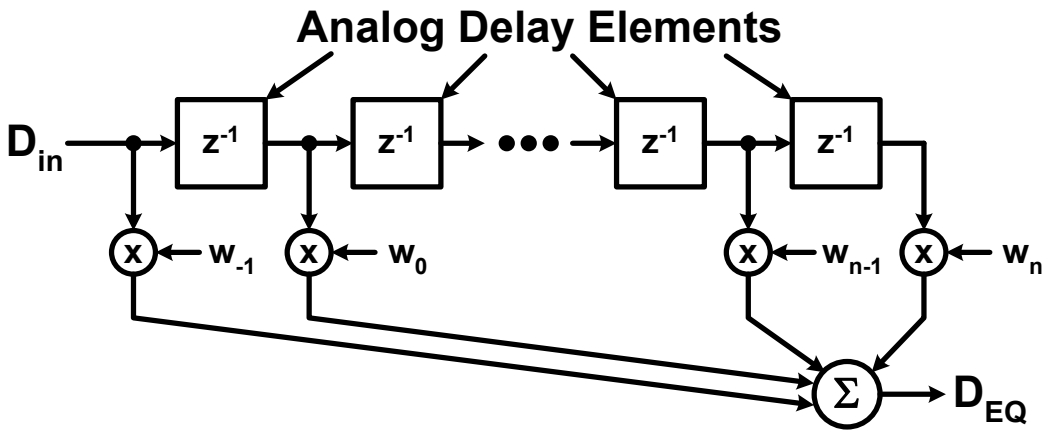


Figure 17 RX equalization with an FIR filter.

Linear receiver equalization can also be implemented with a continuous-time amplifier, as shown in Figure 18. Here, programmable RC-degeneration in the differential amplifier creates a high-pass filter transfer function which compensates the low-pass channel. While this implementation is a simple and low-area solution, one issue is that the amplifier has to supply gain at frequencies close to the full signal data rate. This gain-bandwidth requirement potentially limits the maximum data rate, particularly in time-division demultiplexing receivers.

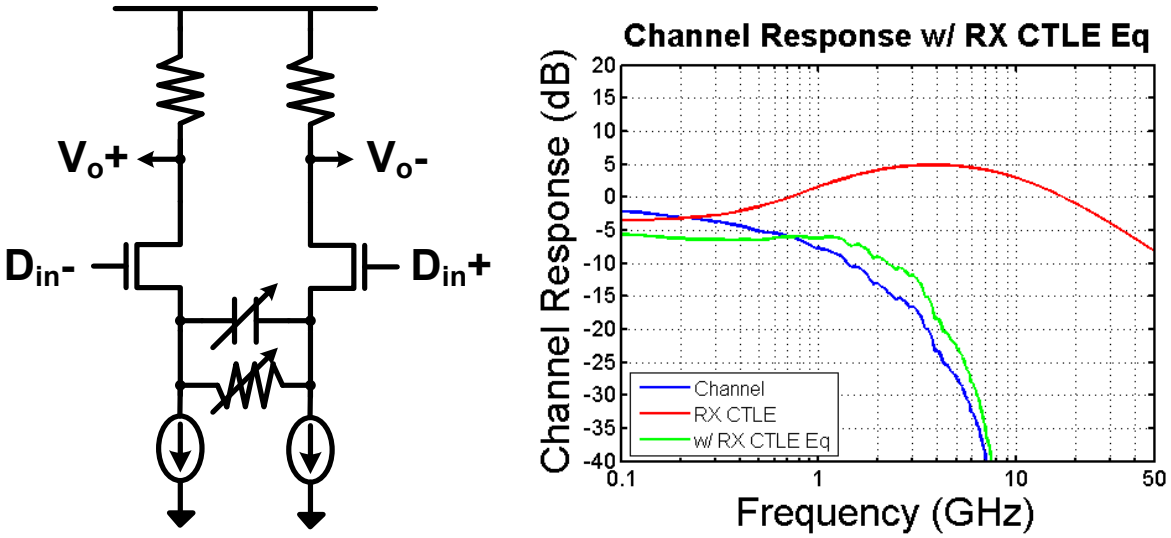


Figure 18 Continuous-time equalizing amplifier.

The final equalization topology commonly implemented in high-speed links is receiver-side decision feedback equalizer (DFE). A DFE, shown in Figure 19, attempts to directly subtract ISI from the incoming signal by feeding back the resolved data to control the polarity of the

equalization taps. Unlike linear receive equalization, a DFE doesn't directly amplify the input signal noise or cross-talk since it uses the quantized input values. However, there is the potential for error propagation in a DFE if the noise is large enough for a quantized output to be wrong. Also, due to the feedback equalization structure, the DFE cannot cancel pre-cursor ISI. The major challenge in DFE implementation is closing timing on the first tap feedback since this must be done in one bit period or unit interval (UI). Direct feedback implementations [3] require this critical timing path to be highly optimized. While a loop-unrolling architecture eliminates the need for first tap feedback [45], if a multiple tap implementation is required the critical path simply shifts to the second tap which has a timing constraint also near 1UI [5].

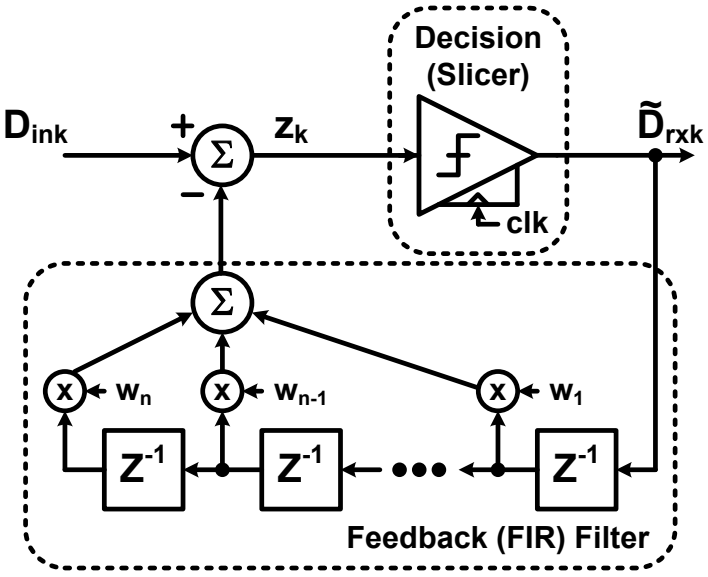


Figure 19 RX equalization with a DFE.

Advanced Modulation Techniques

Modulation techniques which provide spectral efficiencies higher than simple binary signaling have also been implemented by link designers in order to increase data rates over band-limited channels. Multi-level PAM, most commonly PAM-4, is a popular modulation scheme which has been implemented both in academia [46] and industry [47,48]. Shown in Figure 20, PAM-4 modulation consists of two bits per symbol, which allows transmission of an equivalent amount of data in half the channel bandwidth. However, due to the transmitter's peak-power limit, the voltage margin between symbols is 3x (9.5dB) lower with PAM-4 versus simple binary PAM-2 signaling. Thus, a general rule of thumb exists that if the channel loss at the PAM-2 Nyquist

frequency is greater than 10dB relative to the previous octave, then PAM-4 can potentially offer a higher signal-to-noise ratio (SNR) at the receiver. However, this rule can be somewhat optimistic due to the differing ISI and jitter distribution present with PAM-4 signaling [37]. Also, PAM-2 signaling with a non-linear DFE at the receiver further bridges the performance gap due to the DFE's ability to cancel the dominant first post-cursor ISI without the inherent signal attenuation associated with transmitter equalization [4].

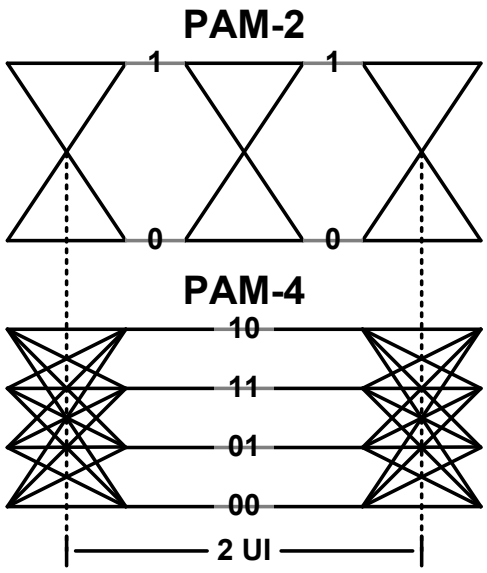


Figure 20 Pulse amplitude modulation – simple binary PAM-2 (1bit/symbol) and PAM-4 (2bits/symbol).

Another more radical modulation format under consideration by link researchers is the use of multi-tone signaling. While this type of signaling is commonly used in systems such as DSL modems [49], it is relatively new for high-speed inter-chip communication applications. In contrast with conventional baseband signaling, multi-tone signaling breaks the channel bandwidth into multiple frequency bands over which data is transmitted. This technique has the potential to greatly reduce equalization complexity relative to baseband signaling due to the reduction in per-band loss and the ability to selectively avoid severe channel nulls. Typically, in systems such as modems where the data rate is significantly lower than the on-chip processing frequencies, the required frequency conversion in done in the digital domain and requires DAC transmit and ADC receive front-ends [50,51]. While it is possible to implement high-speed transmit DACs [52], the excessive digital processing and ADC speed and precision required for multi-Gb/s channel bands results in prohibitive receiver power and complexity. Thus, for power-

efficient multi-tone receivers, researchers have proposed using analog mixing techniques combined with integration filters and multiple-input-multiple-output (MIMO) DFEs to cancel out band-to-band interference [53].

Serious challenges exist in achieving increased inter-chip communication bandwidth over electrical channels while still satisfying I/O power and density constraints. As discussed, current equalization and advanced modulation techniques allow data rates near 10Gb/s over severely band-limited channels. However, this additional circuitry comes with a power and complexity cost, with typical commercial high-speed serial I/O links consuming close to 20mW/Gb/s [33,54] and research-grade links consuming near 10mW/Gb/s [14, 55]. The demand for higher data rates will only result in increased equalization requirements and further degrade link energy efficiencies without improved low-power design techniques. Ultimately, excessive channel loss motives investigation into the use of optical links for chip-to-chip applications.

Low-Power Design Techniques

In order to support the bandwidth demands of future systems, serial link data rates must continue to increase. While the equalization and modulation techniques discussed in the previous section can extend data rates over bandwidth limited channels, the power costs incurred with this additional complexity can be prohibitive without improvements in power efficiency. This section discusses low-power I/O architectures and circuit techniques to improve link power efficiency and enable continued inter-chip bandwidth scaling in a manner compliant with limited chip power budgets.

Figure 21 shows the power breakdown of a 4.8Gb/s serial link designed for a fully buffered DIMM system [56], which achieves a power efficiency of 15mW/Gb/s in a 90nm technology. For the design, more than 80% of the power is consumed in the transmitter and the clocking system consisting of the clock multiplication unit (CMU), clock distribution, and RX-PLL used for timing recovery.

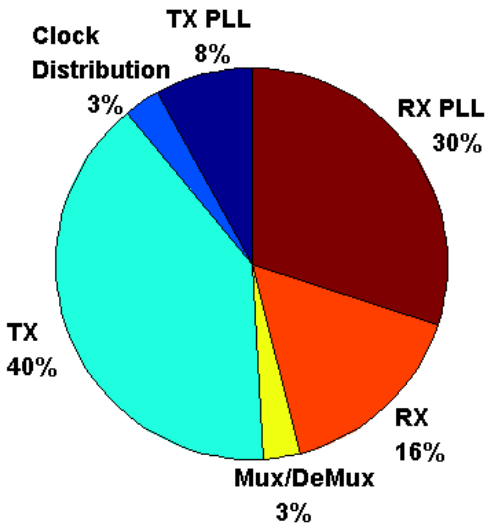


Figure 21 4.8Gb/s serial link power breakdown [56].

Receiver sensitivity and channel loss set the required transmit output swing, which has a large impact on transmitter power consumption. The receiver sensitivity is set by the input referred offset, V_{offset}, minimum latch resolution voltage, $V_{min,}$ and the minimum signal-to-noise ratio (SNR) required for a given BER. Combining these terms results in a total minimum voltage swing per bit of

$$\Delta V_b = \sqrt{SNR}\sigma_n + V_{offset} + V_{\min}, \tag{5}$$

where σ^2_n is the total input voltage noise variance. Proper circuit biasing can ensure that the input referred rms noise voltage is less than 1mV and design techniques can reduce the input offset and latch minimum voltage to the 1-2mV range. Thus, for a BER of 10^{-12} ($\sqrt{SNR} \cong 7$), a sensitivity of less than $20mV_{ppd}$ ($2\Delta V_b$) is achievable.

While this sensitivity is possible, given the amount of variations in nanometer CMOS technology, it is essential to implement offset correction in order to approach this level. Figure 22 shows some common implementations of correction DACs, with adjustable current sources producing an input voltage offset to correct the latch offset and digitally-adjustable capacitors skewing the single-ended regeneration time constant of the latch output nodes to perform an equivalent offset cancellation [30,57].

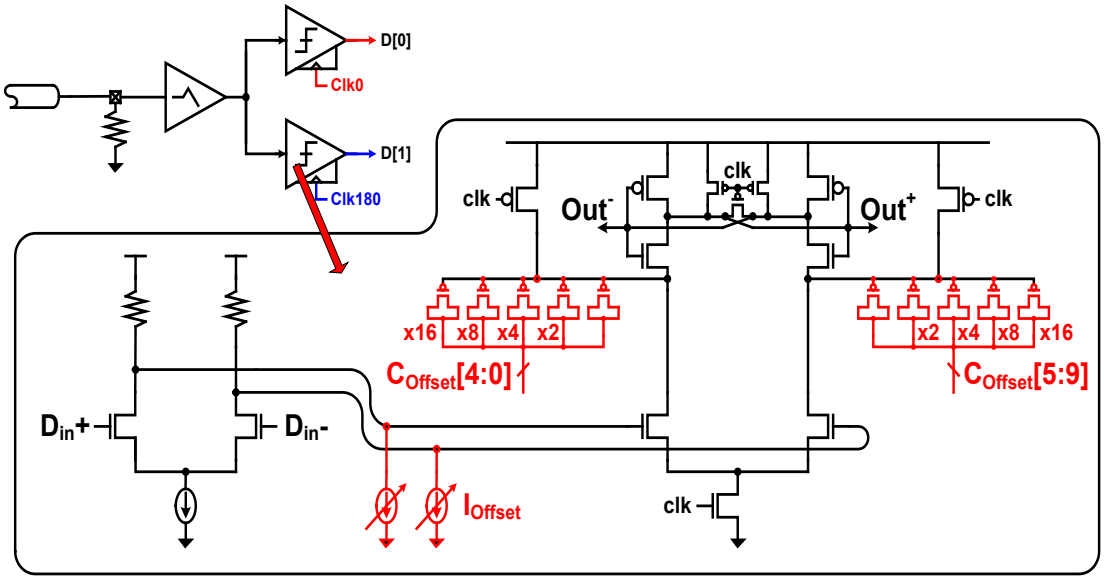


Figure 22 Receiver input offset correction circuitry.

Leveraging offset correction at the receiver dramatically reduces the required transmit swing, allowing the use of low-swing voltage-mode drivers. As discussed in the link overview section, a differentially-terminated voltage-mode driver requires only one-fourth the current of a current-mode driver in order to produce the same receiver voltage swing. Increased output stage efficiency propagates throughout the transmitter, allowing for a reduction in predriver sizing and dynamic switching power. One of the first examples of the use of a low-swing voltage-mode driver was in [9], which achieved 3.6Gb/s at a power efficiency of 7.5mW/Gb/s in a 0.18µm technology. A more recent design [15] leveraged a similar design and achieved 6.25Gb/s at a total link power efficiency of 2.25mW/Gb/s in a 90nm technology.

While low-swing voltage-mode drivers can allow for reduced transmit power dissipation, one drawback is that the swing range is limited due to output impedance control constraints. Thus, for links which operate over a wide data rate or channel loss range, the flexibility offered by a current mode driver can present a better design choice. One example of such a design is [30], which operated over 5-15Gb/s and achieved 2.8-6.5mW/Gb/s total link power efficiency in a 65nm technology while using a scalable current-mode driver. In this design, efficiency gains were made elsewhere by leveraging low-complexity equalization, low-power clocking, and voltage-supply scaling techniques.

With equalization becoming a necessity as data rates scale well past electrical channel bandwidths, the efficiency at which these equalization circuits are implemented becomes paramount. Circuit complexity and bandwidth limitations are present in both transmitter and receiver equalization structures. At the transmitter, increasing FIR equalization tap number presents power costs due to increased flop count and extra staging logic. While at the receiver, decision feedback equalizers implemented in a direct feedback manner [3] have a 1UI critical timing path which must be highly optimized. This critical timing path can be relieved with a loop-unrolling architecture for the first tap [45], but it doubles the number of input comparators and also complicates timing recovery.

Low-complexity active and passive structures are viable equalization topologies that allow for excellent power efficiency. As a receiver input amplifier is often necessary, implementing frequency peaking in this block through RC degeneration is a way to compensate for channel loss with little power overhead. This is observed in recent low-power I/O transceivers [15,30] which have input CTLE structures. Passive equalization structures, such as shunt inductive termination [30] and tunable passive attenuators [58], also provide frequency peaking at close to zero power cost.

While CTLE and passive equalization efficiently compensate for high frequency channel loss, a key issue with these topologies is that the peaking transfer function also amplifies input noise and crosstalk. Also, the input referred noise of cascaded receiver stages is increased when the DC gain of these input equalizers is less than one. This is always the case in passive structures which offer no gain, and occurs often in CTLE structures which must trade-off peaking range, bandwidth, and DC gain. In order to fully compensate for the total channel loss, systems sensitive to input noise and crosstalk may need to also employ higher complexity DFE circuits, which avoid this noise enhancement by using quantized input values to control the polarity of the equalization taps.

A considerable percentage of total link power is consumed in the clocking systems. One key portion of the clocking power is the receiver timing recovery, where significant circuit complexity and power is used in PLL-based and dual-loop timing recovery schemes. High

frequency clock distribution over multiple link channels is also important, with key trade-offs present between power and jitter performance.

As explained earlier in the link overview section, dual-loop CDRs provide the flexibility to optimize for both CDR bandwidth and frequency synthesis noise performance. However, phase resolution and linearity considerations cause the interpolator structures used in these architectures to consume significant power and area. As conventional half-rate receivers using 2x-oversampled phase detectors require two differential interpolators to generate data and quadrature edge sampling clock phases, direct implementations of dual-loop architectures lead to poor power efficiency performance. This has led to architectures which merge the interpolation block into the frequency synthesis loop [15,59,60].

Figure 23 shows an implementation where one interpolator is placed in the feedback path of a frequency synthesis PLL in order to produce a high resolution phase shift simultaneously in all of the VCO multi-phase outputs used for data sampling and phase detection. This is particularly beneficial in quarter or eight-rate systems, as only one interpolator is required independent of the input demultiplexing ratio. Recent power efficient links have further optimized this approach by combining the interpolator with the CDR phase detector [15,60].

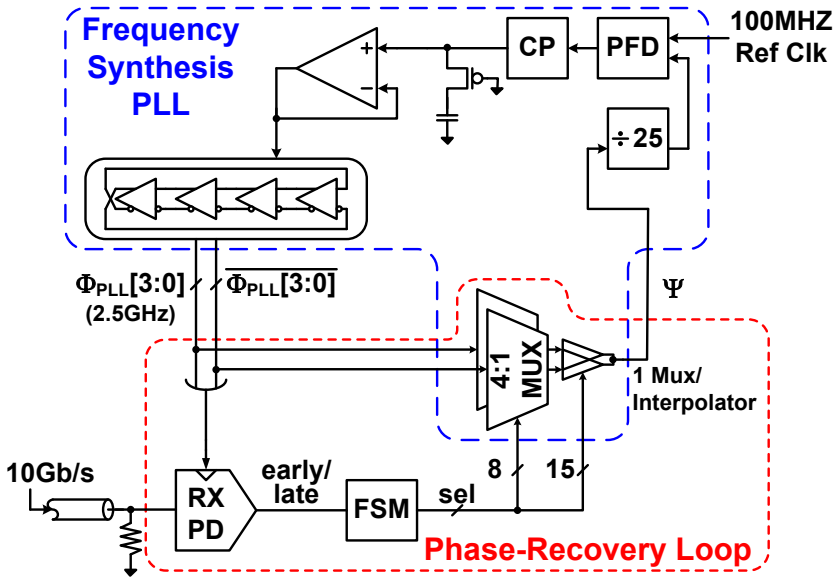
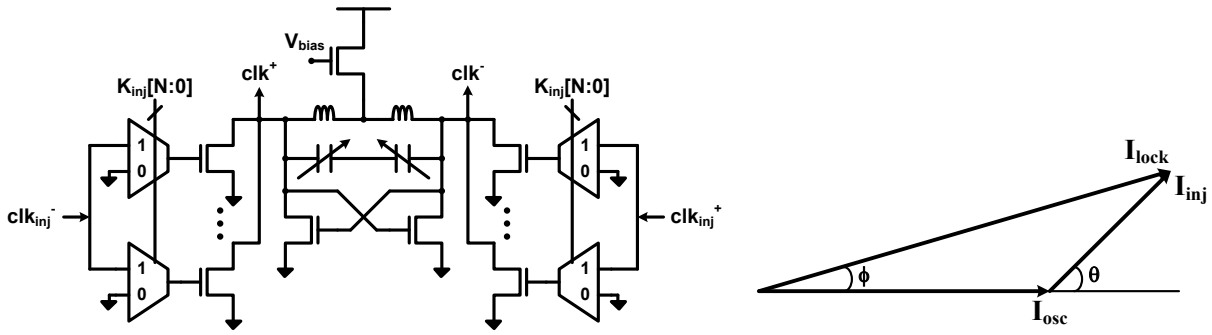


Figure 23 Dual-loop CDR with feedback interpolation [59].

Another low complexity receiver timining recovery technique suitable for forwarded clock links involves the use of injection locked oscillators (Figure 24). In this deskew technique, the forwarded clock is used as the injection clock to a local oscillator at each receiver. A programmable phase shift is developed by tuning either the injection clock strength or the local receiver oscillator free-running frequency. This approach offers advantages of low complexity, as the clock phases are generated directly from the oscillator without any interpolators, and high frequency jitter tracking, as the jitter on the forwarded clock is correlated with the incoming data jitter. Recent low-power receivers have been implemented up to 27Gb/s at an efficiency of 1.6mW/Gb/s with a local LC oscillator [61] and 7.2Gb/s at an efficiency of 0.6mW/Gb/s with a local ring oscillator [62].



(a) (b)

Figure 24 Injection locked oscillator (ILO) deskew: (a) LC ILO [61], (b) ILO vector diagram.

Clock distribution is important in multi-channel serial link systems, as the high frequency clocks from the central frequency synthesis PLL, and the forwarded receive clock in source-synchronous systems, are often routed several mm to the transmit and receive channels. Multiple distribution schemes exist, ranging from simple CMOS inverters to carefully engineered resonant transmission-line structures, which trade-off jitter, power, area, and complexity. There is the potential for significant power cost with active clock distribution, as the buffer cells must be designed with adequate bandwidth to avoid any jitter amplification of the clock signal as it propagates through the distribution network.

One approach to save clock distribution power is the use of on-die transmission lines (Figure 25), which allows for repeaterless clock distribution. An example of this is in the power efficient link architecture of [30], which implemented transmission lines that have roughly 1dB/mm loss and are suitable for distributing the clock several mm. Further improvements in clock power

efficiency are obtained through the use of resonant transmission-line structures. Inductive termination was used to increase transmission-line impedance at the resonance frequency in the system of [15]. This allows for reduced distribution current for a given clock voltage swing. One downside of this approach is that it is narrow-band and not suitable for a system which must operate over multiple data rates.

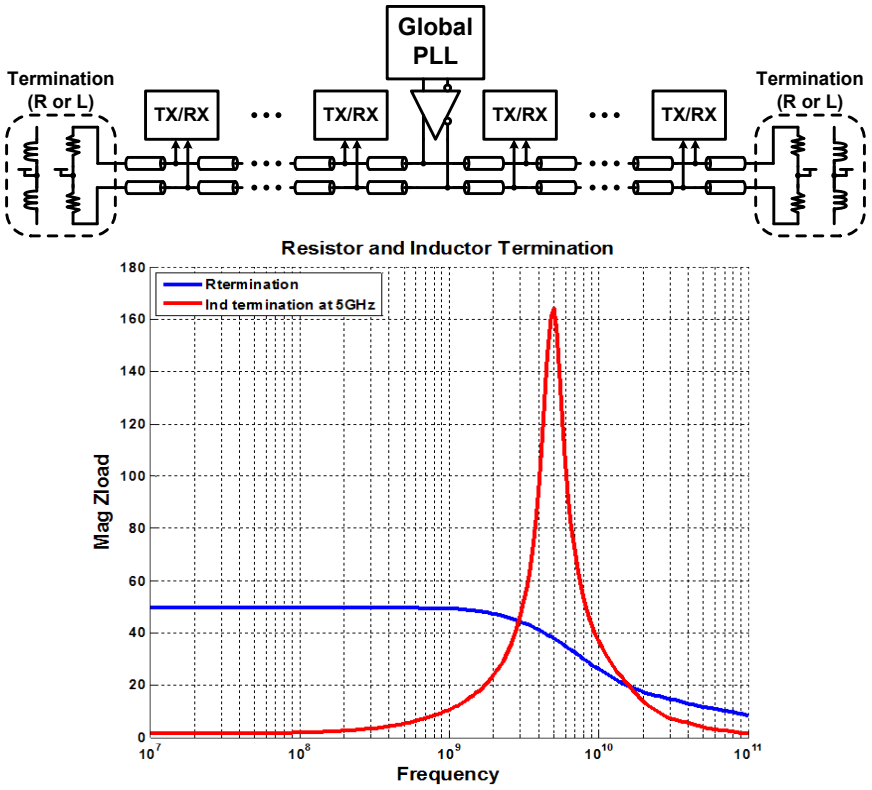


Figure 25 Global transmission-line-based clock distribution.

Serial link circuitry must be designed with sufficient bandwidth to operate at the maximum data rate necessary to support peak system demands. However, during periods of reduced system workload, this excessive circuit bandwidth can translate into a significant amount of wasted power. Dynamic power management techniques extensively used in core digital circuitry [63] can also be leveraged to optimize serial link power for different system performance.

Adaptively scaling the power supply of the serial link CMOS circuitry to the minimum voltage V required to support a given frequency f can result in quadratic power savings at a fixed frequency or data rate since digital CMOS power is proportional to V^2f. Non-linear power savings result with data rate reductions, as the supply voltage can be further reduced to remove

excessive timing margin present with the lower frequency serialization clocks. Increasing the multiplexing factor, e.g. using a quarter-rate system versus a half-rate system, can allow for further power reductions, as the parallel transmitter and receiver segments utilize a lower frequency clock to achieve a given data rate. This combination of adaptive supply scaling and parallelism was used in [17], which used a multiplexing factor of five and a supply range of 0.9-2.5V to achieve 0.65-5Gb/s operation at power efficiency levels ranging from 15-76mW/Gb/s in a 0.25µm process.

Power scaling techniques were extended to serial link CML circuitry in [30], which scaled both the power supply and bias currents of replica-based symmetric load amplifiers [20] to achieve linear bandwidth variation at near constant gain. This design achieved 5-15Gb/s operation at power efficiency levels ranging from 2.7-5mW/Gb/s in a 65nm process.

In order to fully leverage the benefits of these adaptive power management techniques, the regulator which generates the adaptive supply voltage must supply power to the link at high efficiency. Switching regulators provide a high efficiency (>90%) means of accomplishing this [17,29]. Figure 26 shows a block diagram of a general adaptive power-supply regulator, which sets the link supply voltage by adjusting the speed of a reference circuit in the feedback path. This reference circuit should track the link critical timing path, with examples including replicas of the VCO in the link PLL [17] and the TX serializer and output stage pre-driver [30].

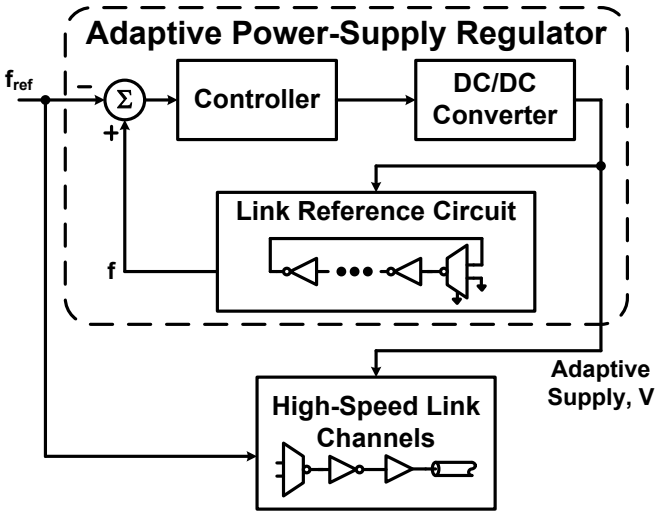


Figure 26 Adaptive power-supply regulator for high-speed link system.

Optical Interconnects

Increasing inter-chip communication bandwidth demand has motivated investigation into using optical interconnect architectures over channel limited electrical counterparts. Optical interconnects with negligible frequency-dependent loss and high bandwidth [64] provide viable alternatives to achieving dramatic power efficiency improvements at per-channel data rates exceeding 10Gb/s. This has motivated extensive research into optical interconnect technologies suitable for high density integration with CMOS chips.

Conventional optical data transmission is analogous to wireless AM radio, where data is transmitted by modulating the optical intensity or amplitude of the high-frequency optical carrier signal. In order to achieve high fidelity over the most common optical channel – the glass fiber, high-speed optical communication systems typically use infrared light from source lasers with wavelengths ranging from 850-1550nm, or equivalently frequencies ranging from 200-350THz. Thus, the potential data bandwidth is quite large since this high optical carrier frequency exceeds current data rates by over three orders of magnitude. Moreover, because the loss of typical optical channels at short distances varies only fractions of dBs over wide wavelength ranges (tens of nanometers) [65], there is the potential for data transmission of several Tb/s without the requirement of channel equalization. This simplifies design of optical links in a manner similar to non-channel limited electrical links. However, optical links do require additional circuits that interface to the optical sources and detectors. Thus, in order to achieve the potential link performance advantages, emphasis is placed on using efficient optical devices and low-power and area interface circuits.

This section gives an overview of the key optical link components, beginning with the optical channel attributes. Optical devices, drivers, and receivers suited for low-power high-density I/O applications are presented. The section concludes with a discussion of optical integration approaches, including hybrid optical interconnect integration where the CMOS I/O circuitry and optical devices reside on different substrates and integrated CMOS photonic architectures which allow for dramatic gains in I/O power efficiency.

Optical Channels

The two optical channels relevant for short distance chip-to-chip communication applications are free-space (air or glass) and optical fibers. These optical channels offer potential performance advantages over electrical channels in terms of loss, cross-talk, and both physical interconnect and information density [64].

Free-space optical links have been used in applications ranging from long distance line-of-sight communication between buildings in metro-area networks [66] to short distance inter-chip communication systems [67-69]. Typical free-space optical links use lenses to collimate light from a laser source. Once collimated, laser beams can propagate over relatively long distances due to narrow divergence angles and low atmospheric absorption of infrared radiation. The ability to focus short wavelength optical beams into small areas avoids many of the crosstalk issues faced in electrical links and provides the potential for very high information density in free-space optical interconnect systems with small 2D transmit and receive arrays [67,71]. However, free-space optical links are sensitive to alignment tolerances and environmental vibrations. To address this, researchers have proposed rigid systems with flip-chip bond chips onto plastic or glass substrates with 45º mirrors [68] or diffractive optical elements [69] that perform optical routing with very high precision.

Optical fiber-based systems, while potentially less dense than free-space systems, provide alignment and routing flexibility for chip-to-chip interconnect applications. An optical fiber, shown in Figure 27, confines light between a higher index core and a lower index cladding via total internal reflection. In order for light to propagate along the optical fiber, the interference pattern, or mode, generated from reflecting off the fiber's boundaries must satisfy resonance conditions. Thus, fibers are classified based on their ability to support multiple or single modes.

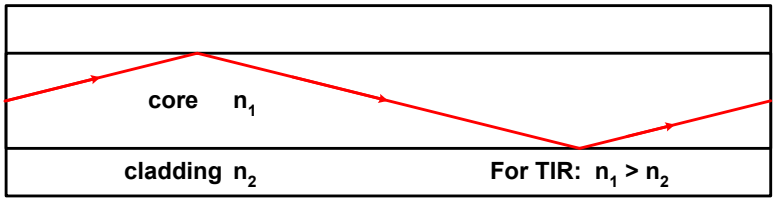


Figure 27 Optical fiber cross-section.

Multi-mode fibers with large core diameters (typically 50 or 62.5μm) allow several propagating modes, and thus are relatively easy to couple light into. These fibers are used in

short and medium distance applications such as parallel computing systems and campus-scale interconnection. Often relatively inexpensive vertical-cavity surface-emitting lasers (VCSEL) operating at wavelengths near 850nm are used as the optical sources for these systems. While fiber loss (~3dB/km for 850nm light) can be significant for some low-speed applications, the major performance limitation of multi-mode fibers is modal dispersion caused by the different light modes propagating at different velocities. Due to modal dispersion, multi-mode fiber is typically specified by a bandwidth-distance product, with legacy fiber supporting 200MHz-km and current optimized fiber supporting 2GHz-km [72].

Single-mode fibers with smaller core diameters (typically 8-10μm) only allow one propagating mode (with two orthogonal polarizations), and thus require careful alignment in order to avoid coupling loss. These fibers are optimized for long distance applications such as links between internet routers spaced up to and exceeding 100km. Fiber loss typically dominates the link budgets of these systems, and thus they often use source lasers with wavelengths near 1550nm which match the loss minima (~0.2dB/km) of conventional single-mode fiber. While modal dispersion is absent from single-mode fibers, chromatic (CD) and polarization-mode dispersion (PMD) exists. However, these dispersion components are generally negligible for distances less than 10km, and are not issues for short distance inter-chip communication applications.

Fiber-based systems provide another method of increasing the optical channel information density – wavelength division multiplexing (WDM). WDM multiplies the data transmitted over a single channel by combining several light beams of differing wavelengths that are modulated at conventional multi-Gb/s rates onto one fiber. This is possible due to the several THz of low-loss bandwidth available in optical fibers. While conventional electrical links which employ baseband modulation do not allow this type of wavelength or frequency division multiplexing, WDM is analogous to the electrical link multi-tone modulation mentioned in the previous section. However, the frequency separation in the optical domain uses passive optical filters [73] rather than the sophisticated DSP techniques required in electrical multi-tone systems.

In summary, both free-space and fiber-based systems are applicable for chip-to-chip optical interconnects. For both optical channels, loss is the primary advantage over electrical channels.

This is highlighted by comparing the highest optical channel loss, present in multi-mode fiber systems (~3dB/km), to typical electrical backplane channels at distances approaching only one meter (>20dB at 5GHz). Also, because pulse-dispersion is small in optical channels for distances appropriate for chip-to-chip applications (<10m), no channel equalization is required. This is in stark contrast to the equalization complexity required by electrical links due to the severe frequency dependent loss in electrical channels.

Optical Transmitter Technology

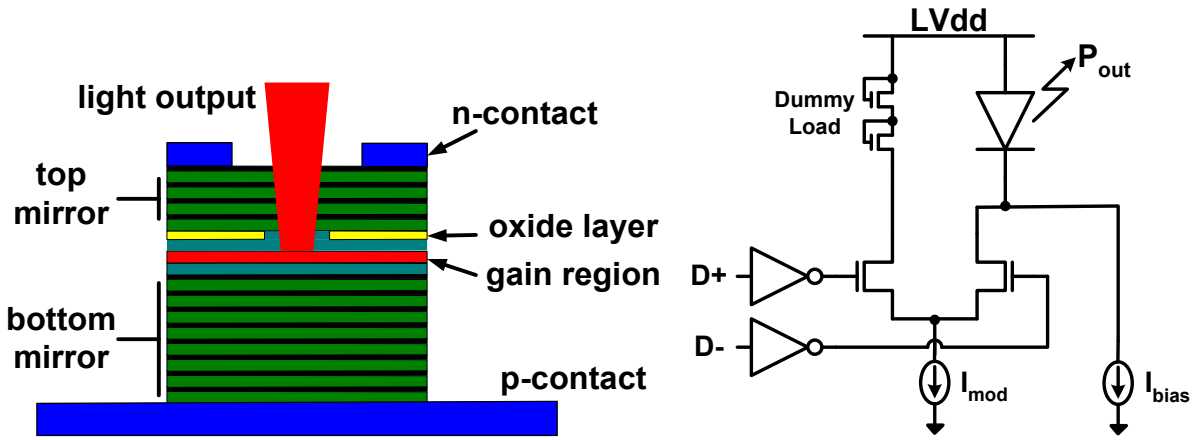
Multi-Gb/s optical links exclusively use coherent laser light due to its low divergence and narrow wavelength range. Modulation of this laser light is possible by directly modulating the laser intensity through changing the laser's electrical drive current or by using separate optical devices to externally modulate laser light via absorption changes or controllable phase shifts that produce constructive or destructive interference. The simplicity of directly modulating a laser allows a huge reduction in the complexity of an optical system because only one optical source device is necessary. However, this approach is limited by laser bandwidth issues and, while not necessarily applicable to short distance chip-to-chip I/O, the broadening of the laser spectrum, or "chirp", that occurs with changes in optical power intensity results in increased chromatic dispersion in fiber systems. External modulators are not limited by the same laser bandwidth issues and generally don't increase light linewidth. Thus, for long haul systems where precision is critical, generally all links use a separate external modulator that changes the intensity of a beam from a source laser, often referred to as "continuous-wave" (CW), operating at a constant power level.

In short distance inter-chip communication, cost constraints outweigh precision requirements, and systems with direct and externally modulated sources have both been implemented. Directly modulated lasers and optical modulators, both electroabsorption and refractive, have been proposed as high bandwidth optical sources, with these different sources displaying trade-offs in both device and circuit driver efficiency. Vertical-cavity surface-emitting lasers [74] are an attractive candidate due to their ability to directly emit light with low threshold currents and reasonable slope efficiencies; however their speed is limited by both electrical parasitics and carrier-photon interactions. A device which doesn't display this carrier speed limitation is the electroabsorption modulator, based on either the quantum-confined Stark effect

(QCSE) [75] or the Franz-Keldysh effect [76], which is capable of achieving acceptable contrast ratios at low drive voltages over tens of nm optical bandwidth. Ring resonator modulators [77,78] are refractive devices that display very high resonant quality factors and can achieve high contrast ratios with small dimensions and low capacitance, however their optical bandwidth is typically less than 1nm. Another refractive device capable of wide optical bandwidth (>100nm) is the Mach-Zehnder modulator [79], however this comes at the cost of a large device and high voltage swings. All of the optical modulators also require an external source laser and incur additional coupling losses relative to a VCSEL-based link.

Vertical-Cavity Surface-Emitting Laser (VCSEL)

A VCSEL, shown in Figure 28a, is a semiconductor laser diode which emits light perpendicular from its top surface. These surface emitting lasers offers several manufacturing advantages over conventional edge-emitting lasers, including wafer-scale testing ability and dense 2D array production. The most common VCSELs are GaAs-based operating at 850nm [80-82], with 1310nm GaInNAs-based VCSELs in recent production [83], and research-grade devices near 1550nm [84]. Current-mode drivers are often used to modulate VCSELs due to the device's linear optical power-current relationship. A typical VCSEL output driver is shown in Figure 28b, with a differential stage steering current between the optical device and a dummy load, and an additional static current source used to bias the VCSEL sufficiently above the threshold current, I_{th}, in order to ensure adequate bandwidth.



(a) (b)

Figure 28 Vertical-Cavity Surface-Emitting Laser: (a) device cross-section (b) driver circuit.

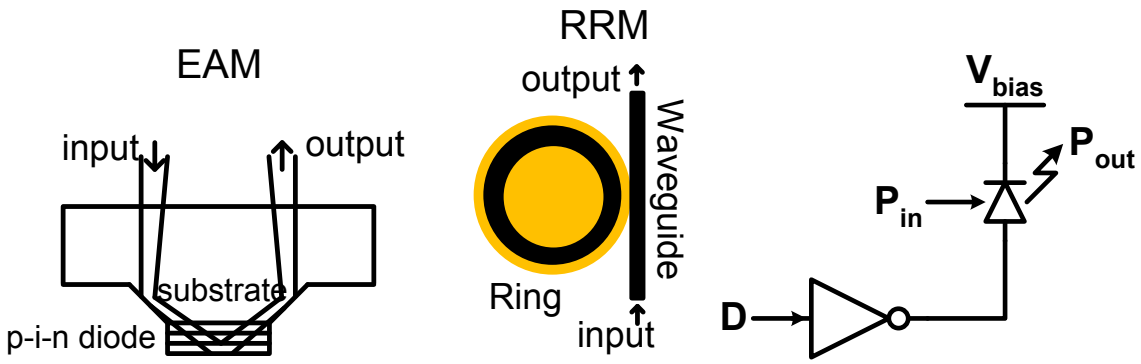
While VCSELs appear to be the ideal source due to their ability to both generate and modulate light, serious inherent bandwidth limitations and reliability concerns do exist. A major constraint in the VCSEL is its bandwidth BW_{VCSEL} which is dependent on the average current I_{avg} flowing through it.

$$BW_{VCSEL} \propto \sqrt{I_{avg} - I_{th}} \tag{6}$$

Unfortunately, output power saturation due to self-heating [85] and also device lifetime concerns [86] restrict excessive increase of VCSEL average current levels to achieve higher bandwidth. As data rates scale, designers have begun to implement simple transmit equalization circuitry to compensate for VCSEL electrical parasitics and reliability constraints [87-89].

Electro-Absorption Modulator (EAM)

An electroabsorption modulator is typically made by placing an absorbing quantum-well region in the intrinsic layer of a *p-i-n* diode. In order to produce a modulated optical output signal, light originating from a continuous-wave source laser is absorbed in an EA modulator depending on electric field strength through electro-optic effects such as the quantum-confined Stark effect or the Franz-Keldysh effect. These devices are implemented either as a waveguide structure [76,90,91] where light is coupled in and travels laterally through the absorbing multiple-quantum-well (MQW) region, or as a surface-normal structure [75,92-94] where incident light performs one or more passes through the MQW region before being reflected out. While large contrast ratios are achieved with waveguide structures, there are challenges associated with dense 2D implementations due to poor misalignment tolerance (due to difficulty coupling into the waveguides) and somewhat large size (>100um) [91]. Surface-normal devices are better suited for high-density 2D optical interconnect applications due to their small size (~ 10x10μm active area) and improved misalignment tolerance [75,93] (Figure 29a). However, as the light only travels a short distance through the absorbing MQW regions, there are challenges in obtaining required contrast ratios.



(a) (b) (c)

Figure 29 External modulators: (a) Electro-absorption modulator [94] (b) Ring-resonator modulator (c) driver circuit.

These devices are typically modulated by applying a static positive bias voltage to the *n*-terminal and driving the *p*-terminal between *Gnd* and *Vdd*, often with simple CMOS buffers (Figure 29c). The ability to drive the small surface-normal devices as an effective lumped-element capacitor offers a huge power advantage when compared to larger waveguide structures, as the CV^2f power is relatively low due to small device capacitance, whereas waveguide structures are typically driven with traveling wave topologies that often require low-impedance termination and a relatively large amount of switching current [91]. However, due to the light only traveling a limited distance in the MQW region, the amount of contrast ratio that surface-normal structures achieve with CMOS-level voltage swings is somewhat limited, with a typical contrast ratio near 3dB for 3V swing [94]. While recent work has been done to lower modulator drive voltages near 1V [75], robust operation requires swings larger than predicted CMOS supply voltages in future technology nodes [6]. One circuit which allows for this in a reliable manner is a pulsed-cascode driver [95], which offers a voltage swing of twice the nominal supply while using only core devices for maximum speed.

Ring Resonator Modulator (RRM)

Ring resonator modulators (Figure 29b) are refractive devices which achieve modulation by changing the interference of the light coupled into the ring with the input light on the waveguide. They use high confinement resonant ring structures to circulate the light and increase the optical path length without increasing the physical device length, which leads to strong modulation effects even with ring diameters less than 20um [96]. Devices which achieve a refractive index change through carrier injection into p-i-n diode structures [96] allow for integration directly

onto silicon substrates with the CMOS circuitry. An alternative integration approach involves fabricating the modulators in an optical layer on top of the integrated circuit metal layers [97], which saves active silicon area and allows for increased portability of the optical technology nodes due to no modifications in the front-end CMOS process. One example of this approach included electro-optic (EO) polymer cladding ring resonator modulators [78] which operate by shuttling carriers within the molecular orbital of the polymer chromophores to change the refractive index.

In addition to using these devices as modulators, ring resonators can also realize optical filters suitable for wavelength division multiplexing (WDM) which allows for a dramatic increase in the bandwidth density of a single optical waveguide [97]. Due to the ring resonator modulator's high selectivity, a bank of these devices can independently modulate several optical channels placed on a waveguide from a multiple-wavelength source. Similarly at the receiver side a bank of resonators with additional "drop" waveguides can perform optical de-multiplexing to separate photdetectors.

Ring resonator modulators are relatively small and can be modeled as a simple lumped element capacitor with values less than 10fF for sub-30μm ring radius devices [78]. This allows the use of non-impedance controlled voltage-mode drivers similar to the EA modulator drivers of Figure 29c. While the low capacitance of the ring resonators allows for excellent power efficiency, the devices do have very sharp resonances (~1 nm) [98] and are very sensitive to process and temperature variations. Efficient feedback tuning circuits and/or improvements in device structure to allow less sensitivity to variations are needed to enhance the feasibility of these devices in high density applications.

Mach-Zehnder Modulator (MZM)

Mach-Zehnder modulators are also refractive modulators which work by splitting the light to travel through two arms where a phase shift is developed that is a function of the applied electric field. The light in the two arms is then recombined either in phase or out of phase at the modulator output to realize the modulation. MZMs which use the free-carrier plasma-dispersion effect in p-n diode devices to realize the optical phase shift have been integrated in CMOS

processes and have recently demonstrated operation in excess of 10Gbps [79,99]. The modulator transfer characteristic with half wave voltage V_π is,

$$\frac{P_{out}}{P_{in}} = 0.5\left(1 + \sin\frac{\pi V_{swing}}{V_\pi}\right) \tag{7}$$

Figure 30 shows a MZM transmitter schematic [99]. Unlike smaller modulators which are treated as lumped capacitive loads, due to MZM length (~1 mm) the differential electrical signal is distributed using a pair of transmission lines terminated with a low impedance. In order to achieve the required phase shift and reasonable contrast ratio, long devices and large differential swings are required; necessitating a separate voltage supply MV_{dd}. Thick-oxide cascode transistors are used to avoid stressing driver transistors with the high supply.

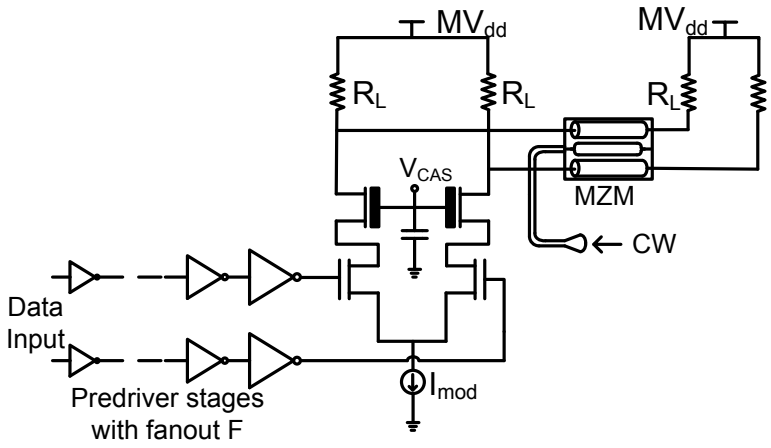


Figure 30 Mach-Zehnder Modulator driver circuit.

Optical Receiver Technology

Optical receivers generally determine the overall optical link performance, as their sensitivity sets the maximum data rate and amount of tolerable channel loss. Typical optical receivers use a photodiode to sense the high-speed optical power and produce an input current. This photocurrent is then converted to a voltage and amplified sufficiently for data resolution. In order to achieve increasing data rates, sensitive high-bandwidth photodiodes and receiver circuits are necessary.

High-speed *p-i-n* photodiodes are typically used in optical receivers due to their high responsivity and low capacitance. In the most common device structures, normally incident light is absorbed in the intrinsic region of width W and the generated carriers are collected at the reverse bias terminals, thereby causing an effective photocurrent to flow. The amount of current generated for a given input optical power P_{opt} is set by the detector's responsivity

$$\rho = \frac{I}{P_{opt}} = \frac{\eta_{pd}\lambda q}{hc} = 8\times 10^{5}\left(\eta_{pd}\lambda\right) \quad (\text{mA/mW}), \tag{8}$$

where λ is the light wavelength and the detector quantum efficiency η_{pd} is

$$\eta_{pd} = 1 - e^{-\alpha W}, \tag{9}$$

where here α is the detector's absorption coefficient. Thus, an 850nm detector with sufficiently long intrinsic width W has a responsivity of 0.68mA/mW. In well designed photodetectors, the bandwidth is set by the carrier transit time τ_{tr} or saturation velocity v_{sat}.

$$f_{3dBPD} = \frac{2.4}{2\pi\tau_{tr}} = \frac{0.45 v_{sat}}{W} \tag{10}$$

From Equations (9) and (10), an inherent trade-off exists in normally incident photodiodes between responsivity and bandwidth due to their codependence on the intrinsic region width W, with devices designed above 10GHz generally unable to achieve maximum responsivity [100]. Therefore, in order to achieve increased data rates while still maintaining high responsivity, alternative photodetector structures are proposed such as the trench detector [101] or lateral metal-semiconductor-metal MSM detectors [102].

In traditional optical receiver front-ends, a transimpedance amplifier (TIA) converts the photocurrent into a voltage and is followed by limiting amplifier stages which provide amplification to levels sufficient to drive a high-speed latch for data recovery (Figure 31). Excellent sensitivity and high bandwidth can be achieved by TIAs that use a negative feedback amplifier to reduce the input time constant [73,103]. Unfortunately, while process scaling has been beneficial to digital circuitry, it has adversely affected analog parameters such as output resistance which is critical to amplifier gain. Another issue arises from the inherent

transimpedance limit [104], which requires the gain-bandwidth of the internal amplifiers used in TIAs to increase as a quadratic function of the required bandwidth in order to maintain the same effective transimpedance gain. While the use of peaking inductors can allow bandwidth extension for a given power consumption [103,104], these high area passives lead to increased chip costs. These scaling trends have reduced TIA efficiency, thereby requiring an increasing number of limiting amplifier stages in the receiver front-end to achieve a given sensitivity and leading to excessive power and area consumption.

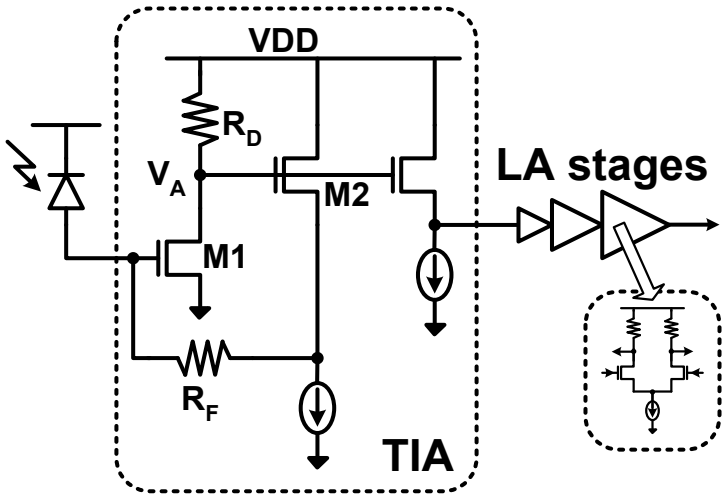


Figure 31 Optical receiver with transimpedance amplifier (TIA) input stage and following limiting amplifier (LA) stages.

A receiver front-end architecture that eliminates linear high gain elements, and thus is less sensitive to the reduced gain in modern processes, is the integrating and double-sampling front-end [105]. The absence of high gain amplifiers allows for savings in both power and area and makes the integrating and double-sampling architecture advantageous for chip-to-chip optical interconnect systems where retiming is also performed at the receiver.

The integrating and double-sampling receiver front-end, shown in Figure 32, demultiplexes the incoming data stream with five parallel segments that include a pair of input samplers, a buffer, and a sense-amplifier. Two current sources at the receiver input node, the photodiode current and a current source that is feedback biased to the average photodiode current, supply and deplete charge from the receiver input capacitance respectively. For data encoded to ensure DC balance, the input voltage will integrate up or down due to the mismatch in these currents. A differential voltage, ΔV_b, that represents the polarity of the received bit is developed by sampling

the input voltage at the beginning and end of a bit period defined by the rising edges of the synchronized sampling clocks Φ[n] and Φ[n+1] that are spaced a bit-period, T_b, apart. This differential voltage is buffered and applied to the inputs of an offset-corrected sense-amplifier [12] which is used to regenerate the signal to CMOS levels.

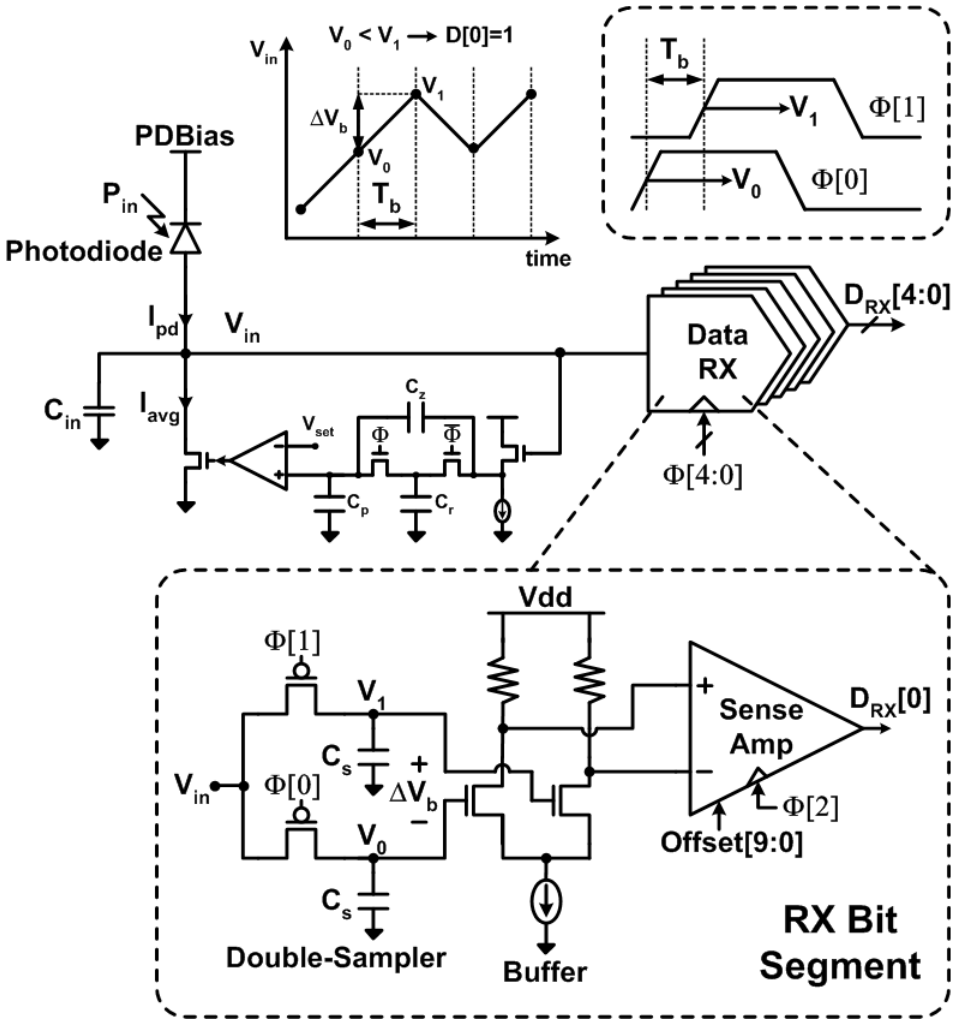


Figure 32 Integrating and double-sampling receiver front-end.

The optimum optical receiver front-end architecture is a function of the input capacitance breakdown between lumped capacitance sources, such as the photodetector, bond pads, and wiring, and the amplifier's input capacitance for a TIA front-end or the sampling capacitor sizes for an integrating front-end. In a TIA front-end, optimum sensitivity is achieved by lowering the input referred noise through increasing the amplifier's input transistors' transconductance up to the point at which the input capacitance is near the lumped capacitance components [106]. While for an integrating front-end, increasing the sampling capacitor size reduces *kT/C* noise at

the cost of also reducing the input voltage swing, with an optimum sampling capacitor size less than the lumped input capacitance [107]. The use of waveguide-coupled photodetectors [102,108] developed for integrated CMOS photonic systems, which can achieve sub-10fF capacitance values, can result in the circuit input capacitance dominating over the integrated photodetector capacitance. This results in a simple resistive front-end achieving optimum sensitivity [108].

Optical Integration Approaches

Efficient cost-effective integration approaches are necessary for optical interconnects to realize their potential to increase per-channel data rates at improved power efficiency levels. This involves engineering the interconnect between the electrical circuitry and optical devices in a manner that minimizes electrical parasitic, optimizing the optical path for low loss, and choosing the most robust and power efficient optical devices.

Hybrid Integration

Hybrid integration schemes with the optical devices fabricated on a separate substrate are generally considered the most feasible in the near-term, as this allows optimization of the optical devices without any impact on the CMOS process. With hybrid integration, methods to connect the electrical circuitry to the optical devices include wirebonding, flip-chip bonding, and short in-package traces.

Wirebonding offers a simple approach suitable for low-channel count stand-alone optical transceivers [109,110], as having the optical VCSEL and photodetctor arrays separated from the main CMOS chip allows for simple coupling into a ribbon fiber module. However, the wirebonds introduce inductive parasitics and adjacent channel crosstalk. Moreover, the majority of processors today are packaged with flip-chip bonding techniques.

Flip-chip bonding optical devices directly to CMOS chips dramatically reduce interconnect parasitics. One example of a flip-chip bonded optical transceiver is [111], which bonds an array of 48 bottom-emitting VCSELs (12 VCSELS at 4 different wavelengths) onto a 48-channel optical driver and receiver chips. The design places a parallel multi-wavelength optical subassembly on top of the VCSEL array and implements WDM by coupling the 48 optical beams into a 12-channel ribbon fiber. A drawback of this approach occurs because the light is

coupled out of the top of the VCSEL array bonded on the CMOS chip. This implies that the CMOS chip cannot be flip-chip bonded inside a package and all signals to the main chip must be connected via wirebonding which can limit power delivery. A solution to this is the approach of [112], which flips the main CMOS chip with bonded VCSELs onto a silicon carrier. This carrier includes thru-silicon vias to connect signals to the main CMOS chip and a cavity for the optical element array. Light is coupled out normally from the chip via 45° mirrors that perform an optical 90° turn into optical waveguides embedded in a board for chip-to-chip routing. While flip-chip bonding does allow for minimal interconnect parasitic, VCSEL reliability concerns necessitate careful thermal management due to close thermal coupling of the VCSEL array and the CMOS chip.

An alternative approach that is compatible with flip-chip bonding packaging is to flip-chip bond both the main CMOS chip and the optical device arrays to the package and use short electrical traces between the CMOS circuitry and the optical device array. This technique allows for top-emitting VCSELs [113], with the VCSEL optical output coupled via 45° mirrors that perform an optical 90° turn into optical waveguides embedded in the package for coupling into a ribbon fiber for chip-to-chip routing.

Integrated CMOS Photonics

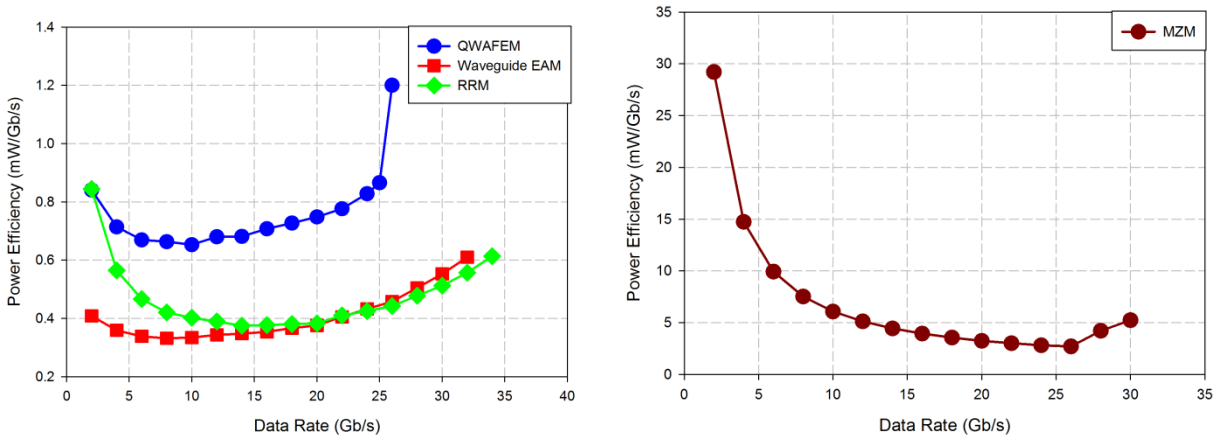
While hybrid integration approaches allow for per-channel data rates well in excess of 10Gbps, parasitics associated with the off-chip interconnect and bandwidth limitations of VCSELs and p-i-n photodetectors can limit both the maximum data rate and power efficiency. Tightly integrating the photonic devices with the CMOS circuitry dramatically reduces bandwidth-limiting parasitics and allows the potential for I/O data rates to scale proportionally with CMOS technology performance. Integrated CMOS photonics also offers a potential system cost advantage, as reducing discrete optical components count simplifies packaging and testing.

Optical waveguides are used to route the light between the photonic transmit devices, filters, and photodetectors in an integrated photonic system. If the photonic devices are fabricated on the same layer as the CMOS circuitry, these waveguides can be realized in an SOI process as silicon core waveguides cladded with SiO_2 [73,96] or in a bulk CMOS process as poly-silicon core waveguides with the silicon substrate etched away to produce an air gap underneath the

waveguide to prevent light leakage into the substrate [114]. For systems with the optical elements fabricated in an optical layer on top of the integrated circuit metal layers, waveguides have been realized with silicon [115] or SiN cores [97].

Ge-based waveguide-coupled photodetectors compatible with CMOS processing have been realized with area less than $5\mu m^2$ and capacitance less than 1fF [102]. The small area and low capacitance of these waveguide photodetectors allows for tight integration with CMOS receiver amplifiers, resulting in excellent sensitivity [108].

Integrated CMOS photonic links are currently predominately modulator-based, as an efficient silicon-compatible laser has yet to be implemented. Electro-absorption [75,76], ring-resonator [96,97], and Mach-Zehnder [73,79] modulators have been proposed, with the different devices offering trade-offs in optical bandwidth, temperature sensitivity, and power efficiency. Figure 33 compares the power efficiency of integrated CMOS photonic links modeled in a 45nm CMOS process [116]. The small size of the electro-absorption and ring-resonator devices results in low capacitance and allows for simple non-terminated voltage mode-drivers which translates into excellent power efficiency. An advantage of the RRM over the EAM is that the RRM may also serve as optical filters for WDM systems. However, the RRM is extremely sensitive to temperature due to the high-Q resonance, necessitating additional thermal tuning. Mach-Zehnder modulators achieve operation over a wide optical bandwidth, and are thus less sensitive to temperature. However, the size of typical devices necessitates transmitter circuitry which can drive transmission lines terminated in a low impedance, resulting in high power. Significant improvements are required for the MZM to obtain lower $V\pi$ and better power efficiency.



(a) (b)

Figure 33 Integrated optical transceiver power efficiency modeled in 45nm CMOS (clocking excluded) (a) QWAFEM EAM [75], Waveguide EAM [76], Polymer RRM [78], (b) MZM [79].

The efficient integration of photonic devices provides potential performance benefits for not only chip-to-chip interconnects, but also for on-chip networks. In order to efficiently facilitate communication between cores in future many-cores systems, on-chip interconnect networks are employed. Electrical on-chip networks are limited by the inverse scaling of wire bandwidth, resulting in shorter repeater distances with CMOS technology scaling, which can severely degrade the efficiency of long global interconnects. Monolithic silicon photonics, which offers high-speed photonic devices, THz-bandwidth waveguides, and wavelength-division-multiplexing (WDM), provides architectures suitable to efficient scale to meet future many-core systems' bandwidth demands. Furthermore, an opportunity for a unified interconnect architecture exists, with the same photonic technology providing both efficient on-chip core-to-core communication and off-chip processor-to-processor and processor-to-memory communication.

Conclusion

Enabled by CMOS technology scaling and improved circuit design techniques, high-speed electrical link data rates have increased to the point where the channel bandwidth is the current performance bottleneck. Sophisticated equalization circuitry and advanced modulation techniques are required to compensate for the frequency dependent electrical channel loss and

continue data rate scaling. However, this additional equalization circuitry comes with a power and complexity cost, which only grows with increasing pin bandwidth.

Future many-core microprocessors are projected to have aggregate I/O bandwidth in excess of 1TBps based on current bandwidth scaling rates of 2-3X every two years [117]. Unless I/O power efficiency is dramatically improved, I/O power budgets will be forced to grow above the typical 10-20% total processor budget and/or performance metrics must be sacrificed to comply with thermal power limits. In the mobile device space, processing performance is projected to increase 10x over the next five years in order to support the next generation of multi-media features [118]. This increased processing translates into aggregate I/O data rates in the hundreds of Gbps, requiring the I/O circuitry to operate at sub-mW/Gbps efficiency levels for sufficient battery lifetimes. It is conceivable that strict system power and area limits will force electrical links to plateau near 10Gb/s, resulting in chip bump/pad pitch and crosstalk constraints limiting overall system bandwidth.

Optical inter-chip links offer a promising solution to this I/O bandwidth problem due to the optical channel's negligible frequency dependent loss. There is the potential to fully leverage CMOS technology advances with transceiver architectures which employ dense arrays of optical devices and low-power circuit techniques for high-efficiency electrical-optical transduction. The efficient integration of photonic devices provides potential performance benefits for not only chip-to-chip interconnects, but also for on-chip networks, with an opportunity for a unified photonic interconnect architecture.

[1] S. R. Vangal *et al.*, "An 80-Tile Sub-100W TeraFLOPS Processor in 65-nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 1, Jan. 2008, pp. 29-41.

[2] B. Landman and R. L. Russo, "On a Pin vs. Block Relationship for Partitioning of Logic Graphs," *IEEE Transactions on Computers*, vol. C-20, no. 12, Dec. 1971, pp. 1469-1479.

[3] R. Payne *et al*., "A 6.25-Gb/s Binary Transceiver in 0.13-μm CMOS for Serial Data Transmission Across High Loss Legacy Backplane Channels," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, Dec. 2005, pp. 2646-2657.

[4] J. F. Bulzacchelli *et al*., "A 10-Gb/s 5-Tap DFE/4-Tap FFE Transceiver in 90-nm CMOS Technology," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, Dec. 2006, pp. 2885-2900.

[5] B. S. Leibowitz *et al*., "A 7.5Gb/s 10-Tap DFE Receiver with First Tap Partial Response, Spectrally Gated Adaptation, and 2nd-Order Data-Filtered CDR," *IEEE International Solid-State Circuits Conference*, Feb. 2007.

[6] Semiconductor Industry Association (SIA), International Technology Roadmap for Semiconductors 2008 Update, 2008.

[7] W. Dally and J. Poulton, *Digital Systems Engineering*, Cambridge University Press, 1998.

[8] K. Lee *et al.*, "A CMOS serial link for fully duplexed data communication," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 4, Apr. 1995, pp. 353-364.

[9] K.-L. J. Wong *et al.*, "A 27-mW 3.6-Gb/s I/O Transceiver," *IEEE Journal of Solid-State Circuits*, vol. 39, no. 4, Apr. 2004, pp. 602-612.

[10] C. Menolfi *et al.*, "A 16Gb/s Source-Series Terminated Transmitter in 65nm CMOS SOI," *IEEE International Solid-State Circuits Conference*, Feb. 2007.

[11] S. Sidiropoulos and M. Horowitz, "A Semidigital Dual Delay-Locked Loop," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 11, Nov. 1997, pp. 1683-1692.

[12] J. Montanaro *et al.*, "A 160MHz, 32b, 0.5W CMOS RISC Microprocessor," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, Nov. 1996, pp. 1703-1714.

[13] A. Yukawa *et al.*, "A CMOS 8-bit high speed A/D converter IC," *IEEE European Solid-State Circuits Conference*, Sep. 1988.

[14] B. Casper *et al*., "A 20Gb/s Forwarded Clock Transceiver in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2006.

[15] J. Poulton *et al.*, "A 14mW 6.25Gb/s Transceiver in 90nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 42, no. 12, Dec. 2007, pp. 2745-2757.

[16] C.-K. Yang and M. Horowitz, "A 0.8-μm CMOS 2.5Gb/s Oversampling Receiver and Transmitter for Serial Links," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 12, Dec. 1996, pp. 2015-2023.

[17] J. Kim and M. Horowitz, "Adaptive-Supply Serial Links with Sub-1V Operation and Per-Pin Clock Recovery," *IEEE Journal of Solid-State Circuits*, vol. 37, no. 11, Nov. 2002, pp. 1403-1413.

[18] M. Horowitz, C.-K. Yang, and S. Sidiropoulos, "High-Speed Electrical Signaling: Overview and Limitations," *IEEE Micro*, vol. 18, no. 1, Jan.-Feb. 1998, pp. 12-24.

[19] A. Hajimiri and T. H. Lee, "Design Issues in CMOS Differential LC Oscillators," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 5, May 1999, pp. 717-724.

[20] J. G. Maneatis, "Low-jitter Process-Independent DLL and PLL based in Self-Biased Techniques," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 11, Nov. 1996, pp. 1723-1732.

[21] S. Sidiropoulos *et al.*, "Adaptive Bandwidth DLLs and PLLs using Regulated Supply CMOS Buffers," *IEEE Symposium on VLSI Circuits*, June 2000.

[22] J. G. Maneatis *et al.*, "Self-Biased High-Bandwidth Low-Jitter 1-to-4096 Multiplier Clock Generator PLL," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 11, Nov. 2003, pp. 1795-1803.

[23] C. R. Hogge, Jr., "A Self Correcting Clock Recovery Circuit," *Journal of Lightwave Technology*, vol. 3, no. 6, Dec. 1985, pp. 1312-1314.

[24] J. D. H. Alexander, "Clock Recovery From Random Binary Signals," *Electronics Letters*, vol. 11, no. 22, Oct. 1975, pp. 541-542.

[25] Y. M. Greshichchev and P. Schvan, "SiGe Clock and Data Recovery IC with Linear-Type PLL for 10-Gb/s SONET Application," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 9, Sep. 2000, pp. 1353-1359.

[26] Y. M. Greshichchev *et al.*, "A Fully Integrated SiGe Receiver IC for 10-Gb/s Data Rate," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 12, Dec. 2000, pp. 1949-1957.

[27] A. Pottbacker, U. Langmann, and H.-U. Schreiber, "A Si Bipolar Phase and Frequency Detector IC for Clock Extraction up to 8Gb/s," *IEEE Journal of Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1747-1751.

[28] K.-Y. K. Chang *et al.*, "A 0.4-4Gb/s CMOS Quad Transceiver Cell Using On-Chip Regulated Dual-Loop PLLs," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 5, May 2003, pp. 747-754.

[29] G.-Y. Wei *et al.*, "A Variable-Frequency Parallel I/O Interface with Adaptive Power-Supply Regulation," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 11, Nov. 2000, pp. 1600-1610.

[30] G. Balamurugan *et al.*, "A Scalable 5-15Gbps, 14-75mW Low-Power I/O Transceiver in 65nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 4, Apr. 2008, pp. 1010-1019.

[31] N. Kurd *et al.*, "Next Generation Intel® Core™ Micro-Architecture (Nehalem) Clocking," *IEEE Journal of Solid-State Circuits*, vol. 44, no. 4, Apr. 2009, pp. 1121-1129.

[32] J. Jaussi *et al*., "A 20Gb/s Embedded Clock Transceiver in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2006.

[33] Y. Moon *et al.*, "A Quad 6Gb/s Multi-rate CMOS Transceiver with TX Rise/Fall-Time Control," *IEEE International Solid-State Circuits Conference*, Feb. 2006.

[34] H. Johnson and M. Graham, *High-Speed Digital Design: A Handbook of Black Magic*, Prentice Hall, 1993.

[35] K. Mistry *et al*., "A 45nm Logic Technology with High-k+ Metal Gate Transistors, Strained Silicon, 9 Cu Interconnect Layers, 193nm Dry Patterning, and 100% Pb-free Packaging," *IEEE International Electron Devices Meeting*, Dec. 2007.

[36] G. Balamurugan *et al*., "Modeling and Analysis of High-Speed I/O Links," *IEEE Transactions on Advanced Packaging*, vol. 32, no. 2, May 2009, pp. 237-247.

[37] V. Stojanović, "Channel-Limited High-Speed Links: Modeling, Analysis and Design," *PhD. Thesis*, Stanford University, Sep. 2004.

[38] A. Sanders, M. Resso, and J. D'Ambrosia, "Channel Compliance Testing Utilizing Novel Statistical Eye Methodology," *DesignCon 2004*, Feb. 2004.

[39] K. Fukuda *et al*., "A 12.3mW 12.5Gb/s Complete Transceiver in 65nm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2010.

[40] W. Dally and J. Poulton, "Transmitter equalization for 4-Gbps signaling," *IEEE Micro*, vol. 17, no. 1, Jan.-Feb. 1997, pp. 48-56.

[41] J. Jaussi *et al.*, "8-Gb/s Source-Synchronous I/O Link With Adaptive Receiver Equalization, Offset Cancellation, and Clock De-Skew," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 1, Jan. 2005, pp. 80-88.

[42] H. Wu *et al.*, "Integrated Transversal Equalizers in High-Speed Fiber-Optic Systems," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 12, Dec. 2003, pp. 2131-2137.

[43] D. Hernandez-Garduno and J. Silva-Martinez, "A CMOS 1Gb/s 5-Tap Transversal Equilzer Based on Inductorless 3rd-Order Delay Cells," *IEEE International Solid-State Circuits Conference*, Feb. 2007.

[44] A. Ho *et al.*, "Common-mode Backplane Signaling System for Differential High-Speed Links," *IEEE Symposium on VLSI Circuits*, June 2004.

[45] K. K. Parhi, "Pipelining in Algorithms with Quantizer Loops," *IEEE Transactions on Circuits and Systems*, vol. 38, no. 7, July 1991, pp. 745-754.

[46] R. Farjad-Rad *et al.*, "A 0.3-μm CMOS 8-Gb/s 4-PAM Serial Link Transceiver," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 5, May 2000, pp. 757-764.

[47] J. Zerbe *et al.*, "1.6Gb/s/pin 4-PAM signaling and circuits for a multidrop bus," *IEEE Journal of Solid-State Circuits*, vol. 36, no. 5, May 2001, pp. 752-760.

[48] J. Stonick *et al.*, "An adaptive pam-4 5-Gb/s backplane transceiver in 0.25-μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 3, Mar. 2003, pp. 436-443.

[49] G. Cherubini *et al.*, "Filter Bank Modulation Techniques for Very High-Speed Digital Subscriber Lines," *IEEE Communications Magazine*, vol. 38, no. 5, May 2000, pp. 98-104.

[50] S. B. Weinstein and P. M. Ebert, "Data Transmission by Frequency-Division Multiplexing Using the Discrete Fourier Transform," *IEEE Transactions on Communications*, vol. 19, no. 5, Oct. 1971, pp. 628-634.

[51] B. Hirosaki, "An Orthogonally Multiplexed QAM System using the Discrete Fourier Transform," *IEEE Transactions on Communications*, vol. 29, no. 7, July 1981, pp. 982-989.

[52] A. Amirkhany *et al.*, "A 24Gb/s Software Programmable Multi-Channel Transmitter," *IEEE Symposium on VLSI Circuits*, June 2007.

[53] A. Amirkhany *et al.*, "Analog Multi-Tone Signaling for High-Speed Backplane Electrical Links," *IEEE Global Telecommunications Conference*, Nov. 2006.

[54] R. Farjad-Rad *et al.*, "0.622-8.0Gbps 150mW Serial IO Macrocell with Fully Flexible Preemphasis and Equalization," *IEEE Symposium on VLSI Circuits*, June 2003.

[55] E. Prete, D. Scheideler, and, A. Sanders, "A 100mW 9.6Gb/s Transceiver in 90nm CMOS for Next-Generation Memory Interfaces," *IEEE International Solid-State Circuits Conference*, Feb. 2006.

[56] D. Pfaff *et al.*, "A 1.8W, 115Gb/s Serial Link for Fully Buffered DIMM with 2.1ns Pass-Through Latency in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2008.

[57] M.-J. E. Lee, W. J. Dally, and P. Chiang, "Low-Power Area-Efficient High-Speed I/O Circuit Techniques," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 11, Nov. 2000, pp. 1591-1599.

[58] D. H. Shin *et al.*, "A 1mW 12Gb/s Continuous-Time Adaptive Passive Equalizer in 90nm CMOS," *IEEE Custom Integrated Circuits Conference*, Oct. 2009.

[59] P. Larsson, "A 2-1600-MHz CMOS Clock Recovery PLL with Low-Vdd Capability," *IEEE Journal of Solid-State Circuits*, vol. 34, no. 12, Dec. 1999, pp. 1951-1960.

[60] T. Toifl *et al.*, "A 72mW 0.03mm^2 Inductorless 40Gb/s CDR in 65nm SOI CMOS," *IEEE International Solid-State Circuits Conference,* Feb. 2007.

[61] F. O'Mahoney *et al.*, "A 27Gb/s Forwarded-Clock I/O Receiver using an Injection-Locked LC-DCO in 45nm CMOS," *IEEE International Solid-State Circuits Conference,* Feb. 2008.

[62] K. Hu *et al.*, "A 0.6mW/Gb/s, 6.4-7.2Gb/s Serial Link Receiver Using Local Injection-Locked Ring Oscillators in 90nm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, Apr. 2010, pp. 899-908.

[63] A. Dancy and A. Chandrakasan, "Techniques for Aggressive Supply Voltage Scaling and Efficient Regulation," *IEEE Custom Integrated Circuits Conference*, May 1997.

[64] D. A. B. Miller, "Rationale and Challenges for Optical Interconnects to Electronic Chips," *Proc. IEEE*, vol. 88, no. 6, June 2000, pp. 728-749.

[65] L. Kazovsky, S. Benedetto, and A. Wilner, *Opitcal Fiber Communication Systems*, Artech House, 1996.

[66] H. A. Willebrand and B. S. Ghuman, "Fiber Optics Without Fiber," *IEEE Spectrum*, vol. 38, no. 8, Aug. 2001, pp. 40-45.

[67] G. A. Keeler *et al*., "The Benefits of Ultrashort Optical Pulses in Optically Interconnected Systems," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, no. 2, Mar.-Apr. 2003, pp. 477-485.

[68] H. Thienpont *et al.*, "Plastic Microoptical Interconnection Modules for Parallel Free-Space Inter- and Intra-MCM Data Communication," *Proc. IEEE*, vol. 88, no. 6, June 2000, pp. 769-779.

[69] M. Gruber, R. Kerssenfischer, and J. Jahns, "Planar-Integrated Free-Space Optical Fan-Out Module for MT-Connected Fiber Ribbons," *Journal of Lightwave Technology*, vol. 22, no. 9, Sep. 2004, pp. 2218-2222.

[70] J. J. Liu *et al*., "Multichannel Ultrathin Silicon-on-Sapphire Optical Interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 9, no. 2, Mar.-Apr. 2003, pp. 380-386.

[71] D. V. Plant *et al*., "256-Channel Bidirectional Optical Interconnect Using VCSELs and Photodiodes on CMOS," *Journal of Lightwave Technology*, vol. 19, no. 8, Aug. 2001, pp. 1093-1103.

[72] AMP NETCONNECT: XG Optical Fiber System (http://www.ampnetconnect.com/documents/XG_Optical_Fiber_Systems_WhitePaper.pdf)

[73] A. Narasimha *et al.*, “A Fully Integrated 4x10Gb/s DWDM Optoelectronic Transceiver in a standard 0.13μm CMOS SOI,” *IEEE International Solid-State Circuits Conference*, Feb. 2007.

[74] K. Yashiki *et al.*, “1.1-μm-range High-speed Tunnel Junction Vertical Cavity Surface Emitting Lasers," *IEEE Photonics Technology Letters*, vol. 19, no. 23, Dec. 2007, pp. 1883–1885.

[75] N.C. Helman *et al.*, “Misalignment-tolerant surface-normal low-voltage modulator for optical interconnects," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 11, no. 2, March/April 2005, pp. 338–342.

[76] J. F. Liu *et al.*, "Ultralow energy, Integrated GeSi electro-absorption modulators on SOI," *5th IEEE Int'l Conf. on Group IV Photonics*, Sep. 2008.

[77] Q. Xu *et al.*, “Micrometre-scale silicon electro-optic modulator,” *Nature*, vol. 435, May 2005, pp. 325-327.

[78] B. A. Block *et al.*, "Electro-optic polymer cladding ring resonator modulators," *Optics Express*, vol. 16, no. 22, Oct. 2008, pp. 18326-33.

[79] A. Liu *et al.*, “High-speed optical modulation based on carrier depletion in a silicon waveguide,” *Optics Express*, vol. 15, issue 2, Jan. 2007, pp. 660-668.

[80] D. Bossert *et al.*, “Production of high-speed oxide confined VCSEL arrays for datacom applications,” *Proc. SPIE*, vol. 4649, June 2002, pp. 142-151.

[81] M. Grabherr *et al.*, “Fabricating VCSELs in a high tech start-up,” *Proc. SPIE*, vol. 4942, Apr. 2003, pp. 13-24.

[82] D. Vez *et al.*, “10 Gbit/s VCSELs for datacom: devices and applications,” *Proc. SPIE*, vol. 4942, Apr. 2003, pp. 29-43.

[83] J. Jewell *et al.*, “1310nm VCSELs in 1-10Gb/s commercial applications,” *Proc. SPIE*, vol. 6132, Feb. 2006, pp. 1-9.

[84] M. A. Wistey *et al.*, "GaInNAsSb/GaAs vertical cavity surface emitting lasers at 1534nm," *Electronics Letters*, vol. 42, no. 5, Mar. 2006, pp. 282-283.

[85] Y. Liu *et al*., "Numerical Investigation of Self-Heating Effects of Oxide-Condfined Vertical-Cavity Surface-Emitting Lasers," *IEEE Journal of Quantum Electronics*, vol. 41, no. 1, Jan. 2005, pp. 15-25.

[86] K. W. Goossen, "Fitting Optical Interconnects to an Electrical World: Packaging and Reliability Issues of Arrayed Optoelectronic Modules," *IEEE Lasers and Electro-Optics Society Annual Meeting*, Nov. 2004.

[87] D. Kucharski *et al.*, "A 20Gb/s VCSEL Driver with Pre-Emphasis and Regulated Output Impedance in 0.13μm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2005.

[88] S. Palermo, A. Emami-Neyestanak, and Mark Horowitz, "A 90nm CMOS 16Gb/s Transceiver for Optical Interconnects," *IEEE Journal of Solid-State Circuits*, vol. 43, no. 5, May 2008, pp. 1235-1246.

[89] A. Kern, A. Chandrakasan, and I. Young, "18Gb/s Optical I/O: VCSEL Driver and TIA in 90nm CMOS," *IEEE Symposium on VLSI Circuits*, June 2007.

[90] H. C. Neitzert *et al*., "InGaAs-InP Superlattice Electroabsorption Waveguide Modulator," *IEEE Photonics Technology Letters*, vol. 7, no. 8, Aug. 1995, pp. 875-877.

[91] H. Fukano *et al.*, "Very-Low-Driving-Voltage Electroabsorption Modulators Operating at 40Gb/s," *Journal of Lightwave Technology*, vol. 24, no. 5, May 2006, pp. 2219-2224.

[92] D. A. B. Miller *et al*., "The Quantum Well Self-Electrooptic Effect Device: Optoelectronic Bistability and Oscillation, and Self-Linearized Modulation," *IEEE Journal of Quantum Electronics*, vol. QE-21, no. 9, Sep. 1985, pp. 1462-1476.

[93] A. L. Lentine *et al*., "High-Speed Optoelectronic VLSI Switching Chip with >4000 Optical I/O Based on Flip-Chip Bonding of MQW Modulators and Detectors to Silicon CMOS,"

IEEE Journal of Selected Topics in Quantum Electronics, vol. 2, no. 1, Apr. 1996, pp. 77-84.

[94] G. A. Keeler, "Optical Interconnects to Silicon CMOS: Integrated Optoelectronic Modulators and Short Pulse Systems," *PhD. Thesis*, Stanford University, Dec. 2002.

[95] S. Palermo and M. Horowitz, "High-Speed Transmitters in 90nm CMOS for High-Density Optical Interconnects," *IEEE European Solid-State Circuits Conference*, Feb. 2006.

[96] M. Lipson, "Compact Electro-Optic Modulators on a Silicon Chip," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 12, no. 6, Nov./Dec. 2006, pp. 1520-1526.

[97] I. A. Young *et al.*, "Optical I/O Technology for Tera-Scale Computing," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 1, Jan. 2010, pp. 235-248.

[98] H.-W. Chen *et al.*, "High speed silicon modulators," *OptoElectronics and Communications Conference*, July 2009.

[99] B. Analui *et al.*, "A Fully Integrated 20-Gb/s Optoelectronic Transceiver Implemented in a Standard 0.13-μm CMOS SOI Technology," *IEEE Journal of Solid-State Circuits*, vol. 41, no. 12, Dec. 2006, pp. 2945-2955.

[100] M. Jutzi *et al.*, "40-Gbit/s Ge-Si Photodiodes," *Silicon Monolithic Integrated Circuits in RF Systems*, Jan. 2006.

[101] M. Yang *et al.*, "A High-Speed, High-Sensitivity Silicon Lateral Trench Photodetector," *IEEE Electron Device Letters*, vol. 23, no. 7, July 2002, pp. 395-397.

[102] M. R. Reshotko *et al.*, "Waveguide Coupled Ge-on-Oxide Photodetectors for Integrated Optical Links," *IEEE/LEOS International Conference Group IV Photonics*, Sep. 2008.

[103] C.-F. Liao and S.-I. Liu, "A 40Gb/s Transimpedance-AGC Amplifier with 19dB DR in 90nm CMOS," *IEEE International Solid-State Circuits Conference*, Feb. 2007.

[104] S. S. Mohan *et al.*, "Bandwidth Extension in CMOS with Optimized On-Chip Inductors," *IEEE Journal of Solid-State Circuits*, vol. 35, no. 3, Mar. 2000, pp. 346-355.

[105] A. Emami-Neyestanal *et al.*, "A 1.6Gb/s, 3mW CMOS Receiver for Optical Communication," *IEEE Symposium on VLSI Circuits*, June 2002.

[106] E. Sackinger, *Broadband Circuits for Optical Fiber Communication*, Wiley-Interscience, 2005.

[107] A. Emami-Neyestanal *et al.*, "CMOS Transceiver with Baud Rate Clock Recovery for Optical Interconnects," *IEEE Symposium on VLSI Circuits*, June 2004.

[108] D. Kucharski *et al.*, "10Gb/s 15mW Optical Receiver with Integrated Germanium Photodetector and Hybrid Inductor Peaking in 0.13µm SOI CMOS Technology," *IEEE International Solid-State Circuits Conference*, Feb. 2010.

[109] D. Kuchta *et al.*, "120Gb/s VCSEL-Based Parallel-Optical Interconnect and Custom 120Gb/s Testing Station," *Journal of Lightwave Technology*, vol. 22, no. 9, Sep. 2004, pp. 2200-2212.

[110] C. Kromer *et al.*, "A 100mW 4x10Gb/s Transceiver in 80nm CMOS for High-Density Optical Interconnects," *IEEE Journal of Solid-State Circuits*, vol. 40, no. 12, Dec. 2005, pp. 2667-2679.

[111] B. Lemoff *et al.*, "MAUI: Enabling Fiber-to-the-Processor with Parallel Multiwavelength Optical Interconnects," *Journal of Lightwave Technology*, vol. 22, no. 9, Sep. 2004, pp. 2043-2054.

[112] C. Schow *et al.*, "A Single-Chip CMOS-Based Parallel Optical Transceiver Capable of 240Gb/s Bidirectional Data Rates," *Journal of Lightwave Technology*, vol. 27, no. 7, Apr. 2009, pp. 915-929.

[113] E. Mohammed *et al.*, "Optical Hybrid Package with an 8-Channel 18GT/s CMOS Transceiver for Chip-to-Chip Optical Interconnect," *Photonics West*, Feb. 2008.

[114] C. Holzwarth *et al.*, "Localized Substrate Removal Technique Enabling Strong-Confinement Microphotonics in Bulk Si CMOS Processes," *Conference on Lasers and Electro-Optics*, May 2008.

[115] L. C. Kimerling *et al.*, "Electronic-Photonic Integrated Circuits on the CMOS Platform," *Silicon Photonics*, Jan. 2006.

[116] A. Palaniappan and S. Palermo, "Power Efficiency Comparisons of Inter-Chip Optical Interconnect Architectures," Accepted for publication in *IEEE Transactions on Circuits and Systems II*, 2010.

[117] F. O'Mahoney *et al.*, "The Future of Electrical I/O for Microprocessors," *IEEE VLSI-DAT*, July 2009.

[118] G. Delagi, "Harnessing Technology to Advance the Next-Generation Mobile User-Experience," *IEEE Internetional Solid-State Circuits Conference*, Feb. 2010.