IEEE Asian Solid-State Circuits Conference
November 14-16, 2011 / Jeju, Korea

Digitally-assisted analog circuits for a 10 Gbps, 395 fJ/b optical receiver in 40 nm CMOS

Philip Amberg, Frankie Liu, Michael Dayringer, Jon Lexau, Dinesh Patil†, Jon Gainsley, Hesam Fathi Moghadam, Elad Alon‡, Xuezhe Zheng*, John E. Cunningham*, Ashok V. Krishnamoorthy* and Ron Ho
Oracle Labs, Redwood Shores, CA, USA; *Oracle Labs, San Diego, CA, USA
†Now with Rambus, Sunnyvale, CA, USA; ‡Also with the University of California, Berkeley, CA, USA

***Abstract*—Digital "assist" circuits can improve the efficiency of traditionally analog circuit blocks, especially as technologies scale to the detriment of analog blocks. We apply some of these techniques to a 10 Gbps optical reciever, and demonstrate 395 fJ/b energy efficiency. Digital calibration blocks wrapped around a simple analog core enabled offset compensation, TIA biasing, and DLL re-timing, and cost negligible performance and power overhead. The assist circuits cost around 40% area overhead.**

I. Introduction

Recent publication activity reflects a growing interest in optical interconnect for large-scale, data-center computing [1]–[3]. Optics over fibers have seen broad commercial adoption for long-distance and high-bitrate communications, and in fact have historically overtaken electronic communications in applications requiring greater than 100 Gbps-m in bandwidth and reach [4]. However, very low-power and short-distance optical interconnects over silicon waveguides and using CMOS-compatible optical devices can also offer signifcant power and integration advantages for silicon-based systems [5]–[7].

To better understand some of the advantages and challenges of silicon nanophotonics, we have been exploring designs that array processors and memories on a single multi-chip package that communicates using optics over waveguides [8], [9]. In our system concept, an array of computing nodes and DRAM chips, called "sites," sit face-up in a silicon lattice, and are electrically bonded to face-down optical device "bridge" chips (see Figure 1). These optical bridge chips, which are driven by CMOS interface circuits in the sites, straddle the sites and the lattice, and steer modulated laser signals into waveguides embedded into the silicon lattice [10], [11]. Because this "macrochip" densely concentrates both processors and memories, and connects them with a high-bandwidth optical network, it should be able to sustain significant performance increases on a wide range of applications [12].

In this application, circuits driving the optical devices must be very low energy, so that high aggregate bandwidth results in manageable total power dissipation. At the same time they must be relatively high performance, to avoid a bandwidth mismatch between on-chip wires and off-chip links. Finally, because they will reside on the processor sites, the circuits must also be scalable across future technology generations.

A. Scaling and optical receiver circuits

Technology scaling has served digital circuits well. Each generation, smaller feature sizes allow designers to build faster circuits for less area and power. However, the impact of scaling on analog circuits has not been as positive. Shorter channel lengths reduce intrinsic gain and supply headroom, complicating the design of power-efficient high-gain amplifiers. Also, manufacturing variability increases transistor mismatch between adjacent devices. In response, designers are building digitally assisted analog circuits, a hybrid approach combining simple analog circuits with digital circuits that correct for non-ideal analog behavior [13]. This enables simple, power efficient circuits whose designs can scale with technology.

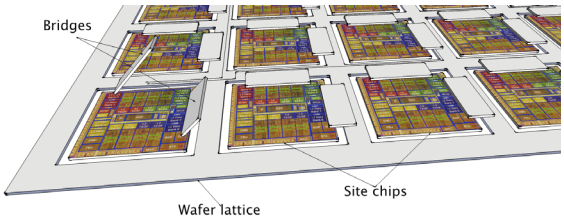

Fig. 1. A "macrochip" array of processors and memories with optical interface bridge chips. Taken from [1].

Optical link circuits in our architecture are conceptually simple. Receivers take a photocurrent from an electrically bonded optical device chip, convert that current to a voltage using a trans-impedance amplifier (TIA), and then compare that voltage to a reference. However, significant optical losses in waveguides, optical couplers, and wavelength-division mux/demux devices [9] increase the receiver's required amplification gain. Moreover, system interconnect energy targets under 0.5 mW/Gbps constrain total energy consumption, and design requirements for future scalability rule out many traditional analog amplifier techniques.

Optical receivers, therefore, present an interesting application for digitally-assisted analog design. In this paper we describe our energy-efficient optical receiver circuits, and how wrapping digital control and calibration blocks around the analog transceiver circuits enabled high-performance and low-power results. The optimizations presented in this paper fit into a broader system-level energy-efficient design [4], [9].

II. Circuits

Figure 2 shows the architecture of our 10 Gbps optical receiver [5]. Higher data rate optical receivers have been demonstrated in older process technologies [14] but 10 Gbps

978-1-4577-1785-7/11/$26.00 ©2011 IEEE

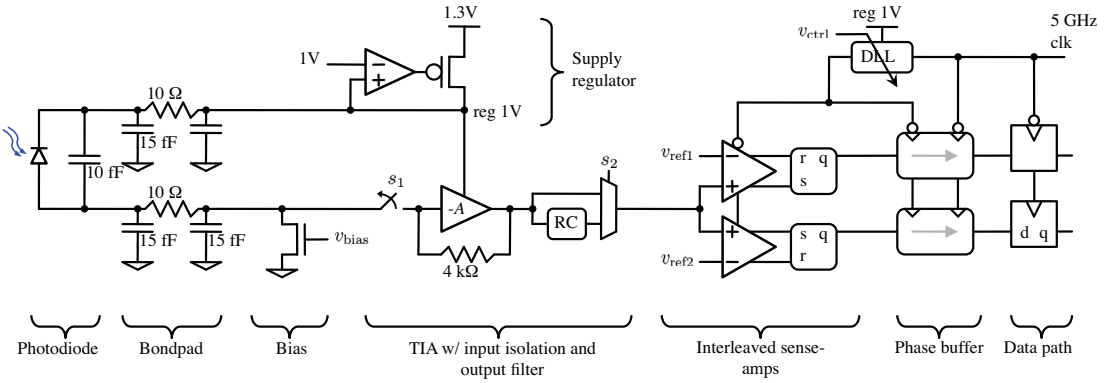

Fig. 2. Optical receiver architecture. An external photodiode (at left) is bonded to a CMOS chip, which uses a TIA and two interleaved sense amplifiers clocked by a DLL. A phase buffer minimizes metastability problems at the interface to a data path.

represents a practical, low-power design point. Simple, power-efficient analog circuits can be used when operating at frequencies much less than f_t of the technology. This data rate also enables the use of relatively inexpensive optical channels. Additionally, the data rate is a low multiple (2-4x) of typical CPU clock rates, so power hungry Ser-Des circuits can be avoided in favor of DDR clocking. An external photonics bridge chip containing photodiodes is bonded to a CMOS chip using fine-pitch (25 µm) solder [11]. The diode is biased by the CMOS chip's supply, and its photocurrent, nominally swinging between 10 to 30 µA, is amplified by a TIA of roughly 4 kΩ gain. The TIA's voltage output is sliced by two interleaved sense amplifiers, each running at 5 Gbps. A delay-locked-loop (DLL) controls the timing of the clock to the sense amps, and the resulting retimed output passes through a metastability phase buffer before moving into a digital data path.

For optimal energy efficiency, the analog circuits in this design have very simple implementations and were instrumented to enable calibration and trimming with digital circuits.

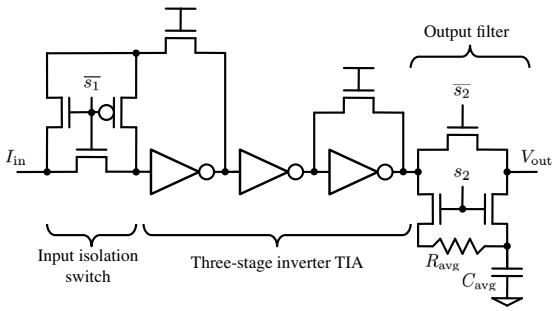

Fig. 3. The transimpedance amplifier and its input and output interfaces

The implementation of the TIA, shown in Figure 3, consists of three cascaded inverters with resistive feedback on the outer two stages. The effect of the feedback is to self-bias the inverters into their high-gain operating region. Additional simple circuitry has been added around the TIA to enable calibration: a three-transistor isolation switch disconnects the TIA input from the photodiode input current during sense amplifier calibration, and an output analog mux passes either the TIA output voltage or an averaged version of the it. The averaged voltage facilitates calibration of the photodiode bias, which helps to mitigate the effects of photodiode dark current.

Following the TIA are two sense amplifiers, based on an edge-triggered latch design [15]. One sense amplifier captures the TIA output on the clock's positive edge, and the other on the negative edge, thus allowing for sense amplifier precharge with full bitrate operation. Both sense amplifiers connect to the TIA output on their positive input, and each sense amplifier has its own reference voltage connected to its negative input. Providing two different reference voltages, each from a dedicated digital-to-analog converter (DAC), enables the sense amplifiers' offset voltages to be individually cancelled.

The sense amplifiers, in deciding whether the input is a logical 0 or 1, must be strobed at the center of the data eye. This timing control is set by a DLL consisting of twenty current-starved inverters, and has a delay adjustment just under 200 ps. For 10 Gbps data, this provides sufficient range to cover any skew between the system clock and the receive data. The DLL control voltage comes from its own DAC.

III. Calibration and trimming

A digital calibration block, shown in Figure 4, wraps around the analog circuits and periodically assists them with calibration and trimming. These tasks consist of compensating each sense amplifier's offset, setting the DC bias on the TIA input, and centering the DLL output in the data eye. The periodicity of this "refresh" operation must be sufficient to avoid circuit drift (*e.g.* from leakage or thermal changes) between calibrations, but infrequent enough so that its performance and energy impact can be economically amortized.

Initiating a refresh cycle requires putting a transmitter into a repeated 0101... sequence, and then running each DAC on the receiver through its calibration process. Because this requires cooperation between all pairs of transmitters and receivers in the system, refresh will be signaled by a system

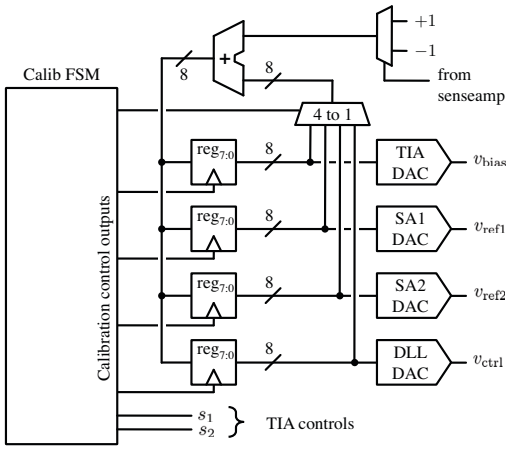

Fig. 4. Control and calibration block

service processor that broadcasts a global command to all sites in a macrochip, using a relatively slow and low-bandwidth electrical control path interconnect on the silicon lattice. This global synchronization process is feasible largely due to a macrochip's relatively small size, around 20 cm on edge.

The calibration block in Figure 4 outputs four control voltages from separate 8b resistor ladder DACs, each powered from a regulated supply to minimize noise. A key observation for our system is that any adjustment to the analog circuits–correcting input offset voltage, adjusting TIA bias, or changing DLL output phase–will be reflected in the digital output of the TIA's sense amplifiers. Therefore the calibration block adjusts each DAC according to the sense amplifier output: if the output is a logical 1, the DAC increments one step; if the output is a logical 0, the DAC decrements one step. Each increment is about 4 mV, as the 8b DAC splits a 1 V supply into 256 steps.

Each calibration process is thus an iterative sequence that successively increments a control voltage until the sense amplifier changes its output value, and begins to dither in a limit cycle. The calibration finite state machine (FSM) can detect this change in output and complete the calibration task.

The full calibration procedure consists of several steps. In our system we plan to run it every 1 ms. It takes 8.2 µs, and hence presents less than 1% overhead on the link throughput.

A. Compensate sense amplifier offsets

Compensating input-referred offset on a differential sense amplifier requires first setting the signal input to an average level, and then sweeping the reference input until the sense amplifier sees no effective difference between inputs.

To this, we first disconnect the photodiode from the TIA by driving switch s_1 low, forcing the TIA into its high-gain region. The output of the TIA, which drives both sense amplifiers, should therefore be held at its switching threshold.

We now set the reference input to 0 V (when the sense amplifier output will be a logical 1) and successively increase it until the output flips to a logical 0. At this limit cycle, the reference voltage is set to properly compensate for the amplifier's input-referred offset (within 4 mV). Because there are two sense amplifiers, we run through this process twice.

B. Calibrate TIA input level

In normal operation, the photodiode output current has a DC level, due in part to the transmitter's finite optical extinction ratio, and in part to dark current in the photodiode itself. Ideally, the bias NMOS shunt device will remove all of this DC current, leaving only the AC signal to reach the TIA. This requires calibrating the control voltage v_{bias}.

We do this by first reconnecting the photodiode to the TIA (by setting s_1 high), so that the transmitter's fixed 0101... pattern can be received. We next filter the output through a large RC pole (by setting switch s_2 high) to generate an average value from the TIA.

We now set the control voltage v_{bias} to 0 V. Because the TIA is seeing the signal swing overlaid on a DC offset, its average output value will be higher than its switching threshold, driving the sense-amplifier to a logical 1. As we successively increment the bias voltage v_{bias}, more and more of the DC offset current will be steered away from the TIA input. Eventually, at the ideal bias voltage, the sense amplifier will flip to a logical 0 output.

C. Calibrate sense amplifier timing

To pick the correct DLL delay to center the sense amplifier strobe in the data eye, we first we select the non-averaged TIA output (by setting s_2 low) and then shift the sense amplifier clock into quadrature. The delay line starts at its minimum delay and is gradually increased. While the transmitter is sending a 0101... pattern, each of the two sense amplifiers sees a constant input due to their interleaving.

When the clock edge reaches the data edge, the sense amplifier output switches state and stops the timing calibration. Due to random jitter, the first output change may not reflect the true center of the data edge, and so a more robust algorithm would use a running average of the sense amplifier outputs. The clock is then switched out of quadrature, moving the clock edge to the approximate middle of the data interval.

IV. Results

The receiver circuit outlined above in Figure 2 was built in a 40 nm CMOS technology. The chip was bonded to an optical chip containing a Ge photodiode with a measured responsivity of 0.8 A/W, and tested using a 2^{31}-1 PRBS sequence. The packaged CMOS and optics chip is shown in Figure 5. As reported previously, at -15 dBm sensitivity, the receiver consumed a total of 3.95 mW while running at 10 Gbps, and with reasonable measured data eyes as shown in Figure 6 [6]. This measured energy efficiency of 395 fJ/b is due in large part to the simple analog signal path and surrounding digital trimming and calibration circuits.

The area for the digital assist circuitry is dominated by registers in the FSM, and was 2800 µm^2, or 40% of the total receiver area. Assuming it scales with technology, while the analog circuit area will not, the area overhead for the assist

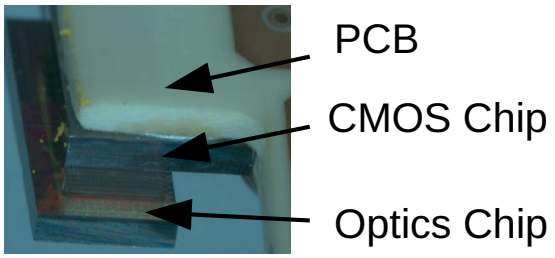

Fig. 5. Photo of packaged CMOS and optics chips.

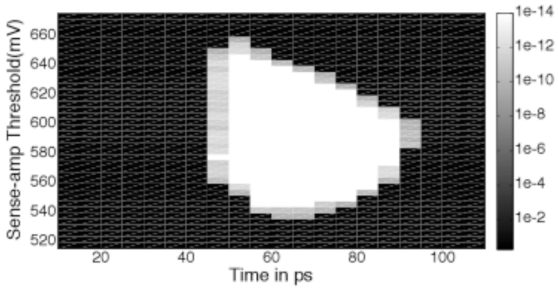

Fig. 6. Measured data eye at -15 dBm sensitivity (24 µA average current). Different shades represent different bit error rates. Adapted from [6].

circuits can drop to 25% at 28 nm and 15% at 20 nm. However, as such scaled technologies become increasingly adversarial to analog circuits, we may choose to add more calibration and trimming circuitry to compensate, resulting in a slower improvement in areal efficiency.

The digital assist circuitry's power overhead is determined by the calibration time constant, because that sets how often the FSM operates. For example, if the sense amplifier's input-referred offsets and photodiode DC currents never changed, calibration would be necessary only once. However, temperature variations will significantly alter the photodiode DC current, and may also affect the sense amplifier offset voltage (by changing the clock's slew rate). Hence, recalibration will need to happen substantially faster than the dominant thermal time constant, motivating the 1 ms recalibration cycle.

At a sufficiently slow calibration time constant, the digital assist circuitry is used infrequently, and power savings can be realized compared to an all-analog implementation. In addition, the digital assist circuitry can also be clocked slower than the core clock. In this design, the calibration block was clocked at 156 MHz, 32× slower than the core 5 GHz clock rate. This significantly relaxed timing constraint allows the design to lie on the lowest energy region of the pareto-optimal energy-delay curve, resulting in lower energy (dynamic and static) usage for calibration.

Table I shows the power used in the different blocks of the calibration system. The scalable power column indicates the dissipation during the 8.2 µs calibration process; over a 1 ms calibration time constant the effective added power is 1.12 nW. The static power column indicates the power drawn by this block continuously due to leakage or power used during normal operation. Note that at a 1 ms calibration period the digital calibration loop's static power greatly exceeds the dynamic power, implying a large benefit to power gating those blocks during normal (non-calibration) operation. The delay line consumes the vast majority of the power in the receiver, because it is running at the full 5 GHz clock rate and is passing a clock signal with full activity factor. Its 1.78 mW is already counted in the 3.95 mW total power discussed above.

TABLE I
POWER OVERHEAD FOR DIGITAL CALIBRATION

Circuit	Power (scalable)	Power (static)
Digital calibration loops	136 µW	100 µW
DAC (x4)	n/a	376 µW
Delay line	n/a	1.78 mW

V. SUMMARY

We built a high-performance optical receiver using simple, low-power analog parts. These analog designs were enabled through the use of digital calibration circuitry that compensates for non-ideal analog behavior. Using digital circuits to periodically tune analog circuit performance is an energy-efficient way to increase performance and will likely scale with technology.

VI. ACKNOWLEDGEMENTS

This work was funded in part by DARPA under contract HR0011-08-9-0001 with Oracle. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. Approved for public release, distribution unlimited.

REFERENCES

[1] R. Ho *et al.*, "Optical interconnect for high-end computer systems," *IEEE Design and Test of Computers*, no. 4, pp. 10–19, 2010.
[2] C. Schow *et al.*, "225 Gb/s bi-directional integrated optical PCB link," in *OFC*, Mar. 2011.
[3] M. Tan *et al.*, "Low cost injection molded 120 Gbps optical backplane," in *OFC*, Mar. 2011.
[4] A. Krishnamoorthy *et al.*, "Progress in low-power switched optical interconnects," *IEEE JSTQE*, vol. 17, no. 2, pp. 357–376, Apr. 2011.
[5] F. Liu *et al.*, "10 Gbps, 530 fJ/b optical transceiver circuits in 40 nm CMOS," in *IEEE Symp. VLSI Ckts.*, Jun. 2011.
[6] X. Zheng *et al.*, "Ultra-efficient 10 Gb/s hybrid integrated silicon photonic transmitter and receiver," *Optics Express*, vol. 19, no. 6, pp. 5172–5186, Mar. 2011.
[7] ——, "Ultra-low power arrayed CMOS silicon photonic transceivers for an 80 Gbps WDM optical link," in *OFC*, Mar. 2011.
[8] R. Ho *et al.*, "Circuits for silicon photonics on a macrochip," in *IEEE ASSCC*, Nov. 2009, pp. 17–20.
[9] A. Krishnamoorthy *et al.*, "Computer systems based on silicon photonic interconnects," *Proc. IEEE*, vol. 97, no. 7, pp. 1337–1361, Jul. 2009.
[10] J. Cunningham *et al.*, "Integration and packaging of a macrochip with silicon nanophotonic links," *IEEE JSTQE, in press*, 2011.
[11] H. Thacker *et al.*, "Hybrid integration of silicon nanophotonics with 40nm-cmos vlsi drivers and receivers," in *IEEE ECTC*, Jun. 2011.
[12] P. Koka *et al.*, "Silicon-photonic network architectures for scalable, power-efficient multi-chip systems," in *ISCA*, Jun. 2010, pp. 117–128.
[13] B. Murmann, "Digitally assisted analog circuits," *IEEE Micro*, pp. 38–47, Mar. 2006.
[14] C.-F. Liao *et al.*, "40 Gb/s transimpedance-AGC amplifier and CDR circuit for broadband data receivers in 90 nm CMOS," *IEEE JSSC*, vol. 43, no. 3, pp. 642–655, Mar. 2008.
[15] D. Dobberpuhl, "Design of a high-performance low-power microprocessor," in *ISLPED*, Aug. 1996, pp. 11–16.