A 1.8-pJ/bit 16x16-Gb/s Source Synchronous Parallel Interface in 32nm SOI CMOS with Receiver Redundancy for Link Recalibration

Timothy O. Dickson, Yong Liu, Ankur Agrawal, John F. Bulzacchelli, Herschel Ainspan,
Zeynep Toprak-Deniz, Benjamin D. Parker, Mounir Meghelli, and Daniel J. Friedman
IBM T. J. Watson Research Center, Yorktown Heights, NY, USA

***Abstract*— A 16x16-Gb/s source-synchronous I/O is reported in 32nm SOI CMOS. The bus-level receiver includes redundant RX lanes to enable lane recalibration with reduced power and area overhead. The I/O also includes an 8:1 TX serializer with 8-phase clocking, and an active-inductor-based RX CTLE whose outputs form current mirrors with the inputs of the RX samplers. A phase rotator based on current-integrating phase interpolator cores is described, with architecture and circuit improvements to performance as compared to prior art. 16-Gb/s link measurements over Megtron-6 traces demonstrate efficiencies of 1.8pJ/bit (0.75" traces) and 1.9pJ/bit (10" traces) with >30% timing margin, with the TX, RX, and PLL operating from 1V supplies.**

I. Introduction

High-density, power efficient serial transceivers continue to be a critical need for high-performance computing systems. One obstacle in achieving high power efficiency and low area is the overhead associated with calibration circuitry. In dynamically managed systems, I/O circuitry must be periodically recalibrated to make the links robust against changes in operating conditions. This often requires additional receiver samplers in parallel with the primary data samplers dedicated to background calibration. Lane redundancy techniques have been introduced to amortize this overhead across a parallel bus interface [1] but require an extra channel to send known patterns during periodic recalibrations. This paper presents a two byte 16-Gb/s parallel source synchronous interface with a receiver calibration scheme that does not require redundant links.

II. Architecture and Circuit Details

A. Link Overview and Receiver Calibration

The bus-level link architecture is depicted in Fig. 1. The transmitter follows the architecture described in [2] and includes 16 data lanes, a data spare, and a half-rate forwarded clock. Differential 8-GHz half-rate clocks are derived from a 16-GHz hybrid analog/digital PLL [3] and distributed across 1.5-mm long transmission lines using a low-skew (~2ps) distribution scheme [2]. In the receiver, data is routed from the input termination network to up to three RX lanes to enable periodic recalibration of each data lane as follows. As an RX lane is taken off-line for recalibration, a neighboring RX lane is reconfigured to accept its input data such that both RX lanes receive the same data input. During recalibration of the first lane, slice levels and clock phases to that lane can be adjusted without disturbing data reception in the neighboring lane. Data from both receivers is sent to a logic macro such that correct data decisions from the neighboring RX lane can be used to qualify data decisions in the RX lane being calibrated. Routing input data to a third RX lane makes the scheme robust in the presence of mechanical failures, as the TX driving a faulty link can be disabled and the spare switched in. While the spare TX can be disabled, all 17 data receivers are always operational: 16 receive data, while one is being recalibrated. Lane recalibration is performed in a round-robin fashion.

The forwarded clock is detected by a CML receiver and converted to CMOS levels. I/Q clock signals are then applied to two phase rotators. The phase rotators are based on the current-integrating interpolators described in [2, 4], but incorporate design enhancements to improve power supply rejection, linearity, and power efficiency. The dual phase rotator architecture generates a single global sampling phase for the 16 active RX lanes, plus an adjustable clock phase to measure timing margins for the lane being recalibrated. Similar to the approach described in [1], the role of the two rotators can be swapped at the end of a bus calibration cycle. The two differential clock phases C2A

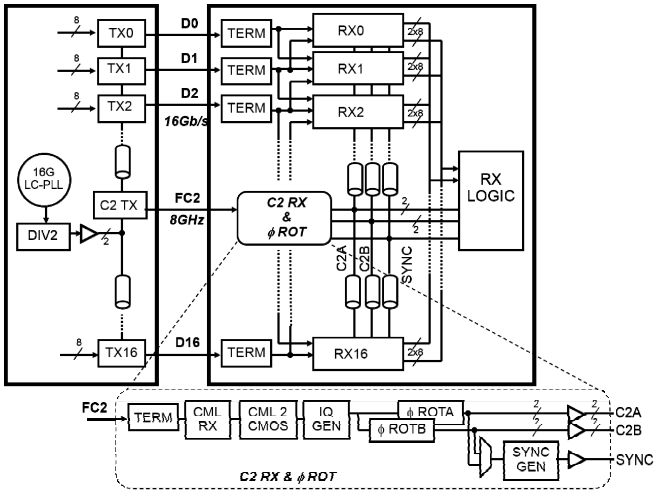

Figure 1: TX & RX bus-level architectures, and clock receiver/rotation details.

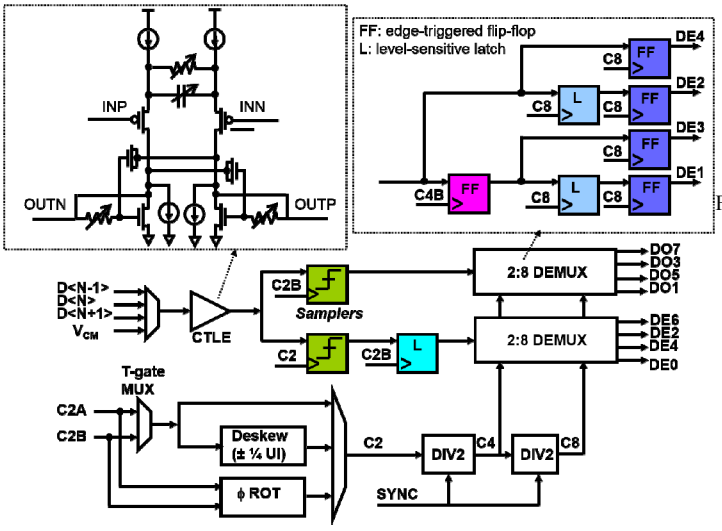

Figure 2: RX data path architecture.

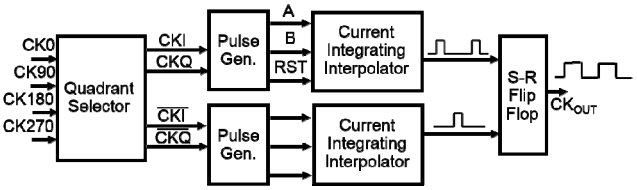

Figure 3: Phase rotator architecture.

and C2B are distributed to the RX lanes using the same low-skew distribution network as in the TX. In the proposed recalibration scheme, data from the two neighboring RX lanes that receive the same data input must be synchronized. The dual global phase rotators are time-aligned at start-up, and dividers within the RX lanes are synchronized as detailed below by sending a SYNC pulse across the bus using the same low-skew transmission line distribution techniques employed in the clock distribution.

B. Receiver Data Path

To enable the proposed recalibration scheme, a 4-input MUX at the lane input selects data from one of three adjacent channels as described in the receiver data lane diagram in Fig. 2. A common-mode voltage V_{CM} is applied to the fourth input of the MUX such that a differential zero can be applied to the data path for offset calibration. The data path employs a CTLE with active inductor and shunt degeneration, which achieves 9dB of peaking at 8GHz. The active inductor forms a current mirror with the input NFETs used in the half-rate samplers. This processing of signals in the current domain mitigates the loss of DC voltage gain inherent to the active inductor topology. The DC gain of the data path is also enhanced by placing current sources in parallel with the active inductor load, which increases the transconductance of the CTLE input stage relative to the load impedance. After sampling, data is deserialized by a 2:8 DEMUX. Low latency is achieved in the DEMUX by directly sampling the latest data bit. As in our previous work [2], three lane deskew modes are implemented to support a range of links from low-skew MCM applications to higher skew PCB links. The SYNC pulse mentioned earlier flushes the CMOS dividers with known bit values such that all dividers are started in the same state. During a handoff between the two clocks at the beginning of a lane calibration, glitchless switching is achieved by first aligning the phase from the adjustable (calibration) phase rotator to match that of the rotator producing the data sampling clock. Once the clock phases are aligned, a transmission gate MUX passes both clocks into the receiver macro for a short time before finally switching completely to the calibration clock. A similar technique is employed after the lane has been recalibrated to restore the sampling clock. This procedure preserves the phase integrity of the clock signal into the divider such that divider synchronization is maintained across the bus.

C. Phase Rotator

The RX phase rotator is based on the current-integrating interpolator cores described in [2, 4] and incorporates design enhancements to improve both power efficiency and linearity. The phase rotator operates on quadrature input clocks, as seen from the high-level block diagram in Figure 3. After quadrant selection, clocks CKI and CKQ are sent to a pulse generation circuit, which derives early and late signals A and B along with a reset signal RST. These signals are applied to the current-integrating phase interpolator core depicted in Figure 4. The early and late signals trigger the integration of currents from a current-steering DAC, along with small quiescent currents I_Q, through a steering network as will be described shortly. As with our previous work [2], the early pulse edge triggers the integration of a variable current I_1 from the DAC, while the late pulse triggers integration of the remainder of the DAC current $I_1 + I_2$. This action results in an integrator output waveform that crosses a threshold level V_{TH} with a programmable amount of delay, as seen in the timing diagram of Figure 5. In contrast to our prior work that relied on an inverter threshold, this work improves robustness by using a differential comparator to detect the threshold crossing and ultimately trigger the S-R flip-flop of Figure 3. A replica integrator is also employed that only operates on the quiescent currents, such that their contribution to the differential signal (Fig. 5) is cancelled. Signals developed across the integration and replica load capacitors C_{INT} are level-shifted via switched series capacitors C_S [5] to ensure proper common-mode levels at the comparator input. The use of a differential comparator and replica integrator also reduces the sensitivity of the circuit to power supply noise. In order to maximize the integrator reset time after the triggering of the comparator, the clock output signal CKO is also used to self-reset the integrators via pull-up devices M1 and M2. A second reset signal

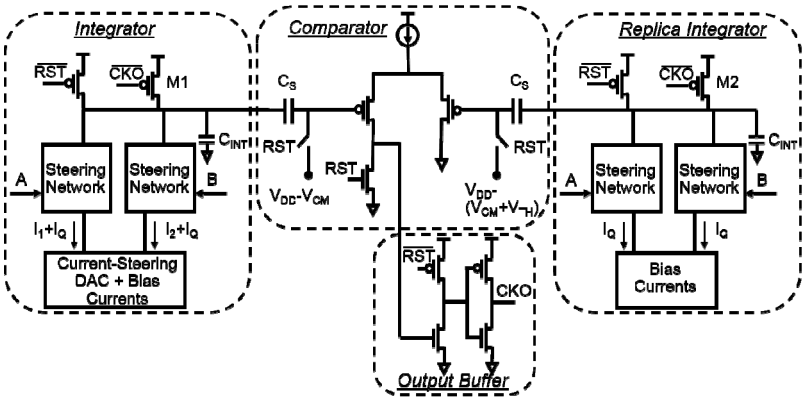

Figure 4: Current-integrating phase interpolator.

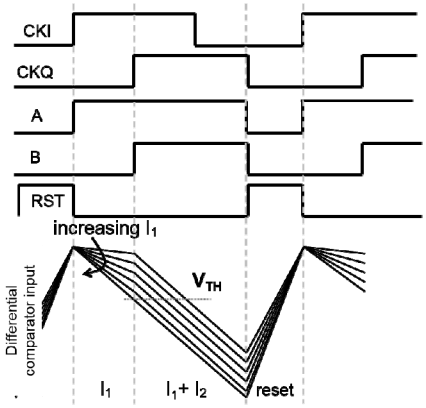

Figure 5: Simplified timing diagram of the current-integrating phase interpolator.

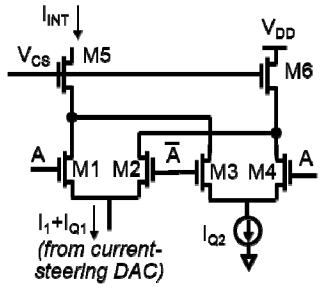

Figure 6: Details of the steering network within the phase interpolator.

RST is derived by the pulse generator from the falling edge of CKQ to enhance the integrator resetting. Note that for clarity, Figure 5 only depicts this resetting action, and not the self-resetting action due to feedback from CKO.

Details of the steering network within the interpolator core are illustrated in Figure 6. The figure shows the operation of a steering network triggered by early signal A, although identical steering networks are employed in the path triggered by late signal B, as well as in the replica integrator. Once triggered by A, current from the DAC as well as a small quiescent current I_{Q1} is steered to the integration node. The inclusion of I_{Q1} avoids non-idealities

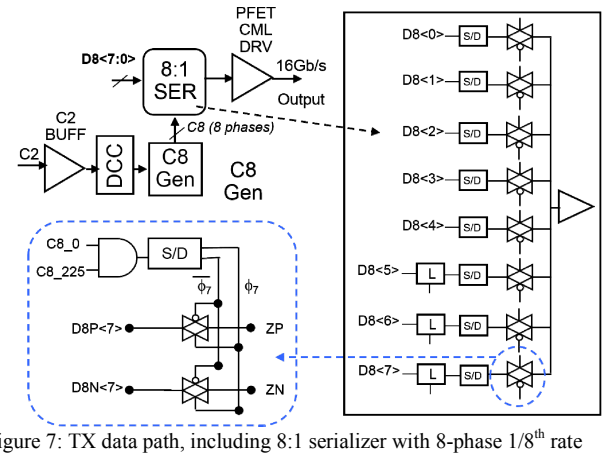

Figure 7: TX data path, including 8:1 serializer with 8-phase 1/8th rate (2GHz) clocks.

associated with the integration of near-zero currents at codes near the extremes of the interpolator range [2]. A cascode device M5 isolates the integration node from parasitic capacitances in the rest of the steering network. Once signal A goes low, current from the DAC is steered to a dummy load M6. During this reset period, a second quiescent current I_{Q2} is steered to the integration node. Maintaining this current through the cascode transistor M5 avoids slow transients associated with the charging and discharing of its source node during each integration cycle.

D. Transmitter Data Path

Figure 7 shows details of the 16-Gb/s data transmitter which includes an 8:1 serializer, output driver, and local clocking circuitry. Duty-cycle correction is applied to the input half-rate clock before dividing to produce 8-phases of a C8 (2 GHz) clock. These clock phases are used to directly serialize the 2-Gb/s input data using a transmission-gate serializer. Each transmission gate is enabled for 1-UI by ANDing two phases of the C8 clocks. While transmission gate structures have been used for 4:1 serializers [2, 6], the use of eighth-rate clocks in this work enables a very low-latency 8:1 serializer. Serialized data is transmitted by a PFET CML driver with adjustable launch amplitude of between 100mVppd and 400mVppd [2].

III. Experimental Results

A C4-limited test chip occupying 2.5x3.7mm^2 was fabricated in 32nm SOI CMOS technology. Test chips in organic SLCs were mounted to Megtron-6 test boards with matched channel lengths varying from 0.75” to 10”. A board with 10” channels is shown in Figure 8. Figure 9 shows aggregate bathtub curves for all 16 lanes in the parallel interface when running 16Gb/s PRBS31 data over the 0.75” and 10” traces. Bus-level timing margin of 30% at a BER of 1E-9 is achieved launching 100mVppd data across the 0.75” links, and is limited by underdamped ref-

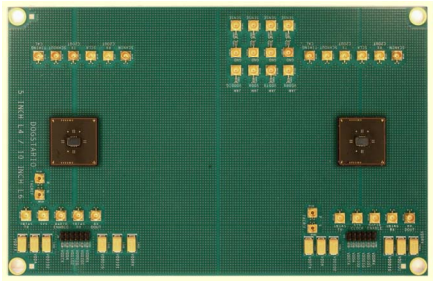

Figure 8: Megtron-6 test board with two packaged test chips interconnected via 10" traces.

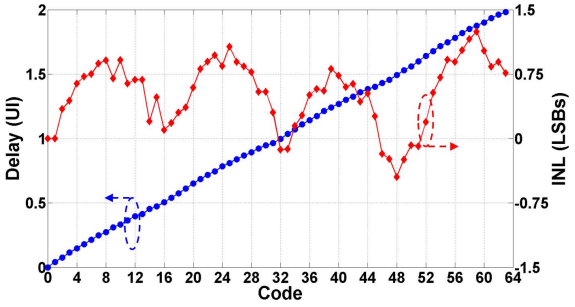

Figure 10: Phase rotator transfer characteristic and integral nonlinearity.

lections. The measured link efficiency is 1.81pJ/bit, including 0.55pJ/bit from the 1V TX, 1.18pJ/bit from the 1V RX, and 0.08pJ/bit from the amortized 1V PLL. Data chronization between two RX lanes was experimentally verified, and glitchless handoff between the two clocks C2A and C2B entering an RX lane was demonstrated without introducing bit errors or losing divider synchronization between lanes. Across 10" links with 10dB C4-to-C4 loss, 200mVppd data is recovered with a bus-level timing margin of 40% at an efficiency of 1.93pJ/bit. The measured phase rotator transfer characteristic is depicted in Figure 10. The measured INL is better than 1.2LSBs, making this phase rotator suitable for a digital-CDR-based Rx in addition to the source-synchronous links targeted in this work.

IV. Conclusions

A 16-lane, 16-Gb/s per lane source-synchronous I/O is reported in 32nm SOI CMOS technology. The architecture employs a receiver redundancy scheme that enables periodic Rx lane recalibration while supporting continuous data transmission, without excessive power or area penalty. Several innovative circuit and architecture features have been described, including an active-inductor CTLE that directly interfaces to the sampling latch, low-latency serializer and deserializer, and a phase rotator based on current-integrating phase interpolators with improved linearity and supply rejection as compared to prior art [2]. Table 1 compares this work with recent I/Os operating

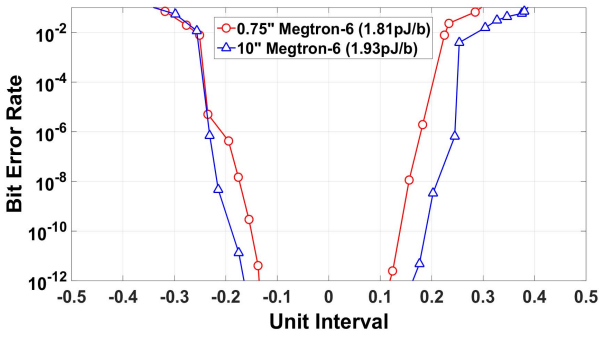

Figure 9: Aggregate bathtub curves for 16 lanes sending 16Gb/s PRBS31 data

	[7]	[8]	[9]	[10]	This Work	
Technology	65nm	40nm	40nm	32nm	32nm SOI	
Per-Lane Data Rate (Gb/s)	16	16	16/20	16	16	
Equalization	TX FFE-5, RX CTLE (CTRL)	Coded differential signaling	TX FFE-2 RX CTLE RX DFE-1	TX FFE-3 RX CTLE	RX CTLE	
Channel Loss (dB)	15	10	15	11	6.2	10
Power Efficiency (pJ/bit)	13 (CTRL)/ 8 (DRAM)	4.1	5.3/6.1	2.6	1.8	1.9

Table 1: Comparison with recently reported 16-Gb/s I/Os.

near 16Gb/s [7-10]. This work achieves better than 2pJ/b power efficiency while equalizing 10dB of channel loss, making this suitable for source-synchronous power critical applications with short-to-moderate reach.

References

[1] T. Dickson *et al.*, "An 8x10Gb/s Source Synchronous I/O System Based on High-Density Silicon Carrier Interconnects," *Symp. VLSI Circuits Dig. Tech. Papers*, pp. 80-81, June 2011.

[2] T. Dickson *et al.*, "A 1.4-pJ/b, Power-Scalable 16x12-Gb/s Source-Synchronous I/O with DFE Receiver in 32nm SOI CMOS Technology," to appear in *IEEE J. Solid-State Circuits*, Aug. 2015.

[3] M. Ferriss *et al.*, "A 28GHz Hybrid PLL in 32nm SOI CMOS," *Symp. VLSI Circuits Dig. Tech. Papers,* pp. 198-199, June 2013.

[4] A. Agrawal *et al.*, "A 19-Gb/s Serial Link Receiver with both 4-Tap FFE and 5-Tap DFE Functions in 45nm SOI CMOS," *ISSCC Dig. Tech. Papers*, pp. 134-135, Feb 2012.

[5] J. Bulzacchelli *et al.*, "A 28Gb/s 4-tap FFE/15-tap DFE Serial Transceiver in 32nm SOI CMOS Technology," *ISSCC Dig. Tech. Papers,* pp. 324-325, Feb. 2012.

[6] J. Kim *et al.*, "A 16-to-40Gb/s Quarter-Rate NRZ/PAM4 Dual-Mode Transmitter in 14nm CMOS," *ISSCC Dig. Tech. Papers,* pp. 60-61, Feb. 2015.

[7] K. Chang *et al.*, "A 16Gb/s/link, 64GB/s bidirectional link asymmetric memory interface cell," *Symp. VLSI Circuits Dig. Tech. Papers,* pp. 126-127, June 2008.

[8] A. Amirkhany *et al.*, "A 4.1pJ/b 16Gb/s Coded Differential Bidirectional Parallel Electrical Link," *ISSCC Dig. Tech. Papers,* pp. 138-139, Feb. 2012.

[9] K. Kaviani *et al.*, "A 0.4mW/Gb/s 16Gb/s Near-Ground Receiver Front-End with Replica Transconductance Termination Calibration," *ISSCC Dig. Tech. Papers,* pp. 132-133, Feb. 2012.

[10] M. Mansuri *et al.*, "A Scalable 0.128-to-1Tb/s 0.8-to-2.6pJ/b 64-Lane Parallel I/O in 32nm CMOS," *ISSCC Dig. Tech. Papers,* pp. 402-403, Feb. 2013.