A 20 Gb/s 0.3 pJ/b Single-Ended Die-to-Die Transceiver in 28 nm-SOI CMOS

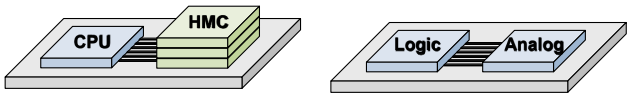
Behzad Dehlaghi and Anthony Chan Carusone
Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada
Email: behzad.dehlaghi@isl.utoronto.ca, tony.chan.carusone@isl.utoronto.ca

***Abstract*—A low-power transceiver architecture for die-to-die applications is presented. The proposed transceiver employs CMOS logic-style circuits and a passive equalizer in the transmitter to reduce the power consumption. Single-ended signaling without a shared reference voltage is used to minimize the number of required signal traces and packaging bumps. A transceiver prototype is fabricated in 28 nm STM FD-SOI CMOS technology and it operates at 20 Gb/s and 16.4 Gb/s data rates over different channels with 5.9 and 7.1 dB of loss relative to DC (10.7 and 12.9 dB total loss) at the Nyquist frequency while consuming 0.30 and 0.33 pJ/bit excluding clocking circuits, respectively.**

I. Introduction

The demand for higher aggregate bandwidth at all levels of communication infrastructure has been driving research into chip-to-chip communication over short printed circuit board (PCB) traces and very short traces on a common packaging substrate. The latter is sometimes referred to as die-to-die communication and has gained more interest in recent years to meet the requirements of next generation high performance computing (HPC) systems. Fig. 1 shows two common applications of these die-to-die links. In Fig. 1a, two chips are realized in different technologies and therefore cannot be integrated onto a single chip, for example a CPU and hybrid memory cube (HMC). In Fig. 1b, a logic/FPGA die is interfaced with high aggregate bandwidth to a die for analog/SerDes functions, which may be difficult to integrate. In these links, multiple dies share a common packaging substrate which may be either a silicon interposer or an organic substrate. Silicon interposers tend to be more expensive and lossy (up to 15dB) than organic substrates but can offer finer bump pitch and trace spacing. Recent work has shown that 10 Gb/s links can be realized on a silicon interposer while using differential signaling and an IIR-DFE to compensate for up to 11.1 dB of loss [1]. In [2], 20 Gb/s single-ended ground-referenced links have been implemented on an organic substrate with 1 dB loss.

In this paper, a low-power small-area transceiver is proposed for die-to-die communication. Single-ended signaling is used to maximize the density of the IOs using fewer bumps and wires than fully-differential signaling. CMOS circuits are used in the building blocks of both transmitter and receiver to minimize the power consumption. To compensate for channel loss, a transmitter passive equalizer is used. The receiver is able to generate a voltage reference signal from the incoming data using an internal reference generator. The transceiver prototype is implemented in 28 nm FD-SOI CMOS technology and it operates at 20 Gb/s consuming only 0.3 pJ/bit of energy including a 4-to-1 multiplexer, transmitter, receiver and 1-to-4 demultiplexer.



(a) (b)

Fig. 1. Examples of die-to-die links (a) CPU to hybrid memory cube (HMC) (b) logic/FPGA to analog/serDes.

The remainder of this paper describes the proposed die-to-die transceiver in detail. Section II explains the overall link architecture. Section III discusses the details of the transmitter and the receiver along with the circuit implementation of their key building blocks. In section IV, the experimental results are presented, and section V concludes the paper.

II. Link Architecture

Fig. 2 shows the proposed link front-end. A passive RC equalizer is used in the transmitter to compensate for the frequency dependent loss that may exist in the channel. In very low power links, a significant fraction of energy consumption arises due to the swing on the die-to-die interconnect. The equalization is implemented on the transmitter side to decrease the signal swings on the channel thereby lowering dynamic power consumption. Moreover, having the equalization in the transmitter – or in the receiver and before the termination impedance – decreases the amount of low frequency current being drawn from the transmitter power supply resulting in significant power savings. For example, adding a 100-Ohm DC resistance to the transmitter of a doubly-terminated 50-Ohm per side link with a 1-V supply voltage, reduces the average DC power from 5 mW to 2.5 mW, and the dynamic power by a factor of 4. The passive equalizer also mitigates one of the main issues of single-ended signaling which is power supply noise. Normally the power supply distribution network has a peak impedance around hundreds of MHz and drawing currents at those frequencies may result in a lot of power supply noise. By using a passive equalizer and shaping the transmitter current spectrum mostly to higher frequencies, the noise on the power supply is reduced.

The disadvantage of having the passive equalizer in the transmitter is the resulting termination impedance mismatch may cause reflections. However having a good termination impedance on the receiver side prevents reflections on that end. In die-to-die links on silicon interposers or organic substrates, there are no significant discontinuities in the channel such as connectors or vias. Therefore, the termination impedance mismatch at the transmitter does not degrade the performance of the link substantially. Moreover, the impedance of the

978-1-4799-8682-8/15/$31.00 ©2015 IEEE

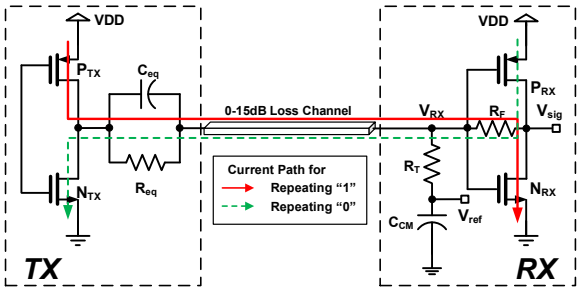


Fig. 2. Proposed link front-end.

transmitter at high frequencies is reduced due to the presence of capacitor C_{eq} providing a better match and less reflections at high frequencies.

One of the problems of single-ended signaling is generating the reference voltage on the receiver side. In this work, the reference voltage is extracted from the incoming signal in the receiver. As shown in Fig. 2, when the transmitter sends a repeating "1" or "0", the capacitor acts as an open circuit and V_{RX} can be found as follows, assuming R_{ch} is the DC resistance of the channel and V_{DD} is the supply voltage on both transmitter and receiver:

$$V_{RX1} = \left(\frac{R_{NRX} + R_F}{R_{NRX} + R_F + R_{ch} + R_{eq} + R_{PTX}}\right) \times V_{DD}, \quad (1)$$

$$V_{RX0} = \left(\frac{R_{ch} + R_{eq} + R_{NTX}}{R_{NTX} + R_F + R_{ch} + R_{eq} + R_{PRX}}\right) \times V_{DD}. \quad (2)$$

When the incoming signal is DC balanced, the DC voltage at V_{RX} is equal to the average of (1) and (2). Therefore, the DC voltage is $V_{DD}/2$ when $R_{PTX} = R_{NTX}$ and $R_{PRX} = R_{NRX}$. One can assure that mismatch between the NMOS and PMOS transistor impedances does not change the DC voltage significantly by designing the on-resistance of the transistors to be smaller than R_F and R_{eq}. This DC reference signal V_{ref} is stored on C_{CM} and used as the reference voltage for the receiver sampler. Hence, C_{CM} must be large enough to ensure small ripples on the reference voltage and also maintain a constant receiver termination impedance, R_T, at high frequencies.

III. Transmitter/Receiver Design

Fig. 3 illustrates the complete block diagram of the proposed transceiver based on the link architecture discussed in section II. An on-chip parallel PRBS7 generator is used to generate quarter-rate data in the transmitter. A half-rate differential clock signal is provided from off-chip and is divided into quadrature clock signals on-chip. The quarter rate PRBS7 signals are then serialized in a 4-to-1 multiplexer using the quadrature clock signals. The full-rate signal goes through a pre-driver and is delivered to the channel using the output driver. The passive equalizer is programmable and can be bypassed using EQ_BP switch. The receiver front-end comprises a termination impedance and a pre-amplifier stage that also extracts the reference signal from the incoming signal. Finally, the received signal is deserialized by a factor 4 and sent off-chip through a 50-Ω output driver for bit error rate (BER) measurements. All the blocks in the transceiver use CMOS logic circuitry in order to save power compared with

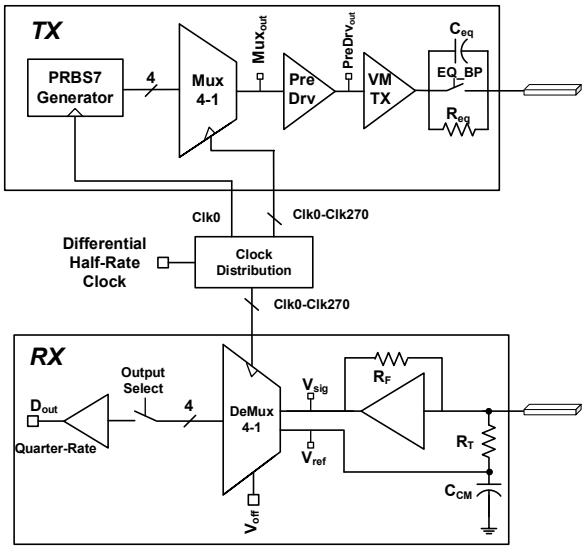


Fig. 3. Block diagram of transceiver prototype.

CML logic. All latches in the PRBS7 generators and 1-to-4 demultiplexer are double-tail latches [3]. In the rest of this section, some of the key building blocks of the transceiver are discussed.

A. 4-to-1 Multiplexer

The schematic of the CMOS 4-to-1 multiplexer is shown in Fig. 4. The multiplexer consists of 4 unit cells where each unit cell comprises a pull-up and a pull-down network. The pull-down network is a modified CMOS version of the CML 4-to-1 multiplexer presented in [4]. When the clock signal (Clk0 in the first cell) goes high the input data is sampled on the "Np" and "Nn" nodes; the parasitic capacitance on these nodes holds its value until the delayed clock (Clk90 in the first cell) goes high. In the time window when clock is high and the delayed clock is low, the output is equal to the input data. Once the delayed clock goes low, "Np" and "Nn" nodes go high and low respectively and they disconnect the output from the first unit cell. Therefore, each unit cell controls the output signal when its clock is high and its delayed clock is low. By connecting the appropriate clock phases to each of the pull-up and pull-down cells, the output is driven by only one cell at a time and each clock phase sees the same capacitive load.

The proposed 4-to-1 CMOS multiplexer consumes less power than its CML counterpart but has significant parasitic capacitance on its output from both the pull-up and pull-down networks preventing the multiplexer from operating at 20 Gb/s. In the next section, a technique to alleviate this problem is presented.

B. Transmitter Pre-driver

Fig. 5 shows a pre-driver circuit that is proposed to overcome the bandwidth limitation of the 4-to-1 CMOS multiplexer. Using a conventional CMOS inverter as the pre-driver and neglecting the input capacitance of the inverter, the bandwidth of the 4-to-1 MUX is limited to $1/(R_{mux} \cdot C_{mux})$. This bandwidth limitation causes inter-symbol interference

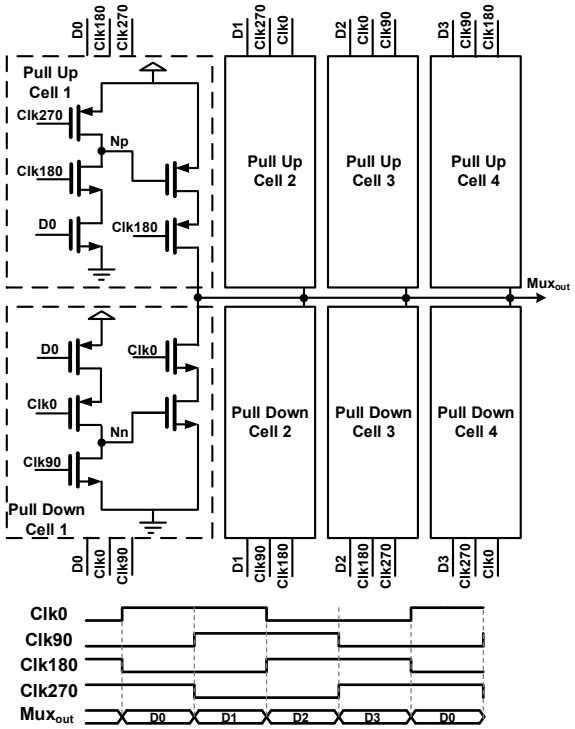


Fig. 4. 4-to-1 CMOS multiplexer.

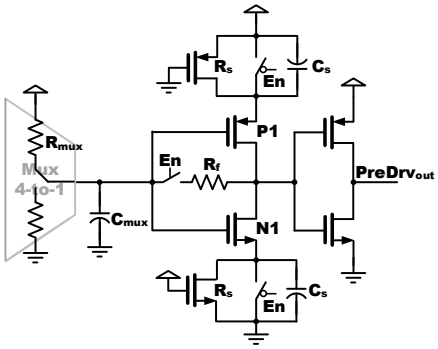
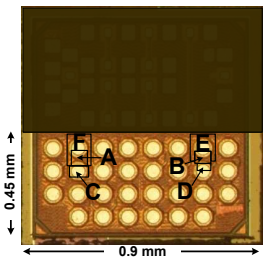


Fig. 5. Pre-driver circuit with bandwidth enhancement technique.

(ISI) in the signal and ultimately translates into jitter in the transmitter output. The pre-driver circuit employs a feedback resistor (R_f) that extends its bandwidth. In addition to that, degeneration capacitors (C_s) appear in the sources of both P1 and N1 to boost gain at high frequencies similar to continuous time linear equalizers (CTLEs) except using CMOS logic. The only disadvantage of using this bandwidth enhancement technique in the pre-driver is that the signals at the output of the multiplexer and pre-driver do not swing rail-to-rail anymore. For larger levels of boosts, this results in higher sensitivity to supply noise, which can introduce jitter. Therefore, this technique is only useful to introduce approximately 0 to 5 dB of high-frequency boost. In the prototype, the switches En are included to disable pre-driver bandwidth enhancement and show the effectiveness of this technique.



	Block Description	Area (μm^2)
A	Transmitter	4556
B	Receiver	4374
C	PRBS Gen. & TX Clock	3442
D	RX Clock	1700
E	TX Decoupling Cap.	8019
F	RX Decoupling Cap.	7290

Fig. 6. Die photo and area breakdown of transceiver prototype.

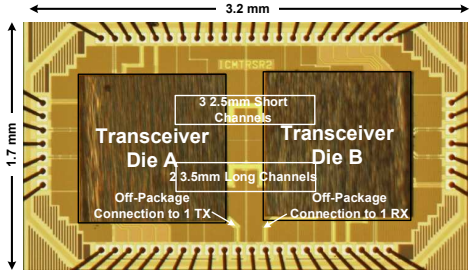


Fig. 7. The system prototype package where two transceiver dies are flip-chip mounted onto a silicon interposer.

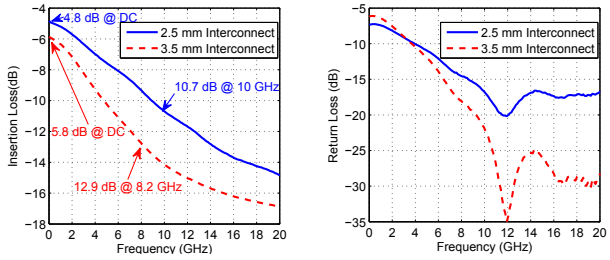
IV. Experimental Results

The transceiver prototype was fabricated in 28-nm STM FD-SOI CMOS. The die photo and area breakdown of the transceiver is shown in Fig. 6. The pad pitch is 100 μm and the transmitter and receiver occupy 67.5×67.5 μm^2 and 81×54 μm^2 of silicon area respectively. Therefore, both transmitter and receiver including all their decoupling capacitors and clocking circuits easily fit under 4 pads (V_{DD}, Gnd, TX_{out}, RX_{in}) which is an important feature for high-density IOs.

Fig. 7 shows the system package that is used for the measurements. Two identical transceiver dies are flip-chip mounted onto a silicon interposer. The interposer chip includes the interconnects between two dies as well as more decoupling capacitors for the power signals. There are three identical transmitter/receiver pairs on each transceiver die. All the transmitters on die B are connected to receivers on Die A through a 2.5-mm interconnect. Two of the transmitters on die A are connected to receivers on die B through a 3.5-mm interconnect. The last transmitter on die A and the last receiver on die B are brought off-package to be able to measure the performance of the transmitter and receiver circuits independently.

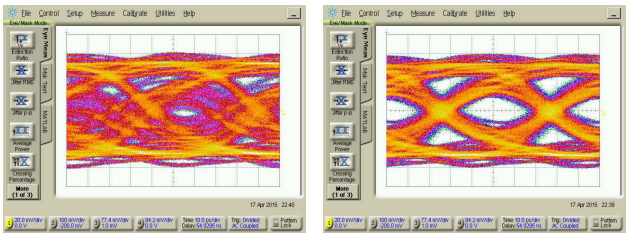
The measured insertion loss and return loss of the channels on the interposer are shown in Fig. 8. The return loss of the channels degrades below several GHz because these interconnects have large series resistance which is dominant at low frequencies. Since the interposer interconnects are aluminum with only 0.64 μm metal thickness, the channel loss at Nyquist frequency relative to the loss at DC in our $3.5mm$ interconnect lays between the 2-cm to 4-cm customized (thick copper) interposer channels reported in [1]. This is also comparable to the loss of a 10-inch channel on an organic substrate based on the results reported in [2].

Fig. 9 shows the transmitter output at 20 Gb/s with and without the pre-driver bandwidth enhancement technique showing its effectiveness to overcome the bandwidth limitation



(a) (b)

Fig. 8. The silicon interposer channels' (a) insertion loss (b) return loss.



(a) (b)

Fig. 9. The transmitter output eye diagram taken off-package at 20 Gb/s (a) without pre-driver bandwidth enhancement (b) with pre-driver bandwidth enhancement.

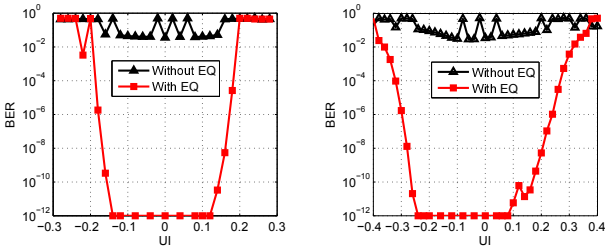
of the CMOS multiplexer. It must be noted that the transmitter signal in these figures goes through package discontinuities such as wirebonds, which would not be encountered on the interposer channels. Therefore, the only purpose of these eye diagrams to illustrate the effectiveness of the pre-driver bandwidth enhancement technique.

Fig. 10 shows the measured bathtub curves of the received signal on both short and long interconnects with and without the passive equalizer for one of the links. In these measurements, the half-rate differential clock signals for both transmitter and receiver are provided from a Keysight 4960A clock synthesizer and the quarter-rate received signal is sent to a Centellax TG1B1-A BERT for BER measurements. As shown in Fig. 10a, the eye opening is 0.26UI with 10^{-12} BER at 16.4 Gb/s over the long channel with 5.8 dB and 12.9 dB loss at DC and Nyquist frequency respectively. The power consumption of the transmitter and receiver are 2.6 mW and 2.9 mW respectively. Also, the eye opening is 0.24UI at 20 Gb/s over the short channel with 4.8 dB and 10.7 dB loss at DC and Nyquist frequency. while the transmitter and receiver power consumption are 2.8 mW and 3.3 mW, respectively.

Table I summarizes the performance of the proposed transceiver in comparison with recently published works for short reach applications. The energy efficiency of this transceiver is comparable to [2], but this work can also compensate for losses in the channel using the passive equalizer.

V. Conclusion

A 0.3 pJ/bit 0.009 mm^2 single-ended transceiver has been introduced for die-to-die applications. The proposed transceiver can be used on both organic substrates and silicon interposers. The performance of the transceiver was measured by flip-chip mounting two prototypes onto a silicon interposer. The proposed transceiver consumes 6.1 mW power at 20 Gb/s



(a) (b)

Fig. 10. Bathtub curves with and without the TX passive equalizer for a PRBS7 pattern and: (a) 3.5 mm channel at 16.4 Gb/s; (b) 2.5 mm channel at 20 Gb/s.

TABLE I. Performance Summary Comparison.

	JSSC12 [1]	JSSC13 [2]	JSSC12 [5]	**This work**	
Signaling	Diff. VM	SE Ground Referenced	SE VM	**SE VM**	
Channel	Si Int. 3μm-thick Cu	Organic Substrate	FR4	**Si Int. 0.64μm-thick Al**	
Equalization	IIR-DFE	-	TX FIR + RX CTLE + DFE	**TX Passive EQ**	
Technology	45-nm SOI	28-nm	40-nm	**28-nm SOI**	
Area [mm^2]	0.023[a]	0.005	-	**0.009**[a]	
Supply [V]	1	0.9	1	**1**	
Channel Length	4cm	6mm	3inch	**3.5mm**	**2.5mm**
Atten. @ Nyquist (Relative to DC) [dB]	- (11.1)	1 (1)	10 (10)	**12.9 (7.1)**	**10.7 (5.9)**
Data Rate [Gb/s]	10	20	12.8	**16.4**	**20**
Energy Efficiency [pJ/bit]	2.6[b]	0.255[b]	2.15[b]	**0.33**	**0.30**

[a] The transmitter area is calculated in a way to fit under 6 μC4 bumps [1].
[a] Does not include the decoupling capacitors and clocking and PRBS generator circuits.
[b] The energy efficiency of transceiver excluding the clocking is calculated based on the reported power breakdown.

over a 2.5 mm interconnect with 10.7 dB loss at Nyquist frequency and 4.8 dB loss at DC. The excellent energy efficiency of the proposed transceiver is mainly due to the use of CMOS building blocks, minimizing current in the transmitter to receiver signal path, and lowering the signal swings on the channel.

Acknowledgment

The authors would like to acknowledge CMC for fabricating and packaging the chips, provision of equipment and CAD tools, and Huawei Canada for financial support and technical discussions.

References

[1] T. Dickson *et al.*, "An 8x 10-Gb/s source-synchronous I/O system based on high-density silicon carrier interconnects," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 4, pp. 884–896, April 2012.

[2] J. Poulton *et al.*, "A 0.54 pJ/b 20 Gb/s ground-referenced single-ended short-reach serial link in 28 nm CMOS for advanced packaging applications," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 12, pp. 3206–3218, Dec 2013.

[3] D. Schinkel *et al.*, "A double-tail latch-type voltage sense amplifier with 18ps setup+hold time," in *IEEE International Solid-State Circuits Conference 2007*, Feb 2007, pp. 314–605.

[4] A. Hafez *et al.*, "A 32-48 Gb/s serializing transmitter using multiphase serialization in 65 nm CMOS technology," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 3, pp. 763–775, March 2015.

[5] A. Amirkhany *et al.*, "A 12.8-Gb/s/link tri-modal single-ended memory interface," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 4, pp. 911–925, April 2012.