A 16-Gb/s/pin 277-fJ/bit Edge-Modulated On-Chip Interconnect Transceiver in 28-nm CMOS

Jiyong Park*, Jusung Park*, and Jintae Kim

Department of Electrical and Electronics Engineering, Konkuk University, Seoul, Korea

Email: {jy.park, js.park}@msel.konkuk.ac.kr, jintkim@konkuk.ac.kr *: Two authors contributed equally to this work

***Abstract*— This paper presents a 16-Gb/s/pin edge-modulated on-chip transceiver achieving 0.277 pJ/bit energy efficiency in a 28-nm CMOS process. Unlike the traditional on-chip interconnect that encodes data on the voltage-level, a new signaling scheme that encodes information on the signal-edge is described. A novel self-clocked edge combiner and an inverter-based reference-less time-based transceiver enable a low-power 16-Gb/s edge-modulated signaling on an 8-GHz clock. Fabricated in 28-nm CMOS, the prototype 16-Gb/s transceiver achieves BER<10^{-12} over 1.5-mm on-chip interconnect. Measured FoM of 277 fJ/bit at data rate of 16 Gb/s achieves state-of-the-art energy efficiency.**

Keywords*— *On-chip Interconnect, Chiplet Interface, Edge-Modulation, High-speed I/O

I. Introduction

Insatiable performance demand for computing AI workload requires ever-increasing I/O bandwidth at all levels of interconnects in AI computing systems. To cope with this challenge, traditional chip-to-chip wireline interface standards have evolved by adopting multi-level signaling schemes such as PAM3 in DDR6 [1] and PAM4 in PCIe 6.0 to keep increasing the per-lane data rate. On the other hand, parallel and short-reach interconnect such as high-bandwidth memory (HBM) or universal chiplet interconnect express (UCIe) commonly uses non-return-to-zero (NRZ) signaling. The merit of NRZ signaling is its simplicity, which is crucial for very wide I/Os. Modern on-chip interconnect is massively wide. For example, there are 2048 parallel I/O pins for HBM4. Addressing random and systematic mismatches for these parallel lanes is a huge design challenge. In particular, given that the on-chip interconnect is single-ended by nature, the voltage offset mismatch at the receiver significantly degrades the bit error rate performance. Addressing this issue mandates per-pin receiver offset calibration. For instance, in [2], authors use a two-step coarse-fine offset calibration scheme to reduce the reference mismatch for HBM3 receivers. While this technique can be effective, it certainly adds substantial design complexity given that such a calibration should be applied to all I/O channels, i.e., 2048 receivers in HBM4.

This paper presents a 16-Gb/s on-chip I/O transceiver using a new signaling scheme in which the information is encoded on the *phase* or the signal *edge* rather than the voltage level. Phase-modulation is a classic concept in the field of digital pass-band communication, but it is rarely used in I/O transceivers because design complexity can be prohibitively high. This work presents a novel time-domain transmitter and a receiver circuit that are optimized for an edge-modulated on-chip link transceiver in advanced technology. Fabricated in 28-nm CMOS, the prototype chip demonstrates per-pin data rate of 16 Gb/s with BER<10^{-12} over 1.5-mm channel, achieving state-of-the-art energy efficiency of 277 fJ/bit.

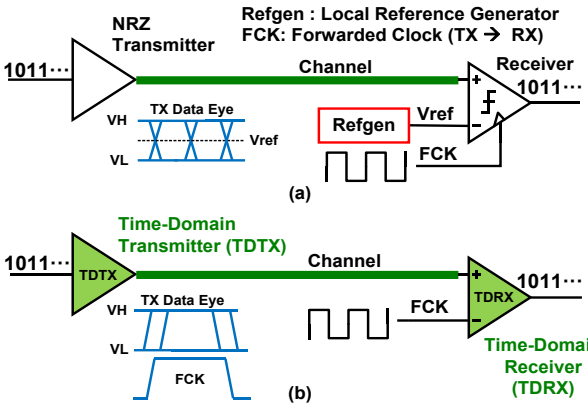

Fig. 1. (a) Conventional NRZ-based link with on-chip reference voltage. (b) The proposed edge-modulated link without reference voltage.

II. Edge-Modulated Signaling

Fig. 1 displays the key concept of the signaling and transceiver schemes in this work. Conventional NRZ signaling in Fig. 1(a) requires a reference voltage (V_{ref}) for the receiver, which is locally generated. Generating reference voltages for massively wide links increases hardware complexity and consumes significant power and area. Also, it is impossible to achieve a perfect agreement between the actual reference at RX and the desired reference from the TX, leading to the reference offset and reduced voltage margin. Furthermore, because TX and RX dies are usually physically separate, uncorrelated supply noise will further degrade the voltage margin and bit error rates.

Fig. 1(b) shows the proposed edge-modulated signaling in which the time-domain transmitter (TDTX) encodes the information on the displacement of both the rising and falling edge of the clock. Since our new signaling scheme encodes information on the clock edge location, the receiving comparator in conventional NRZ receiver is replaced by the time-domain comparator as a receiver (TDRX). An immediate advantage is that it does not need a reference voltage. In particular, since the time domain comparator compares the edge-modulated data signal with the forwarded reference clock (FCK), the presented signaling scheme does not need a reference voltage at all, thereby eliminating all concerns related to the reference voltage in conventional NRZ signaling scheme.

III. Transceiver Architecture

Fig. 2 shows the overall transceiver architecture. The transmitter and receiver are implemented on the same die with 1.5-mm on-chip channel routed on metal 6 layer. There are 8 data channels alongside a shared clock channel. An 8-GHz clock is forwarded via the clock channel, which is used as a reference clock in the RX time comparator.

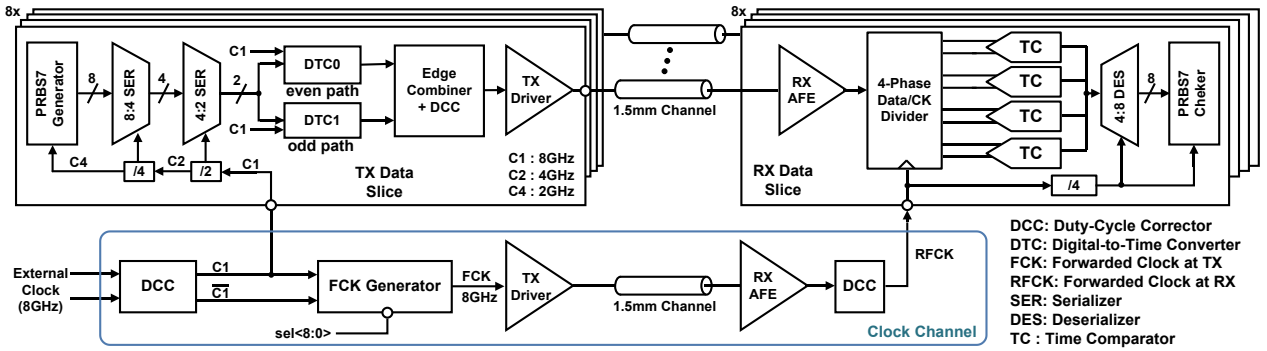

Fig. 2. Overall architecture of the edge-modulated I/O link transceiver.

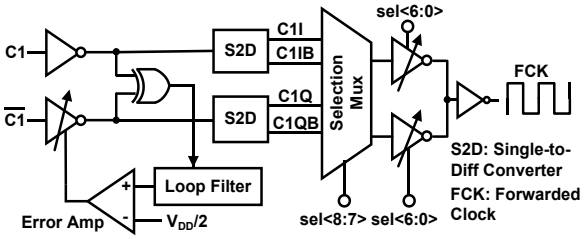

Fig. 3. The phase-interpolator digital-to-time converter (PI-DTC) and analog quad-clock generator in the forward-clock (FCK) generator.

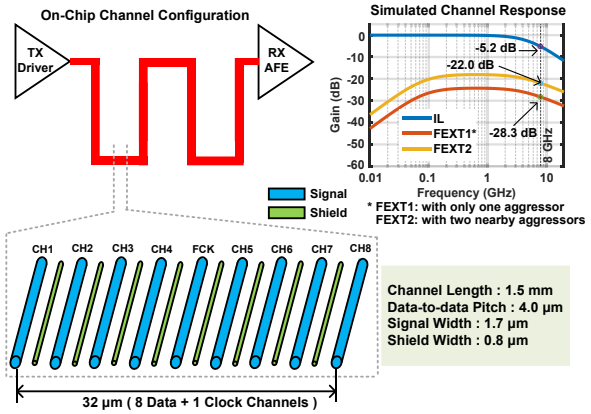

Fig. 4. On-chip channel configuration and simulated S21 + far-end crosstalk characteristics.

In the transmitter, an 8-GHz clock (C1) drives on-chip dividers to generate multi-phase 4-GHz (C2) and 2-GHz clocks (C4) for the PRBS generator and serializers. An on-chip 8-bit parallel PRBS generator running at 2-GHz generates 8-bit parallel PRBS7 pattern. The 8×2 Gb/s parallel data stream is serialized by cascaded serializers, generating 2×8 Gbps NRZ data stream, which then encodes the data into the delay of C1 clock by modulating two high-speed digital-to-time converters (DTCs). The edge combiner combines two single-edge modulated 8-Gbps data streams from DTC0 and DTC1 into a single 16-Gbps dual-edge-modulated signal by alternately choosing the rising and falling edge of the two inputs. The receiver architecture is based on quad-rate clocking, consisting of RX analog front-end (AFE), quad-phase data and clock divider, and four time-interleaved time comparators (TCs). The outputs of the four TCs are deserialized, generating an 8-bit parallel data that drive an on-chip parallel PRBS checker.

When the transmitter forwards 8-GHz clock (FCK) to the receiver alongside the data, fine phase adjustment of the FCK is necessary to measure the bit error rate over the 1UI time span. Fig. 3 displays the details of FCK generator. A 9-bit phase-interpolator based DTC (PI-DTC), consisting of the selection mux and the inverter-based phase-interpolator, is used for fine phase adjustment. Since the PI-DTC needs four-phase C1 clocks for proper phase interpolation, an XOR-based analog quad clock generator is used as shown in Fig. 3. The duty cycle of an XOR phase detector is 50% when two inputs are 90° apart. The error amplifier compares the RC-filtered XOR output with $V_{DD}/2$ to control supply voltage of the delay cell in the quadrature generator to close the feedback loop. Fig. 4 illustrates the channel implementation alongside the simulated insertion loss and crosstalk characteristic. The entire channel consists of 8 data lanes + 1 clock lane with alternating ground shield wires. Implemented in metal-6 layer, each channel is 1.5 mm long, 1.7 μm wide with 4-μm channel pitch. The calculated beachfront bandwidth is 4 Tb/s/mm. The simulated S21 is -5.2dB at 8 GHz, and the worst case FEXT, denoted as FEXT2 in Fig. 4, is -22.0dB at 8 GHz.

IV. Key Building Blocks

A. DTC-based Data Encoder and Edge Combiner

The DTC in Fig. 2 is a key building block that encodes NRZ data into the delay of the clock. However, implementing a programmable delay without degrading jitter performance is challenging. To achieve both low-jitter and high-speed while providing a full-scale range of 10 ps, a two-stage DTC is used to maintain a sharp clock edge that minimizes jitter performance degradation in the DTC. Fig. 5 illustrates the detailed circuit implementation. Each DTC incorporates a 1-bit capacitor DAC (CDAC) at the encoding node so that the digital input determines the time delay. The PMOS device in the DTC is degenerated to increase the encoding delay for the rising edge, while the non-degenerated NMOS device quickly resets the encoding node to minimize unnecessary delay and ensure rapid recovery.

Fig. 6 displays a novel self-aligning edge combiner that time-multiplexes two DTC outputs into a single dual-edge-modulated output. The edge combiner takes even and odd path output from DTC0 and DTC1, respectively, with an extra inversion in the odd path that converts the edge-encoding from rising to falling edge. The passive mux serializes rising and falling-edge encoded signals, ck0 and ck1, into a dual-edge-modulated signal. The selection signals for the mux, sel0 and sel1, are self-generated.

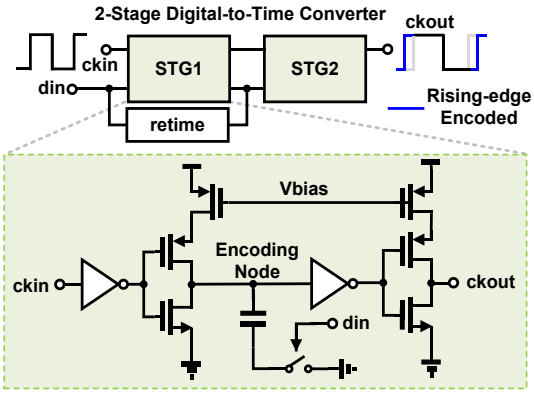

Fig. 5. Circuit implementation of the 2-stage digital-to-time converter.

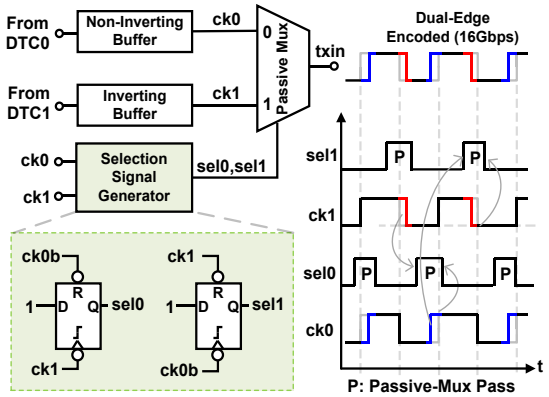

Fig. 6. Circuit implementation of the self-aligning edge combiner.

Specifically, the selection signal for ck0 (even path) is set at the falling edge of the ck1 (odd path) and is reset when ck0b is low. The selection signal for ck1 is similarly generated so that the edge combining is self-aligned without an external clock signal. The passive multiplexer is based on the transmission-gate with matched on-resistance to ensure identical rise and fall time, generating a merged 16-Gb/s output with data encoded on both rising and falling edges, which is delivered to the inverter-based TX driver.

B. Inverter-based TX driver and RX Analog Frontend

Fig. 7 shows an inverter based AC-coupled TX driver and RX analog frontend (AFE). In on-chip or short-reach links, signaling power is a substantial part of the total power. Since the insertion-loss of the on-chip link is not excessive (usually smaller than -5dB), reducing signal swing is an attractive way to lower the total power. In our scheme, a series capacitor, C_{AC}, is introduced at the TX output, forming a capacitive divider with the channel capacitance (C_{CH}) to lower the signal swing [3]. This reduces signal swing and allows us to use a smaller driver to save power. In an RX, an inverter-based active-inductor [3] is used as a continuous-time linear equalizer (CTLE) to compensate for the channel loss. The peaking frequency is chosen by resistor (R_{IND}) value and the inverter I3 sizing. The design target is to achieve ≈5 dB peaking at 8 GHz. In addition, an inverter-based transimpedance amplifier (TIA) restores the reduced signal swing to rail-to-rail. The proposed inverter-based TRX structure is highly compact, offering competitive advantage in terms of silicon area and power consumption. The simulated AC response in Fig. 7 shows that the raw channel loss (in red), -5.2 dB at 8 GHz, is effectively compensated by the active inductor-based CTLE

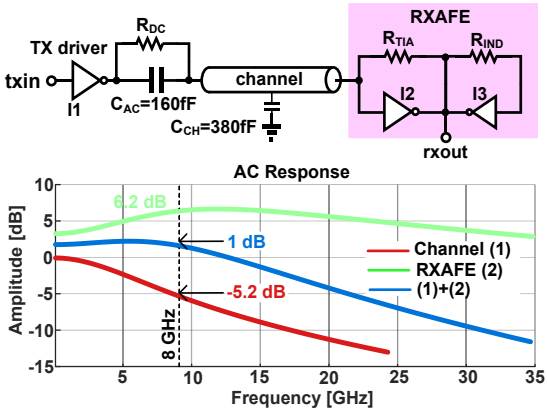

Fig. 7. Inverter-based AC-coupled transceiver.

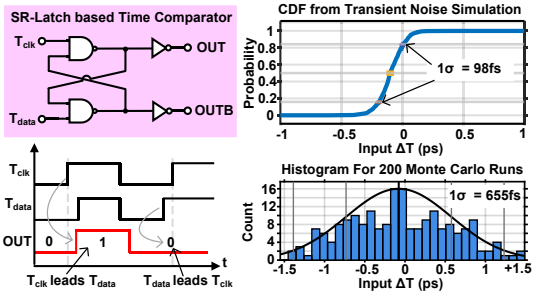

Fig. 8. Time-comparator with simulated noise and offset.

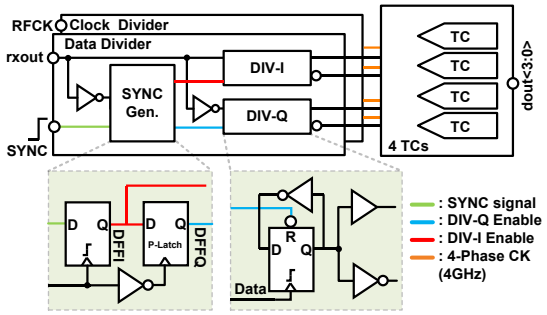

Fig. 9. Receiver data/clock divider and its initialization scheme.

in the RXAFE, yielding nearly flat combined frequency response (in blue) up to 8 GHz.

C. Time-Domain Comparator and Data Divider

In edge-modulated links, the data is encoded on the timing of the signal edge. Therefore, a time-comparator (TC) is used in the receiver instead of a conventional voltage-mode comparator. Unlike a voltage-mode receiver that relies on a reference voltage to determine logic levels, the time-domain receiver compares the relative timing between the received data edge and the forwarded reference clock edge. Fig. 8 displays the transistor-level design of the TC, which is essentially an SR-latch. When the rising edge of T_{clk} arrives earlier than T_{data}, the TC output is high. Conversely, if the T_{data} leads the T_{clk}, the TC output is low. Therefore, the TC can determine whether the data edge arrives earlier or later than the reference clock edge. This approach eliminates the need for a reference voltage, which is a huge advantage compared to conventional voltage-based links. The other advantage is that the TC exhibits much smaller input-referred noise and offset compared to the voltage-level comparator. Shown in Fig. 8, the simulated

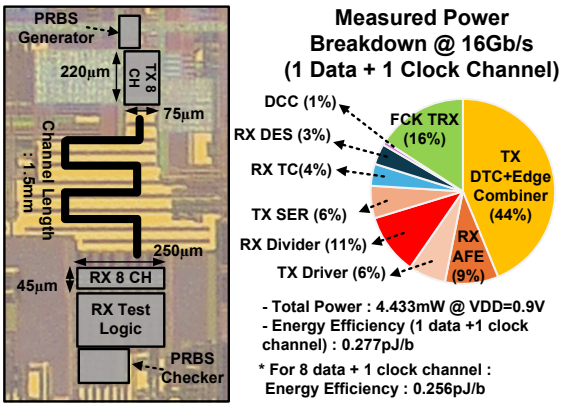

Fig. 10. (Left) Die micrograph and (Right) power breakdown.

input-referred noise and 1-sigma offset are 99 fs,rms and 655 fs, respectively. Therefore, compared to the full-scale of the modulated data edge, which is ~10 ps, the noise and mismatch are much smaller and extra offset calibration is not necessary. However, operating a single TC at full 16 Gb/s would be very tough to meet timing margin. To relax the timing, a 4-phase data and clock divider deserializes the 16-Gb/s dual-edge-modulated signal and clock into four time-interleaved 4-Gb/s rising-edge encoded data and clock signals. Shown in Fig. 9, the data and clock dividers increase the spacing between adjacent edges without altering modulated edge information, improving the TC timing margin. Note that an extra inversion is applied in the Q-path divider clock path to relocate the encoded-edge from falling to rising edge. Hence, all data divider outputs are rising edge encoded. Improper reset alignment of the dividers results in incorrect data sequence. To correctly align the dividers, the sync pulse generator (SYNC Gen) derives divider enable signals using edge-modulated data signal and clock so that I-path and Q-path dividers are correctly aligned.

V. Experimental Results

Shown in Fig. 10, the I/O transceiver is fabricated in 28-nm CMOS. The 8-channel TX and RX core occupies 0.0165 mm^2 and 0.01 mm^2, respectively. The measured FoM is 277 fJ/bit for a single data channel transceiver + a clock-forwarding channel. The projected FoM for 8 data channels + 1 clock channel is 256 fJ/bit. The prototype includes an off-chip 50-ohm driver to measure the eye. Shown on the left of Fig. 11, the edge-modulated transmitter signal, measured by a DCA (Keysight N1000A), shows a clear dual-edge-modulated signal with a modulation full scale of 10 ps. Also shown in Fig. 11 is the measured BER bathtub curve of the proposed transceiver. The measurement demonstrates that the system achieves BER < 10^{-12} within a timing margin of approximately 0.04 UI, indicating reliable operation at 16 Gb/s. Table I shows the comparison with state-of-the-art on-chip transceivers fabricated in 28-nm process with similar data rates. The comparison highlights that this work achieves competitive energy efficiency while not requiring an explicit reference voltage in the receiver, significantly simplifying on-chip I/O design.

VI. Conclusions

This paper introduced a 16 Gb/s/pin edge-modulated on-chip transceiver using a new signaling scheme that encodes information on the signal-edge. Fabricated in 28-nm CMOS,

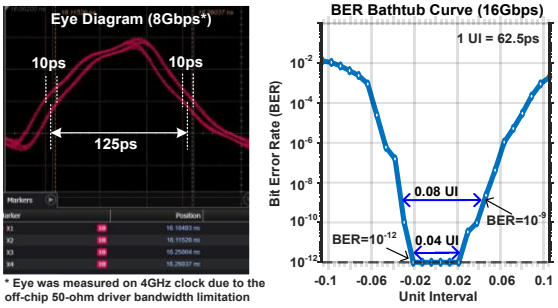

Fig. 11 (Left) Measured eye diagram for an edge-modulated signal on a 4GHz clock and (Right) measured BER bathtub.

TABLE I. Comparison with State-of-the-Art On-Chip Links

	This work	ISSCC '26 [4]	ISSCC '22 [5]	ISSCC '25 [6]	JSSC '23 [7]
Process (nm)	28	28	28	28	28
Signaling	Edge-Modulated	NRZ	Di-code	CR-PAM-3	DECS
Data Rate (Gb/s)	16	16	10	42	20
Channel length (mm)	1.5	2	6	2	1
Reference Voltage Required?	X	O	O	O	O
Energy Efficiency (pJ/bit)	0.277 /0.256*	0.51	0.385	0.275	1.27/ 0.86**
Edge Density (Tb/s/mm)	4	-	2.0	9.16	0.631

* This is calculated assuming 8 data + 1 clock channel
** This is calculated assuming 10 data + 1 clock channel

the 16-Gb/s I/O transceiver achieves BER<10^{-12} over a 1.5-mm on-chip channel with competitive energy efficiency. The presented edge-modulated I/O transceiver, without needing a reference voltage, can be a potential candidate for higher data rate HBM and chiplet interfaces.

Acknowledgment

The authors thank IDEC and KAIST for CAD support. This work is supported by Samsung Research Funding Center under SRFC-IT2302-03. Gemini by Google was used only for grammar correction.

References

[1] S.-Y. Cho et al., "A 16Gb 37Gb/s GDDR7 DRAM with PAM3-Optimized TRX Equalization and ZQ Calibration," *IEEE ISSCC*, Feb. 2024, pp. 242-244.

[2] K. Chae et al., "A 4-nm 1.15 TB/s HBM3 Interface With Resistor-Tuned Offset Calibration and In Situ Margin Detection," *IEEE JSSC*, vol. 59, no. 1, pp. 245-257, Jan. 2024.

[3] Y. Nishi et al., "A 0.190-pJ/bit 25.2-Gb/s/wire Inverter-Based AC-Coupled Transceiver for Short-Reach Die-to-Die Interfaces in 5-nm CMOS," *IEEE VLSI*, June 2023, pp. 1-2.

[4] Y. He et al., "A 16Gb/s/pin 0.51pJ/b Single-Ended NRZ Transceiver with Distributed Dual-Loop VDDQ-Ripple Compensation and Dynamic Clock Duty-Cycle Calibration for Memory Interfaces," *IEEE ISSCC*, Feb. 2024, pp. 134-136.

[5] H. Park et al., "A 0.385-pJ/bit 10-Gb/s TIA-Terminated Di-Code Transceiver with Edge-Delayed Equalization, ECC, and Mismatch Calibration for HBM Interfaces," *IEEE ISSCC*, vol. 65, Feb. 2022, pp. 1-3.

[6] K. Kim et al., "A 0.275pJ/b 42Gb/s/pin Clock-Referenced PAM3 Transceiver Tolerant to Supply Noise, Reference Offset and Crosstalk for Chiplets and Short-Reach Memory Interfaces," *IEEE ISSCC*, Feb. 2025, pp. 394-396.

[7] J. Seo et al., "A 20-Gb/s/Pin Compact Single-Ended DCC-Less DECS Transceiver With CDR-Less RX Front-End for On-Chip Links," *IEEE JSSC*, vol. 58, no. 11, pp. 3253-3265, Nov. 2023.