11.1 A 1.7pJ/b 112Gb/s XSR Transceiver for Intra-Package Communication in 7nm FinFET Technology

Ramy Yousry*[1], Ehung Chen*[1], Yu-Ming Ying[1], Mohammed Abdullatif[1], Mohammad Elbadry[1], Ahmed ElShater[1], Tsz-Bin Liu[2], Joonyeong Lee[1], Dhinessh Ramachandran[1], Kaiz Wang[2], Chih-Hao Weng[2], Mau-Lin Wu[2], Tamer Ali[1]

[1]MediaTek, Irvine, CA
[2]MediaTek, Hsinchu, Taiwan

*Equally-Credited Authors (ECAs)

2021 IEEE International Solid- State Circuits Conference (ISSCC) | 978-1-7281-9549-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/ISSCC42613.2021.9365752

COVID-19 sparked a paradigm shift for businesses, education, and social life. The measures taken have emphasized, more than ever, the importance of on-line communication platforms. Online streaming services, education tools, and workplace communication tools all experienced a multifold increase in demand in 2020 as the world population was under a lockdown. Service providers' ability to meet these demands relies on the capacity of mega-scale data centers (MSDC). Hence, capacity and bandwidth of MSDCs are expected to grow exponentially. The transition from 12.8 to 51.2Tb/s switches for data centers leaf and spine is underway, and is expected to deploy starting in 2021~2022 [1]. Unfortunately, the increase of switch throughput comes with serious area, yield, and power limitations. To mitigate the problem, recent advances in chiplets and multi-die technologies for high throughput switches show a promising increase in yield and lower solution cost. Another direction is silicon photonic and ASIC co-integration, which allows for package throughput well beyond the electrical interface capability.

This work presents a 112Gb/s extra-short reach (XSR) transceiver built in 7nm FinFET technology. The transceiver complements the 112G long-reach (LR) solution [1], and targets next generation 51.2Tb/s multi-die switch chips of 512 lanes. It comprises a common lane with 2 PLLs for multi-rate operation, 8 TX and RX lanes, and a digital part that contains the DSP adaptation and calibration engines. Figure 11.1.1 shows the receiver architecture. A continuous-time linear equalizer (CTLE) and a variable-gain amplifier (VGA) are combined in one stage for power efficiency. The receiver front-end (RXFE) drives a total of 18-slicers for data and edge sampling, and eye-scanning. The data samplers are clocked by 4 phases of a 14GHz clock to recover the 112Gb/s PAM-4 data, where each phase uses one redundant sampler to enable background calibration. An edge sampler and an eye sampler are used for the bang-bang CDR and to perfrom eye-scanning, respectively. The digital CDR loop directs the sampling clocks to their optimum positions by digitally controlling the phase interpolators (PI) on the analog side. The samplers are implemented as two-stage StrongArm slicers, as shown in Fig. 11.1.1 (bottom). This pipelining allows for both high-speed operation and minimal RXFE loading. The slicer input common mode is dynamically adjusted to ensure optimal performance over different process and temperature corners. To minimize sampling aperture asymmetry, all slicers are built identically, and a wide-range DAC is used to control the threshold and calibrate the offset of each slicer. An 8b switched-capacitor (SC) DAC is used, per slicer, to provide best linearity and area efficiency. The entire RX data path and back-end DSP are powered from a 0.8V supply.

Figure 11.1.2 shows the RXFE block diagram. It consists of a delay-line-based CTLE, VGA, and low-frequency linear equalizer (LFEQ). The input RLC network provides a ~1UI-delayed version of the input signal, which is subtracted from the main path with programmable weight. Similar to transversal FFE equalization [2], the two paths will destructively interfere at low-frequency, and constructively interfere at Nyquist frequency. This results in high-frequency peaking in the signal transfer function. To compensate for parasitic poles along the signal path, the delay-line is tuned to less than 1UI to push the peak frequency slightly beyond Nyquist. A low-pass filtered copy is used to provide low-frequency equalization. All paths are directly combined at the RXFE output and the load resistance is digitally controlled to adjust the swing at the sampler input. The CTLE can provide 8dB of high-frequency peaking and 4dB of amplitude control. The RX is AC-coupled to allow for independently optimized RX and TX performance. A baseline wander cancellation DAC is added at the CTLE output. The RXFE is implemented as pseudo-differential to enable low-voltage operation. The RX tolerance to TX common-mode and power supply ripples is enhanced by two techniques: 1) the delay-line is implemented differentially with a relatively high coupling coefficient, which results in a much shorter common-mode delay. This provides wideband common-mode "destructive interference" at the CTLE output, with no peaking at Nyquist; 2) a wideband common-mode feedback (CMFB) loop is directly wrapped around the input AC-coupling capacitor. This results in two poles (AC-coupling and op-amp) well below unity-gain frequency. To improve the loop stability, a zero is introduced in the transfer function by adding a small common-mode feedforward path (CMFF) as shown in Fig. 11.1.2 (bottom). This path helps improve stability and increases the loop bandwidth, thus providing wideband CM and power-supply rejection.

At 112Gb/s, a precise spacing between interleaved sampling clocks can be detrimental to receiver performance. A digitally assisted skew-calibration loop is applied to align each clock phase by controlling the bias of current-starved inverters (Fig. 11.1.3). A 9b SC-based delay control, applied to two successive stages within the AC clock buffer, provides enough skew calibration range while maintaining the clock duty cycle close to 50%. This duty cycle provides a good balance between regeneration and reset time for the data slicers.

The TX is an SST DAC driver (Fig. 11.1.4), a simplified version of the LR TX in [1], with several modifications to tailor it to the needs of the XSR application and achieve more than 50% power reduction. The DAC resolution is only 5b and the final 4:1 MUX is implemented in the last stage of the driver to reduce high-speed path power. The TX is powered from a 0.8V supply. The high-speed clocking path and the final driver are powered by LDOs to provide a clean output swing of 600mV$_{ppd}$. The TX output quality is determined by the duty cycle and skew between 14GHz quarter-rate clocks which drive the final 4:1 MUX directly. The 25% quadrature clock phases are produced by pulse generators and their duty cycles are adjusted precisely through a digitally-assisted calibration loop, illustrated in the inset of Fig. 11.1.4 [1]. The clock duty-cycle is sensed and a DC voltage generated by a 9b SC-DAC is applied in the AC-coupled buffer to control the duty cycle with wide adjustable range and minimal impact on jitter performance.

The DSP includes a 5-tap FIR in the TX side for pre-emphasis with 2 floating taps for reflection cancelation. The DSP adapts the appropriate CTLE gear and slicer level to facilitate CDR locking. After locking, the DSP fine-tunes the CTLE gear, slicer level, and other calibration algorithms to optimize performance and to track voltage and temperature variation.

To evaluate link performance, TX-to-RX loopback with a PRBS-31 pattern is performed over 20 package traces ranging from 5mm to 50mm. Figure 11.1.5 (top left) shows the received eye for a 5mm channel. The vertical margin for 10^{-9} BER is 44mV. The bit-error-rate (BER) bathtub plot for the same channel is shown in Fig. 11.1.5 (top right), where the horizontal margin for 10^{-9} BER is 2.5ps. Figure 11.1.5 (bottom) shows the BER of the worst 5/15/50mm channels. All link measurements are performed with near-end and far-end cross-talk aggressors. The PAM-4 transmitter output eye diagram in Fig. 11.1.6 is captured using a scope with de-embedded channel response. The measured SNDR is 39.9dB. The table in Fig. 11.1.6 compares transceiver performance with recently published parts of high-speed transceivers for short reach applications. The transceiver occupies an area of 0.228mm^2/lane and consumes 191mW/lane (1.7pJ/b), which is at least 2× more power efficient than [3-6]. The 8-lane transceiver photo is shown in Fig. 11.1.7.

References:

[1] T. Ali et al., "6.2 A 460mW 112Gb/s DSP-Based Transceiver with 38dB Loss Compensation for Next-Generation Data Centers in 7nm FinFET Technology," *ISSCC*, pp. 118-120, Feb. 2020.

[2] E-Hung Chen et al., "A 40-Gb/s serial link transceiver in 28-nm CMOS technology," *IEEE Symp. VLSI Circuits*, pp. 1-2, June 2014.

[3] Y. You et al., "A 25×50Gb/s 2.22pJ/b NRZ RX with Dual-Bank and 3-Tap Speculative DFE for Microprocessor Application in 7nm FinFET CMOS," *IEEE Symp. VLSI Circuits*, pp. 1-2, June 2020.

[4] M. Erett et al., "A 2.25pJ/bit Multi-lane Transceiver for Short Reach Intra-package and Inter-package Communication in 16nm FinFET," *IEEE CICC*, pp. 1-8, Apr. 2019.

[5] A. Cevrero et al., "A 100Gb/s 1.1pJ/b PAM-4 RX with Dual-Mode 1-Tap PAM-4 / 3-Tap NRZ Speculative DFE in 14nm CMOS FinFET," *ISSCC*, pp. 112-114, Feb. 2019.

[6] C. Menolfi et al., "A 112Gb/S 2.6pJ/b 8-Tap FFE PAM-4 SST TX in 14nm CMOS", *ISSCC*, pp. 112-114, Feb. 2018.

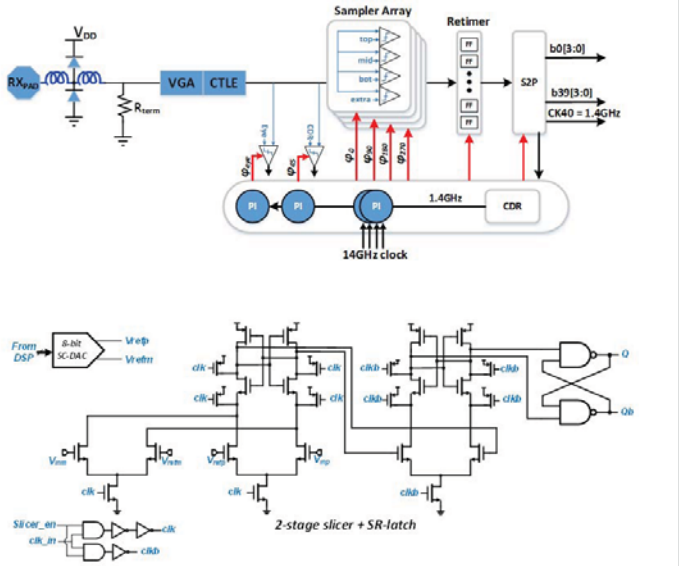


Figure 11.1.1: RX block diagram (top) and 2-stage slicer with wide-range threshold control (bottom).

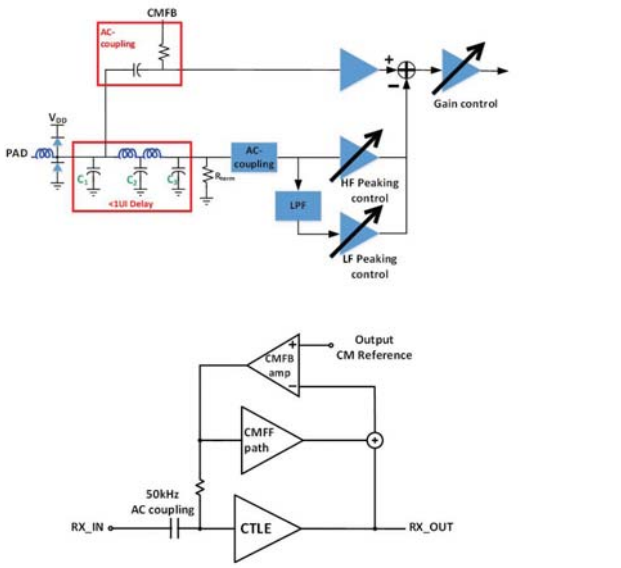


Figure 11.1.2: RX front-end architecture: delay-line based CTLE (top), CMFB (bottom).

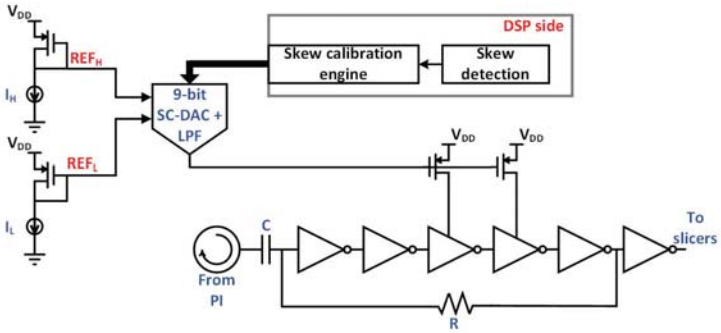


Figure 11.1.3: RX SC-based skew calibration.

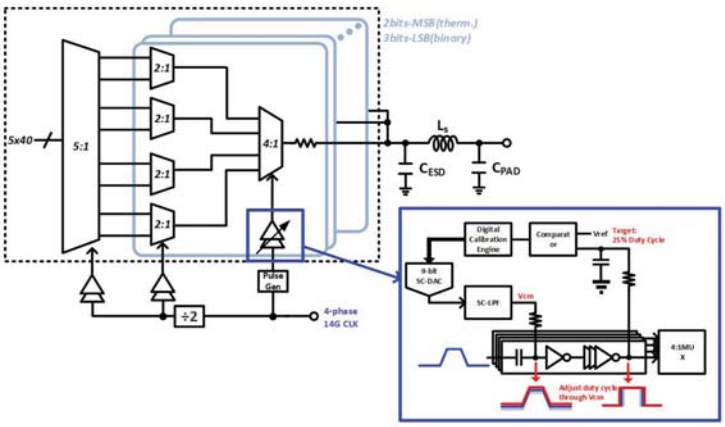
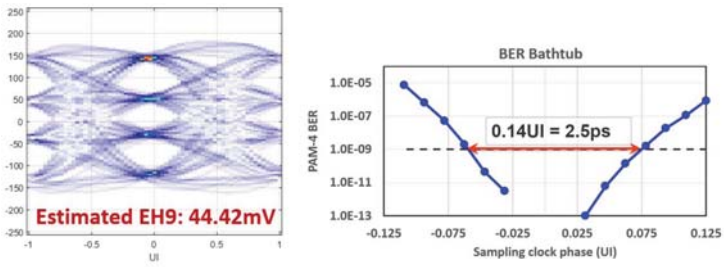
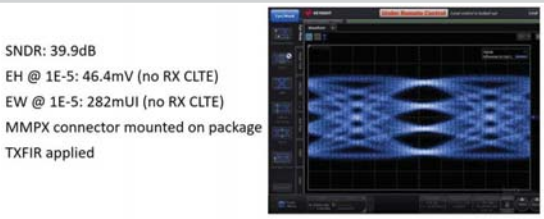


Figure 11.1.4: TX block diagram, showing duty-cycle calibration.



Channel	Loss @ Nyquist (dB)	Tested lanes	Worst BER
5mm	3.7	8	7.00E-11
15mm	4	8	6.00E-11
50mm	7	4	1.30E-10

Figure 11.1.5: Transceiver performance summary: measured PAM-4 eye-diagram captured with on-die scope (top left), measured bathtub curve (top right) and worst-case BER per channel category (bottom).



	[3]	[5][6]	[4]	This Work
Tech.	7nm	14nm	16nm	7nm
DR (Gb/s)	50	100	56	112
Loss at Nyq. (dB) / BER	20 @ BER<1e-12	19.2 @ BER<1e-12	8 @ BER<1e-15	4 @ BER<6e-11 7 @ BER<1.3e-10
Power (mW/Lane)	111 (RX analog only)	340 (RX & TX analog) 56 (Digital)	126 (Analog & digital)	153 (Analog) 38 (Digital)
Area (mm²/lane)	0.067 (RX analog only)	0.148 (Analog, active area)	0.33	0.167 (Analog) 0.061 (Digital)
TX Arch.	NA	SST 8-tap FIR	SST	5-bit DAC 5-tap FIR
RX Arch.	CTLE/ FFE/3-tap DFE	CTLE/VGA/DFE	CTLE/VGA/CDR	CTLE/VGA/CDR

Figure 11.1.6: TX eye diagram at 112Gb/s (top) and performance comparison table (bottom).

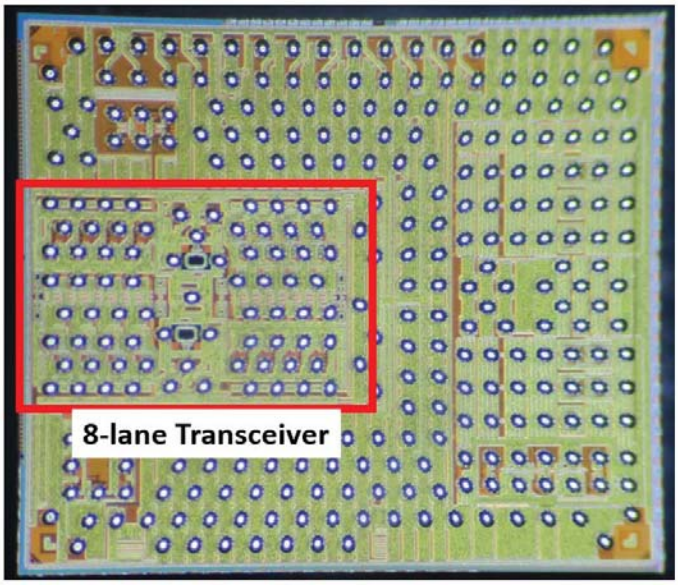


Figure 11.1.7: Transceiver photograph.

978-1-7281-9549-0/21/$31.00 ©2021 IEEE