6.5 A 1.6Tb/s Chiplet over XSR-MCM Channels using 113Gb/s PAM-4 Transceiver with Dynamic Receiver-Driven Adaptation of TX-FFE and Programmable Roaming Taps in 5nm CMOS

G. Gangasani[1], D. Hanson[1], D. Storaska[1], H. H. Xu[1], M. Kelly[1], M. Shannon[1], M. Sorna[1], M. Wielgos[1], P. B. Ramakrishna[2], S. Shi[3], S. Parker[1], U. K. Shukla[2], W. Kelly[1], W. Su[3], Z. Yu[4]

[1]Marvell, Hopewell Junction, NY
[2]Marvell, Bangalore, India
[3]Marvell, Shanghai, China
[4]Marvell, Santa Clara, CA

2022 IEEE International Solid-State Circuits Conference (ISSCC) | 978-1-6654-2800-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/ISSCC42614.2022.9731636

Hyperscale data center applications are driving the need for high bandwidth, high throughput per chip-edge, ultra-low-power serial-IO solutions over extremely short-reach (XSR) MCM connections. This paper demonstrates an MCM chiplet with >1.6Tb/s throughput over a range of CEI-112G-XSR standard-compliant channels using a fully integrated and adaptive PAM-4/NRZ transceiver supporting data rates from 9.8-to-113Gb/s. Key features of the transceiver include an architecture designed to lower TX signal launch amplitude and RX data path gain to relieve power/area/bandwidth constraints. It achieves required equalization while minimizing residual-ISI and quantization errors for improved performance. It has the capability for RX-driven dynamic adaptation of TX and RX equalization settings for seamless bring-up over hundreds of lanes in an MCM application, as well as to track supply and temperature drifts. It also features programmable TX-FFE roaming taps to boost performance in case of severe reflections and includes precision clock phase generation and correction.

The transceiver equalization strategy is geared towards achieving a low power and area signature while targeting <10^{-9} BER performance across XSR channels in PAM-4 mode. To that end, most of the 1st post-cursor (post-1) ISI is nulled using the RX-CTLE peaking. Residual-ISI is reduced by using fine peaking steps (<0.2dB). Remaining residual-ISI terms at tap positions (-1,1,2) are targeted by TX-FFE4 taps with small quantization errors (<1mV/LSB). As the TX-FFE4 by design is targeting residual-ISI terms after CTLE equalization, the signal de-emphasis at the RX-input is limited over XSR MCM channels. This lowers power and area, not only in the TX by lowering the launch amplitudes, but also in the RX by requiring less DC-gain to maintain signal envelop at the samplers. Impedance mismatches in some XSR applications could be unavoidable, for example due to congested C4 bump-via regions to achieve high throughput/mm or due to a silicon interposer in a co-packaged optics application. These mismatches could lead to substantial reflections and a sharp performance roll-off. To mitigate this the design has the capability of using two independent programmable roaming taps in TX-FFE between tap positions 3-to-22. The transceiver has the capability for all TX/RX equalization settings to be RX-driven using suitable data qualifications and adapted dynamically using a shared token-based backchannel across multiple lanes. As shown in Fig. 6.5.1, the transceivers in the chiplet are configured as an 8-port IO-core over two chips in an MCM, leading to 16-TX/RX lanes for the chiplet. A single PLL per core generates quarter-rate clocks which are distributed over a resonant clock tree to save power.

To minimize power in the TX, a quarter-rate clocking architecture is adopted along with a tailless 4:1 multiplexer and CML output driver [1]. The tailless CML driver approach takes advantage of the reduced TX launch amplitudes in the architecture (~550mV_{ppd}) to reduce FET sizing, power, and parasitic capacitance. The TX architecture implementing FFE-4 (pre-1, post-1, and post-2), along with 2-programmable roaming taps (between tap positions 3-to-22) is shown in Fig. 6.5.1. The data path includes a 64:4×3 serializer followed by three equally weighted driver slices. Thermometer-encoded quarter-rate data is routed to the 3 driver slices in PAM-4 mode, while all 3 slices receive same data in NRZ mode. Each of these ×3 4-bit data streams are offset by 1UI in the FIFO to enable FFE-4 operation. A dedicated data path exists for each roaming tap in the serializer and in the slices. Programmable tap delay between position 3-to-22 is achieved through a combination of 0-3UI delay in the logic RLM and modulo-4 delay in the respective 64:4×3 serializers. Each of the slices have a maximum of 12 segments for C_0, 2 for C_{-1}, 4 for C_1, 2 for C_2, and 1 each for the two roaming taps. In each segment, the 4UI-wide data bits are carved into 1UI wide pulses such that FFE segments are using appropriate tap bits corresponding to the cursor, the details of which are shown in the UI carver timing section of Fig. 6.5.2. The 1UI-wide pulses are converted to full-rate data by the 4:1 MUX in the pre-driver before reaching the final tailless CML driver stage. The current from all segments is summed at the output of the tailless CML driver, and coarse FFE is achieved by turning ON/OFF the segments for each tap and cursor. Fine tuning of the tap weights is done by adjusting tap currents using the V_{cas} gate voltage of the driver segments. This combination of coarse-fine settings enables the required tap range while ensuring <1mV/LSB quantization error. The FFE-DAC biasing network, shown in Fig. 6.5.2, is designed to ensure dynamic glitch-free adaptation of coarse-fine tap settings. To achieve better power efficiency and finer tap control only the highest ON segment of the taps are controlled by their respective programmable DACs. The other ON segments use a V_{rail} voltage from a replica driver segment. All the DAC updates can be dynamically adapted over a single-ended backchannel from the RX-side. The backchannel is a token-based Manchester encoded communication channel shared across the IO-core lanes. The updates are inc/dec to tap weights based on RX-driven data qualified accumulators for each tap or inc/dec to the cursor based on power envelop calculations. These updates trigger the FFE engine to enter one of the shown modes to ensure glitch-free addition/deletion of segments and DAC updates

The RX architecture is shown in Fig. 6.5.1. The DC-coupled transmitted signal meets the analog-frontend at the RX with its ESD protection, 100Ω differential termination and broadband Tcoil-network. Following this is an on-chip AC-coupled stage with a programmable passive attenuator for AGC of the signal envelop, driving a trans-admittance trans-impedance (TAS-TIS) based CTLE. The RX equalization is primarily done using this CTLE and its F_{baud}/2 boost (min-max of ~5dB) with fine step (<0.2dB) increments. The PAM-4/NRZ data at the CTLE output is sampled by array of 20 quarter-rate samplers, 5 per quarter. Each quarter has 3 data samplers clocked by a quarter-rate data clock phase to sample +2, 0, -2 data levels, an edge sampler clocked with quarter-rate edge clock to sample edge information, and a swappable sampler off data clock for sampling ±3, ±1 levels. The data pipeline for the swappable sampler or the data samplers in NRZ mode through the 4:32 deserializer are set to low-to-zero data activity using a saturated level from the DAC to minimize power. The levels for each of the samplers are generated by corresponding metal resistor-ladder-based shift-register-controlled voltage DAC. The DAC is 8b, signed-differential with necessary range to span the input data levels to reach ±3 levels and apply appropriate local offset correction for the samplers while maintaining <1.5mV/LSB resolution. The sampler swap is done seamlessly in the digital domain and makes possible dynamic adaptation algorithms for global and local voltage offset correction, AGC gain control, CTLE peaking, and FFE tap calibration. The data from edge and data samplers from all the quarters is aggregated to generate early and late timing information using ±3 and ±1 edge qualifications to drive a low-latency 2nd-order BB-CDR loop to control the phase rotator in the quarter-rate clock phase generation macro.

The 45°-spaced quarter-rate edge and data clocks are controlled by a single 8b phase-interpolator (PI). It achieves superior INL (<1.5LSB, 64-step/UI) using eight 45°-spaced input phases and a DAC that is intrinsically octal-symmetric. The input phases are generated using a delay-line (DLL1), shown in Fig. 6.5.3, using the quarter-rate global clock from the PLL. The average differential IQ errors are dynamically corrected using a 10b resistor-ladder-based voltage DAC controlling the regulator driving the supply for the delay elements. The single differential output from the PI enters another delay-line (DLL2) to generate eight 45°-spaced quarter-rate edge/data clock phases with requisite output buffering to drive the sampler fanout. The average IQ-error and differential DCD are sensed at the output to drive the dynamic adaptation of the delay-line supply regulator using a 10b voltage DAC, and the DCD correction for each differential clock phase. The end result of these dynamic correction loops is a 3-sigma clock phase position error of <500fs at maximum rates over mismatch and drift. The TX IQ clock phase generation is similar, with a default option to bypass the PI to save power.

The transceiver is implemented in 5nm bulk CMOS FinFET technology. All measurements are done with a PRBS31 pattern. Figure 6.5.4 shows the measured TX output (equalized) PAM-4 data eye at 113Gb/s, and the on-chip BER contour map measurements done at the RX for the transceiver operating at 113Gb/s in PAM-4 mode. At BER <10^{-9}, the vertical-horizontal eyes (Veye-Heye) are at least 20mVp-5%$_{UI}$ open. The clock pattern measured at the TX output, shows an integrated RJ of 139fs_{rms}, as shown in Fig. 6.5.5. The chiplet in Fig. 6.5.7 shows two dies with 8 TX/RX ports each. Thirteen of these lanes are over die-to-die channels with insertion losses ranging from 0.8dB to 11.5dB at 28GHz. The BER performance over these 13-lanes for 54-to-128Gb/s PAM-4 and 16-to-56.5Gb/s NRZ modes shows error-free operation within the gating window for rates less than 103Gb/s. Figure 6.5.6 shows the measured BER performance and comparison. Roaming taps boost performance for short channels, likely due to reflection cancellation. This chiplet achieves a net throughput of >1.6Gb/s over these 13-lanes at 128Gb/s. The performance demonstrates that over the full range of XSR channels the presented solution achieves the highest reported data rates and margins.

References:

[1] Z. Toprak-Deniz, et al., "A 128Gb/s 1.3pJ/b PAM-4 Transmitter with Reconfigurable 3-tap FFE in 14nm CMOS", *ISSCC*, pp. 122-123, Feb. 2019.
[2] R. Shivnaraine et al., "A 26.5625-to-106.25Gb/s XSR SerDes with 1.55pJ/b Efficiency in 7nm CMOS", *ISSCC*, pp. 182-183, Feb. 2021.
[3] R. Yousry et al., "A 1.7pJ/b 112Gb/s XSR Transceiver for Intra-Package Communication in 7nm FinFET Technology", *ISSCC*, pp. 180-181, Feb. 2021.

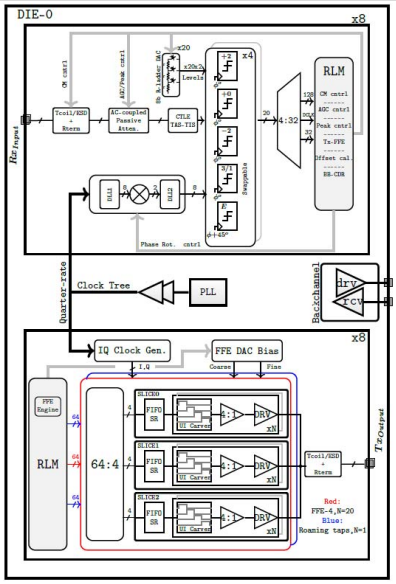


Figure 6.5.1: Block diagram of a single 8-port transceiver core, two of which are used to configure the 16-lane MCM chiplet.

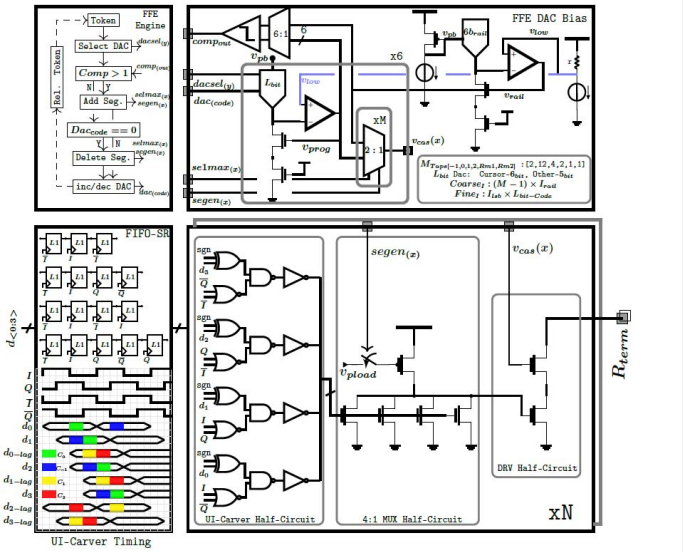


Figure 6.5.2: Details of the Segments in the Slice (bottom right), along with UI carver timing diagram (bottom left), and the glitch-free fine/coarse FFE DAC control of segments and bias currents (top).

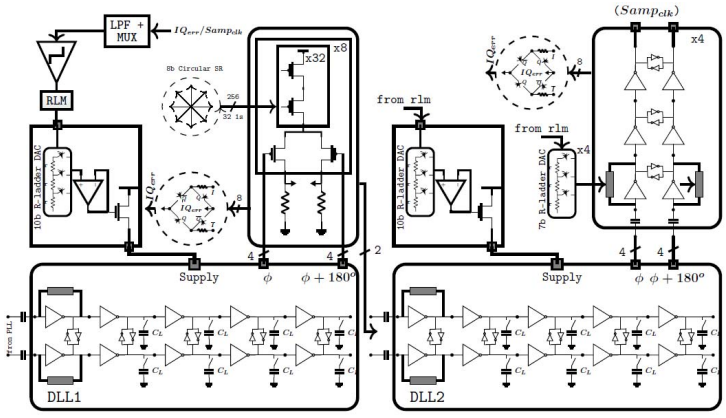


Figure 6.5.3: Details of 45°-spaced quarter-rate clock generation with dynamic adaptation of phase errors and 360° phase-rotation.

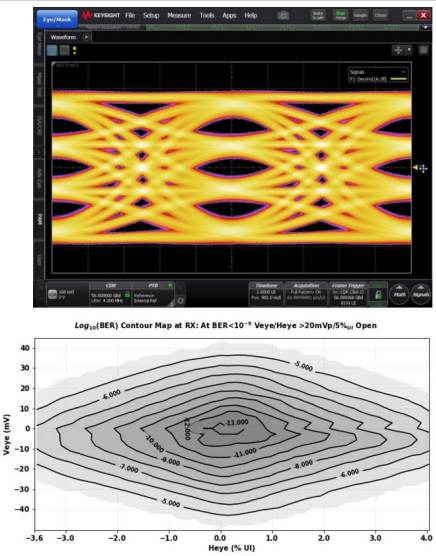


Figure 6.5.4: Measured TX output data eyes at 113Gb/s (with equalization) in PAM-4 mode (top). Measured on-chip BER contour map at RX at 113Gb/s in PAM-4 mode over 54mm channel length (bottom).

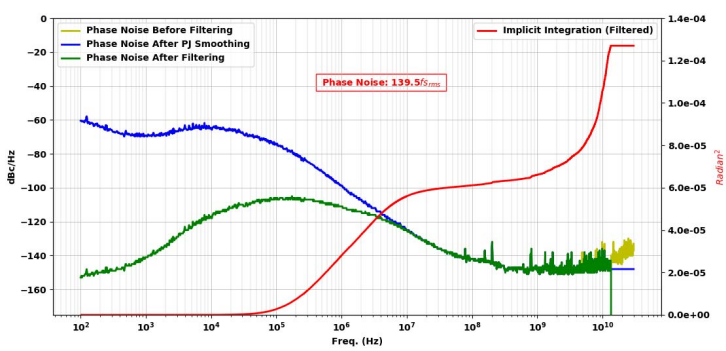
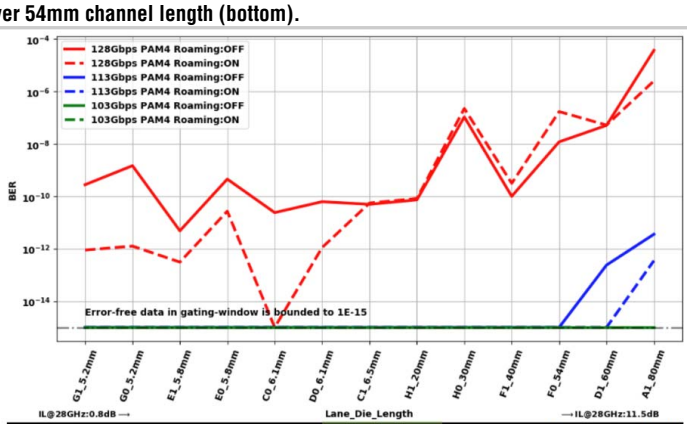


Figure 6.5.5: Measured RJ at TX output with clock pattern.



	This Work	[2]	[3]
CMOS FinFET	5nm	7nm	7nm
Data Rate (Gb/s)	113	106.25	112
Channel lengths (mm)	5-to-80	5-to-50	5-to-50
Max. Channel loss@28GHz (dB)	11.5	9.6	7
Throughput/mm (Gb/s/mm)	475	722	NA
Power (pJ/b)	$1.55_{@113Gb/s}$	$1.55_{@106.25Gb/s}$	$1.7_{@112Gb/s}$
Area ($mm^2/lane$)	0.264	0.225	0.228

Figure 6.5.6: Measured BER for full transceiver over the MCM channels (top), along with a table of comparison (bottom).

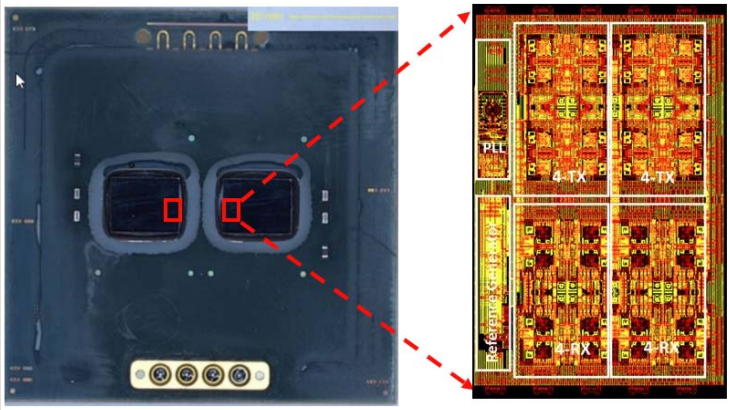


Figure 6.5.7: An optical image of the two chip MCM chiplet with 8-TX/RX each (left).

978-1-6654-2800-2/22/$31.00 ©2022 IEEE