

Design of 224Gb/s DSP-based transceiver in CMOS technology: signal integrity, architecture, circuits, and packaging

Jihwan Kim, Ariel Cohen, Mike Peng Li, Ajay Balankutty, Sandipan Kundu, Ahmad Khairi, Yoel Krupnik, Yoav Segal, Marco Cusmai, Dror Lazar, Ari Gordon, Noam Familia, Kai Yu, Yutao Liu, Matthew Beach, Priya Wali, Hsinho Wu, Masashi Shimanouchi, Jenny Xiaohong Jiang, Zhiguo Qian, Kemal Aygun, Itamar Levin, Frank O'Mahony

Intel Corporation

Abstract

In this article, we present analyses and implementation detail of 224Gb/s DSP-based electrical transceiver system fabricated in modern CMOS technologies for long-reach serial data communication. We investigate signal integrity performance at 224Gb/s using various modulation schemes with associated error correction codes for realistic channel characteristics, and discuss implementation techniques of the DAC-based transmitter (TX) and ADC-based receiver (RX) from architecture and circuits perspectives. Design techniques for high performance packaging is also discussed. Demonstrations of PAM-4 224Gb/s TX and RX fabricated in Intel 7 (formerly 10nm) and TSMC N5 (5nm) CMOS technologies are presented.

Introduction

Thanks to aggressive scaling of CMOS technology, both performance and energy efficiency of wireline SerDes transceiver circuits and systems have improved by more than 20x in the past 20 years as presented in Fig. 1. This trend of improved performance/power will continue with more advancement in the CMOS technology for the wireline transceivers benefiting from the technology scaling for digital signal processing (DSP) units which is becoming bigger portion in the overall system.

On top of 5G wireless, AI/ML has risen as a major force to increase global data communication throughput in mega data centers. To accommodate the explosively increasing throughput demand, aggregate bandwidth of networking switch SOC has reached 51.2Tb/s exercising 64 ports of 800GbE that consists of 8 lanes of 100Gb/s/pin SerDes. Field programmable gate arrays (FPGAs) also started adopting 116Gb/s SerDes PHY for high-bandwidth external data communication. Recently deployed networking ASICs and IPs have demonstrated high-performance and energy-efficient 100-116Gb/s Ethernet SerDes PHYs developed in advanced process technologies such as 10, 7 and 5nm CMOS [1-10]. The next focus toward 1.6TbE networking is naturally shifted to developing 224Gb/s/pin SerDes PHY optimizing performance-per-watt, physical form-factor, and ultimately, implementation/maintenance cost for all applications.

To implement the high-performance low-power 224Gb/s transceiver, many critical decisions need to be made. This includes choosing the optimal modulation scheme, robust/versatile (and power efficient) equalization schemes, architecture of the system, and topology of analog/mixed-signal circuit components to meet the system requirements in various aspects. This process becomes extremely challenging in designing a 224Gb/s transceiver because of the facts that (1) the unit interval (UI) of data becoming comparable to a single gate delay of the transistor (putting challenges onto timing/delay sensitive circuits) and (2) analog circuits' bandwidth being scaled linearly while requiring improved noise, linearity, and power-efficiency (putting challenges onto clock buffers, amplifiers, driver, sampling switches, and comparators). Therefore, thorough investigation and optimization of the overall link system as well as individual circuit blocks must be performed.

In this article, we present the overview/summary of development process of 224Gb/s PAM-4 transceiver in modern CMOS processes, including the background of PAM-4 modulation selection, the TX and the RX architecture, cutting-edge design techniques of key circuit/system components such as PLL, clock distribution, TX DAC,

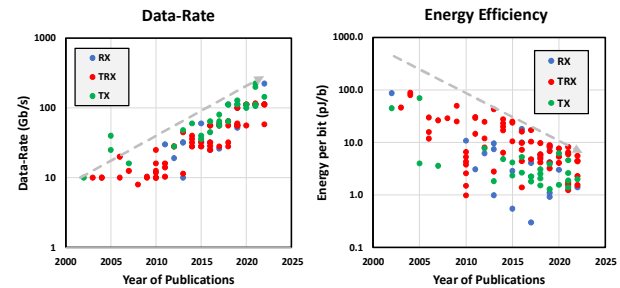


Fig. 1. Trend of published CMOS SerDes in the past 20 years.

RX analog front-end (AFE), RX ADC, and high-performance package. Measurement result of the TX and the RX is followed.

224 Gb/s Modulation and Channel Characteristics

Modulation schemes determine not only the way the binary data payload is encoded/decoded before being transmitted through the channel medium, but also the bandwidth requirements for the components in the link system. Currently, four-level pulse amplitude modulation (PAM-4), which encodes two bits data into one symbol, which increases spectral efficiency (and reduces Nyquist frequency by half) compared to the non-return-to-zero (NRZ) modulation, is used in 50G/100G Ethernet and 56G/112G OIF-CEI electrical links. Naturally, it is highly desired to continue using PAM-4 for the 224 Gb/s SerDes considering benefits in backward compatibility, electrical-to-optical interface compatibility, utilization of established and proven testing methodology and equipment. However, it is necessary to carefully investigate all possible options to understand their implications on the system performance and required circuit specifications.

A. PAM_L Modulation Considerations

Table 1 summarizes the key characteristics for PAM_L, where L is 4, 5, 6, 7, 8, 16. Higher order modulations require lower analog bandwidths for the same data-rate due to relaxed symbol rates, which makes designing process for critical circuit components relatively easier. However, lower signal-to-noise ratio (SNR) caused by reduced eye height becomes a great challenge for the link's performance especially when L is greater than 4.

Data rate (Gb/s)	224	224	224	224	224	224
PAMn levels	4	5	6	7	8	16
Bit per symbol	2.00	2.32	2.58	2.81	3.00	4.00
Symbol rate (GS/s)	112.00	96.47	86.66	79.79	74.67	56.00
UI (ps)	8.93	10.37	11.54	12.53	13.39	17.86
Nyquist freq (GHz)	56.00	48.24	43.33	39.90	37.33	28.00
SNR delta	0.00	-2.50	-4.44	-6.02	-7.36	-13.98

Table 1. A comparison of 224Gb/s modulation schemes.

We have investigated performance of the link using PAM-4, 6, and 8 modulation schemes (which are the most viable options for the next generation SerDes) using behavioral models of the TX/RX with circuit impairments included over various channels [11]. For a realistic end-to-end channel characteristic (including package, PCB, connector, and cable) with ball-to-ball insertion loss (IL) of around 30dB at 56GHz (Nyquist frequency for 224Gb/s in PAM-4) and insertion loss deviation (ILD) of less than 3dB, the PAM-4 modulation outperformed PAM-6 and 8 showing larger eye openings at the targeted bit-error-rate (BER) as shown in Fig. 2. The TX/RX characteristics and the channel profiles are also presented in Fig. 2.

B. FEC Performance and Modulations

The forward error correction (FEC) has become essential part of SerDes using PAM-4 modulation starting from 50G Ethernet and

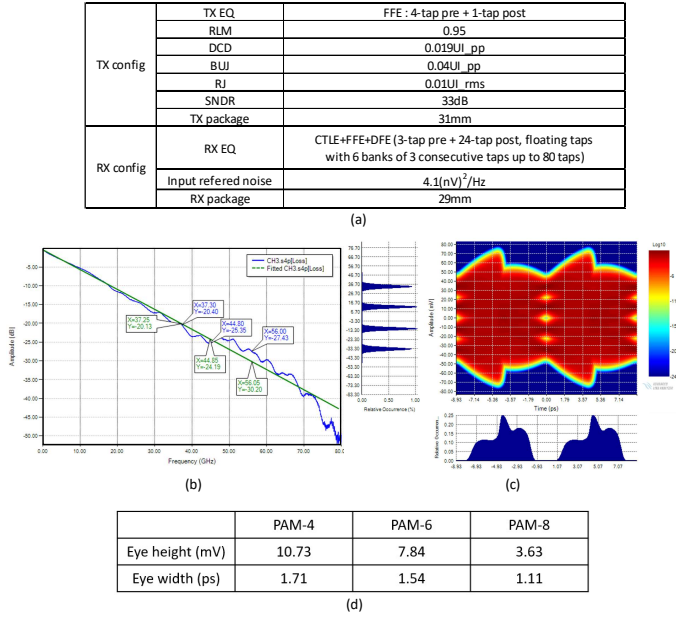


Fig 2. 224Gb/s link simulation results: (a) TX/RX configurations, (b) realistic end-to-end channel characteristic, (c) simulation results of PAM-4, and (d) performance comparison of PAM-4, 6, and 8.

Level (L)	$\log_2(L)$	Mod_Symb/FEC_Symb
2	1	10
4	2	5
6	$2.58 \approx 2.5$	4
8	3	$3.33 \approx 4$

Table 2. Number of modulation symbols per FEC symbol of PAM_L.

56G OIF/CEI. The number of the modulation symbols per FEC symbol depends on the order of the PAM_L modulation as summarized in Table 2. This information is used to translate the modulation symbol error to the FEC symbol error. For example, while 5UI-long modulation symbol burst error may cause only one FEC symbol error with PAM-4, it may cause two FEC symbol errors with PAM-6 or PAM-8. The relation between symbol error ratio (SER) and signal-to-noise ratio (SNR) is shown in Fig. 3(a) for PAM-4/6/8. Note that higher SNR is required for higher order modulation to achieve the same raw SER. Achieving higher SNR generally requires larger device power in implementation. If the device power is assumed to be the same, higher order modulation would result in higher SER.

Another important performance metric to consider for FEC is the burst error probability. Decision feedback equalizer (DFE) error propagation is often considered as the primary source of the burst error. Example error propagation probability of 1-tap DFE model (simple but very useful to study the burst error mechanism) is shown in Fig. 3(b) for PAM-4/6/8. The higher the order of the modulation for a given DFE tap coefficient, the more chance to have longer burst error exists.

Combining the random error model and the burst error probability model together, FEC performance model can be generated. Figure 3(c) shows the comparison of Reed Solomon Ethernet PAM-4 KP FEC (RS(544,514)) performance among PAM-4, 6, and 8 with certain error propagation condition. This example data indicates that a stronger FEC would be needed for PAM-6 and PAM-8 while KP FEC still works for PAM-4. For instance, while PAM-4 can achieve $1e-15$ post-FEC BER with $1e-4$ pre-FEC BER using KP FEC, PAM-6 and 8 require $1e-6$ and $1e-8$ pre-FEC BERs to achieve the same post-FEC BER, respectively. Using stronger FEC is not desired because of extended latency and increased power consumption of the system.

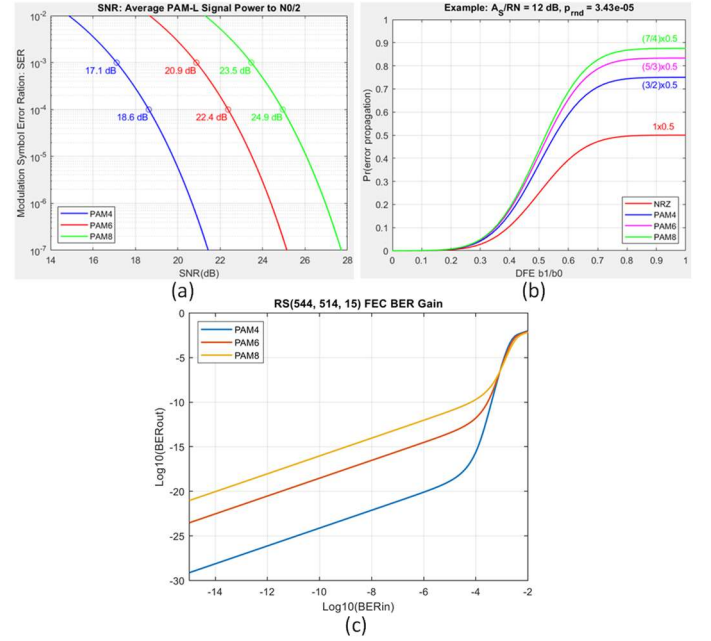


Fig. 3. (a) Modulation SER vs. SNR, (b) error propagation probability of PAM-4/6/8, and (c) FEC performance comparison of PAM-4/6/8.

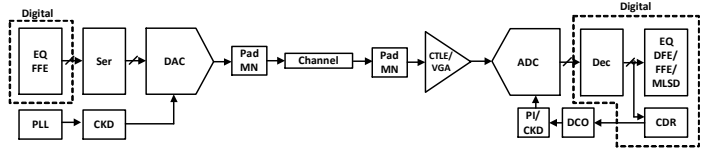


Fig. 4. A 224Gb/s transceiver architecture.

Based on the investigation result of the link performance with associated FEC over the realistic channel characteristic, we concluded that the PAM-4 modulation is the optimal selection for the modulation scheme to be used for 224Gb/s SerDes. Now, developing the system/circuits to enable 56GHz (Nyquist of PAM-4) analog bandwidth is the next focus.

Transceiver Architecture

The transceiver architecture for a power efficient 224Gb/s PAM-4 is shown in Fig. 4. Basic structure follows the trend of 100-116Gb/s DSP based transceivers that demonstrated compelling performance and power efficiency as in [1-4, 6-8, 10], but with architecture and design optimization for 2x higher bandwidth requirement. The TX consists of a digital FFE, a data serializer, a 7-bit DAC with a CML driver, an LC-PLL, clock distribution, and an output pad network. The RX includes the pad network, a front-end continuous-time linear equalizer (CTLE) and a variable gain amplifier (VGA) stage, a time-interleaved SAR-ADC, clock distribution with phase interpolator (PI), a DCO-based CDR, and a digital equalizer (FFE and optional DFE). The partitioning between the analog and the digital equalization and performance requirement of each component are determined by system simulations. Design and implementation details will be discussed in following sections.

Circuit design for 224Gb/s TX in 10nm

The TX architecture for 224Gb/s PAM-4 is illustrated in Fig. 5 [12-13]. The data-path includes an FSM based clock phase selection loop (controlling PI) to ensure robust timing at the 8:4 serialization path across PVT variation. The DAC is 7-bit CML based topology incorporating a high-bandwidth 4:1 serializer. The LC-PLL generates

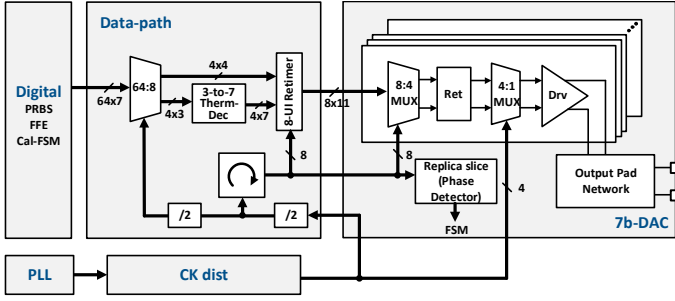


Fig. 5. A 10nm TX architecture.

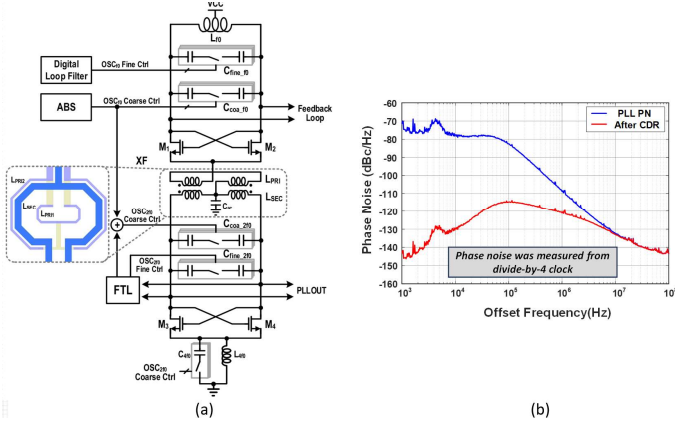


Fig. 6. (a) A schematic diagram of LC-DCO used in PLL and (b) measured phase noise plot (through divide-by-4 clock).

low jitter 28GHz clock and send it to the DAC and the data-path through high bandwidth and jitter-filtering clock distribution.

A. Clocking architecture and jitter budgeting

The first consideration in TX design is to decide an optimal data serialization (final stage) and a clocking architecture. There are three most realistic options: (1) 2:1 serialization using two-phase 56GHz clock, (2) 4:1 serialization using four-phase 28GHz clock, and (3) 8:1 serialization using eight-phase 14GHz clock. The first option is challenging because of difficulty in implementing a 56GHz PLL and clock distribution. Reasonable power consumption and required jitter performance will not be easily achieved in modern CMOS technology. The third option relaxes design burden in clocking circuits, but comes along with other challenges such as lower output bandwidth of the serializer due to large circuit/wire parasitic capacitance, large systematic error in clock phases due to crowded/asymmetric routings, and complicated clock spacing calibration method. The best trade-off in power and performance is achieved by choosing the second option, a 4:1 serialization and a quadrature clocking architecture in 10nm technology. Many 112-128Gb/s PAM-4 TXs (56-64Gbaud) chose this architecture while accurately calibrating clock spacing error, and achieved high enough output bandwidth [5, 10, 14-15] with good power efficiency. This can be a good candidate for 224Gb/s TX as well with bandwidth optimization of the 4:1 serializer. Although the same clock calibration method can be used, the oscillator frequency and the clock distribution bandwidth need to be doubled.

Design target for the jitter is $J_{3\sigma}$ of 1.053ps, J_{rms} of 210fs, and even-odd jitter (EOJ) of 160fs with UI of 8.93ps. This is translated to 90fs_{rms} of random jitter (RJ) and 385fs_{pp} of deterministic jitter (DJ) including periodic jitter) using conventional jitter breakdown. With linear-model based system simulations, we budgeted 60fs RJ to the PLL and 60fs RJ to the rest of the clock distribution assuming root-square-summing of them to be less than 90fs (also 50/50 allocation for DJ). This turns into the 28GHz PLL phase noise target of -103dBc/Hz at 1MHz offset with around 100KHz bandwidth, which is challenging to achieve with conventional architectures in the 10nm

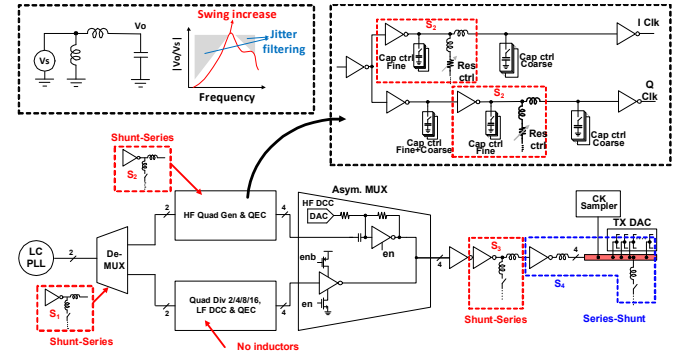


Fig. 7. TX clock distribution.



Fig. 8. Measured 56GHz clock pattern jitter.

CMOS. The main issue comes from low quality factor (Q) of the switchable capacitor (SW-cap) bank that is limited by channel resistance of the switches (in lower side of frequency range). The LC-tank Q is degraded limiting the performance of the oscillator. To achieve lower phase noise at the same power, a DCO based on coupled oscillators is used for the digital PLL as shown in Fig. 6 [16]. Higher SW-cap Q at 14GHz together with the coupling mechanism provides 4.6dB of phase noise improvement at the same power. A frequency tracking loop (FTL) ensures that the locking frequency of the 2f₀ oscillator matches the second harmonic frequency of the f₀ oscillator maximizing the phase noise improvement. Using this LC DCO, a digital PLL with 7ps-resolution TDC, digital loop-filter, and multi-modulus divider with the first order $\Sigma\Delta$ modulator achieved -102.3dBc/Hz at 1MHz offset while consuming 18.4mA from 1.8V supply. Bandwidth is set to 100KHz to suppress TDC quantization noise. The DCO is regulated by an LDO to minimize DJ as well.

B. Clock distribution with jitter filtering

High operating frequency means wide bandwidth for noise integration in the clock distribution network. A traditional approach to reduce the integrated noise (or jitter) is to lower down the noise floor of each stage at the cost of power. The RJ is also amplified through bandwidth limited circuits [17] (e.g., cascaded CMOS clock buffers with large fan-outs), which also limits power-jitter co-optimization in the high frequency clock distribution. For the 224Gb/s TX in 10nm CMOS, we used an inductor-based jitter filtering technique to distribute 28GHz clock through simple CMOS clock buffers. The shunt-series peaking inductors were used to generate a bandpass filter (Fig. 7) providing (1) amplitude peaking by the LC network, and (2) rejecting the high-frequency noise out of the band to reduce the jitter. By implementing these inductors with fairly low-Q (<5) using less conductive metals in back-end, a wide operation range of jitter filtering with minimal area overhead is achieved. In simulation with fan-out-3 clock buffers, 43% reduction in a jitter amplification factor at 28GHz was achieved. Due to its nature in shaping the voltage/current waveforms (similar to a class-F amplifier), dynamic

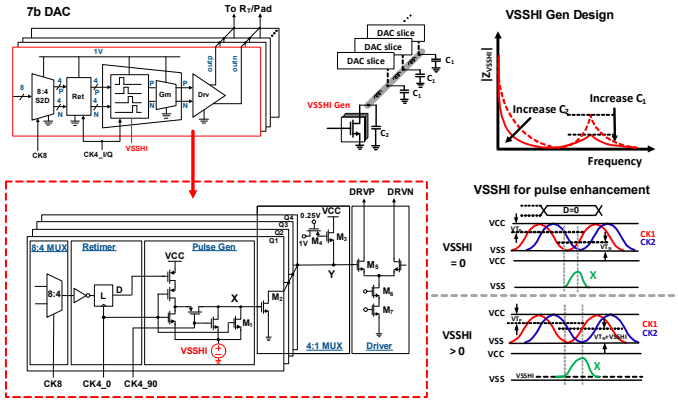


Fig. 9. Diagrams of 7b DAC based TX driver.

power consumption is also reduced by 20~30% in the frequency range of 16-32GHz. Another benefit of using inductively-peaked CMOS clock buffer is its negating effect on overall DJ caused by power supply noise. The phase of the output voltage changes in the opposite direction to that of the inductor-less CMOS buffer, and any supply voltage modulation results in delay cancellation to some extent when the inductor-less and the inductively-peaked CMOS buffers are mixed in the distribution. Simulation shows DJ is reduced by 20% by one inductively-peaked buffer used within a 6-stage buffer chain. As presented in Fig. 7, the full clock distribution network has four inductively-peaked CMOS buffers (three shunt-series and one series-shunt) along with quadrature clock generator/corrector and a clock sampler that detects the clock spacing error and the clock amplitude exceeding the overstress limit (FSM loop calibrates both). The dual-path (high frequency and low frequency) clock distribution supports multiple data-rates. In measurement, 56GHz output clock pattern shows 56fs (second-order CDR) or 65fs (first-order CDR) of RJ with 4MHz CDR bandwidth proving effectiveness of jitter filtering technique (Fig. 8).

C. TX DAC: 4:1 serializer and driver

The TX is based on 7-bit DAC (112GS/s) where the resolution was determined from the equalization step size and signal-to-noise/distortion ratio (SNDR) requirement. Key performance metric of the DAC is bandwidth (ISI/RJ amplification), static/dynamic linearity (level mismatch ratio, SNDR), and output reflection (effective return loss). The 4:1 serializer is the most critical block that sets the maximum bandwidth of the TX, and poor performance of it will deteriorate all DAC metrics. Therefore, choosing the optimal circuit topology and exercising the best layout practice is critical. The pulse generator based 4:1 serializer [14-15] was proven to be effective and power efficient for high data-rate applications where the bandwidth at the output of transconductance (Gm) stage can be extended with help of active/passive inductive peaking. Additional care was taken to choose the right pulse generator topology in this TX design to minimize uncorrelated jitter at the differentially taken TX output. Combining the rising and falling edge of two phases of the clock provides averaging of uncorrelated jitter from each clock edge and attenuate the overall (uncorrelated) jitter by 30% (or 3dB) [13].

The DAC takes 4-bit binary and 3-bit thermometer coded data from the 64:8 serializer (partitioning was done in a way to minimize the static linearity). Each slice includes 8:4 serializer, retimer, pulse generator based 4:1 serializer, and CML diff-pair based NMOS driver as presented in Fig. 9. The NMOS diff-pair driver with a current source provides high output impedance and current (swing) controllability. The CML MUX performs data serialization by switching the current generated by the 1-UI pulse to the voltage at the CML diff-pair input through a NMOS load acting as an active inductor. Thus, the bandwidth at the serializer output is improved. The proper common-mode voltage is set by the gate bias voltage of the NMOS load through a large PMOS linear-mode resistor. The height of the pulse that switches the NMOS based MUX is increased to further improve the bandwidth and the swing at the slow process corner. This is performed by elevating the local VSS (VSSH) of the pulse generator with a controllable NMOS linear-mode resistor. The

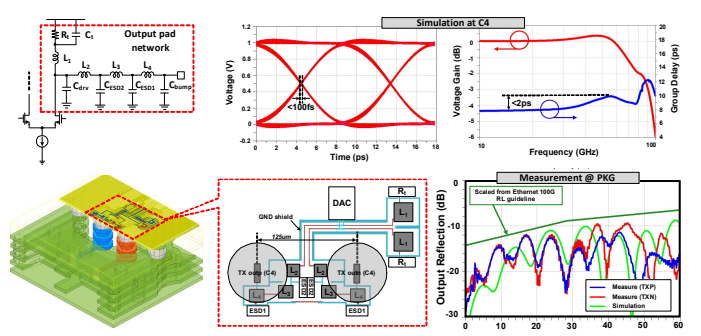


Fig. 10. Diagrams of 9-th order LC based output pad network.

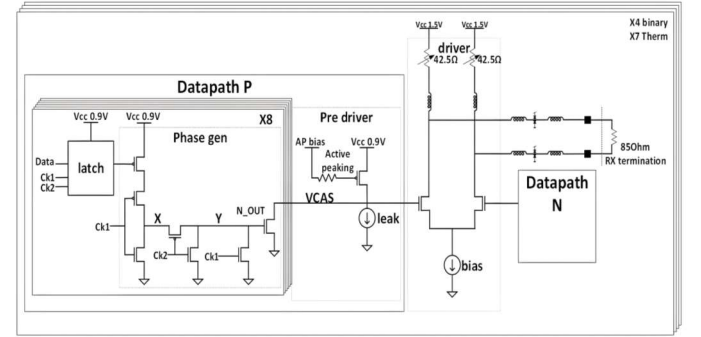


Fig. 11. A schematic diagram of TX DAC in 5nm CMOS

process monitor can set the optimal voltage for the VSSH. The high-frequency impedance of the VSSH generator should be minimized by placing large enough decoupling capacitors to ensure that voltage ripples caused by routing inductance is minimized not to degrade the linearity of the DAC.

D. 9-th order LC based pad matching network

Due to extreme bandwidth requirement, the higher order LC based matching network than conventional T-coil or Pi-coil was used in this TX design. The 9-th order LC network (four inductors and five capacitors) provides 68-GHz 1-dB bandwidth with less than 2-ps group delay variation up to Nyquist at C4 bump and superior impedance matching property compared to other designs using simpler LC networks. The high order LC network also provides less sensitivity on variation of the inductance and the capacitance because of its broad-band matching characteristic. Simulation and measurement results are shown in Fig. 10 along with 3D diagram of the pad network. The design is optimized to provide good phase linearity as well to maximize the horizontal eye opening. Large ESD protection diodes are split into two locations while the total protection capability is good for handling 2.5A ESD current.

Design of 224Gb/s TX in 5nm CMOS

The DAC based TX was implemented in TSMC 5nm technology with a few design changes to improve the energy efficiency. Instead of coupled oscillator based 28GHz LC-DCO, a 14GHz CMOS Gm based LC-DCO was used along with 8-phase 14GHz clock distribution to get benefit of better phase noise at lower frequency and reduced power consumption for the clock distribution. This choice leads to relatively complex clock calibration (for all 8 phases of the clock), which is handled effectively in digitally controlled loop. One challenge in this architecture is to optimize the performance of 8:1 serializer unit (8:1 MUX) before the final CML driver. Leveraging lower parasitic of the technology and careful layout optimization (to reduce data dependent jitter and deterministic clock spacing error) enabled this architecture to meet the performance requirements. The TX DAC schematic diagram is shown in Fig. 11. It is worth mention that reduced parasitic capacitance at the signal output (pad) also relaxed the design order of the pad matching network.

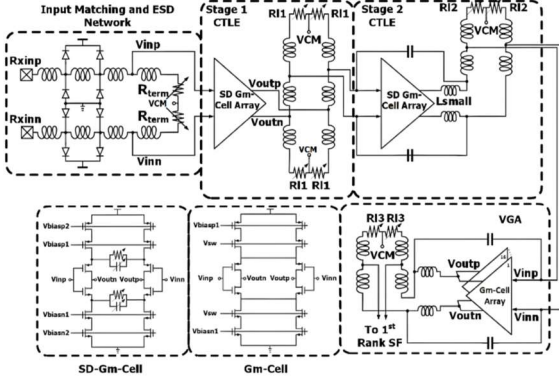


Fig. 12. A schematic diagram of RX AFE in 10nm CMOS.

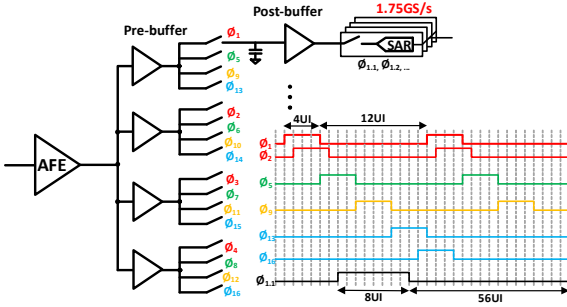


Fig. 13. A simplified diagram of 64-way interleaved ADC based RX and clock timing in 5nm.

Circuit design for 224Gb/s RX-AFE in 10nm CMOS

The first prototype of 224Gb/s PAM-4 RX-AFE was implemented in 10nm CMOS [18]. This design employs the 9-th order LC based input matching network followed by three stages of amplifier chain to drive the ADC. The first two stages feature CTLE using complementary-Gm based amplifiers with inductive peaking and the source degeneration using controllable RC. This hybrid CTLE provides good linearity, good mid-frequency zero-location controllability, and good high-frequency channel matching capability [1].

To meet the AFE noise target ($<2\text{mVrms}$ output referred), the first stage CTLE needs to provide the highest gain, which requires the highest Gm and relatively high load resistance. The output bandwidth is extended by using 3-segment inductors (π network) whose Q is controlled by modulating the load resistor (to modulate the peak gain). Following stages provide additional boosting (CTLE) and broad-band gain control (VGA). To further extend the bandwidth, cross-coupling capacitors (anti-Miller capacitors) were used in these stages.

Circuit design for 224Gb/s RX in 5nm CMOS

A. ADC based RX architecture

Choosing the optimal interleaving architecture for ADC based RX is the most important step in whole design process. Majority of 100-116Gb/s ADC based RXs employed 64-way interleaving to relax the convergence time requirement for ADC slices (i.e. 0.875GS/s for unit ADC) [1, 3-4, 6]. For doubling the overall sampling rate to 112GS/s , scaling to 128-way interleaving keeping the same unit ADS speed would not be an attractive solution due to many issues rising from multi-phase clock distribution, complex calibration requirements, and more importantly power. In this design, 64-way interleaving was used with 16-way interleaved sampling switches in the first stage (TH1) followed by 4-way interleaved sampling switches per each TH1 in the second stage (TH2) [19-20]. As shown in Fig. 13, this architecture requires 7GHz 16-phase clocks (1UI separated) with 25% duty-cycle for TH1 and $1/4^{\text{th}}$ frequency of that for TH2 with 12.5% duty-cycle. Therefore, each SAR-ADC operates at 1.75GS/s given 56UI of convergence time. This architecture provides good trade-off for overall complexity, power, and performance. Four TH1 switches are grouped and driven by a single pre-buffer to minimize channel-to-

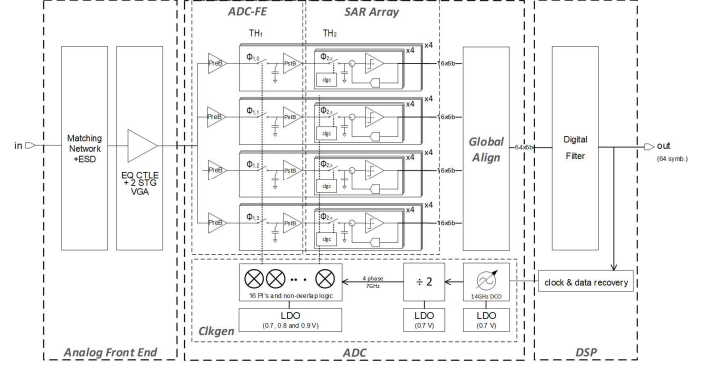


Fig. 14. Architecture of 224-Gb/s ADC based RX in 5nm CMOS.

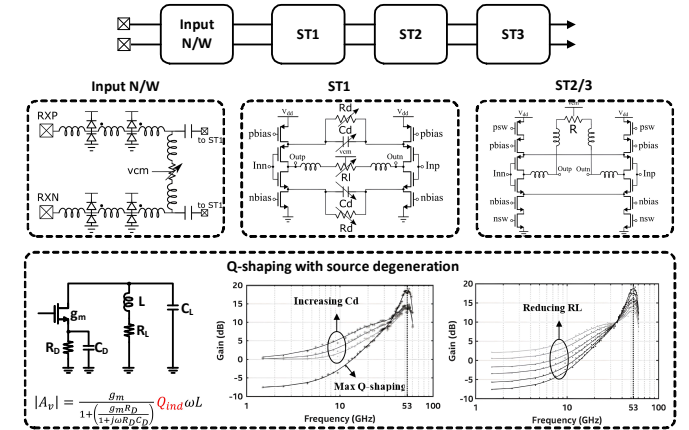


Fig. 15. The RX AFE in 5nm CMOS.

channel coupling (and ISI) ensuring only one TH1 in the grouped four is turned on at a time. Timing alignment between TH1 and TH2 clocks is managed by a logic within 16 PIs to guarantee non-overlapping clocks with required duty cycle. The entire RX architecture is presented in Fig. 14. Output of ADC (either directly or through the digital FFE) is used for clock-data recovery (CDR) loop using Mueller-Muller baud-rate phase detectors implemented in digital that optimizes the clock sampling point through an LC-DCO.

B. Analog front-end (AFE) design

Following the 9-th-order LC based input matching network that provides over 60GHz of 1-dB bandwidth at C4 bump, the CTLE performs major ISI cancellation. The hybrid CTLE similar to 10nm design was used. As presented in measured results in Fig. 15, the mid-band channel characteristic is matched by changing the degeneration cap (shifting zero frequency) while the high frequency characteristic is matched by the inductive peaking of the amplifier (Q-shaping [1]). The first CTLE has a simple shunt-peaking topology to make Q-shaping more effective. The second stage features VGA instead of CTLE as in 10nm design, which provides simpler calibration of the AFE to match the channel characteristic. The VGA output bandwidth is around 70GHz with help of the series-shunt peaking inductors. The entire AFE including the pad matching network provides 25dB of boosting with 18dB of peak gain at 53GHz in measurement.

C. ADC design

The ADC front-end circuits are presented in Fig. 16. The TH1 is implemented by PMOS switch with optimal size to minimize signal feedthrough and clocking power. The clock slope and amplitude are carefully optimized to minimize distortion from TH1 when switching. Following TH1 is the post-buffer that drives four SAR-ADC slices where the bandwidth and input-output coupling should be optimized to minimize ISI. If the bandwidth of the post-buffer is not sufficiently

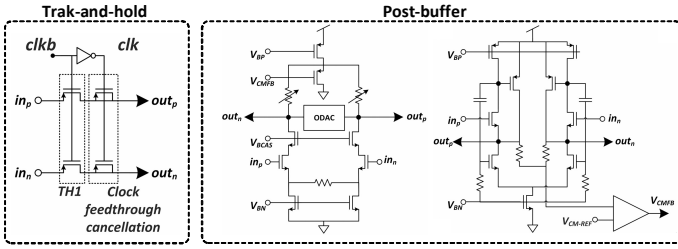


Fig. 16. The ADC front-end circuits.

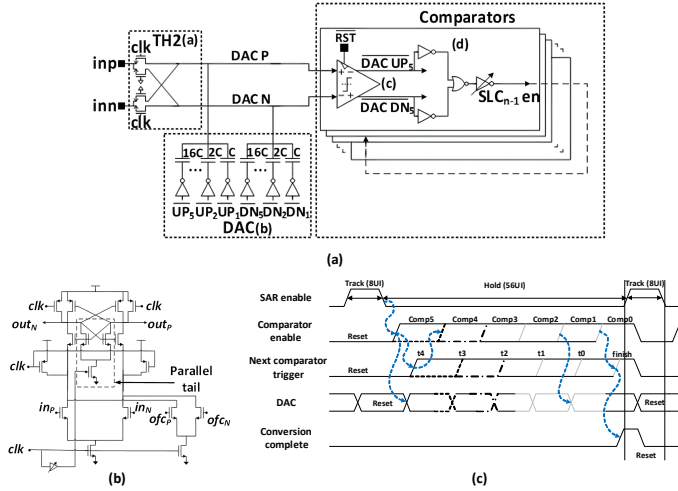


Fig. 17. Diagrams of (a) SAR-ADC (b) double-tail strong-arm comparator, and (c) ADC timing.

high to completely settle during the tracking time of 4UI, finite voltage error is generated in the next tracking phase due to input-output capacitive coupling of the post-buffer during the hold time of the TH1 [21]. Two-stage post-buffer is used in this design to improve the performance: a cascode amplifier with tunable gain (0-3dB) in the first stage and a semi-differential super-source follower in the second stage. The first stage provides gain tunability to calibrate the gain mismatch between channels and good input-output isolation, and the second stage provides high bandwidth driving capability to the next TH2 with low enough output impedance (by negative feedback) and optimal common-mode level.

Strong ISI cancellation through the AFE relaxes a resolution constraint of the ADC, and 6-bit resolution is found to be sufficient to achieve the target BER in the system simulation. The 6-bit sub-ADC shown in Fig. 17 operates at 1.75GS/s in a 64-way interleaved architecture. To achieve this high sampling rate with low power and area, a loop unrolled (LU) SAR ADC is implemented [22]. Each bit has a unique comparator that triggers the next bit's comparator when its decision is made. In this LU ADC, the comparator is not reset during the conversion, which leads to removal of a memory unit between the comparator and the DAC resulting in reduced delay, power, and complexity. However, lack of reset induces a common-mode drift at the input due to the comparator kickback, and it is compensated by having the capacitor DAC operate in a single-ended manner that generates opposite phase of common-mode drift. The comparator is based on a strong-arm topology with the addition of a parallel tail. The NMOS cross-coupled pair that clocks the parallel tail generates low-resistance discharge path for the comparator's latch. With this topology the clock-to-output time and the output slope of the comparator are improved that overall improves the ADC convergence time and performance.

D. RX clocking

A clock generation system is used to generate the sixteen of 7GHz clocks for TH1 and the sixty-four 1.75GHz clocks for TH2. As shown in Fig. 18, the clock generation system consists of a 14GHz DCO a

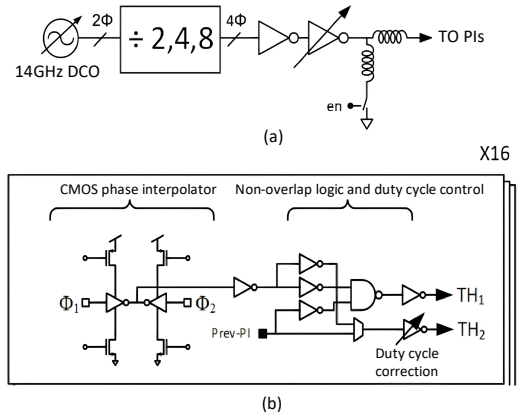


Fig. 18. Diagrams of (a) ADC clock generation and (b) CMOS PI with 25% duty cycle for TH1.

divider, jitter-filtering clock distribution network driving 16 PIs, combinational logic to generate 25% duty-cycle TH1 clocks, and shift register-based clock divider to generate TH2 clocks. The clock generation system is locked by a digital CDR that detects the digital phases at the ADC output and controls the DCO and the PIs to the required frequency and phase. Relative phases with 1-UI spacing of the interleaved samplers are determined by the PIs. Since those PIs are not part of the main CDR loop and do not need to be continuously rotating, their linearity is not significant.

The 14-GHz clock goes through a quadrature divider to be fed into a set of 16 CMOS PIs with a maximum resolution of 100fs that are used to generate the 16 phases used to drive the TH1 of the ADC. A shunt-series peaking network provides jitter filtering to achieve <100fs RJ (Fig. 18(b)) for this clock. The PI is controlled using a fine control of the NMOS/PMOS strengths to achieve the required resolution.

Each PI output is fed into a combinational logic block that creates the 25% duty-cycle TH1 clock and the 50% duty-cycle main TH2 clock. The main TH2 clock is divided by 4 to generate the 12.5% duty-cycle TH2 clock for each individual SAR sub-ADC.

E. Digital equalizer

The digital equalizer is composed of 30-tap FFE and optional 1-tap DFE. The FFE is adapted by the sign-sign least mean square algorithm operated by on-die micro-controller. The CDR path takes the input at either direct ADC output or FFE output depending on configuration. While most of channel ISI is canceled by the AFE, flexible and fine-controllable FFE and DFE provides better pulse shaping of the overall system and improve the SNR because they effective cancels out ISI and reflections caused by channel discontinuity.

Packaging for 224Gb/s SerDes

In this section, we discuss several advancements in the package to enable low insertion loss at the Nyquist frequency of 56Gz. Thorough analysis and description of design technique are presented in [23].

A. BGA ball pitch

General BGA ball arrangements are configured in a way that signal balls (differential) are located at center surrounded by ground balls that generate common return paths. Forming similar pattern to the coaxial connector/cable, the BGA ball pitch must be chosen in such a way that the cut-off frequencies of the higher order modes are well above the operation frequency. Figure 19 shows that finer ball pitch is desired to ensure this condition is met along with better group delay variation and smaller insertion loss.

B. Skip-layer design for package routings

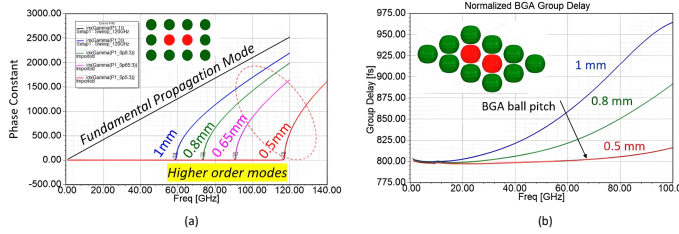


Fig. 19. (a) Cutoff frequency and (b) normalized group delay vs BGA ball pitch.

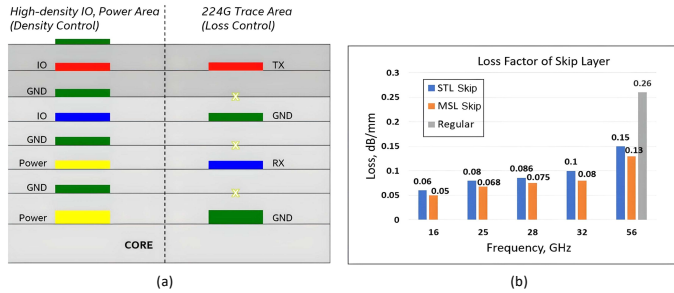


Fig. 20. (a) Localized skip-layer configuration, and (b) loss factors of different skip-layer routing configurations.

Frequency dependent dielectric/conductor loss is addressed by using localized skip-layer design for the package routing. This technique reduces the package loss effectively by reducing E-field stored in thicker dielectric (by skipping layer) and implement less than 0.15dB/mm of loss depending on its configuration. The skip-layer is used only on selected signals not to disturb power routing. Illustration of the skip-layer routing and comparison of the loss per unit length is presented in Fig. 20. This package routing performance have further improved with more advancement to enable 0.125dB/mm loss in the later development as presented in [24].

C. 224Gb/s PAM-4 package design

A 224G package design was carried out on a 12-layer organic package. The skip-layer configuration and the 0.5mm pitch BGA ball pattern are utilized. The package trace to PTH-BGA ball transition was optimized to diminish the discontinuity and target -15dB return loss up to 56GHz. Simulated performance of the package trace-via-PTH-BGA ball is shown in Fig. 21. The package achieved IL of ~2dB at 56GHz with RL of better than -10dB up to 80GHz. The effective ball/pad capacitance is smaller than 30fF.

Measurement Results

A. 10nm 224Gb/s TX measurement

Figure 22 shows the TX chip micrograph fabricated in Intel 7 (formerly 10nm) FinFET. The TX with PLL occupies an area of 0.1875mm². Measured QPRBS-13 eye diagrams for 112Gb/s and 224Gb/s PAM-4 are also presented. Measured RLM is 0.99 and SNDR is 37.5dB and 33.3dB at 112Gb/s 224Gb/s, respectively. Using 0.85V (digital) /1V (data-path and clock distribution) /1.5V (driver), energy efficiency of the TX is 1.88pJ/b (including PLL) or 1.74pJ/b (excluding PLL).

B. 5nm 224Gb/s TX measurement

Figure 23 shows the eye diagram and measurement summary for 224Gb/s TX implemented in TSMC N5 (5nm) CMOS. Measured RLM and SNDR are >0.977 and >37dB. Measured DJ and RJ are 120fs_{pp} and 113fs_{rms}, respectively in QPRBS-13 pattern. Analog power consumption is 290mW (1.3pJ/b) showing ~30% of power reduction thanks to technology scaling and architecture changes.

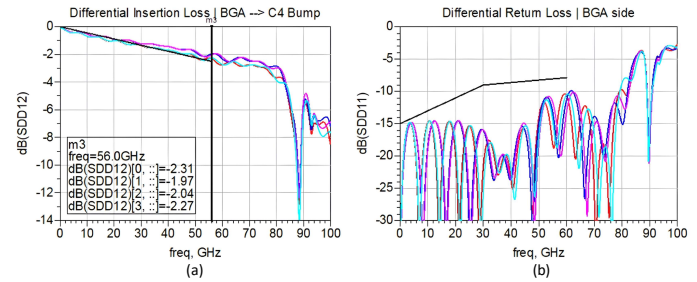


Fig. 21. Simulated package performance: (a) differential insertion loss, (b) differential return loss.

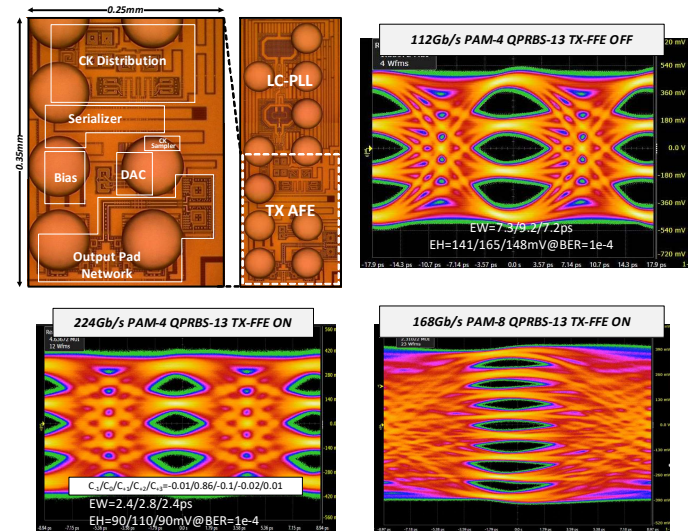


Fig. 22. A chip micrograph of 10nm TX and measured eye diagrams.

C. 10nm 224Gb/s RX-AFE measurement

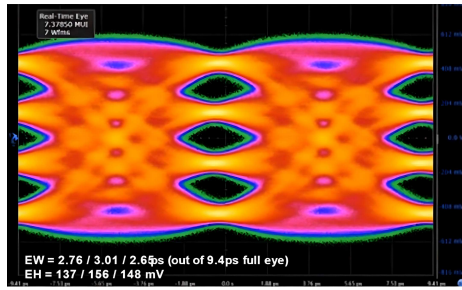
The chip micrograph of the RX-AFE implemented in Intel 7 and measured frequency responses exercising DC/HF gain control is shown in Fig. 24. The RX-AFE achieved 54-GHz peak frequency, 19-dB gain boosting, 36.8-dB HD3 (low frequency) with 60-mW power (0.27-pJ/b energy efficiency) using 1.5V supply.

D. 5nm 224Gb/s RX Measurement

The RX was fabricated in TSMC 5nm FinFET technology. Figure 24 shows the chip micrograph of the RX with size of 0.3397mm². Lane width is 0.43mm to maximize die-edge (shoreline) bandwidth density. Figure 25 shows measurements result of the RX tested with 224Gb/s PAM-4 PRBS-31 using 5nm TX-RX external loopback (both the TX and RX are on-die). Test platform includes the package, connectors, and ISI board with a ball-to-ball loss of 31.4dB at 56GHz. At this channel loss, BER of 1.2e-8 was achieved. The RX achieves BERs of 1.9e-11, 5.2e-10, and 6e-7 for 20.6, 27, and 38dB of channels, respectively. The jitter tolerance is measured with 5MHz of CDR loop bandwidth including FFE enabled, but DFE disabled, showing 0.1UI of eye opening for 1e-4 BER. This RX achieves 1.41pJ/b energy efficiency at 224Gb/s excluding DSP power.

Conclusion

We discussed 224Gb/s PAM-4 transceiver design process including modulation selection, architecture and circuit optimization, measurement results of the TX and the RX implemented in 10nm and 5nm CMOS technologies. For maximized link performance,



(a)

Parameter	Spec	Units	Results		Notes
			Min	Max	
VI	0.40.6	V	0.474	0.548	Steady-state output levels
FPR	0.7		0.74		Linear Fit Pulse Peak (Peak-to-DC ratio)
RLM	0.95		0.977	0.989	Relative level mismatch (Pam4 Level separation mismatch ratio)
SNDR	32.5	dB	37	38.5	Signal to noise and distortion ratio
VCMrms	30	mV	26		With supply noise, rms(VCM-avg(VCM))
DJ		Fs/mV	120		Measured using 8UI/2UI patterns (spec checked with PRBS13)
Rj		fs	113		RMS 1 sigma (spec checked with PRBS13)
Power		mW	290		TX work at full rate 224Gb/s (1.3 p/b)

(b)

Fig. 23. (a) Measured eye diagram of 5nm 224Gb/s TX and (b) summary of key performance metrics.

PAM-4 modulation was selected (based on thorough simulations with various channel characteristics) and required transceiver system/circuits implementations were realized by employing innovative design techniques. These include process-specific optimal clocking architecture (PLL and clock distribution), data serialization (TX) and interleaving (RX) architecture, bandwidth optimization for high-speed circuit blocks such as the input/output pads/ESD/termination, RX AFE data path, and ADC slice. We also discussed high-performance package design details. To authors' best knowledge, this work demonstrated the first fully functioning high performance long-reach SerDes transceiver for 224Gb/s/pin PAM-4 implemented in CMOS technology.

References

- [1] Y. Krupnik et al., "112Gb/s PAM4 ADC Based SERDES Receiver for Long-Reach Channels in 10nm Process," *Symp. VLSI 2019*.
- [2] J. Im et al., "A 112Gb/s PAM-4 Long-Reach Wireline Transceiver Using a 36-Way Time-Interleaved SAR-ADC and Inverter-Based RX Analog Front-End in 7nm FinFET," *ISSCC 2020*.
- [3] T. Ali et al., "A 460mW 112Gbps DSP-Based Transceiver with 38dB Loss Compensation for Next Generation Data Centers in 7nm FinFET technology," *ISSCC 2020*.
- [4] E. Groen et al., "A 10-to-112Gb/s DSP-DAC-Based Transmitter with 1.2Vppd Output Swing in 7nm FinFET," *ISSCC 2020*.
- [5] M. Kossel et al., "An 8b DAC-based SST TX using metal gate resistors with 1.4pJ/b efficiency at 112Gb/s PAM4 and 8-taps FFE in 7nm CMOS," *ISSCC 2021*.
- [6] M. -A. LaCroix et al., "A 116Gb/s DSP based Wireline Transceiver in 7nm CMOS achieving 6pJ/bit at 45dB Loss in PAM-4/Duo-PAM-4 and 52dB in PAM-2," *ISSCC 2021*.
- [7] P. Mishra et al., "A 112Gb/s ADC-DSP-Based PAM-4 Transceiver for Long-Reach Applications with >40dB Channel Loss in 7nm FinFET," *ISSCC 2021*.
- [8] Z. Guo et al., "A 112.5Gb/s ADC-DSP-Based PAM-4 Long-Reach Transceiver with >50dB Channel Loss in 5nm," *ISSCC 2022*.
- [9] B. Zhang et al., "A 112Gb/s Serial Link Transceiver With 3-tap FFE and 18-tap DFE Receiver for up to 43dB Insertion Loss Channel in 7nm FinFET Technology," *ISSCC 2023*.
- [10] H. Park et al., "A 4.63pJ/b 112Gb/s DSP-Based PAM-4 Transceiver for a Large-Scale Switch in 5nm FinFET," *ISSCC 2023*.
- [11] M. Li et al., "PAMn vs Channel and FEC Investigations for 224 Gbps," available online : https://www.ieee802.org/3/df/public/22_03/mli_3df_03a_220316.pdf.

mli_3df_03a_220316.pdf

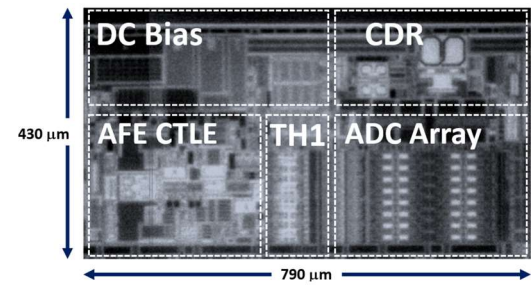


Fig. 24. A chip micrograph of 5nm 224Gb/s ADC based RX.

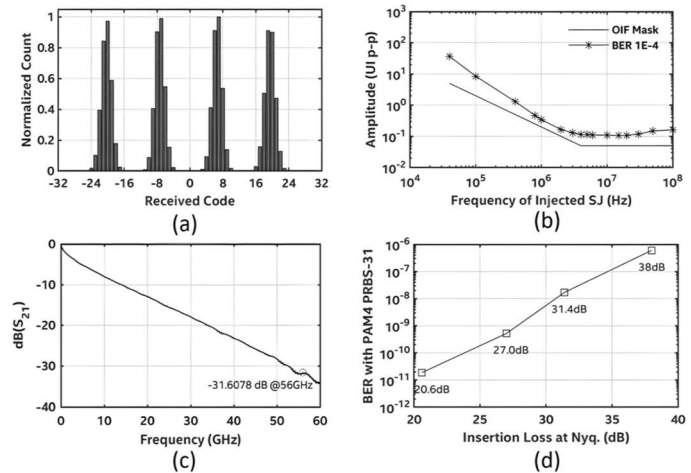


Fig. 25. The RX Measurement results: (a) PAM-4 histogram, (b) jitter tolerance, (c) channel loss, and (d) BER for different channels.

- [12] J. Kim et al., "A 224Gb/s DAC-Based PAM-4 Transmitter with 8-Tap FFE in 10nm CMOS," *ISSCC 2021*.
- [13] J. Kim et al., "A 224-Gb/s DAC-Based PAM-4 Quarter-Rate Transmitter With 8-Tap FFE in 10-nm FinFET," in *JSSC* vol. 57, no. 1, pp. 6-20, Jan. 2022.
- [14] J. Kim et al., "A 112Gb/s PAM-4 Transmitter with 3-Tap FFE in 10nm CMOS," *ISSCC 2018*.
- [15] Z. Toprak-Deniz et al., "A 128Gb/s 1.3pJ/b PAM-4 Transmitter with Reconfigurable 3-Tap FFE in 14nm CMOS," *ISSCC 2019*.
- [16] D. Shin et al., "A 23.9-29.4GHz Digital LC-PLL with a Coupled Frequency Doubler for Wireline Applications in 10nm FinFET," *ISSCC 2021*.
- [17] B. Casper et al., "Clocking Analysis, Implementation and Measurement Techniques for High-Speed Data Links—A Tutorial," *IEEE Transactions on Circuits and Systems I*: vol. 56, no. 1, Pp. 17-39, Jan. 2009.
- [18] S. Kiran et al., "A 56GHz Receiver Analog Front End for 224Gb/s PAM-4 SerDes in 10nm CMOS," *Symp. VLSI 2021*.
- [19] Segal Y. et al., "A 1.41pJ/b 224Gb/s PAM-4 SerDes receiver with 31dB loss compensation," *ISSCC 2022*.
- [20] Khairi A. Y. et al., "A 1.41-pJ/b 224-Gb/s PAM4 6-bit ADC-Based SerDes receiver with hybrid AFE capable of supporting long reach channels" *JSSC* vol. 58, no. 1, pp. 8-18, Jan. 2023.
- [21] Y. Shifman et al., "A 1.64mW Differential Super Source-Follower Buffer with 9.7GHz BW and 43dB PSRR for TimeInterleaved ADC Applications in 10nm," *A-SSCC 2019*.
- [22] L. Kull et al., "A 24–72-GS/s 8-b time-interleaved SAR ADC With 2.0–3.3-pJ/conversion and >30 dB SNDR at Nyquist in 14-nm CMOS FinFET," *JSSC* vol. 53, no. 12, pp. 3508-3516, Dec. 2018.
- [23] X. Jiang et al., "224Gbps-PAM4 End-to-End Channel Solutions for High-Density Networking System", *Signal Integrity Journal*, July 14, 2022. public/22_03/mli_3df_01_220316.pdf.
- [24] D. Ye et al., "Package Technology Enabling for 224 Gbps Electrical Signaling," *IEEE 32nd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)*, 2023.