12.1 A 0.96pJ/b 7×50Gb/s-per-Fiber WDM Receiver with Stacked 7nm CMOS and 45nm Silicon Photonic Dies

Mayank Raj[1], Chuan Xie[1], Ade Bekele[1], Adam Chou[1], Wenfeng Zhang[1], Ying Cao[1], Jae Wook Kim[1], Nakul Narang[2], Hongyuan Zhao[2], Yipeng Wang[2], Kee Hian Tan[2], Winson Lin[1], Jay Im[1], David Mahashin[1], Santiago Asuncion[1], Parag Upadhyaya[1], Yohan Frans[1]

[1]AMD, San Jose, CA
[2]AMD, Singapore, Singapore

Emerging applications such as machine learning, high-performance computing, and cloud storage continue to push compute demands at the data center. To keep up, distributed computing architectures are being increasingly adopted where the physical locations of the CPU, GPU, FPGA, memory, and storage may span over several meters. In package silicon-photonics-based optical links with wavelength division multiplexing (WDM) and non-return-to-zero (NRZ) signaling provides a power-efficient, high-bandwidth, and low-latency interface between these components. In this paper, we present a low-power (0.96pJ/b), high-sensitivity (-11.1dBm median), high-bandwidth (7×50Gb/s NRZ WDM) receiver (RX) that achieves <1e-12 bit-error-rate (BER) without forward-error-correction (FEC).

Figure 12.1.1 shows the multi-chip module (MCM) prototype consisting of an electrical chip (EIC), a Si photonics chip (PIC), integrated fan-out (InFO), an organic interposer, and an organic laminate substrate. To achieve low power and high bandwidth, the EIC RX input is placed directly on top of the PIC using copper-pillar bumps with 55µm bump pitch. This allows the TIA to connect to the photodetector (PD) with minimal parasitic effects. InFO is used to connect the EIC to the proxy Core IC through a parallel interface (not used in this work). Light is coupled in and out of the PIC using a 15-fiber array with a 250µm pitch via V-grooves.

The PIC RX architecture is detailed Fig. 12.1.1. To enable WDM channel selection a cascaded ring resonator (CRR) is used. A CRR filter is formed by coupling two ring resonators to create a second-order flat-top filter centered around a given wavelength and low crosstalk from adjacent wavelengths. The free spectral range (FSR) of the CRRs is designed to be 8× the channel spacing of 1.5nm. The drop port for each of the RX1-RX7 CRRs is connected to a SiGe-based waveguide photo detector (PD) with a responsivity of 0.9A/W. The drop port of the last CRR (RX8) is not connected to a PD and is used for debugging purposes in this test chip. The CRRs are designed to have an insertion loss of <1.5dB and crosstalk of <15dB. A racetrack design with a straight coupling section is chosen to increase coupling distances and thereby improve process variation. The second order filter provides lower crosstalk between adjacent wavelengths compared to a first-order filter (a single ring resonator) owing to -40dB/dec gain roll-off. Additionally, the flat-top characteristic is also less susceptible to laser wavelength spacing variations. This allows for collective thermal tuning of the CRR array to match the laser wavelengths. For the best overall BER performance, each CRR passband must be matched to its corresponding laser wavelength while maintaining a 1.5nm grid spacing. To optimize the laser grid and filter alignment, the through port power (which comes for free in this design) for RX1-RX7 is summed and the lowest loss middle value of 1313nm is used as the wavelength for RX1 (Fig. 12.1.2). Laser wavelengths for RX2-RX7 are determined by decrementing them by multiples of 1.5nm. The same operation can be achieved by collectively tuning the CRR array with fixed laser wavelengths.

Figure 12.1.3 shows the complete RX architecture. A low-power, highly sensitive TIA is essential to an optical RX design. To utilize gm/C scaling benefits offered by the 7nm technology node, a 3-stage inverter-based TIA [1] design with t-coil peaking [2] is chosen. The TIA output (Voutb) is directly fed to the strong-arm latch-based slicers. The positive terminal of the slicers (Vcm) is connected to a fixed slicer DC potential generated by a self-biased inverter with PMOS/NMOS strength control. To match the kickback from the sampling clocks on the Vcm and Voutb node, a programmable, transmission-gate-based switch is added to the Vcm node. Instead of a conventional continuous time linear equalizer (CTLE) circuit, NMOS-based switches are added to resistors R1 and R3 of the TIA to implement programmable bandwidth and peaking. A DC current cancellation loop with 10MHz bandwidth is added to ensure that the input common-mode of the slicer is maintained at Vcm regardless of the average photo-diode current level. The DC loop can be disabled by powering off the operation transconductance amplifier (OTA) (Fig. 12.1.3) and operating the TIA in an open loop. In this mode, the output voltage of the TIA after the RC filter is monitored using a static analog probe (SAP) to characterize the insertion loss of the CRR filters. The TIA is designed for transimpedance (Zt) between 4kΩ to 6kΩ, > 35GHz bandwidth, and 3µA integrated input-referred noise. The power supply sensitivity of the TIA is improved by using an on-chip regulated supply, which provides >25dB of power supply rejection ratio (PSRR). To avoid (mid-band 10-200MHz) differential PSRR peaking through the Vcm path due to tracking from the DC loop, a low pass filter is added (Fig. 12.1.3).

The 4 data and 1 error slicer are clocked by four phases of a quarter-rate clock. The error slicer samples the peak of the input signal and is used in combination with the data samples to perform a baud-rate clock and data-recovery (CDR) [2]. The CDR loop directs the sampling clocks to their optimum positions by digitally controlling an 8b phase interpolator (PI). The 8-phase clock to the PI is generated by a ring-based injection-locked oscillator (ILO1) (Fig. 12.1.3) which uses the 12.5GHz clock from the common PLL as a reference. The 4-phase clock for slicers is generated by ILO2, which uses the PI output as a reference. A quadrature locked loop (QLL) is used to correct the phase mismatch of the 4-phase sampling clock and generate the supply voltage for the PI and the two ILOs [3]. In this design, the quadrature error detector (QED) driven voltage to current (V-to-I) circuit directly drives the NMOS source follower of the regulator. This allows us to remove the OTA and the filter capacitor (C2). Eliminating C2 pushes the second pole out, thus C1 can be reduced further, thereby saving silicon area.

The EIC is fabricated in a 7nm FinFET process with a per-channel RX active area of 165×186µm^2. The PIC is fabricated in a 45nm process. The measurement setup is shown in Fig. 12.1.4. All 7 receivers are characterized using a reference optical modulator, which generates a 50Gb/s PRBS7 pattern with a 3.7dB extinction ratio (ER). The wavelength is set by a tunable laser. A tunable optical attenuator is used to control the power sent to the PIC. The RX1-RX7 measured sensitivity varies from -11.4dBm (OMA) to -10.1dBm (OMA) (Fig. 12.1.4). The median sensitivity is -11.1dBm (OMA). The variation in optical sensitivity is due to the loss variation in the CRR filters. This is verified by disabling the dc loop of the TIA and measuring the output voltage of the TIA for different optical attenuation. Figure 12.1.4 shows that the worst channel (RX1) response matches that of the best channel (RX7) with 1.5dBm additional power input. This corresponds to the 1.3dB OMA loss in RX1. Figure 12.1.5 shows the measured BER bathtub curves for RX1-RX7 with 1dB link margin. All channels meet BER<1e-12 for >14%UI opening at 50Gb/s. Thus, in aggregate we achieve 350Gb/s BER <1e-12 operation in a single fiber without any FEC. The measured 2-D eye scans (Fig. 12.1.5) with 1e10 bits show the internal eye shapes. The horizontal resolution for the 2-D eye scan is one PI code (312.5fs at 50Gb/s) and the vertical resolution is 4.4mV. The advantage of using a second-order flat top CRR filter is demonstrated by RX5 which is <1e-12 error-free despite its laser wavelength being off from the filter center by 200pm (Fig. 12.1.2). To quantify the impact of optical crosstalk among neighboring channels, bathtub BER, and 2-D BER measurements are done for RX2 and RX3 with a 2- laser source. Measurements show minimal impact on horizontal and vertical eye-openings (Fig. 12.1.6), despite the narrow channel spacing of 1.5 nm. The RX consumes 40mW while the common PLL consumes 8.2mW (amortized across 7 lanes). The combined energy efficiency for the RX is measured to be 0.96pJ/b at 50Gb/s. Figure 12.1.6 summarizes the system performance and compares it to previously reported CMOS receiver frontends. We achieve the highest aggregate data-rate while being most power efficient when compared to works with integrated SerDes. Figure 12.1.7 shows the die photos of the CRR array in the PIC, the RX circuits in the EIC, and the packaged module.

Acknowledgement:
The work is funded by DARPA PIPES contract HR0011-19-3-0004. The views, opinions and/or findings expressed are those of the authors and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

References:
[1] M. Rakowski et al., "A 4×20Gb/s WDM ring-based hybrid CMOS silicon photonics transceiver," *ISSCC*, pp. 408-409, Feb. 2015.
[2] M. Raj et al., "Design of a 50-Gb/s Hybrid Integrated Si-Photonic Optical Link in 16-nm FinFET," *IEEE JSSC*, vol. 55, no. 4, pp. 1086-1095, Apr. 2020.
[3] C. F. Poon et al., "A 1.24-pJ/b 112-Gb/s (870 Gb/s/Mm) Transceiver for In-Package Links in 7-nm FinFET," *IEEE JSSC*, vol. 57, no. 4, pp. 1199-1210, Apr. 2022.
[4] H. Li et al., "11.6 A 100Gb/s-8.3dBm-Sensitivity PAM-4 Optical Receiver with Integrated TIA, FFE and Direct-Feedback DFE in 28nm CMOS," *ISSCC*, pp. 190-191, Feb. 2021.
[5] I. Ozkaya et al., "A 64Gb/s 1.4-pJ/b NRZ Optical Receiver Data Path in 14nm CMOS FinFET," *IEEE JSSC*, vol. 52, no. 12, pp. 3458-3473, Dec. 2017.

2023 IEEE International Solid-State Circuits Conference (ISSCC) | 978-1-6654-9016-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISSCC42615.2023.10067617

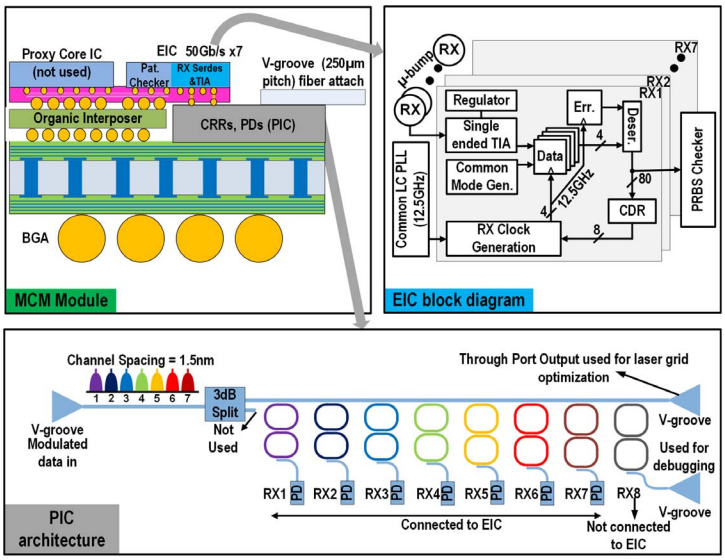


Figure 12.1.1: MCM prototype (top left), EIC block diagram (top right) and PIC architecture (bottom).

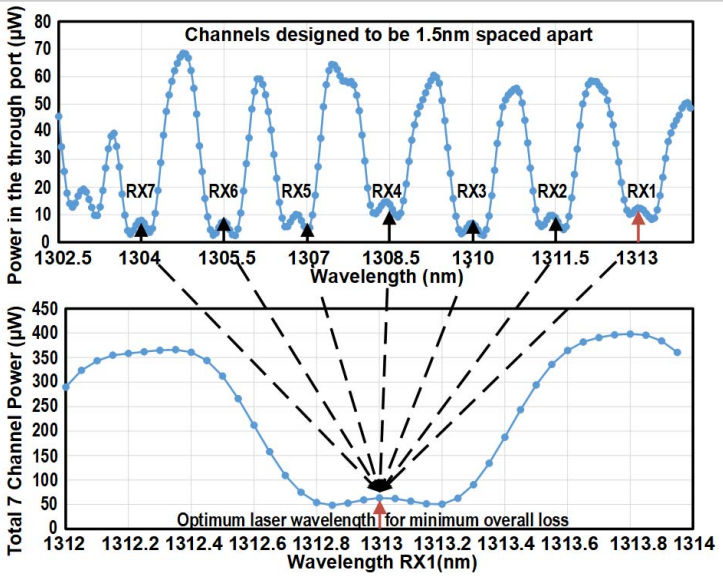


Figure 12.1.2: Measured through port power (top) and optimizing 1.5nm laser grid to minimize overall loss (bottom).

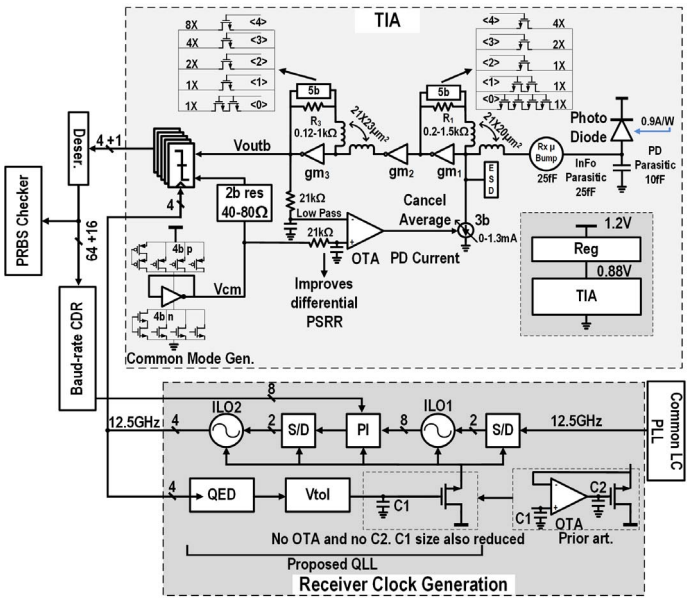


Figure 12.1.3: Detailed receiver architecture.

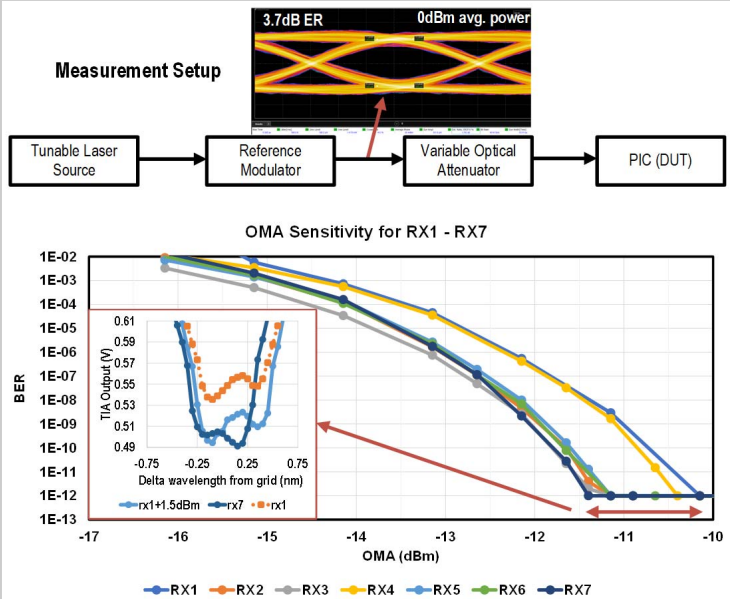


Figure 12.1.4: Measurement setup and measured sensitivity for RX1-RX7 at 50Gb/s.

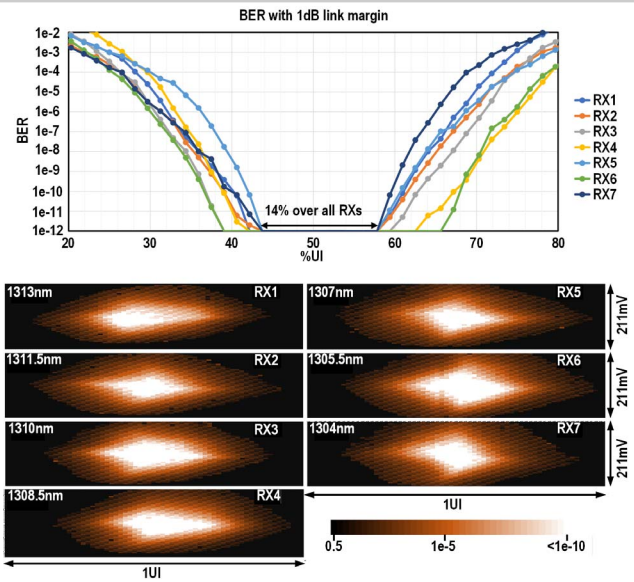
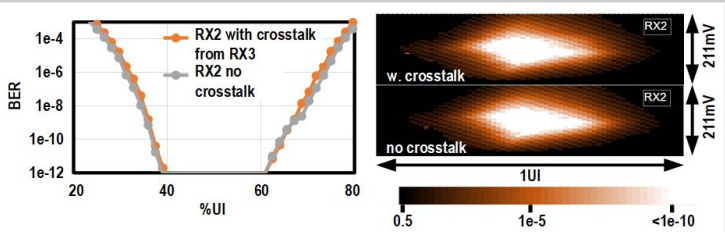


Figure 12.1.5: Bathtub BER (top) and 2-D eye scan (bottom) measurement for RX1-RX7 at 50Gb/s.



	[1]	[2]	[4]	[5]	This Work
Technology	40nm CMOS	16nm FinFET	28nm CMOS	14nm FinFET	**7nm FinFET**
Integration	Hybrid	Hybrid	Hybrid	Hybrid	**Hybrid**
Data Rate	4x20Gb/s	50 Gb/s	100 Gb/s	64 Gb/s	**7x50 Gb/s**
WDM	Yes	No	No	No	**Yes**
Channel Grid	2.4 nm	NA	NA	NA	**1.5 nm**
Wavelength	1540-1550 nm	1570 nm	1310 nm	850 nm	**1304nm - 1313nm**
Modulation	NRZ	NRZ	4-PAM	NRZ	**NRZ**
Input Cap	80 fF	90 fF	70 fF	69 fF	**60 fF**
PD Resp.	-	1 A/W	1 A/W	0.52 A/W	**0.9 A/W**
Sens. (OMA)	-7.2 dBm	-10.9 dBm	-11.1 dBm for 56 Gb/s NRZ	-9dBm at 56Gb/s	**-11.4 dBm***
RX Area	-	0.27mm^2	0.45 mm^2	0.028 mm^2	**0.031 mm^2**
RX Efficiency	0.73 pJ/bit**	1.66 pJ/bit	1.17 pJ/bit	1.4 pJ/bit	**0.96 pJ/bit**

*Best across 7 channels **No SerDes

Figure 12.1.6: BER bathtub and 2-D eye scan measurement with and without crosstalk (top). Performance summary and comparison (bottom).

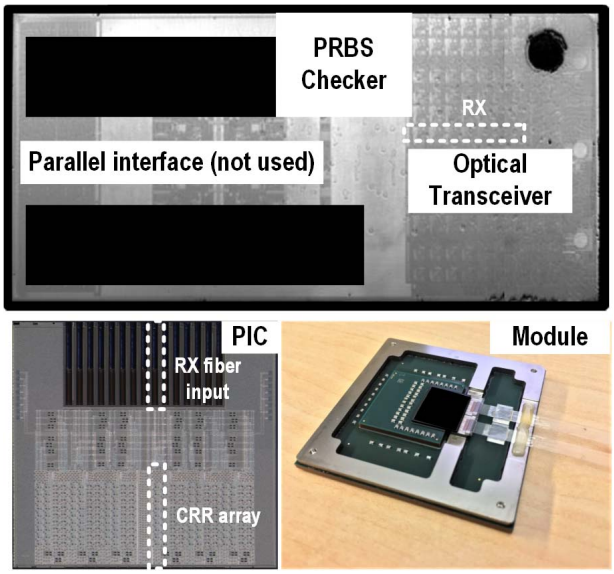


Figure 12.1.7: EIC micrograph (top), PIC micrograph (bottom left), and packaged module (bottom right).

978-1-6654-9016-0/23/$31.00 ©2023 IEEE