Device Engineering and Performance Optimization of Silicon PICs for 800Gb/s Coherent Transmission

Hao Wu*, [1], Chen Zhu[2], Ning Cheng[1], Chenlei Li[1], Min Teng[1], Pan Sun[1], Feng Wang[1], Yinchao Du[1], Gang Cheng[2], Xuezhe Zheng[1]

[1]*Innolight Technology Research Institute (ITRI), No.8, Xiasheng Road, Suzhou Industrial Park, Suzhou, Jiangsu Province, 215000, China*
[2] *Baidu, Network Infrastructure Department, Kuike Technology Building, No.9 Shangdi Information Road , Haidian District, Beijing, China*
hao.wu@innolight.com

Abstract: An end-to-end link model is developed to optimize silicon PIC and evaluate the coherent system performance. 128Gbaud 16-QAM data transmission based on silicon photonics is achieved with -9.1dBm Tx-output power and 26dB Rx OSNR sensitivity.

1. Introduction

Driven by the ever-increasing bandwidth demands from on-demand video, cloud computing, and other emerging applications, the global datacenter interconnect (DCI) network traffic has increased by a factor of one thousand in the past decade [1]. As a compact, cost-effective and highly-scalable solution, coherent 400G ZR transceivers in QSFP-DD or OSFP form factors has been widely deployed for metro-reach (<120km) IPoverDWDM DCI using dual-polarization 64GBaud 16-QAM modulation. It is predicted that an 800G ZR transceiver with 120~130GBaud analog bandwidth while keeping the same 16-QAM modulation format is necessary to support next-generation IPoverDWDM DCI [2]. Study groups, including IEEE Beyond 400G and OIF 800G Coherent, have been formed to explore the technologies and standards for future DCI [3, 4].

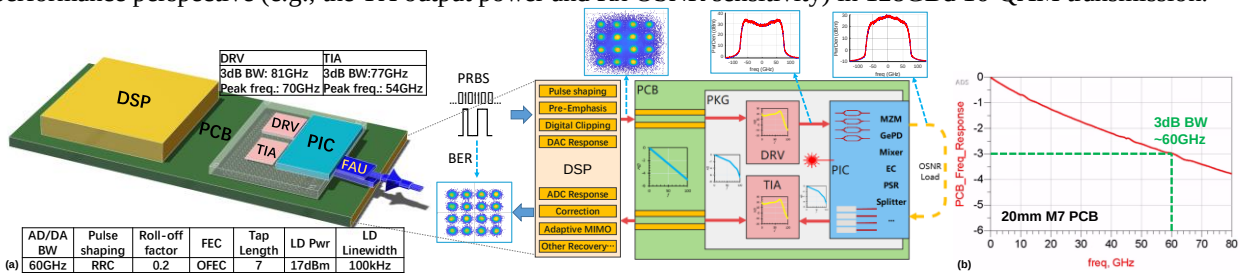
The key challenge for IPoverDWDM is to shrink the coherent transceiver into small form-factor pluggable modules with low power consumption that can be directly fit into the router port which is initially designed for short-reach grey optics. Benefiting from CMOS mass production infrastructure, silicon photonic integrated circuits (PICs) provide significant advantages, such as small footprint, large bandwidth, and high energy efficiency, hence is widely adopted to build the optical engine of 400G ZR coherent transceivers [5]. Although substantial efforts have been made to improve the transceiver to the state-of-art 80~96GBaud through device engineering and performance optimization for modulators, detectors, drivers, and packages [6]-[8], there is still a significant gap in achieving 800G transmission at ~128GBd using silicon photonics. There are a number of key challenges on realizing 128GBaud silicon PIC. At the transmitter side, the carrier-depletion-based modulation in silicon has low efficiency, and it is hard to simultaneously realize high bandwidth and low driving voltage. In addition to meeting the 128GBaud-feasible analog bandwidth, the modulator design has to balance the loss, half-wave voltage (V_{pi}), and frequency response. At the receiver side, the high-speed germanium photodetector (Ge PD) design should overcome both the carrier transit time limitation as well as the sensitive nature with package parasitics. It's also required to determine the best packaging method that would meet the RF loss requirement. As the performance of silicon devices is being pushed to the limit, active and passive silicon components on a transceiver PIC must be carefully engineered. Meanwhile, due to the strong interplay between different parts, it is very hard to decompose the specifications of single devices. Therefore, co-simulation is indispensable to meet the desired system performance. In this work we have developed a comprehensive end-to-end link model, including the electronic integrated circuits (EICs) and the package parasitics, to support optimal design decisions on silicon PICs. The silicon transmitter and receiver can be co-optimized using this link model for the best overall performance. With device engineering and performance optimization, 128Gbaud 16-QAM data transmission using silicon photonics can be achieved with a -9.1dBm Tx-output power and a 26dB Rx OSNR sensitivity, meeting the draft specifications of 800G ZR.

2. Optimization of silicon photonics devices based on a coherent system model

2.1. End-to-end link model

Fig. 1(a) shows the schematic of a silicon photonics 800G ZR transceiver and the end-to-end coherent link model consist of the equivalent compact models of coherent digital signal processing (DSP), printed circuit board (PCB) traces, packaging, EICs, and PIC. A PRBS 2^{17} pattern is fed into the DSP at the transmitter side, the digital signal goes through pulse shaping using a raised cosine filter and a pre-emphasizer to compensate for the limited bandwidth of the silicon modulator. The digital signal is then converted to analog waveform by an analog-to-digital converter (ADC) with a sampling rate of 256GSa/s. Subsequently the analog signal travels through a 20mm-long PCB trace (see Fig. 1(b)), the parasitics of packaging, and a modulator driver (DRV), before driving the Mach-Zehnder modulators (MZM) on the PIC. Noise is loaded on the transmitted optical signal to characterize the OSNR sensitivity. On the receiver side, after going through a 90° hybrid, the optical signal is converted to photocurrent by high-speed balanced detectors. The photocurrent is amplified by a trans-impedance amplifier (TIA) and goes

through the package parasitics and a short PCB trace. Then the receiver DSP processes the I-Q signals with orthogonalization, and carrier recovery. Finally, the bit error rate (BER) is calculated by comparing the transmitted bit sequence with the received data. Constellation diagrams and power spectrum density at each point in the signal path can be extracted to evaluate the signal quality. The characteristics of PCB trace and package parasitics are extracted from a real physical model. Compact circuit models based on simulated and measured results for 128GBd commercial DRV and TIA are used in the simulations with 81GHz and 77GHz 3dB-bandwidths, and the frequency response peaks at 70GHz and 54GHz, respectively. Physical parameters of the modulator, photodetector, and passive devices are extracted from the device design models, and their designs are optimized from the system performance perspective (e.g., the TX output power and Rx OSNR sensitivity) in 128GBd 16-QAM transmission.



AD/DA BW	Pulse shaping	Roll-off factor	FEC	Tap Length	LD Pwr	LD Linewidth
60GHz	RRC	0.2	OFEC	7	17dBm	100kHz

Fig. 1 (a) Schematic of end-to-end link model for coherent module. (b) Frequency response of M7 PCB trace.

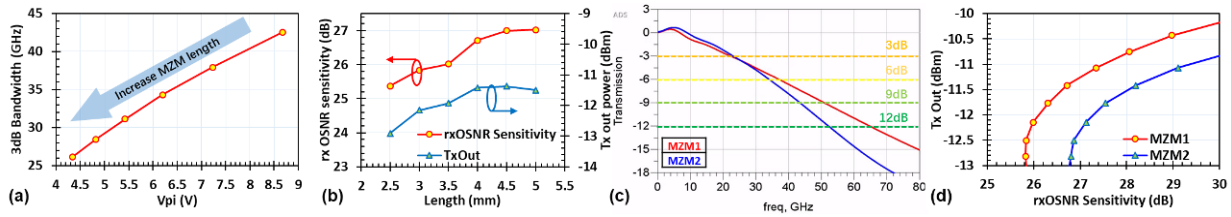


Fig. 2 (a) Bandwidth and V_{pi} of MZM with different lengths. (b) System performance of MZM with different lengths. (c) Frequency responses of two typical MZM designs. (d) System performance of two typical MZM designs.

2.2. *Device engineering and optimization*

Fig. 2(a) shows the tradeoff between bandwidth and Vpi by choosing different lengths of silicon MZM. The design of modulator is based on mature travelling-wave MZM model [9]. When MZM length increases, the bandwidth decreases, which leads to significant inter-symbol interference (ISI) and consequently a worse Rx OSNR sensitivity. On the other hand, the V_{pi} decreases, contributing to a larger Tx output power. However, if length keeps increasing and V_{pi} approaches to the peak-to-peak driving voltage ($V_{p\text{-}p}$), nonlinearity manifests in the output signal, and the benefit of increased output power diminishes. At the same time, the increase in doping absorption reduces the output power. Hence, there exists an optimal MZM length for a certain Tx output power and Rx OSNR sensitivity, as shown in see Fig. 2 (b).

Once the length of MZM is optimized (e.g., 4mm in our case), its frequency response (S21) can be fine-tuned by properly designing its electrode and termination [10]. Fig. 2 (c) and (d) show the S21 and system performance of two 4mm-long slow-wave traveling-wave MZM (TW-MZM) with different electrode designs. MZM1 (red curve) has a lower peaking and a smaller 3dB bandwidth but lower RF loss at high frequencies (40~70GHz). On the contrary, MZM2 (blue curve) has a higher peaking and a larger 3dB bandwidth. System simulation shows that MZM2 leads to more than 1dB OSNR penalty than MZM1. Hence, optimizing the 3dB bandwidth of the MZM as the only figure of merit (FOM) may result in worse system performance. The overall frequency response and RF loss at high frequencies should be taken into account when designing an MZM. The most accurate criterion of optimization is the overall system performance.

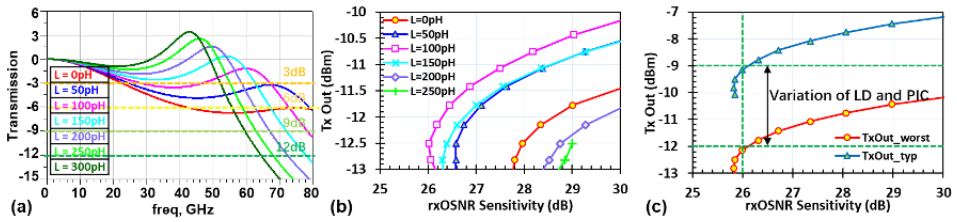


Fig. 3 (a) Frequency responses of a Ge PD with different inductance introduced; (b) Coherent system performance of Ge PD with different inductance. (c) The typical and worst case of overall system performance with optimal PIC design.

The responsivity of the Ge PDs in coherent receivers can be increased to the quantum limit by lengthening the absorption region. In this case, the bandwidth is mainly limited by the transit time for carriers. The bandwidth of a Ge PD can be further extended by the peaking effect induced by the inductance of packages [11]. Fig. 3(a) and (b) show the frequency response and system performance of a typical lateral PIN-based Ge PD derived from a verified photodetector model [12] with different inductances. As the inductance increases, the LC resonance frequency reduces, hence the peaking and roll-off frequency of Ge PD redshifts. Meanwhile, the Rx OSNR sensitivity decreases initially and increases very fast when the inductance is over a certain threshold. The roll-off at lower frequencies causes the OSNR sensitivity degradation with larger inductance. Hence, optimizing the design of Ge PD and controlling the package parasitics is essential to achieve better system performance. The optimal inductance value (~100pH in our case) depends on the characteristic of Ge PD, TIA and package, and it is found through system simulations. However, one should leave enough margin for packaging variations, considering the fast degradation in system performance after the inductance exceeds a certain value.

With the guidance of the system model, we have achieved -9.1dBm (typical) and -12.1dBm (worst) Tx output power and a 26dB Rx OSNR sensitivity with optimal MZM and Ge PD designs (see in Fig. 3(c)). The 3dB difference in output power is caused by the performance variation of the laser diode and devices on PICs.

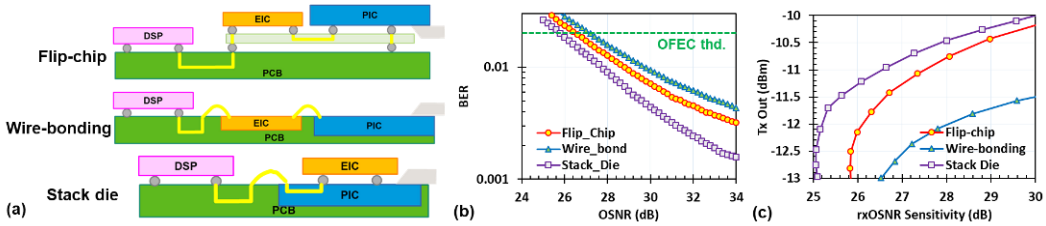


Fig. 4 (a) Different packaging schemes; (b) BER-OSNR curve of different packages. (c) Coherent system performance of different packages.

2.3. Package design and optimization

Fig. 4(a) shows three typical packaging schemes in coherent transceiver modules. The frequency responses of these packaging schemes, obtained from physical modeling using the three dimension finite element method in commercial software (ANSYS HFSS), are used in our system simulations. Since the wire-bonding scheme has the largest RF loss at high frequencies, it has the lowest performance. DRV and TIA are directly connected to the modulators and photodetectors on the silicon chip for the stack die scheme. As a result, it has the best signal integrity and system performance. The flip-chip scheme has intermediate performance and medium complexity among the packaging schemes. As shown in Fig. 4(b) and (c), wire bonding has a large performance penalty in 128GBd transmission. Stack die and flip-chip provide better performance with increased complexity.

3. Conclusions

For next-generation 128GBaud IPoverDWDM DCI transmission, the performance of the silicon PICs must be optimized and stretched to the limits. The silicon PIC and its optoelectronic packaging must be optimized and engineered from a system-level link performance perspective. To better understand the interplay of the device properties, package parasitics, and transmission performance, we built an end-to-end link model for the co-design and optimization of each PIC component and the device package. With optimal designs of silicon TW-MZM, Ge PD, and package scheme, 128Gbaud 16-QAM coherent DCI link can achieve a 26dB Rx OSNR sensitivity with a Tx output power of -9.1dBm (typical) and -12.1dBm (worst), meeting the draft specification of 800G ZR transceiver [2].

4. References

[1] X. Zhou et al., “Beyond 1Tb/s datacenter interconnect technology: Challenges and solutions,” OFC’2019, paper Tu2F.5.
[2] D. Tauber, “Role of Coherent System in the Next DCI Generation,” OFC’2022, paper M1D.1.
[3] https://www.ieee802.org/3/B400G/index.html
[4] https://www.oiforum.com/technical-work/hot-topics/800g-coherent
[5] E. Pincemin, and Y. Loussouarn, “Silicon photonic ZR/ZR+ DCO-CFP2 interface for DCI and metro-regional 400G optical communications,” OFC’2021, paper M3F.2.
[6] J. Zhou et al., “Silicon photonics carrier depletion modulators capable of 85Gbaud 16QAM and 64Gbaud 64QAM,” OFC’2019, Tu2H.2.
[7] S. Yamanaka and Y. Nasu, “Silicon Photonics Coherent Optical Subassembly for High-Data-Rate Signal Transmissions,” OFC’2021, Th5F.2.
[8] T. Sun et al., “Silicon Photonic Mach-Zehnder Modulator Driver for 800+ Gb/s Optical Links,” IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium, 2021.
[9] D. Patel “Design, analysis, and performance of a silicon photonic traveling wave Mach-Zehnder modulator,” McGill University, 2015.
[10] H. Yu et al., “800 Gbps fully integrated silicon photonics transmitter for data center applications.” OFC’2022, paper M2D.7.
[11] A. Novack, et al. “Germanium photodetector with 60 GHz bandwidth using inductive gain peaking,” Optics Express 21, 28387-28393, 2013.
[12] H. Chen. “Advanced germanium pin and avalanche photodetectors for low-power optical interconnects,” Diss. Ghent University, 2016.