The future of multi-terabit datacenter interconnects based on tight co-integration of photonics and electronics technologies

Maria Spyropoulou[1], Giannis Kanakis[1], Giorgos Brestas[1], Yuqing Jiao[2], Salim Abdi[2], Zhaowei Chen[2], Desalegn Feyisa Wolde[2], Ripalta Stabile[2], Nicola Calabretta[2], Kevin Williams[2], Virginie Nodjiadjim[3], Romain Hersent[3], Agnieszka Konczykowska[3], Muriel Riet[3], Richard Schatz[4], Oskars Ozolins[4,5], Xiaodan Pang[4,5], Mahdieh Joharifar[4], Jakub Zvěřina[6], Martin Žoldák[6], Boaz Atias[7], Paraskevas Bakopoulos[8], Elad Mentovich[7], and Hercules Avramopoulos[1]

[1]*Photonics Research Communications Laboratory, School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou str., Zografou campus, 15772, Athens, Greece;*
[2]*Eindhoven Hendrik Casimir Institute (EHCI), Eindhoven University of Technology, Eindhoven, The Netherlands;*
[3]*III-V Lab, Palaiseau Cedex, 91767,France;*
[4]*School of SCI, KTH Royal Institute of Technology, Kista, Sweden;*
[5]*RISE Research Institutes of Sweden AB, Kista, Sweden;*
[6]*Argotech A.S., Trutnov, Czech Republic;*
[7]*NVIDIA, Yokneam, Israel*
[8]*NVIDIA, Athens, Greece.*
mspy@mail.ntua.gr

Abstract: We propose a novel co-packaged optical transceiver architecture capable of operating at 112 Gbaud per lane and scalable to 1.6 Tb/s capacity and beyond for next generation 51.2T and 102.4T digital switches.

1. Introduction

The advent of modern cloud 5G applications, the Internet of Things (IoT) and artificial intelligence (AI) are the main drivers for the explosion of cloud datacenter traffic that imposes major challenges to datacenter operators in terms of bandwidth and latency [1]. What is more, the covid-19 pandemic has induced a large transformation to the everyday communications across various environments with remote ubiquitous access being considered a new standard. Datacenters grow in size, power consumption and complexity to manage this skyrocketing growth of global cloud traffic. In this regime, the sales of 400 GbE optical transceivers is still ramping and the first 800 GbE products have entered the market [2]. 800 GbE transceivers are based on the use of 8 optical lanes operating at 53 Gbaud 4-level pulse-amplitude modulation (PAM4) providing 106 Gb/s per wavelength and rely on pluggable QSFP and OSFP-DD form factors. Recently, a new form factor OSFP-XD was introduced leveraging 16 electrical channels while maintaining the same module size as 800GbE OSFP and enabling 1.6 Tb/s total capacity [3].

Co-packaged optics (CPO) transceivers are placed very close to the core switch chip allowing co-assembly of the optoelectronic engine (OE) with the electrical signal processing IC on the same substrate. First demonstrations of CPO transceivers rely on silicon photonics (SiPh) technologies due to the maturity of CMOS processes and the capability for integration with electronics computing chips [4-7]. Although it is believed that CPO products will dominate the market towards 2028 timeframe, proprietary ethernet switches with CPO are already being announced [8]. In this paper, we propose a novel CPO transceiver architecture based on tight wafer-scale co-integration of Indium Phosphide (InP) membrane photonics and InP double heterojunction bipolar transistor (InP DHBT) electronics technologies operating at 100 Gbaud per lane. Latest advances under the European ICT research project TWILIGHT are presented.

2. Transceiver architecture

The optoelectronic engines (OE) are placed as satellite transceivers around the core switch chip interfacing with the 100 Gb/s SerDes of the ASIC at very close distances (Fig. 1(a)). A 51.2T switch would provide 512 electrical lanes distributed to 16 satellite transceivers assuming 3.2 Tb/s of bidirectional traffic for each using state-of the-art 100 Gb/s technology (16x100Gb/s for the transmit and 16x100Gb/s for the receive path). This would require 16 individual transceivers and 32 fibers per satellite, while scaling to 102.4T switches double the number of transceivers per satellite and their associated optical interfaces would be necessary. In the case of SiPh additional polarization maintaining fibers would be required to transfer the light from the external lasers to the optical modulators which further increase cost and packaging complexity. In the proposed transceiver architecture electrical signal multiplexing and demultiplexing is realized at both the transmitter and the receiver respectively,

resulting in twice the line rate i.e. 200 Gb/s PAM4 per lane and half the number of transceivers to achieve the same aggregate switch capacity. Moreover, optical wavelength multiplexing and demultiplexing on the photonic layer requires only 2 I/O fibers per octal transceiver array and monolithically integrated lasers are considered (Fig. 1(b)).

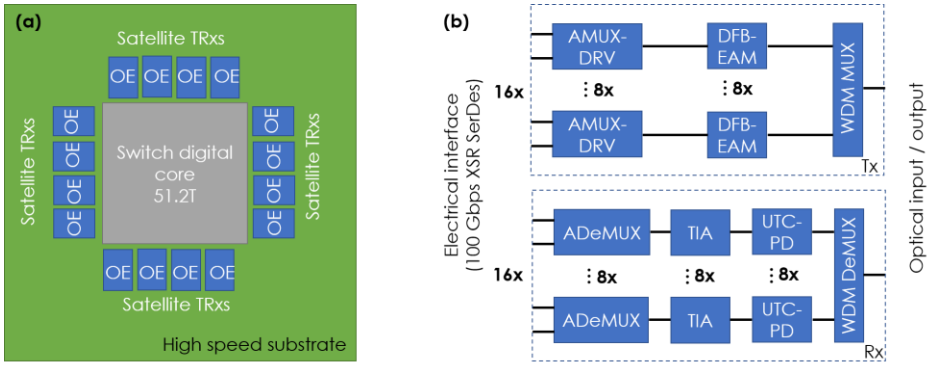

Fig. 1: Conceptual block diagram of CPO transceiver architecture: (a) OE are placed on the same substrate with the switch digital core, (b) electrical and optical multiplexing employed at the transmitter and receiver.

3. Building blocks for 100 Gbaud per lane transmission

High bandwidth optical components and electrical ICs are the fundamental building blocks required for operation at 100 Gbaud. The InP membrane photonic integration platform offers high optical confinement enabling the development of high-speed modulators and detectors. Monolithic integration of actives and passives into large scale photonic integrated circuits (PICs) via selective area growth (SAG) technology in combination with butt-joint process can be achieved [9]. With respect to the electronics, the 0.5 μm InP DHBT process allows the development of ultra high-speed transistors and ICs enabling operation at 160 Gbaud and beyond [10,11].

At the transmitter side, distributed-feedback traveling-wave electro-absorption modulators (DFB-TWEAMs) with segmented electrode sections are developed on the InP membrane platform. Every electro-absorption modulated laser (EML) is driven by an analog multiplexing-driver (AMUX-DRV) that performs both electrical signal interleaving and linear amplification on a single chip. Both the EML design and the AMUX-DRV have demonstrated record performances as standalone devices [12, 11]. To achieve optimum performance for the co-packaged transmitter photonics and electronics co-design is prerequisite. SAG is employed to monolithically integrate the arrayed EMLs with different bandgap-wavelengths in only one epitaxy, which significantly reduces the risk and complexity of the laser/modulator integration. Through the co-integration, the EMLs can be realized directly on top of InP-DHBT electronics. The proposed biasing scheme is depicted in Fig. 2(a).

At the receiver, uni-travelling carrier photodiodes (UTC-PDs) are employed for signal detection followed by high speed InP-DHBT transimpedance amplifiers (TIAs). So far, UTC-PDs have been demonstrated on the InP membrane platform with 110 GHz bandwidth [13] and linear TIAs with 68 GHz 3-dB bandwidth and a wide range of gain and bandwidth controls [14]. Likewise at the transmitter, the UTC photodiodes must be co-designed with the InP-DHBT transimpedance amplifiers (TIAs) to achieve the required receiver performances. The 200 Gb/s PAM4 electrical signal is then demultiplexed into two 100 Gb/s tributaries by means of the analog demultiplexing circuit (ADeMUX).

4. Wafer-scale photonics and electronics co-integration

The photonics and electronics are vertically co-integrated using wafer-scale adhesive bonding as shown in Fig. 2(b-d). The physical properties of the adhesive polymer (Benzocyclobutene or BCB) allow for a high bond strength and low electrical and thermal crosstalk between the vertical devices. Short (<20μm) through polymer-vias gold interconnects are employed to vertically drive the signal, reducing parasitic effects and increasing the integration density. To enable this co-integration, several design and processing considerations need to be taken into account. Boundary conditions for the co-design of the photonic and the electronic components have been defined to ensure optimal performance after integration. Moreover, this co-integration scheme comes with several technological challenges coming naturally with each step of the fabrication process flow as described next.

Before co-integration, the electronics are fully fabricated while the photonics are semi-fabricated since they require double-side processing. Given that soft-baked BCB is used for adhesive bonding for high yield, wafer-scale post-bonding misalignment is controlled by introducing hard anchors to the bonding interface [15]. The wafers are then aligned and bonded. The effect of bonding temperature and BCB residual stress on the performance of

electronics was extensively investigated and safe conditions are defined and used [15]. After bonding, the photonics wafer is wet etched to reach membrane thickness and continue processing. Given that both co-integrated wafers and InP-based, protective coatings with conformal coverage are pre-deposited on the backside of the InP electronics wafer to avoid damaging it during this process [16]. Next, the fabrication of photonics is continued on the freshly made surface. BCB is then opened with a sloped sidewall where 1-2µm thick gold interconnects are fabricated.

The co-integrated OE is flip-chip bonded on the interposer and optical coupling is then realized by means of fiber arrays with standard lid. Any height difference may be accommodated using a spacer. The OE chip is interfaced to the interposer by means of high-speed through vias. Thermal management is addressed in two ways in this co-integrated scheme; heat is dissipated passively towards the interposer which acts also as a heat sink and actively via a thermo-electric cooler (TEC) placed on top of the electronics layer, which generates the most heat (Fig. 2(e)). Critical parameters are currently under investigation such as the stable wavelength operation of the EML under various operating temperatures and the length of the RF lines for supporting the target speed.

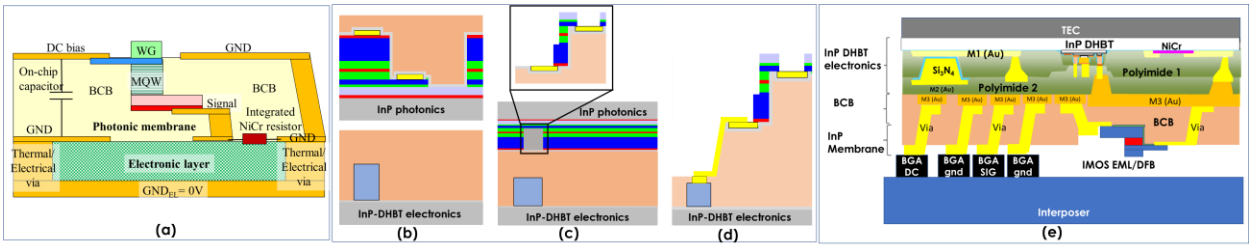

Fig. 2: (a) Proposed biasing scheme of DFB-TWEAM co-designed with AMUX-DRV. (b-d) Schematic of photonics-electronics co-integration process flow: (b) before integration, (c) after bonding and double-side fabrication of the InP photonics, (d) final state after BCB open and interconnect fabrication. (e) cross section of the co-integrated OE on interposer.

5. Conclusions

We propose a novel CPO transceiver scheme based on wafer-level co-integration of InP membrane photonics and InP-DHBT electronics for next generation 51.2T and 102.4T datacenter switches relying on 100 Gbaud per lane transmission via electrical analog multiplexing and demultiplexing. To support this, the photonics and electronics layers are co-designed and interconnected by means of ultra-short (<20 µm) vias through a polymer adhesive layer. The two wafers are processed together and the OE assembly is placed on an interposer. Thermal management is addressed both passively and actively to ensure stable operation.

6. Acknowledgement

This work is funded by the H2020 ICT TWILIGHT project (contract No. 781471) under the Photonics PPP.

7. References

[1] https://www.lightwaveonline.com/data-center/article/14182325/growth-drivers-for-data-center-networks
[2] https://www.lightcounting.com/newsletter/september-2020-high-speed-ethernet-optics-report-26
[3] https://www.lightwaveonline.com/optical-tech/transmission/article/14205029/new-osfpxd-form-factor-could-challenge-copackaged-optics
[4] R. Mahajan et al., JLT **40**(2), 379-392 (2022), doi: https://doi.org/10.1109/JLT.2021.3104725
[5] E. Timurdogan et al., in 2020 Optical Fiber Communications Conference and Exhibition (OFC), 2020, pp. 1-3, paper T3H.2.
[6] D.F. Logan, et al., in 2020 IEEE Photonics Conference (IPC), 2020, pp. 1-2
[7] B. Chou, et al., in 2022 IEEE 72nd Electronic Components and Technology Conference (ECTC), 2022, pp. 394-402, doi: https://doi.org/10.1109/ECTC51906.2022.00071
[8] https://www.electronicdesign.com/technologies/embedded-revolution/article/21248744/electronic-design-broadcoms-first-switch-with-copackaged-optics-is-cloudbound
[9] Z. Chen, et al., in 25th annual Symposium of the IEEE Photonics Benelux Chapter, 25-26 November 2021, Mons, Belgium, 2021
[10] R. Hersent et al., in 2021 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS), 2021, pp. 1-4, doi: https://doi.org/10.1109/BCICTS50416.2021.9682463.
[11] R. Hersent et al., IEEE Microwave and Wireless Components Letters **32(**6), 752-755, (2022), doi: https://doi.org/10.1109/LMWC.2022.3161706
[12] O. Ozolins, et al, in OFC 2022, postdeadline paper Th4A.6.
[13] J. d. Graaf, et al., IEEE JSTQE **28**, 1-10 (2022), doi: https://doi.org/10.1109/JSTQE.2021.3110411
[14] J. -Y. Dupuy, et al., in 2015 IEEE Compound Semiconductor Integrated Circuit Symposium (CSICS), 2015, pp. 1-4, doi: https://doi.org/10.1109/CSICS.2015.7314463
[15] S. Abdi, et al., in 25th annual Symposium of the IEEE Photonics Benelux Chapter, 25-26 November 2021, Mons, Belgium, 2021.
[16] S. Abdi, et al., in 26th annual Symposium of the IEEE Photonics Benelux Chapter, 24-25 November 2022, Eindhoven, The Netherlands, 2022.