Packaging Photonics for AI/ML Systems

Sujit Ramachandra
sujit.r@ieee.org

Abstract—**The rapid emergence of a multitude of machine learning (ML) models with trillions of parameters has highlighted the need for high performance compute systems leveraging Artificial Intelligence (AI) accelerators with disaggregated memory. Considering the demanding bandwidth, density, energy and latency requirements, Silicon photonics is the technology of choice to realize these architectures. Scalable solutions can only be implemented by developing novel and reliable packaging schemes, with emphasis on thermal budgeting, reduced parasitics, and increased bandwidth density. This article covers some challenges seen when packaging photonic circuits for AI/ML systems and some innovative solutions that have been developed in the field.**

Keywords—Packaging, PIC, Integration, Artificial Intelligence, Light Sources

I. INTRODUCTION

With the advent of GPT-4 The rapidly growing size and complexity of Machine learning (ML) and Artificial Intelligence (AI) models has crossed the trillion-mark w.r.t parameters involved [1]. The turn of the decade has seen a large number of ML machine learning models made public, each with billions of parameters. Fig. 1 shows the exponential trend of number of parameters in published ML models over the last five decades. This exponentially growing number of parameters also brings in the need for parallelization of data over tens of thousands of memory and processor nodes. Each of these nodes requires ultra-low latency and power to meet standards, Tb/s optical I/Os and high-speed interconnects between the multiple processing units involved. For instance, early publications report NVIDIA DGX systems consisting of 8 H100 GPUs, designed with a 7.2 Tb/s off chip bandwidth [2].

Despite developments in new interconnect technologies like NVLink and CXL, the bandwidth and energy requirements of the compute-loads of the future cannot be met with this technology since it primarily relies on electrical interconnects [3].

From a standpoint of computing, conventional computers are based on a centralized processing architecture (with a physically separate memory), more suited towards sequential execution. The current AI/ML workloads tend to be distributed and massively parallel and cannot be implemented efficiently on conventional
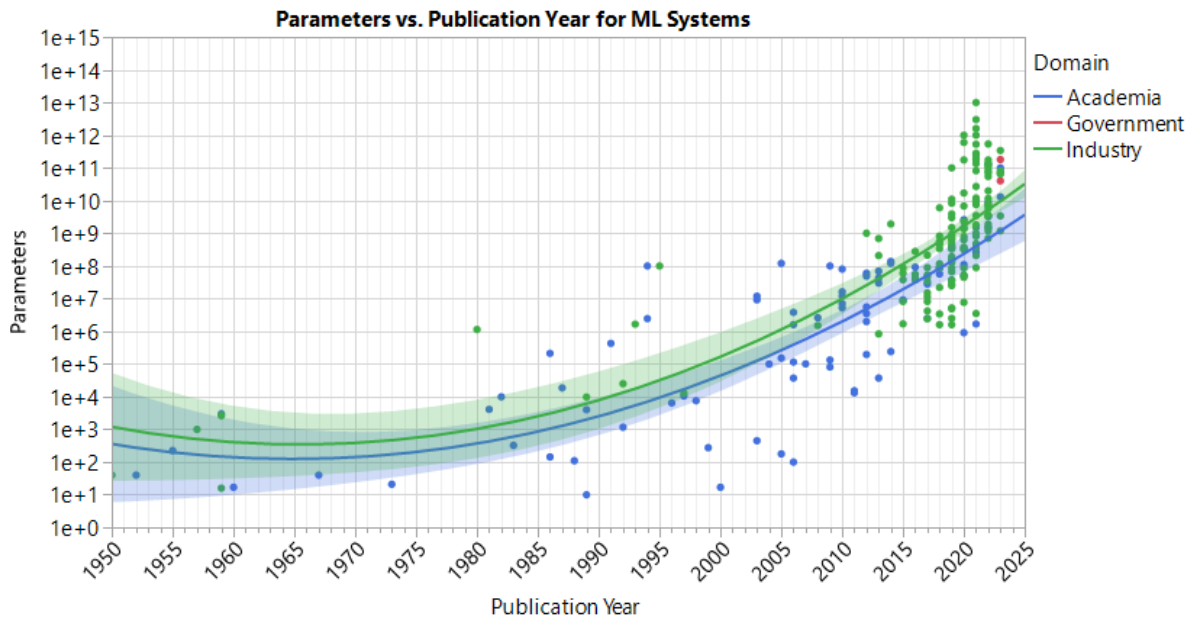

Figure 1. Parameters vs Publication year for ML Systems

processor architectures. Matching hardware to the algorithms themselves would be necessary for faster and energy-efficient processing.

Silicon photonic integrated circuits show great promise in this regard as photons guided by a Silicon waveguide inherently dissipate lower power than electrons guided through metal traces. In addition, higher data rates can be achieved using Opto-Electronic (OE) conversion through modulators, paving the way for co-packaged optical I/O and CPUs/GPUs [4-5].

II. ROLE OF PHOTONICS IN AI/ML

The role of photonics in AI/ML is not to replace conventional computers but to enable applications which require parallelization, low latency and high bandwidth such as nonlinear programming, machine learning acceleration (matrix multiplications etc.) and quantum computing.

Disaggregated architecture for AI/ML computing is one such application that decouples memory and storage (DRAM) from processors and accelerators (e.g., CPU, GPU) [6]. Disaggregation also provides large amounts of memory access at reduced latency for processor nodes using features such as remote direct memory access (RDMA) and GPU-Direct. The signal pathways in such architectures can be implemented by leveraging standard silicon photonic platforms offered by several foundries such as GF, TSMC, Tower Semiconductor, IMEC, AMF and the rest [7-10].

Though AI/ML chips are leveraging photonics to overcome the issues mentioned earlier, they still require a large density of electrical interconnects for a range of functions such as control systems for ring resonators, temperature stabilization, modulator drivers, Trans-Impedance Amplifiers (TIA) for photodiodes, control loops for laser diodes, and in general, biasing electrical or optoelectronic devices on the PIC/ASIC. In fact, the required number of electrical ports is reported to scale quadratically with the number of optical ports [11]. The approach employed by most entities in the field is integration of CMOS electronics and photonics.

III. PHOTONIC INTEGRATION

Typical integration approaches employed in Photonic Integrated Circuits (PICs) for datacenter applications are equally suitable for photonic AI/ML processors. Fig. 2 shows simplified illustrations of a few integration techniques that are historically/commonly used [12].

Fig. 2(a) illustrates Monolithic Integration, an approach that integrates electronics and photonics on the same substrate. A typical example is the 45SPCLO process offered by GF [13]. Though this is theoretically the most appealing integration method and aims to balance electronic circuit performance with that of photonics, the differences in the feature sizes involved often lead to a compromise on the Electrical IC (EIC) with respect to size and power. As an example, Luxtera (now part of Cisco Systems Inc.) initially developed a monolithic transceiver
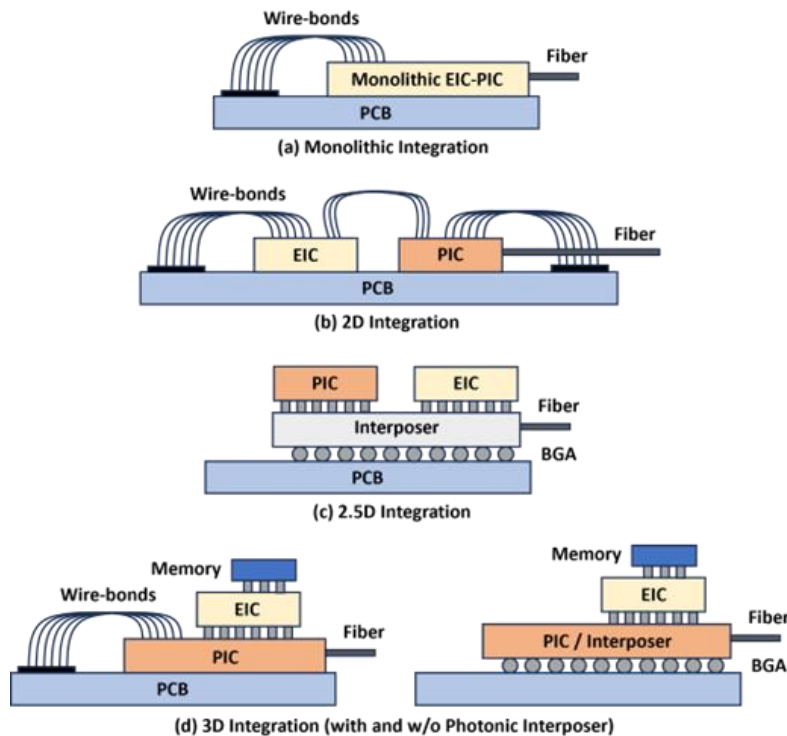

Figure 2. Commonly used Integration approaches

[14], but ultimately switched to hybrid 2.5D integration for the above reason.

One of the most popular approaches to integration in systems involving digitally controlled analog electronic circuits is 2D Integration (Fig. 2(b)), where the application specific integrated circuit (ASIC) (serving as the EIC) and the PIC are mounted on a common PCB, with wire bonds routing signals to both. However, this solution is not readily scalable. As the number of electrical I/O-s increases, the area occupied by wire bonds and associated pads makes routing physically impossible under standard DRC constraints. In addition, larger chip sizes are necessitated, making it prohibitively expensive. Lastly, the wire bonds themselves add significant parasitic inductance and reduce the bandwidth of any high-speed devices that they interface to.

The second popular option that most of the industry has moved towards is flip chip bonding where the ASIC and the Photonics IC are designed, optimized, and sometimes even processed separately in different foundries. The ASIC and the PIC are designed to have matching electrical pads with a common interposer (Fig. 2(c)) onto which they are bonded. This scheme is called 2.5D Integration. As compared to wire bonds, the electrical interconnections have reduced parasitic inductance owing to their short lengths. Since the number of interconnects also scales naturally with the area of the ASIC and PIC, to support higher electrical I/O density, and to further minimize parasitics, the ASIC can also be flipped and bonded or soldered onto the PIC. This technique, termed 3D Integration is represented in Fig. 2(d). The main differentiating factor from 2.5D integration is that a third chip such as memory is now integrated onto the ASIC. However, using a PIC as an interposer between the PCB and ASIC is sometimes still referred to as 3D Integration in literature.

One of the major drawbacks of the 2.5D/3D integration approach is thermal management on the PIC and mechanical stability of the integrated system. Since the ASIC is generally the major source of heat, with the heat dissipated varying under the operating load, some elements on the PIC might require well designed control loops to prevent performance drifts. Typically, modulators based on rings, multiplexers based on rings and any passives that are sensitive to temperature are affected. This is particularly true since PICs for acceleration and computing typically employ several wavelengths across different channels and corresponding Multiplexer (MUX) and Demultiplexer (DEMUX) elements that tend to be temperature sensitive. If the sizes of the ASIC and PIC become large, flexing/warping of dies is another problem that needs to be considered.

As mentioned, solutions such as addition of control loops to maintain devices in the desired operation regimes and designing devices to be athermal are approaches that are employed to overcome the issues outlined above.

Some foundries also offer Through-Dielectric-Vias (TDV) and Through-Silicon-Vias (TSV) to route signals from the ASIC to the PIC, and substrate to the PCB. TSVs bring in high integration density, simplify routing to an extent (as traces need not be routed over extended distances to wire bond pads), and are great for realizing PICs for AI/ML applications. However, they introduce added thermal complexities as they are sources of heat if they carry sizable currents. Depending on TSV density requirements, the mechanical stability of the dies also needs careful consideration. Though TSVs increase the electrical interconnect density, exclusion zones often limit the placement of optical devices and involve more reliability tests under different mechanical stress conditions. Nevertheless, foundries are now invested in offering TSVs as a process capability and it is a scalable path forward w.r.t integration approaches.

High performance compute systems are developed to handle AI/ML workloads by integrating processor dies, high-bandwidth memories (HBM), and co-packaged optics in a single 2.5D/3D package (System in-package (SiP)). Considering the cost and limited number of available wavelengths (8, 16 or 32) in today's Wavelength Division Multiplexed (WDM) sources, it is highly desired to achieve +100Gb/s/wavelength data-rates to increase bandwidth density and reduce the static laser power overhead for direct detect applications [15].

IV. PACKAGING LIGHT SOURCES

One of the continual challenges faced by Silicon photonics is light generation on chip. Though the lack of developed on chip light sources does not gate the development of AI/ML PICs, in many such implementations, there is no need to route light off the chip, making it desirable to have an integrated light source to benefit from the scalability and size benefits it would bring in. Some potential solutions that have been explored include strain engineering in Si/Ge to realize light sources, rare-earth element doping in Si waveguides, and integration of III-V light sources onto PICs. The latter cannot be directly or easily achieved owing to the lattice mismatch between Si and III-V materials. However, approaches typically involve bonding light sources implemented on III-V and evanescently coupled onto the PIC, and micro-packages containing internal light sources.

Integration of group III–V dies with Si has also been achieved by bonding wafers with group III–V quantum wells to silicon waveguides [16,17]. Finally, group III–V quantum dots (QD) have been grown directly on silicon as

unlike quantum well based lasers, QDs can tolerate lattice mismatch without the loss of their optical gain properties.

Intel Labs recently announced a fully integrated 8-channel DFB laser array to serve as Dense Wavelength Division Multiplexed (DWDM) light sources with controlled wavelength spacing and power [18]. The approach involves patterning waveguides on Silicon on Insulator (SOI) prior to a III-V wafer bonding process and marks a significant advancement in terms of the ability to fabricate lasers in high volume and integrate them reliably with Si PICs.

Despite the developments listed above, many entities in the field tend to prefer external laser packages with either edge or grating coupled interfaces on the PIC, where the laser is mounted on the printed circuit board (PCB). Typically, decoupled heat sinks are designed to reduce thermal crosstalk between the laser and the PIC.

Regardless of the light sources used, some added challenges seen in the data center products such as alignment tolerances during fiber array unit (FAU) attach are applicable in this application as well. If large substrates are used alongside high fiber counts, the FAU attach process should have the ability to deal with warpage which can introduce additional loss at the gratings or edge couplers. The epoxy selected for the FAU attach process should have a suitable coefficient of thermal expansion (CTE) and remain stable when subjected to significant thermal cycling with the typical temperature range being 70°C.

V. Summary

Packaging is a complex, interdisciplinary process which requires expertise in materials engineering, fabrication processes, mechanical and thermal design, assembly, and photonics while keeping in mind the end application. The growth of Silicon photonics and the added demands in terms of I/O density, latency and bandwidth density in terms of Tb/s per mm of package edge (referred to as the shoreline), while managing demanding thermal conditions and ensuring a small footprint have served to reinforce the importance of developing scalable packaging solutions.

Packaged PICs should eventually satisfy industry reliability standards in terms of life expectancy, aging, moisture sensitivity and thermal tolerance. In particular, PICs for telecommunications (or in general, PICs used in transceivers) are expected to meet the Telcordia (Bellcore) testing standards [19], with the added customer requirements (generally less demanding). Hence, there is a general move towards the development of standard package designs, reliable integration methods and increased collaboration between design houses, foundries and packaging partners as the proliferation of Silicon

Photonics as a technology in the industry is more governed by packaging than design.

References

[1] M. Bastian, "GPT-4 has more than a trillion parameters - Report," THE DECODER, Mar. 25, 2023. https://the-decoder.com/gpt-4-has-a-trillion-parameters/#:~:text=Further%20details%20on%20GPT%2D4%27s

[2] A. C. Elster and T. A. Haugdahl, "Nvidia Hopper GPU and Grace CPU Highlights," in Computing in Science & Engineering, vol. 24, no. 2, pp. 95-100, 1 March-April 2022, doi: 10.1109/MCSE.2022.3163817.

[3] Mehrdad Khani, Manya Ghobadi, et. al., 2021. SiP-ML: high-bandwidth optical network interconnects for machine learning training. In Proceedings of the 2021 ACM SIGCOMM 2021 Conference (SIGCOMM '21). Association for Computing Machinery, New York, NY, USA, 657–675. https://doi.org/10.1145/3452296.3472900.

[4] Kazanskiy, Nikolay L., Muhammad A. Butt, and Svetlana N. Khonina. 2022. "Optical Computing: Status and Perspectives" Nanomaterials 12, no. 13: 2171. https://doi.org/10.3390/nano12132171.

[5] S. Razdan, M. Traverso, and A. Torza, "Co-Packaged Optics Integration for Hyperscale Networking." Available: https://eps.ieee.org/images/files/Photonics_TC_Co-Packaged_Optics_Final.pdf

[6] Sangjin Han, Norbert Egi, et. al., 2013. Network support for resource disaggregation in next-generation datacenters. In Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks (HotNets-XII). Association for Computing Machinery, New York, NY, USA, Article 10, 1–7. https://doi.org/10.1145/2535771.2535778

[7] M. Rakowski et al., "45nm CMOS — Silicon Photonics Monolithic Technology (45CLO) for Next-Generation, Low Power and High Speed Optical Interconnects," 2020 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2020, pp. 1-3.

[8] J. E. Cunningham et al., "Scaling hybrid-integration of silicon photonics in freescale 130nm to TSMC 40nm-CMOS VLSI drivers for low power communications," 2012 Optical Interconnects Conference, Santa Fe, NM, USA, 2012, pp. 7-7, doi: 10.1109/OIC.2012.6224475.

[9] Philippe P. Absil, Peter De Heyn, Hongtao Chen, Peter Verheyen, Guy Lepage, Marianna Pantouvaki, Jeroen De Coster, Amit Khanna, Youssef Drissi, Dries Van Thourhout, Joris Van Campenhout, "Imec iSiPP25G silicon photonics: a robust CMOS-based photonics technology platform," Proc. SPIE 9367, Silicon Photonics X, 93670V (27 February 2015);

[10] E. Preisler, F. Rezaie, Y. Qamar, M. Pathirane, S. Soltanmohammad, R. Tang, O. Martynov, "Considerations for Silicon Photonics Process Technologies in a Commercial Foundry Environment," 2022 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2022, pp. 1-3.

[11] Shastri, B.J., Tait, A.N., Ferreira de Lima, T. et al. "Photonics for artificial intelligence and neuromorphic computing". Nat. Photonics 15, 102–114 (2021). https://doi.org/10.1038/s41566-020-00754-y

[12] N. C. Abrams et al., "Silicon Photonic 2.5D Multi-Chip Module Transceiver for High-Performance Data Centers," in Journal of Lightwave Technology, vol. 38, no. 13, pp. 3346-3357, 1 July1, 2020, doi: 10.1109/JLT.2020.2967235.

[13] "EUROPRACTICE | IC Service." https://europractice-ic.com/

[14] D. Kucharski and L. Team, "40Gb/s optical active cable using monolithic transceivers implemented in silicon photonics enabled 0.13-µm SOI CMOS Technology," 2009 IEEE Hot Chips 21 Symposium (HCS), Stanford, CA, USA, 2009, pp. 1-24, doi: 10.1109/HOTCHIPS.2009.7478354.

[15] S Moazeni, "Next-generation Co-Packaged Optics for Future Disaggregated AI Systems", Systems and Control Emerging Technologies, March 3, 2023.

[16] Liang, D. & Bowers, J. E. "Highly effi cient vertical outgassing channels for low-temperature InP-to-silicon direct wafer bonding on the silicon-on-insulator (SOI) substrate". J. Vac. Sci. Tech. B 26, 1560–1568 (2008).

[17] Pasquariello, D. & Hjort, K. "Plasma-assisted InP-to-Si low temperature wafer bonding". IEEE J. Sel. Top. Quant. Electron. 8, 118–131 (2002)

[18] "Intel Labs Announces Integrated Photonics Research Advancement," Intel. https://www.intel.com/content/www/us/en/newsroom/news/intel-labs-announces-integrated-photonics-research-advancement.html

[19] GR–468–CORE Reliability Assurance for Optoelectronic Devices