Enabling Bandwidth Scaling for Datacenter and AI/ML Applications Using III-V and Silicon Photonic Devices

Brian Koch, Michael Davenport, John Garcia, and Alan Liu

Quintessent, Inc.
120 Cremona Drive, Suite 155, Goleta, CA, 93117
brian@quintessent.com

Keywords: Photonics, lasers, artificial intelligence

Abstract

Datacenters and AI/Machine Learning applications demand energy efficient optical connectivity solutions with high bandwidth density that can scale in the future as needs increase. To address these needs effectively, we propose an approach using silicon photonic serial micro-ring resonators with integrated multiwavelength lasers.

INTRODUCTION

Datacenters and AI/Machine Learning applications continue to grow in scale and energy consumption at rapid rates. Communication between nodes is a significant portion of this technology and energy. In many cases, the connections between nodes are a bottleneck and require distances much too far for electrical connections. This requires photonic interconnects that have low energy, reach up to 100 meters or more, and an ability to scale to higher bandwidth and bandwidth density as future demands increase. Quintessent is developing approaches to address these applications using multi-wavelength lasers and photonic integrated circuits based on III-V and silicon photonic materials.

MOTIVATIONS FOR NEW OPTICAL SOLUTIONS

The computing requirements of deep learning systems scale dramatically with the error rate (accuracy). A recent published study asserted that the amount of energy that will be required to train a deep learning system within 5% error is equivalent to the energy used by New York City in one month [1]. Furthermore, this system would cost $100B. Recent trends in the allowable error used in actual deep learning systems would indicate that this ~5% error requirement will be commonplace within the next few years. Clearly the energy and cost demands become unreasonable at some point so this may not actually be realized. We might also expect that advancements in software and hardware will continue to improve to enable better and better performance over time with some reduced scaling of the energy and cost requirements. However, it is also clear that energy and cost of all system elements must be minimized in order for AI/ML scaling to continue. Optical communication between compute nodes will be a large portion of this energy and cost as we will explain below.

Figure 1 shows the computation effort required for machine learning training problems in terms of number of PetaFLOPS times days of computation time versus date. Data is collected from [2] and other sources. Up to approximately the year 2010, trends followed Moore's law shown by the green dashed line, a 2x increase in compute needs every 2 years. More recently the trend has exploded to follow the trend of the red dashed line which is 100x every 2 years. Basic electronic elements are not able to scale this fast, so the only way to enable this expansion is by using massively parallel GPUs (or more generally XPUs).

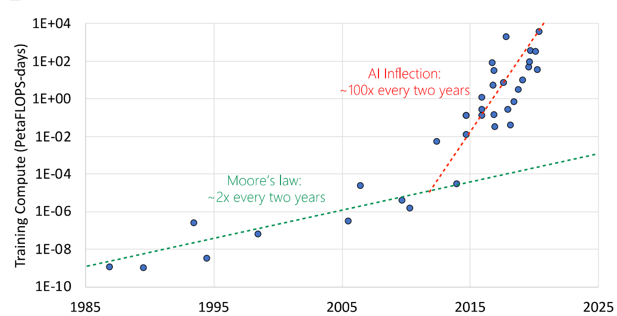

Fig. 1. Trends and extrapolated PetaFLOPS-days in AI systems versus date. The only way to keep up with rapidly expanding computing needs is to use massively parallel XPUs.

This large-scale parallelization requires a huge amount of communication between XPUs. In many, or perhaps most, cases the network becomes the bottleneck [3]. Furthermore, as more and more XPUs are added in parallel, the distances that must be traveled to communicate between them becomes too far for electrical connections, so optical connections are needed. These optical connections require extremely high bandwidth. For example, as quoted in [4], "Training of a Transformer model on a 2x larger system with 5x faster GPUs than today [2020] already requires almost 3 Tbps bandwidth per GPU (i.e. 8x400GigE)."

SUGGESTED OPTICAL LINKS FOR AI/ML SYSTEMS

There are many ways to scale bandwidth for optical links, including higher baud rate, higher symbol count (amplitude/phase), multiple polarizations, parallel spatial lanes, and multiple wavelengths. Various combinations of these approaches can also be used. A detailed analysis of tradeoffs between these approaches is beyond the scope of this paper, but some of the key considerations are noted below. Using higher baud rates has the advantage of requiring fewer components, lowest link loss, high bandwidth density, and small size. However, this approach requires higher speed electronics, may be less energy efficient, and can be harder to scale without other approaches. The parallel spatial lanes approach also has the advantage of lowest link loss, and can potentially use the lowest power. However, it requires more elements, more edge bandwidth density, and more fibers, which will clearly limit the scalability. Higher order amplitude, phase, and polarization approaches can be attractive but often require specialized electronics and higher-quality optical components which may both be expensive to manufacture. Using wavelength division multiplexing (WDM) has the advantages of lower speed electronics for the same overall data rates and bandwidth density, and high scalability. Unfortunately, this approach requires more optical elements in the link and has more optical link loss.

However, newly emerging WDM technology allows the impact of the negative factors to be minimized. A traditional dense WDM optical link architecture is shown in Figure 2 (a). Combining several newer approaches applicable to DWDM, we are developing photonic integrated circuits as shown in Figure 2 (b), consisting of silicon photonic-based serial micro-ring resonator modulators and filters, and integrated III-V multiwavelength lasers. Below is some explanation of the strengths of this approach.

Micro-ring resonators have been in existence for many years, but more recent developments of ring resonators on silicon on insulator platforms have made them more feasible to use in practical applications. Silicon can be implanted to form pn junctions and these junctions can be used to modulate the index of refraction of light, and therefore the optical amplitude of the output of a ring resonator, at very high data rates. The resonator wavelengths can also be fine-tuned using micro heaters. Ring resonators can be cascaded serially on a single bus waveguide, with each ring modulating only a single wavelength while other wavelengths can pass by with very low loss. Ring resonators can similarly pick out single wavelengths from a group of wavelengths at the receive side. This can avoid the need for lossy and complex WDM multiplexers and demultiplexers.

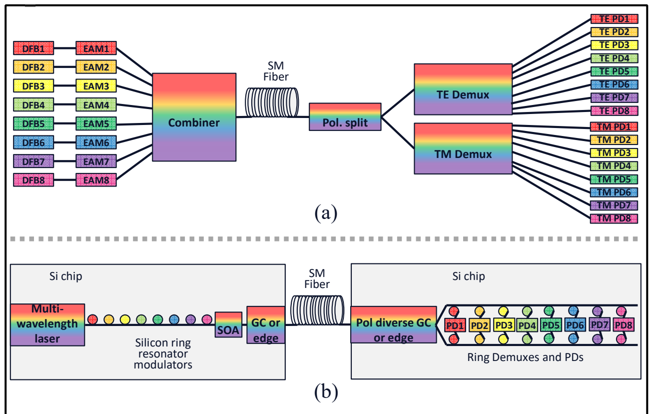

Fig. 2. (a) An example traditional approach to DWDM optical link architecture. (b) Proposed DWDM optical link architecture enabled by silicon photonics with integrated multiwavelength lasers and optical gain.

Silicon photonics has claimed to offer dramatically lower cost compared to its traditional III-V semiconductor process and PIC counterparts. While this is debatable for recently developed PICs and products, and cost is highly dependent on the fabrication methods and facilities that are used, most experts would agree that the yield of silicon photonic components can be much higher than III-V counterparts today. Therefore, for highly complex photonic integrated circuits, silicon photonics will offer a significant advantage over III-V photonics. For example, most of the recent demonstrations of complex PICs with many elements utilize silicon photonics platforms. However, III-V materials are still required to form useful lasers.

Generating the laser wavelengths to be modulated is another problem that must be resolved for this approach to be practical. It is possible to generate single wavelengths from separate lasers and multiplex them together, but this process adds unwanted loss and complexity. It is more desirable to have all light generated from a single laser cavity so that the output of the laser can be sent directly to the modulator bus waveguide without incurring additional loss or circuit complexity. A common problem with multiwavelength lasers is that the relative intensity noise (RIN) of the individual lines is unusable, but by careful device design we have demonstrated multiwavelength lasers with more than 40 wavelengths each having RIN of approximately -135 dB/Hz.

Furthermore, the traditional approach of using a discrete external laser that is fiber coupled to the transmitter PIC (photonic integrated circuit) requires a separately packaged component and fiber coupling of both the laser to fiber and the fiber to the transmitter PIC. This approach may be required in some applications, but it will require additional cost and loss in the link budget. A laser that can be integrated onto the same silicon photonics PIC as the transmitter modulators would be highly desirable. Additionally, with the ability to make a laser on the silicon photonics chip comes the ability to amplify light elsewhere along the photonic link. This approach can be used to close challenging links and/or reduce energy dissipated in the link, providing an additional mechanism for scaling.

CONCLUSIONS

AI/ML systems require massive parallelization of XPUs in order to scale to meet the computing demands of the future. These parallel architectures require communication between XPUs at very high data rates and over significant distances that are too far for electronics. Optical interconnects with high bandwidth density and high energy efficiency are required, and they must be scalable so that demands can be met in the future. DWDM is an attractive scalable solution for these links. The link loss, link complexity, and yield (cost) challenges associated with traditional DWDM systems can be overcome using the serial ring resonator architecture combined with integrated multiwavelength lasers on silicon photonics.

REFERENCES

[1] N. Thompson et al, “Deep Learning’s Diminishing Returns,” IEEE Spectrum, Sept. 2021

[2] https://openai.com/blog/ai-and-compute/

[3] Li, C., et al, Y., “1-bit LAMB: Communication Efficient Large-Scale Large-Batch Training with LAMB's Convergence Speed”, arXiv e-prints, 2021

[4] B. Klenck and L. Dennison, “Why Data Science and Machine Learning Need Silicon Photonics,” OFC 2020

ACRONYMS

AI: Artificial intelligence
GPU: Graphics processing unit
ML: Machine learning
PIC: Photonic integrated circuit
RIN: Relative intensity noise
WDM: Wavelength division multiplexing
XPU: “X” processing unit (CPU, GPU, TPU, etc.)