Optical

AI Optical Processing Technologies Extend Moore's Law

TODD HAUGEN, Chief Executive Officer, Enablence Technologies

New hybrid solutions maximize the capacity and efficiency of planar lightwave circuit technology and combine it with the miniaturization and active integration of SiN and SiPH platforms.

THE SURGE IN DEMAND FOR ARTIFICIAL intelligence (AI) is resulting in enormous computing power and memory requirements. At the same time the industry is experiencing the breakdown of Moore's law. To address these conflicting realities, the industry is investing heavily and increasingly in the transition from electronics to photonics to alleviate these challenges. When successful, photonics will resume the advancement of Moore's law, enabling the computing requirements of AI and the next wave of technological innovations.

The rapidly evolving optical AI processing market focuses on developing hardware and software solutions specifically designed to address the computational challenges of traditional semiconductor architectures. While electronics took decades to enable innovations such as Chat GTP, optical processing is leveraging AI itself to rapidly evolve and overcome limitations in a fraction of the time.

This is creating opportunities for established players like Intel and Nvidia, a host of new technology companies such as Lightmatter, component suppliers like Enablence, as well as traditional datacenter companies such as AOI. These innovators aim to transform the computing landscape by enabling terabyte networking, exascale AI data centers, autonomous vehicles, and robotics.

Limitations of traditional semiconductor technologies

Intensifying demand for AI processes such as pattern recognition and data analytics achieved by matrix computation requires immense computing power. According to OpenAI, the computational power needed Continued on page 32

Characteristic	Planar Lightwave Circuits(PLCs)	Silicon Nitride (SiN)	Silicon Photonics (SiPH)
Material	Silica	Silicon Nitride	Silicon
Cost	Low	Moderate-High	High
Maturity	High (well-established)	Developing	Emerging
Speed	Moderate	High (up to certain limit)	High (w/active electro-optic switching)
Power Handling	Very High	Medium	Low
Energy efficiency	High	Moderate	Moderate-Low
Scale (form factor)	Large	Medium	Small

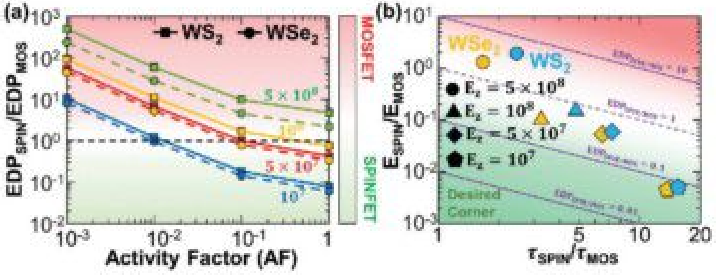

Figure 11. The graphs above show various energy-delay-product (EDP) benchmarking results: left is the EDP ratio of SpinFET-to-MOSFET with p-type WS2 and WSe2 channels, as a function of circuit activity factor (AF) for out-of-plane electric fields (EZ) of 107 V/m, 5×107 V/m, 1×108 V/m, and 5×108 V/m at 100K. The black dashed line (ratio = 1) demarcates the regions where SpinFETs (green) and MOSFETs (red) show more EDP benefit. The right graph compares the energy consumption (E) and switching delay ratios of SpinFET to MOSFET in digital circuits at AF=1 for both p-WS2 and p-WSe2.

to both materials and device engineers to advance the state-of-the-art (**FIGURE 11**). In particular, the authors reveal that the 2D material graphene has spin diffusion lengths >25 μm, and they show that circuits based on optimized 2D spin-FETs outperform their MOSFET counterparts by an order of magnitude in energy delay, and by two orders of magnitude in energy efficiency. **(Paper 3.4, "*A Materials-Device Co-Design Framework for Realizing Ultra Energy-Efficient All-2D-Spin-Logic Circuits with 2D-Materials,*" S. Zhang et al, UC-Santa Barbara/Zhejiang Univ.)**

Optical

Continued from page 22

to train AI has doubled every 3.4 months since 2012 (The computing power needed to train AI is now rising seven times faster than ever before | MIT Technology Review), a roughly 10x increase in computing power per year. The training of an AI model like GPT-3 consumes power equivalent to driving 112 gas powered cars for an entire year (2104.10350.pdf (arxiv.org)). Graphics Processing Units (GPUs) have been the preferred architecture for AI processing due to their parallel processing capabilities. However, they have inherent limitations. Traditional electronic GPU processing methods, which are generally limited to a few GHz, ultimately cannot meet the super-high-speed and low-latency demands of data processing a modern AI ecosystem. Moreover, GPUs face limitations in miniaturization and frequency constraints that result in overheating and energy inefficiencies.

PLC, SiN, SiPH Technologies offer alternative energy efficient paths

The promise of photonic chips for computation is emerging as a significant technological advancement that will overcome the traditional electronics limitations. Photonic chips are superior in speed, energy consumption, and bandwidth while generating minimal heat. Photonic processors operate efficiently at room temperature, minimizing the need for the specialized cooling systems found in data centers employing traditional CPUs and GPUs.

Photonics technology is based on three platforms, Planar Lightwave Circuits (PLC), Silicon Nitride (SiN), and Silicon Photonics (SiPH).

Historically designers and manufacturers have focused on homogenous solutions leveraging a single photonic technology. This limited photonics' ability to solve the most complex problems.

Now, manufacturers can turn to a myriad of hybrid solutions that maximize the capacity and efficiency of PLC technology and combine it with the miniaturization and active integration of SiN and SiPH platforms. These hybrid solutions enable photonics to address the complex computing tasks required by AI, helping to usher in new waves of technological innovations.

About the author

Todd Haugen is CEO of Enablence Technologies, a leading provider of photonics semiconductors based on Planar Lightwave Circuit (PLC) technology for datacom and telecom applications as well as LiDAR technology-based products for automotive and robotics. Prior to joining Enablence, Haugen held senior management positions at Microsoft. Before Microsoft, he advised several technology start-ups in the greater Seattle area, including one project, where he launched one of the first ecommerce platforms enabling companies like Disney, Nike, Ziff-Davis, and NFL to launch their ecommerce businesses. Prior to this, he held leadership positions at Accenture, and Q-Strategies, where he served as CEO, as well as Aventail where he served as vice president of engineering. Todd can be reached at todd.haugen@enablence.com .