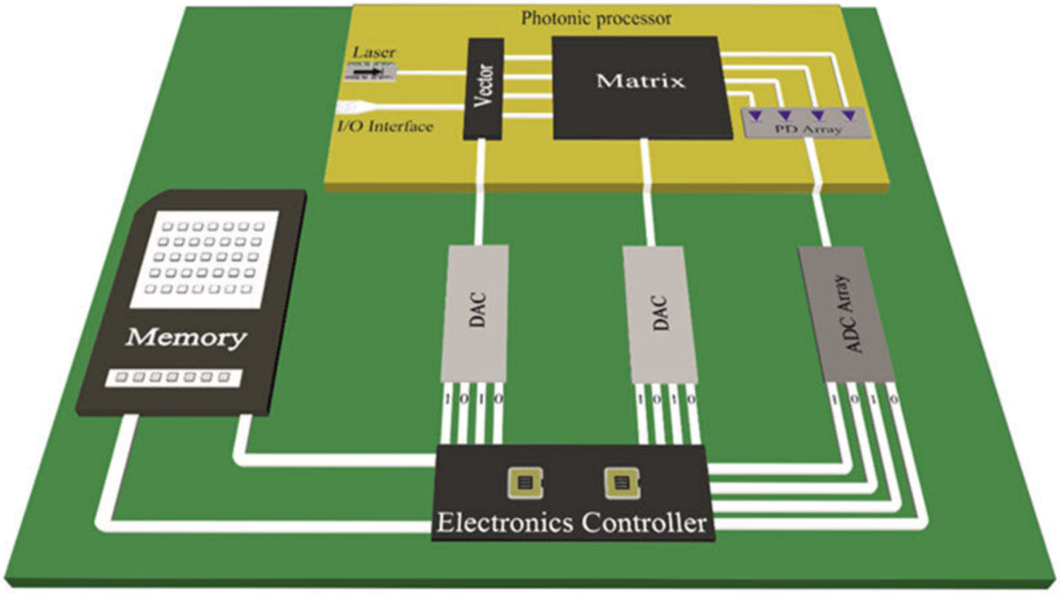
TechXplore

Performing matrix multiplications at the speed of light for enhanced cybersecurity

February 1 2023



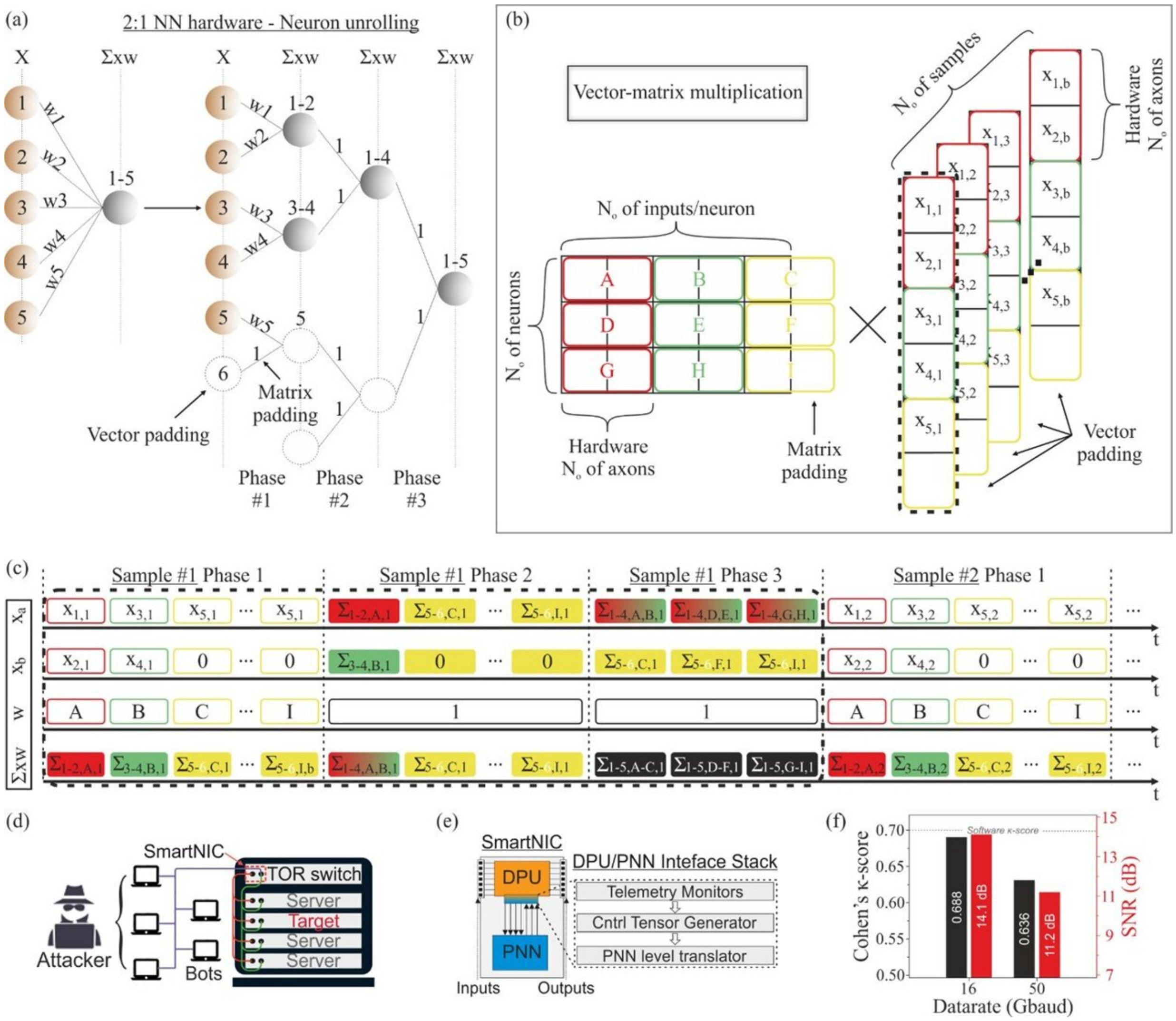
Electro-optic blocks cointegrated for the development of a neuromorphic photonic processor. Credit: Giamougiannis et al., doi 10.1117/1.AP.5.1.016004

"All things are numbers," avowed Pythagoras. Today, 25 centuries later, algebra and mathematics are everywhere in our lives, whether we see them or not. The Cambrian-like explosion of artificial intelligence (AI) brought numbers even closer to us all, since technological evolution allows for parallel processing of a vast amounts of operations.

Progressively, operations between scalars (numbers) were parallelized into operations between vectors, and subsequently, matrices. Multiplication between matrices now trends as the most time- and energy-demanding operation of contemporary AI computational systems. A technique called "tiled matrix multiplication" (TMM) helps to speed computation by decomposing matrix operations into smaller tiles to be computed by the same system in consecutive time slots. But modern electronic AI engines, employing transistors, are approaching their intrinsic limits and can hardly compute at clock-frequencies higher than ~2 GHz.

The compelling credentials of light—ultrahigh speeds and significant energy and footprint savings—offer a solution. Recently a team of photonic researchers of the WinPhos Research group, led by Prof. Nikos Pleros from the Aristotle University of Thessaloniki, harnessed the power of light to develop a compact silicon photonic computer engine capable of computing TMMs at a record-high 50 GHz clock frequency.

As reported in *Advanced Photonics*, they employ silicon-germanium electro-absorption modulators and a novel neuromorphic architectural design capable of encoding and computing data. According to corresponding author George Giamougiannis, "This work paves the way for the resolution of DL-based applications that require line-rate computations," and the work promises to contribute significantly to data center cybersecurity.



(a) Unrolling neuron's linear operations in the time domain. (b)Vector-matrix multiplication operations decomposed into smaller tiles. (c) Real-time classification of the neural network samples. (d) Bots-assisted attacks in Data Center's data. (e) Photonic assisted SmartNIC topology employed in the TOR switch. (f) Accuracy of identification of DDoS attacks in NVIDIA's Data Center's server at 16 and 50 GHz, via the ultra-fast silicon photonic processor presented by researchers of the WinPhos Research group of the Aristotle University of Thessaloniki. Credit: Giamougiannis et al., doi 10.1117/1.AP.5.1.016004

Data center cybersecurity: Light hunting the evil

Undoubtedly, the AI burst has equipped both benign and nefarious users with strong toolkits to speed-up and automate their activities. With the data traveling in data centers (DCs) augmenting by about 13 percent year by year, they have become a major target for malicious individuals who aim to compromise sensitive data, e.g., financial data, personal information, and intellectual property of many organizations, including government agencies, military forces, hospitals, and financial institutions. For that reason, DC cybersecurity is imperative to prevent invaders from accessing classified information.

Indeed, threat detection mechanisms face a new set of requirements resulting from the quantity of data flowing through the vast number of servers and switches within contemporary DCs. Real-time threat detection is imperative: Packet inspection must be processed at ultrahigh speeds. Moreover, threats must be detected as early as possible within the route of the malicious packets: every DC node should be equipped with a powerful cybersecurity toolkit.

Exploiting their ultrafast processor, the researchers from Aristotle University of Thessaloniki, in collaboration with NVIDIA's experts in the field of DC cybersecurity, successfully merged silicon photonics with AI to establish a framework to identify successfully and swiftly one of the most common types of DC attacks, namely distributed denial-of-service (DDoS) attacks, in NVIDIA's servers at line-rates. Thanks to this novel computational scheme, the number may soon be up for DC attacks—at least for the time being.

More information: George Giamougiannis et al, Neuromorphic silicon photonics with 50 GHz tiled matrix multiplication for deep-learning applications, *Advanced Photonics* (2023). DOI: 10.1117/1.AP.5.1.016004

Provided by SPIE

Citation: Performing matrix multiplications at the speed of light for enhanced cybersecurity (2023, February 1) retrieved 27 April 2024 from https://techxplore.com/news/2023-02-matrix-multiplications-cybersecurity.html