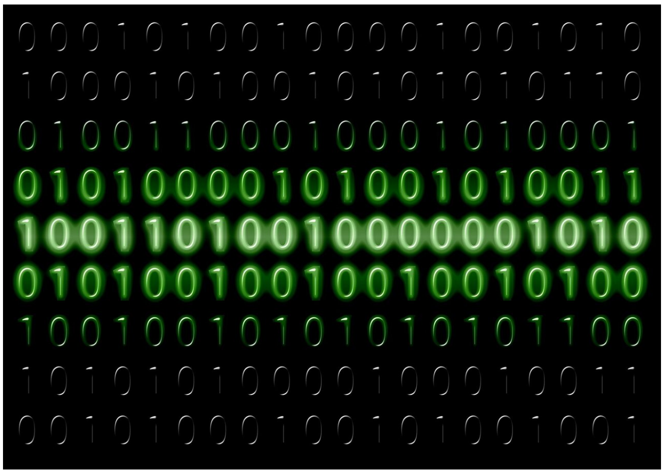
Research brings analog computers just one step from digital

8 December 2021, by Brandie Jefferson



Credit: Pixabay/CC0 Public Domain

The future of computing may be analog.

The digital design of our everyday computers is good for reading email and gaming, but today's problem-solving computers are working with vast amounts of data. The ability to both store and process this information can lead to performance bottlenecks due to the way computers are built.

The next computer revolution might be a new kind of hardware, called processing-in-memory (PIM), an emerging computing paradigm that merges the memory and processing unit and does its computations using the physical properties of the machine—no 1s or 0s needed to do the processing digitally.

At Washington University in St. Louis, researchers from the lab of Xuan "Silvia" Zhang, associate professor in the Preston M. Green Department of Electrical & Systems Engineering at the McKelvey School of Engineering, have designed a new PIM circuit, which brings the flexibility of neural networks to bear on PIM computing. The circuit has the potential to increase PIM computing's performance by orders of magnitude beyond its current theoretical capabilities.

Their research was published online Oct. 27 in the journal *IEEE Transactions on Computers*. The work was a collaboration with Li Jiang at Shanghai Jiao Tong University in China.

Traditionally designed computers are built using a Von Neuman architecture. Part of this design separates the memory, where data is stored; and the processor, where the actual computing is performed.

"Computing challenges today are data-intensive," Zhang said. "We need to crunch tons of data, which creates a performance bottleneck at the interface of the processor and the memory."

PIM computers aim to bypass this problem by merging the memory and the processing into one unit.

Computing, especially computing for today's machine-learning algorithms, is essentially a complex—extremely complex—series of additions and multiplications. In a traditional, digital central processing unit (CPU), this is done using transistors, which basically are voltage-controlled gates to either allow current to flow or not to flow. These two states represent 1 and 0, respectively. Using this digital code—binary code—a CPU can do any and all of the arithmetic needed to make a computer work.

The kind of PIM Zhang's lab is working on is called resistive random-access memory PIM, or RRAM-PIM. Whereas in a CPU, bits are stored in a capacitor in a memory cell, RRAM-PIM computers rely on resistors, hence the name. These resistors are both the memory and the processor.

The bonus? "In resistive memory, you do not have to translate to digital, or binary. You can remain in the analog domain." This is the key to making RRAM-PIM computers so much more efficient.

"If you need to add, you connect two currents," Zhang said. "If you need to multiply, you can tweak the value of the resistor."

But at some point, the information does need to be translated into a digital format to interface with the technologies we are familiar with. That's where RRAM-PIM hit its bottleneck—converting the analog information into a digital format. Then Zhang and Weidong Cao, a postdoctoral research associate in Zhang's lab, introduced neural approximators.

"A neural approximator is built upon a neural network that can approximate arbitrary functions," Zhang said. Given any function at all, the neural approximator can perform the same function, but improve its efficiency.

In this case, the team designed neural approximator circuits that could help clear the bottleneck.

In the RRAM-PIM architecture, once the resistors in a crossbar array have done their calculations, the answers are translated into a digital format. What that means in practice is adding up the results from each column of resistors on a circuit. Each column produces a partial result.

Each of those partial results, in turn, must then be converted into digital information in what is called an analog-to-digital conversion, or ADC. The conversion is energy-intensive.

The neural approximator makes the process more efficient.

Instead of adding each column one by one, the neural approximator circuit can perform multiple calculations—down columns, across columns or in whichever way is most efficient. This leads to fewer ADCs and increased computing efficiency.

The most important part of this work, Cao said, was determining to what extent they could reduce the number of digital conversions happening along the outer edge of the circuit. They found that the neural approximator circuits increased efficiency as far as possible.

"No matter how many analog partial sums generated by the RRAM crossbar array columns—18 or 64 or 128—we just need one analog to digital conversion," Cao said. "We used hardware implementation to achieve the theoretical low bound."

Engineers already are working on large-scale prototypes of PIM computers, but they have been facing several challenges, Zhang said. Using Zhang and Cao's neural approximators could eliminate one of those challenges—the bottleneck, proving that this new computing paradigm has potential to be much more powerful than the current framework suggests. Not just one or two times more powerful, but 10 or 100 times more so.

"Our tech enables us to get one step closer to this kind of computer," Zhang said.

More information: Weidong Cao et al, Neural-PIM: Efficient Processing-In-Memory with Neural Approximation of Peripherals, *IEEE Transactions on Computers* (2021). DOI: 10.1109/TC.2021.3122905

Provided by Washington University in St. Louis

APA citation: Research brings analog computers just one step from digital (2021, December 8) retrieved 28 May 2022 from https://techxplore.com/news/2021-12-analog-digital.html