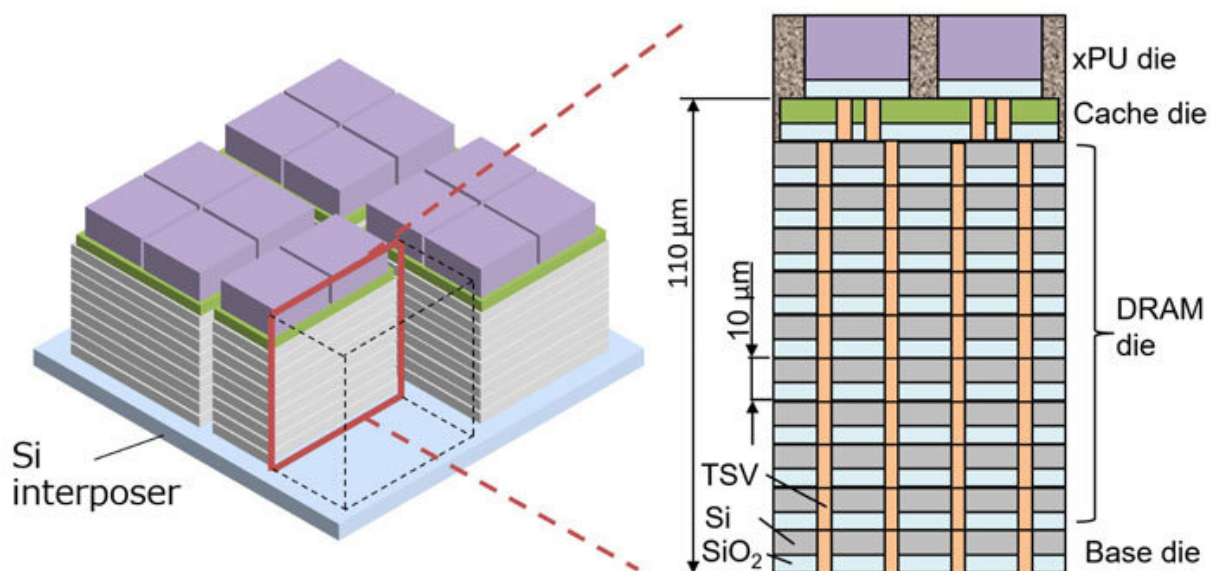


BBCube 3D: Implementing CPU/GPU and memory in a hybrid 3D approach

June 29 2023



The proposed technology uses a stacked design where processing units (xPU) sit atop multiple interconnected memory layers (DRAM). By replacing wires with through-silicon vias (TSVs), the lengths of the connections can be shortened, leading to better overall electrical performance. Credit: Tokyo Institute of Technology

A technology for the three-dimensional integration of processing units and memory, as reported by researchers from Tokyo Tech, has achieved the highest attainable performance in the whole world, paving the way to

faster and more efficient computing. Named "BBCube 3D," this innovative stacked architecture achieves higher data bandwidths than state-of-the-art memory technologies, while also minimizing the energy needed for bit access.

In the present digital age, engineers and researchers keep coming up with new computer-assisted technologies that require higher data bandwidths between the processing units (or PUs, such as GPUs and CPUs) and [memory](#) chips. Some examples of modern bandwidth-extensive applications include [artificial intelligence](#), molecular simulations, climate prediction, and genetic analysis.

However, to increase the data bandwidth, one must either add more wires between the PUs and the memory, or increase the data rate. The first approach is difficult to implement in practice because transmission between the above-mentioned components usually happens in two dimensions, making the addition of more wires tricky. On the other hand, increasing data rate requires increasing the energy needed to access a bit each time, called the "bit access energy," which is also challenging.

Fortunately, a team of researchers at Tokyo Institute of Technology (Tokyo Tech) in Japan may now have found a viable solution to this problem. In a recent [IEEE 2023 Symposium on VLSI Technology and Circuits](#) study, Prof. Takayuki Ohba and colleagues have proposed a technology called "Bumpless Build Cube 3D" or BBCube 3D. This technology holds the potential to resolve the above-mentioned issues for better integration between PUs and dynamic random access memory (DRAM).

As the name suggests, the most notable aspect of BBCube 3D is the realization of connections between PUs and DRAMs in three dimensions, instead of two dimensions. The team was able to achieve

this feat by using an innovative stacked structure in which the PU dies sit atop multiple layers of DRAM, all interconnected via through-silicon vias (TSVs).

BBCube 3D's overall compact architecture, the lack of typical solder microbumps, and the use of TSVs in place of longer wires, together contribute to low parasitic capacitance and low resistance. This improves the electrical performance of the device on various fronts.

Furthermore, the researchers implemented an innovative strategy involving four-phase shielded input/outputs (IOs) to make the BBCube 3D more resistant to noise. They adjusted the timing of adjacent IO lines such that they are always out of phase with each other, meaning that they never change values simultaneously. This reduces crosstalk noise and makes the device operation more robust.

The team evaluated the speed of their proposed architecture and compared it to that of two state-of-the-art memory technologies: DDR5 and HBM2E. "The BBCube 3D has the potential to achieve a bandwidth of 1.6 terabytes per second, which is 30 times higher than DDR5 and four times higher than HBM2E," says Prof. Ohba, while explaining the results of their experiment.

Moreover, BBCube 3D also represents a major breakthrough in terms of the bit access energy. "Due to the BBCube's low thermal resistance and low impedance, the thermal management and [power supply](#) issues typical of 3D integration can be relieved," explains Prof. Ohba, "As a result, the proposed technology could reach a remarkable bandwidth with a bit access energy that is 1/20th and 1/5th of DDR5 and HBM2E, respectively."

More information: Bumpless Build Cube (BBCube) 3D: Heterogeneous 3D Integration Using WoW and CoW to Provide TB/s

Bandwidth with Lowest Bit Access Energy.

[www.vlsisymposium.org/files/pr ... vanceprogram0612.pdf](http://www.vlsisymposium.org/files/pr...vanceprogram0612.pdf)

Provided by Tokyo Institute of Technology

Citation: BBCube 3D: Implementing CPU/GPU and memory in a hybrid 3D approach (2023, June 29) retrieved 11 May 2024 from <https://techxplore.com/news/2023-06-bbcube-3d-cpugpu-memory-hybrid.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.