

Business / Tech

What is high bandwidth memory and why is the US trying to block China's access to it?

By [Wayne Chang](#) and [Rosa de Acosta](#), CNN

🕒 5 min read · Published 12:56 AM EST, Sun December 8, 2024



🗨️ 12 comments



Samsung Electronic's 12-layer HBMs, top, and other modules arranged in Seoul, South Korea on April 4, 2024 SeongJoon Cho/Bloomberg/Getty Images

Editor's Note: *Sign up for [CNN's Meanwhile in China newsletter](#) which explores what you need to know about the country's rise and how it impacts the world.*

(CNN) — The US government has imposed fresh export controls on the sale of high tech memory chips used in artificial intelligence (AI) applications to China.

The rules apply to US-made high bandwidth memory (HBM) technology as well as foreign-produced ones.

Here's everything you need to know about these cutting-edge semiconductors, which have seen demand soar along with the global frenzy for AI.

What is high bandwidth memory?

High bandwidth memory (HBM) are basically a stack of memory chips, small components that store data. They can store more information and transmit data more quickly than the older technology, called DRAM (dynamic random access memory).

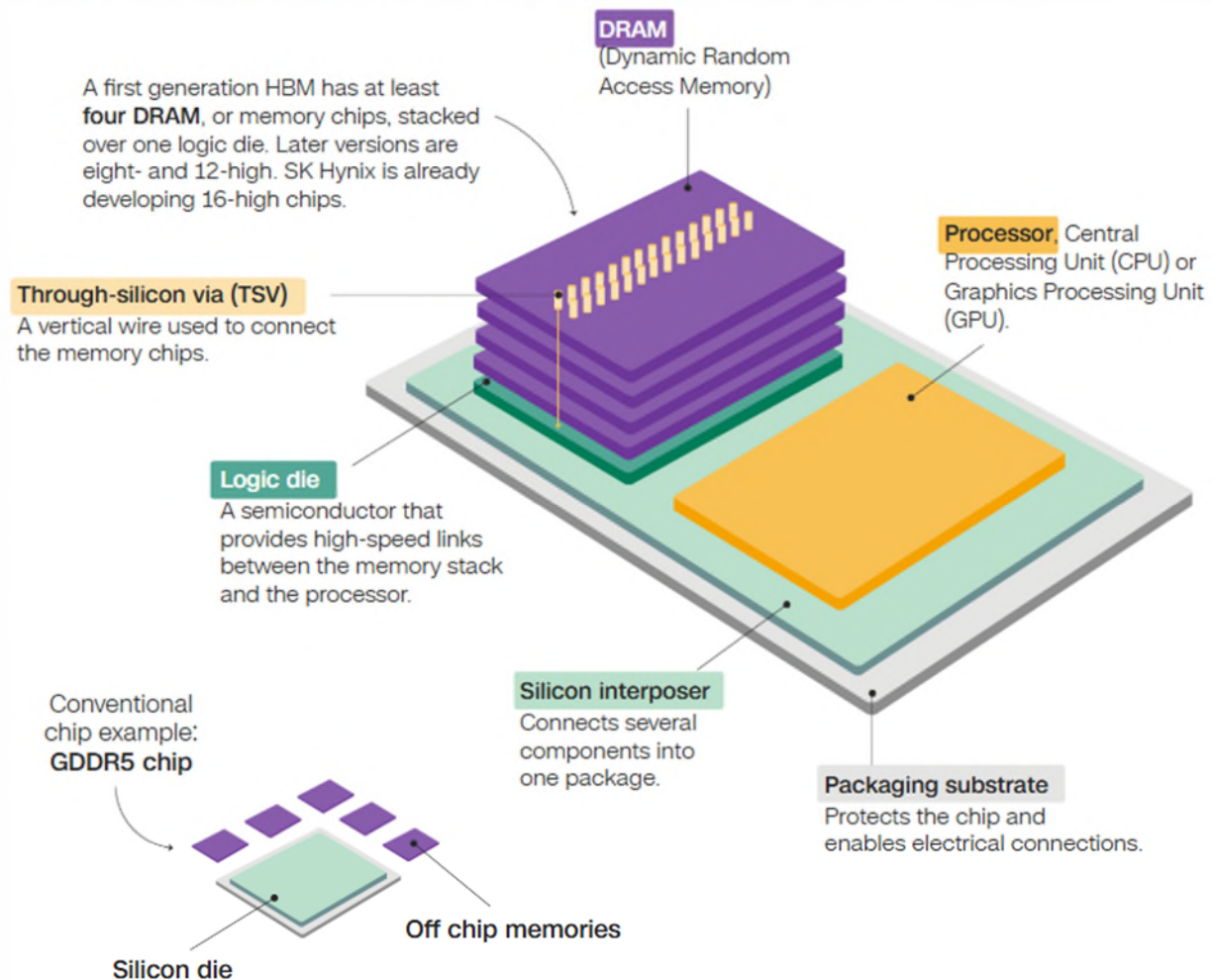
HBM chips are commonly used in graphic cards, high-performance computing systems, data centers and autonomous vehicles.

Most importantly, they are indispensable to running increasingly popular AI applications, including generative AI, which are powered by AI processors, such as the graphic processing units (GPU) produced by Nvidia (NVDA) and Advanced Micro Devices (AMD).

"The processor and the memory are two essential components to AI. Without the memory, it's like having a brain with logic but not having any memory," G Dan Hutcheson, vice chair of TechInsights, a research organization specializing in chips, told CNN.

How HBM chips differ from conventional microchips

HBM chips are gaining popularity due to the rise of generative artificial intelligence. These chips have larger storage space, enhancing the performance of AI applications. They're also faster than conventional memory chips making those applications run more efficiently.



Note: Conventional chip not at scale compared to HBM chip

Sources: SK Hynix, Micron, Science Direct, AMD
Graphic: Rosa de Acosta, CNN

How will the restrictions impact China?

The latest set of export restrictions, announced on December 2, follow two previous rounds of curbs on advanced chips announced by the Biden administration over three years years, with a view to blocking China's access to critical technology that could give it a military edge.

As retaliation, Beijing hit back by imposing fresh curbs on exports of germanium and gallium and other materials elements essential for making semiconductors and other high tech equipment.

Experts say the latest export restrictions will slow China's development of AI chips and, at most stall, its access to HBM. While China's ability to produce HBM currently lags South Korea's SK Hynix and Samsung and America's Micron (MU), it is developing its own capabilities in the area.

"What the US export restrictions would do is cut China's access to HBM of better quality in the short run," said Jeffery Chiu, CEO of Ansforce, an expert network consultancy specializing in tech, told CNN. "In the long run though, China will still be able to produce them independently, albeit with less advanced technologies."

In China, Yangtze Memory Technologies and Changxin Memory Technologies are the leading manufacturers of memory chips. They are purportedly ramping up capacity to build HBM production lines to fulfill its strategic goal of tech self-sufficiency.

Why is HBM so important?

What makes HBM chips so powerful is primarily their larger storage space and much faster speed of passing along data, compared to conventional memory chips.

Because AI applications require a lot of complex computing, such traits ensure that these applications run smoothly, without delays or glitches.



SK Hynix's HBM booth during the SEDEX Semiconductor Exhibition 2024 in Seoul on October 25, 2024.
Steve Cho Kyewoong/Penta Press/Shutterstock

Larger storage space means more data can be stored, transmitted and processed, which enhances the performance of AI applications as the large language models (LLM) enable them to have more parameters to train on.

Think of the faster speed of sending over data, or higher bandwidth in chips parlance, as a highway. The more lanes a highway has, the less likely there will be a bottleneck and thus the more cars it can accommodate.

"It's like the difference between a two-lane highway and a hundred-lane highway. You just don't get traffic jams," said Hutcheson.

Who are the top makers?

Currently, just three companies dominate the global market of HBM.

As of 2022, Hynix accounted for 50% of the total market share for HBM, followed by Samsung at 40% and Micron at 10%, according to a research note published by Taipei-based market research agency TrendForce. Both South Korean firms are expected to take on similar shares in the HBM market in 2023 and 2024, collectively commanding around 95%.

As for Micron, the company aims to grow its HBM market share to somewhere between 20% and 25% by 2025, Taiwan's official Central News Agency reported, citing Praveen Vaidyanatha, a senior executive at Micron.

The high value of HBM has led to all manufacturers dedicating a significant portion of its manufacturing capacity towards the more advanced memory chip. According to TrendForce's Senior Research Vice President Avril Wu, HBM is expected to account for more than 20% of the total market for standard memory chips by value starting in 2024 and potentially exceeding 30% by next year.

How is HBM made?

Imagine stacking multiple standard memory chips in layers like a hamburger. That is essentially the structure of HBM.

On the surface it sounds simple enough, but it is no easy feat to pull off, so much so that it reflects on the price. The unit sales price of HBM is several times higher than that of a conventional memory chip.

That is because the height of an HBM is roughly that of six strands of hair, which means each layer of those standard memory chips that are stacked together need to be extremely thin as well, a feat that calls for top-notch manufacturing expertise known as advanced packaging.

“Each of those memory chips need to be ground to as thin as the height of half a strand of hair before they are stacked together, something very difficult to pull off,” said Chiu.

In addition, holes are drilled on these memory chips before they are mounted on top of each other for electric wires to pass through, and the position and size of these holes need to be extremely precise.

“You have a lot more failure points when you try to make these devices. It’s almost like building a house of cards,” Hutcheson said.



12 comments