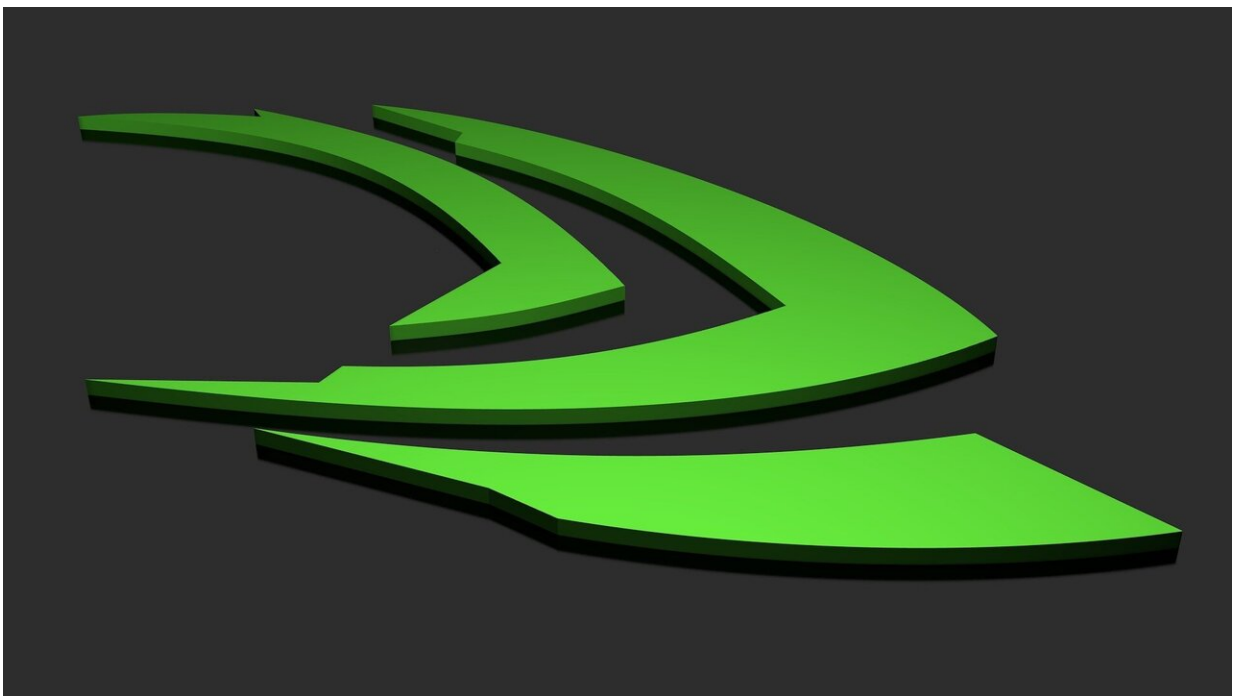


Nvidia and AMD square off in fight to take control of AI

June 4 2024, by Vlad Savov, Jane Lanhee Lee, Takashi Mochizuki,
Bloomberg News



Credit: Pixabay/CC0 Public Domain

Nvidia Corp. and Advanced Micro Devices Inc.'s chiefs showcased new generations of the chips powering the global boom in AI development, deepening a rivalry that may decide the direction of artificial intelligence design and adoption.

Jensen Huang and Lisa Su—both born in Taiwan and now local celebrities for leading U.S. tech powerhouses—employed different tactics in conveying their expertise during back-to-back shows at the world's largest computing conference this week in Taipei.

Nvidia's CEO repeatedly voiced his \$2.7 trillion company's dominance in the accelerators that OpenAI and Microsoft Corp. rely on to build generative AI services like ChatGPT. Huang went as far as to tease a chip envisioned for 2026 he dubbed Rubin—after Vera Rubin, the American woman who helped discover dark matter. The chip, which will succeed the Blackwell family, will be key to sustaining its runaway leadership.

While Huang headlined much of his own two-hour presentation on Sunday, AMD's Su chose to make hers more of a team effort. She brought out a stream of big-name partners from HP Inc. CEO Enrique Lores to Lenovo Group Ltd.'s Luca Rossi to convey the company's focus on designing neural processors—a type of chip that runs AI services directly from laptops.

At one point during her Computex address, Asustek Computer Inc. Chairman Jonney Shih called her "the pride of Taiwan"—a characterization often associated with Huang of Nvidia, whose market valuation is now about 10 times that of AMD's.

"People see Nvidia as a personification of Jensen. And while Lisa is the savior of AMD, she's very clear that it's about everyone around her," said Ian Cutress, chief analyst at the consultancy More Than Moore. "AMD still have that underdog element about the business, and with AI it's very much true."

The back-to-back presentations—attended by hundreds and watched by thousands more across the globe—underscore the growing stakes in a

technology that has the potential to redefine a slew of industries and create new devices that can near-instantly generate video and other content from simple commands.

There's a personal component to the competition too. Su and Huang are not only both Taiwanese, they were born in the same city of roughly 1.8 million on Taiwan's southern coast and are distant relatives. That hasn't made either one more willing to cede ground to the other.

Su was originally tapped by organizers to kick off the week-long conference, one of the most important summits for tech executives around the world. Instead, the Nvidia CEO organized his own presentation the Sunday night before the gathering's official opening to talk about the chipmaker's strategic plans.

Huang, who had keynoted Computex a year earlier, spoke at National Taiwan University just hours before the conference opened, essentially stealing the spotlight from his closest competitor. AMD's Su on Monday referred obliquely to the episode during an OpenAI demo of a chatbot.

"Let's start by letting the tool know that we're interested in Taiwan, and that we're going to be attending Computex," Su said. "And you might ask something about the opening keynote," she quipped, drawing laughter from the audience.

Nvidia sees the rise of generative AI as a new industrial revolution and expects to play a major role as the technology shifts to personal computers, the CEO said in his keynote address at National Taiwan University. He returned to themes he set out a year ago at the same venue, including the idea that those without AI capabilities will be left behind.

Delegates to Computex remarked that Huang's performance drove home

Nvidia's continued dominance—a position that appeared difficult for AMD or any other rival to shake in the short run. One fund manager told Bloomberg News that Huang generated a lot of buzz in particular around Rubin, even though the CEO didn't go into detail about a chip slated for 2026.

Shares of Taiwan Semiconductor Manufacturing Co. and other suppliers rose after the announcement. TSMC's stock climbed 3.1%, while Wistron Corp. gained 1.8%.

Nvidia is selling customers a fully proprietary system, where businesses can buy its chips, networking gear and everything else required to run advanced AI development in data centers he's called "AI factories." AMD, on the other hand, touts open standards that make its hardware interoperable with that of rivals like Intel Corp.

Huang said the upcoming Rubin AI platform will use HBM4, the next iteration of the essential high-bandwidth memory that's grown into a bottleneck for AI accelerator production. Leading manufacturer SK Hynix Inc. is largely sold out through 2025. He otherwise didn't offer detailed specifications for Nvidia's upcoming products.

"Teasing out Rubin and Rubin Ultra was extremely clever and is indicative of its commitment to a year-over-year refresh cycle," said Dan Newman, CEO and chief analyst at Futurum Group. "What I feel he hammered home most clearly is the cadence of innovation, and the company's relentless pursuit of maximizing the limit of technology including software, process, packaging and partnerships to protect and expand its moat and market position."

But it was Su who commanded a broader, star-studded supporting lineup. She was joined onstage by Microsoft Windows chief Pavan Davuluri, a key leader of a company that's regarded as stealing a march on Google

and Baidu Inc. in the generative AI revolution.

She and her partners used their Computex slot to show off new laptops with AI enhancements branded Copilot+. The majority of those devices coming to market are based on NPUs, which are more efficient at handling AI tasks and thus help extend battery life. Qualcomm Inc. was among the first out of the gate with Copilot+ PCs, but AMD's response will be on the market from July. Nvidia has no comparable offering.

Su concluded her keynote by returning to a key sphere of competition: data centers and AI training. She matched Huang's annual upgrade timeline by announcing new AI chips for 2025 and 2026 and promising to speed up the rate of improvement from her company's semiconductors.

"The major difference between the two to me is that Jensen is not only technical and visionary, he's also a salesman. Lisa is much more execution-focused," said Dave Altavilla, principal analyst at HotTech Vision and Analysis. "She has her work cut out for her, and they're picking spots where it matters, and the data center is obviously where the high margin is."

2024 Bloomberg L.P. Distributed by Tribune Content Agency, LLC.

Citation: Nvidia and AMD square off in fight to take control of AI (2024, June 4) retrieved 20 June 2024 from <https://techxplore.com/news/2024-06-nvidia-amd-square-ai.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.