Accessibility Menu

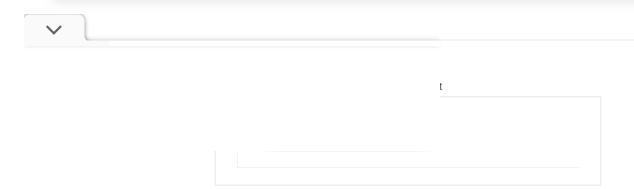


Microsoft Just Showed the Future of AI, and It's Great News for Intel and AMD

By Timothy Green - Apr 24, 2025 at 6:10AM

KEY POINTS

- Microsoft has created an AI model capable of running on a CPU with little memory usage that matches the performance of other models in its size class.
- This continues the trend started by DeepSeek of lowering the computational bar for running advanced AI models.
- CPUs could take share from GPUs in the AI inference market, benefiting Intel and AMD.



The most powerful generative AI models from the likes of OpenAI, **Alphabet**, and Anthropic require costly and power-hungry AI accelerators stuffed into data centers to produce results. OpenAI's recent GPT-4.5 model, for example, was rolled out in

phases to users because it required an immense amount of computational resources.

Al models from Chinese start-up DeepSeek released earlier this year turned some assumptions about the Al infrastructure market on their heads. DeepSeek managed to produce a model that was far cheaper to train and run, while producing results of similar quality, compared to top-tier models from U.S. Al companies. The assumption that Al models would require ever-increasing quantities of computational horsepower, the foundation of the bull case for **Nvidia** stock, started to look a lot less like a sure thing.

Another AI breakthrough

Al features have started to show up on PCs, smartphones, and other devices, but Al models small enough to run on those devices just aren't that capable. Tom's Hardware called **Microsoft**'s Copilot+ PC Al features "a bad joke" when they first launched last year, and The New York Times concluded that **Apple** Intelligence, Apple's suite of Al-powered features, was "still half-baked" in October.

There are multiple problems with on-device AI. First, generative AI is not deterministic, meaning that the same input can produce wildly different outputs. That's fine if you're using AI to write a blog post, but it's not so great if you want it to perform a specific task on your smartphone reliably.

That first problem may never be fully solved, but the second problem could be.

Problem no. 2 is that PCs and smartphones only have so much memory and computational power, which puts a hard limit on how capable an Al model running locally can be. The Al models that power ChatGPT, which run in data centers, require

monstrous amounts of memory, compute, and energy to produce results. Obviously, that's not feasible with a laptop running on battery.

MOTLEY FOOL RETURNS

	STOCK ADVISOR RETURNS	1	S&P 500 RETURNS	
	0070/			
	907%		163%	
5/11/2025.				
iscounted offers are	only available to new members. Stock	k Advisor list price is \$199 p	er year.	
iscounted offers are	only available to new members. Stock	k Advisor list price is \$199 p	er year.	

Microsoft may have an answer. The company recently unveiled a new "1-bit" Al model that is small enough to run on a CPU and uses just 0.4 GB of memory. Amazingly, this new model matches the performance of AI models in its size class that use far more memory. What's more, running on a single CPU, the model can produce output at a speed comparable to human reading, which is fast enough to be useful.

Changing the game

Nvidia dominates the market for AI accelerators, and there's essentially no chance that **AMD (AMD 1.14%)** or **Intel (INTC 2.07%)** will be able to catch up. Playing the same game as Nvidia, AMD and Intel are destined to remain far behind the market leader.



Alert: Top Cash Back Card Now Offers 0% Intro APR Until 2026

Enjoy 0% intro APR on purchases and balance transfers well into 2026—plus up to 5% cash back in popular categories.

Click Here to Start Earning More >

Where to invest \$1,000 right now

When our analyst team has a stock tip, it can pay to listen. After all, *Stock Advisor's* total average return is 907% — a market-crushing outperformance compared to 163% for the S&P 500.

They just revealed what they believe are the **10 best stocks** for investors to buy right now, available when you join *Stock Advisor*.

See the stocks >





Today's Change

(2.07%) \$0.43

Current Price

\$21.43

However, breakthroughs from DeepSeek and Microsoft raise the possibility that AI inference, the act of running a trained AI model to produce results, could eventually be run on CPUs in data centers and on devices without sacrificing quality. The ongoing cost of running AI models goes down dramatically if you can cut Nvidia's expensive GPUs out of the equation, and on-device AI becomes far more compelling for users if more powerful models are able to be stuffed into the memory footprint of a PC or smartphone.

Intel and AMD both sell server and PC CPUs that feature built-in AI accelerators. Some of AMD's EPYC server CPUs excel at certain AI inference tasks, and Intel's Granite Rapids server CPUs can run 70-billion parameter models. On the PC, both AMD and Intel now include dedicated AI processors in their CPUs.

Advertisement

Nvidia appears unstoppable, and it may very well be if you assume the AI infrastructure market remains centered on ultra-powerful data center GPUs. However, the company is vulnerable if powerful AI models no longer require its most powerful GPUs to run. If it starts to make more financial sense to build an AI datacenter filled with CPUs rather than GPUs, that's what's going to start happening. Microsoft's new AI model is a step in that direction.

Spending could start shifting back from GPUs to CPUs in the data center, which would be great news for AMD and Intel. In the PC market, more capable and useful AI features could drive demand for PCs and pull the PC market out of its post-pandemic funk. More powerful hardware is part of the equation, but so are powerful AI models able to fit into smaller memory footprints like Microsoft's latest innovation.

Nvidia can't lose in AI unless the game changes. With Microsoft and others working to lower the computational and memory cost of running powerful AI models, it appears the game is changing as we speak. Intel and AMD, permanently behind Nvidia in the AI accelerator market, could be the big winners as AI inference moves to the CPU.

Suzanne Frey, an executive at Alphabet, is a member of The Motley Fool's board of directors. **Timothy Green** has positions in Intel. The Motley Fool has positions in and recommends Advanced Micro Devices, Alphabet, Apple, Intel, Microsoft, and Nvidia The Motley Fool recommends the following options: long January 2026 \$395 calls on Microsoft, short January 2026 \$405 calls on Microsoft, and short May 2025 \$30 calls on Intel. The Motley Fool has a **disclosure policy**.

STOCKS MENTIONED





Motley Fool Stock Advisor's Latest Pick

Get Access

907% Avg Return*

Microsoft

MSFT

\$438.49 (0.07%) \$0.32

Alphabet

GOOGL

\$152.71 (-1.02%) \$-1.57

Apple

AAPL

\$198.46 (0.49%) \$0.97

Advanced Micro Devices

AMD

\$102.86 (1.14%) \$1.16

Nvidia

NVDA

\$116.64 (-0.62%) \$0.73

Alphabet

GOOG

\$154.32 (-0.92%) \$-1.43

^{*}Average returns of all recommendations since inception. Cost basis and return based on previous market day close.

RELATED ARTICLES

Is Intel Stock a Buy Right Now?				
Intel Slashes Arrow Lake Prices Amid Tough Competition From AMD				
Where Will Intel Stock Be in 5 Years?				
Why Intel Stock Fell 11.5% Last Month				
Better Artificial Intelligence (AI) Stock: Nvidia vs. Intel				
Advertisement				
Making the world smarter, happier, and richer.				
© 1995 - 2025 The Motley Fool. All rights reserved. Market data powered by <u>Xignite</u> and <u>Polygon.io</u> .				

ABOUT THE MOTLEY FOOL OUR SERVICES AROUND THE GLOBE

About Us All Services Fool UK

Careers Stock Advisor Fool Australia

Research Epic Fool Canada

Newsroom Epic Plus

Contact Fool Portfolios

Advertise Fool One

Motley Fool Money

FREE TOOLS AFFILIATES & FRIENDS

CAPS Stock Ratings Motley Fool Asset

Management

Motley Fool Wealth

Calculators Management

Discussion Boards

Financial Dictionary Motley Fool Ventures

Motley Fool Foundation

Become an Affiliate Partner

Terms of Use
Privacy Policy
Disclosure Policy
Accessibility Policy
Copyright, Trademark and Patent Information
Terms and Conditions
Do Not Sell My Personal Information