SHARE

CLUB

3 surprises from Nvidia earnings help justify a one-day $250 billion boost to its market value

PUBLISHED THU, FEB 22 2024•3:28 PM EST | UPDATED AN HOUR AGO

Kevin Stankiewicz
@IN/KEVIN-STANKIEWICZ-B5593466
@KEVIN_STANK

In this article

GOOGL -0.35 (-0.24%)

AMD -0.81 (-0.45%)

NVDA +3.62 (+0.46%)

SHARE

Visitors visit the NVIDIA booth at the 2023 Hangzhou Apsara Conference in Hangzhou, Zhejiang province, China, October 31, 2023. (Photo by Costfoto/NurPhoto via Getty Images)

Costfoto | Nurphoto | Getty Images

Nvidia did not just meet the moment with its latest quarter. It crushed it and sprinkled in a few surprises that brightened an already blazing outlook for the leading artificial intelligence chipmaker.

gain tacked on more than $250 billion to Nvidia's market capitalization, the biggest single-session addition ever. Club name Meta Platforms held the prior record with its nearly $200 billion one-day market cap gain after its earnings earlier this month. Nvidia's Thursday stock gain also pushed the company toward a $2 trillion value market value. The only U.S. companies to go above $2 trillion are Club names Apple and Microsoft, and both of those stocks have briefly topped $3 trillion.

Nvidia's stock was on its way to a fresh record high, erasing all of its losses over the prior four sessions as worries about its earnings report rose and sparked profit-taking in last year's top-performing S&P 500 member. Including Thursday's gains, Nvidia has surged 57% year to date — once again the best of any S&P 500 stock. Nvidia's overall earnings report lived up to the hype, motivating us to raise our price target on NVDA shares to $850. But there also were unexpected elements to the quarter and conference call that deserve more focus.

NVIDIA Corp

RT Quote | Last NASDAQ LS, VOL From CTA | USD

After Hours: Last | 5:03 PM EST

789.00 ▲ +3.62 (+0.46%)

WATCHLIST + QUOTE DETAILS

785.38 ▲ +110.66 (+16.40%)

Close

Nvidia's stock performance over the past 12 months.

1. The growth of inference

Nvidia's dominance in AI training — thanks to its cutting-edge graphics processing units, or GPUs — has been widely understood, with many estimates putting its share of the market above 80%. Less clear has been its position in the "inference" realm because traditional, non-GPU processors can perform the task well. But, on Wednesday, Nvidia estimated that 40% of its data center revenue over the past year was for inference.

inference is the day-to-day use of the model when it puts that "learning" to work and generates outputs.

“The inference part of our business has grown tremendously. We estimate about 40%,” CEO Jensen Huang said. “The amount of training is continuing. Because these models are getting larger and larger, the amount of inference is increasing.”

The estimate is notable for a few reasons. In the short run, it suggests demand for inferencing is significant — in other words, that AI models are actually being used, not simply trained with hopes that they’ll be useful down the road.

A key question swirling around Nvidia has been demand sustainability. Orders for its chips are off the charts right now, but eventually, the companies pouring billions into buying them will need to see evidence that the money is paying off for them, known as “return on investment.” If that return isn’t there, will they keep ordering the latest and greatest Nvidia chips? Will the AI boom turn into a bust?

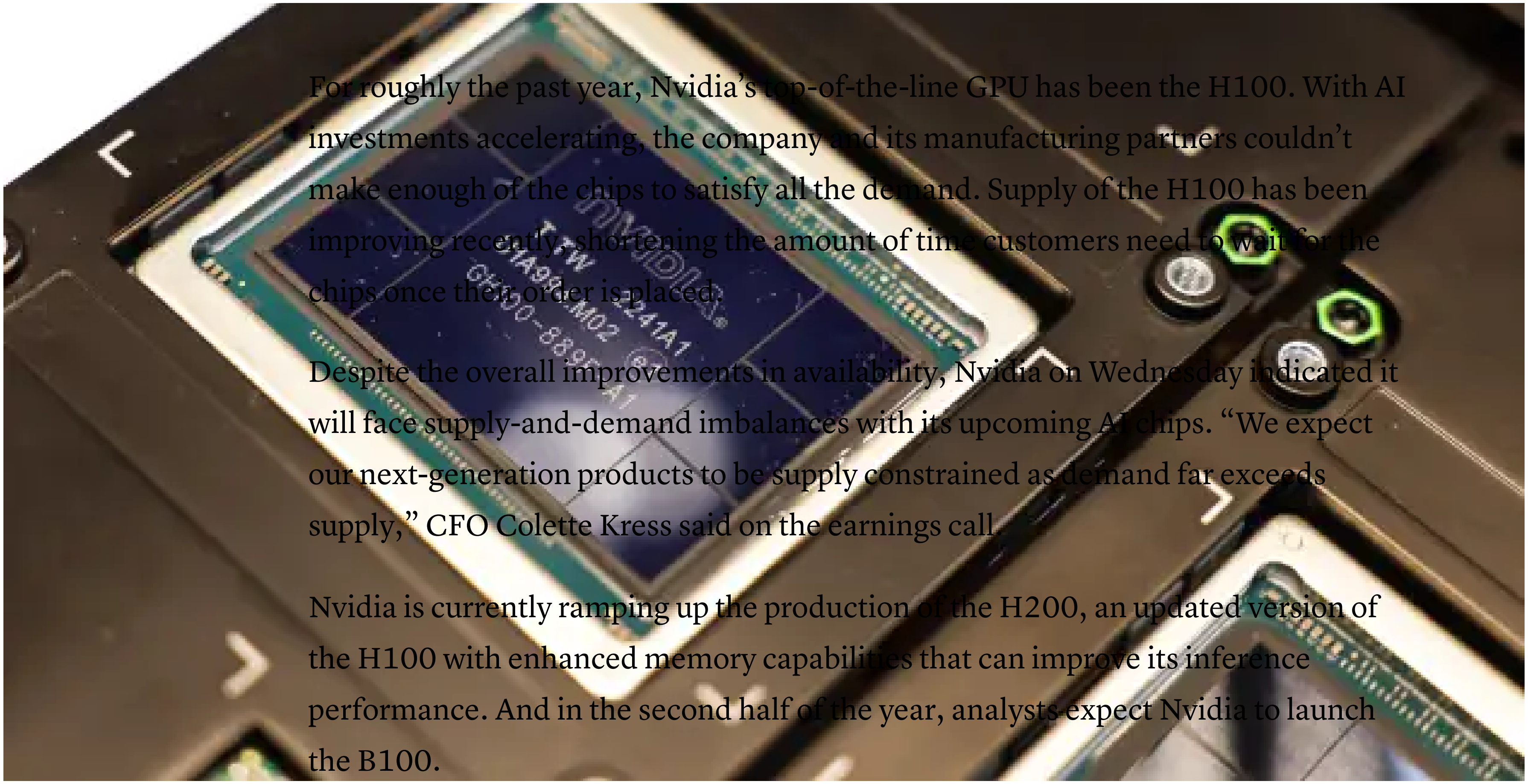
In a note to clients Wednesday night, Raymond James analysts contended that the inference disclosure “highlights the improving monetization trends” of generative AI. That supports the idea that investments in technology may not fall off a cliff anytime soon.

Nvidia’s estimate is positive over the long run, too, because inference is expected to be a more significant portion of AI workloads down the road. The thinking, while the daily use of the existing ones will be more prevalent."

"There is a perception that Nvidia's share will be lower in inferencing vs. training, so this revelation helps shed light on its ability to benefit from the coming inferencing explosion," analysts at Melius Research said in a note.

2. Demand for next-gen chips

CNBC SHARE



For roughly the past year, Nvidia's top-of-the-line GPU has been the H100. With AI investments accelerating, the company and its manufacturing partners couldn't make enough of the chips to satisfy all the demand. Supply of the H100 has been improving recently, shortening the amount of time customers need to wait for the chips once their order is placed.

Despite the overall improvements in availability, Nvidia on Wednesday indicated it will face supply-and-demand imbalances with its upcoming AI chips. "We expect our next-generation products to be supply constrained as demand far exceeds supply," CFO Colette Kress said on the earnings call.

Nvidia is currently ramping up the production of the H200, an updated version of the H100 with enhanced memory capabilities that can improve its inference performance. And in the second half of the year, analysts expect Nvidia to launch the B100.

"Whenever we have new products ... it ramps from zero to a very large number. And you can't do that overnight. Everything is ramped up. It doesn't step up. And so whenever we have a new generation of products – and right now we are ramping H200s – there's no way we can reasonably keep up on demand in the short term as we ramp," Huang added on the call.

mounting competition from rival Advanced Micro Devices and Big Tech companies that make their chips, like Google parent Alphabet. The general goal is to deliver improved AI chips that customers want to upgrade to quicker than before — in an October presentation, Nvidia indicated it plans to release a new version roughly every year, instead of every two years, or so.

The commentary on Wednesday night's earnings call suggests that customer interest in Nvidia's upcoming chips is strong, further quelling the demand sustainability concerns.

"We have heard that most H100 supply through end of life has already been allocated and that products such as B100 are sold out for several quarters," Morgan Stanley analysts wrote in a note to clients.

3. China situation isn't all bad

Nvidia has started shipping modified AI chips to Chinese customers that comply with the latest U.S. government restrictions, which went into effect in October, the company said. Nvidia's plans to develop compliant chips were widely known, but the timing of the rollout had been uncertain.

"We reset our product offering to China and now we're sampling to customers in China," Huang said. "We're going to do our best to compete in that marketplace and succeed in that marketplace within the specifications of the restriction."

AI chips out of the hands of the Chinese military. The controls — which were first implemented in 2022 and then tightened in October — serve as de-facto bans on chips that exceed certain performance thresholds. Nvidia's H100, for example, is well above the limits.

In the November-through-January quarter, China accounted for a mid-single-digit percentage of Nvidia's $18.4 billion in data center revenue, according to the company. Hypothetically, if those sales were 5% of the segment, it would represent $920 million in revenue. Nvidia said it expects sales to China to remain roughly in the same in the current quarter.

Historically, Chinese customers contributed between 20% and 25% of Nvidia's data center revenue, the company has said.

Nvidia's update on its China business carries mixed implications. But, for the most part, it is favorable.

Just as management told Wall Street, Nvidia was more than able to offset the loss of China revenue in the quarter due to growing demand in other parts of the world. Despite China plummeting to mid-single-digit percentages, Nvidia's data center segment still grew a hearty 27% compared with the prior quarter.

The fact Nvidia still booked China revenue in the hundreds of millions of dollars means it could take a further hit in the region. Put another way: There is room to go hurdles of local competition crowding Nvidia out of the market.

The alternative — that Nvidia's China revenue picks up later this year — appears to be the stronger possibility, particularly considering the company and the U.S. government at least publicly maintain a solid working relationship. Indeed, analysts at Truist Securities wrote in a note Wednesday that a recovery in Nvidia's July quarter is "more likely" than further deterioration.

Even if the China business does not quickly jump back to its historic contribution levels, it can still be a source of additional revenue for Nvidia.

(Jim Cramer's Charitable Trust is long NVDA. See here for a full list of the stocks.)

As a subscriber to the CNBC Investing Club with Jim Cramer, you will receive a trade alert before Jim makes a trade. Jim waits 45 minutes after sending a trade alert before buying or selling a stock in his charitable trust's portfolio. If Jim has talked about a stock on CNBC TV, he waits 72 hours after issuing the trade alert before executing the trade.

MORE IN INVESTING CLUB

How Nvidia's AI strengths are enhancing the businesses of Alphabet, Microsoft, Meta, Amazon

This record-high industrial stock is a big winner of Nvidia's robust data center business

Nvidia lives up to the earnings hype and then some, prompting our big price target hike

MORE IN ANALYSIS

SHARE

CLUB How Nvidia's AI strengths are enhancing the businesses of Alphabet, Microsoft, Meta, Amazon

Paulina Likos AN HOUR AGO

CLUB DuPont is turning, plus the Nvidia halo and how to play Palo Alto's sell-off

Jeff Marks 2 HOURS AGO

CLUB Strong earnings results at Bausch Health once again overshadowed by legal battle

Zev Fima 4 HOURS AGO