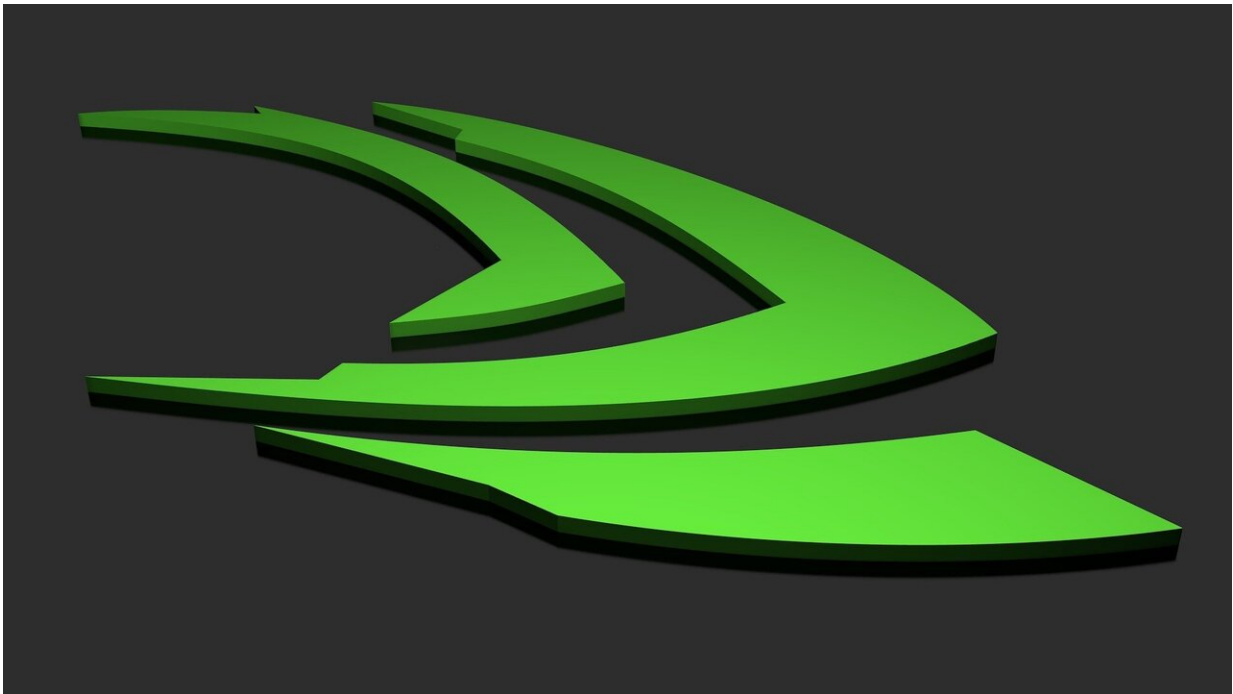


Nvidia backs little-known upstart in India's biggest AI bet yet

March 18 2024, by Saritha Rai, Bloomberg News



Credit: Pixabay/CC0 Public Domain

It's a sultry March evening in the suburbs of Mumbai and a group of men hovers anxiously at the back gate of a startup called Yotta Data Services. They pace, pause and fret. It's approaching midnight, 10 hours late, when a truck pulls up with the precious cargo they've been waiting for: semiconductors from Nvidia Corp.

The company's products are so coveted because they're essential for the development of artificial intelligence, the technology that's set off a frenzy in industries around the world. While companies like OpenAI and Google have poured billions of dollars into such chips in the U.S., Yotta is making India's largest bet yet on the promise of AI.

Sunil Gupta, chief executive officer and co-founder, has gotten a jump on the country's better-known technology players and conglomerates in part because of the relationship he's forged with Jensen Huang, Nvidia's celebrity CEO. Yotta is expected to feature at Nvidia's developer conference Monday in California, an early example of the potential for AI in markets beyond the U.S.

"I'm ambitious, I'm hungry," said Gupta, 52. "I'm willing to take a bet on the future of AI."

Yotta's strategy is to offer high-performance computing capabilities from data centers in India so the country's corporations, startups and researchers will be able to develop their own AI services. Nvidia's chips, the most advanced on the market, are essential for training large language models and building applications like OpenAI's ChatGPT and Microsoft Corp.'s coding assistant, GitHub Copilot.

Gupta figures he's got an edge over cloud computing services outside the country because of latency issues, and he vows to offer the least expensive access to Nvidia AI chips in the world. He's even considering letting Indian startups with tight budgets give him equity instead of cash.

Demand is on his side. The global AI market is projected to grow from \$168.5 billion in 2022 to over \$2 trillion by 2032, according to a report by Spherical Insights & Consulting.

"This is a gold rush," said Stacy Rasgon, an analyst at Sanford C.

Bernstein. "It's still the early days of AI, and companies just can't buy enough of this stuff."

The new era got off to a rocky start this month in India. The country's customs officials were flummoxed by the unusually high value of the Nvidia chips that Yotta had purchased, leading to requisitions of additional paperwork and bureaucratic approval. Back in his data center outside Mumbai, Gupta paced the marble floors of the lobby for the better part of a day, working the phones to get his chips released.

The delivery truck finally pulled up and workers unloaded the first of more than 4,000 H100 chips that Yotta ordered from Nvidia. The beefy graphics processing units, or GPUs, run \$30,000 to \$40,000 each and are called Hoppers in a nod to computer science pioneer Grace Hopper.

The delivery was a religious experience for Gupta, quite literally. A priest adorned the boxes with red vermilion marks and strings of yellow chrysanthemum flowers, while hymns in ancient Sanskrit filled the night air. A camera-carrying drone recorded as Gupta symbolically smashed a coconut on the floor near the truck. "It's a dream moment," he said, amid exploding party poppers.

Yotta's haul of Nvidia chips, which will reach about 20,000 by June, isn't huge by global standards. Tech giants like Microsoft Corp. purchase them by the tens of thousands, and Meta Platforms Inc.'s Mark Zuckerberg said he aims to get 350,000 H100s by year-end. Still, Nvidia's supply is far short of demand so CEO Huang has to calibrate allocations as corporate titans and heads of state press for allotments.

India is getting special attention. In September, Huang met with Prime Minister Narendra Modi and said he would prioritize any orders from data center operators in country. "You have the data, you have the talent," Huang said at the time. "This is going to be one of the largest AI

markets in the world."

The next day, Gupta got a call from the Nvidia team asking him if he could meet the CEO in the western city of Pune. Though it was late evening and the meeting would be the next morning, Gupta quickly agreed. He jumped in his car and drove three and a half hours through the night for the confab. It was a demonstration that Yotta would go above and beyond.

Gupta has serious bona fides in the field. He's been working for decades on data center businesses and co-founded Yotta in 2019 with the backing of real estate billionaire Niranjan Hiranandani. As a cloud computing operator, Yotta offers companies like Wells Fargo & Co. access to data storage and computing power they can scale up or down as needed, without buying and installing their own hardware.

Tata Group and Reliance Industries Ltd., two of the country's largest conglomerates, plan to develop AI infrastructure too, but have yet to order Nvidia's most advanced chips.

An Nvidia spokeswoman declined to comment on the specifics of Yotta's order, pointing out that more will be revealed this week. Gupta is speaking at Nvidia GPU Technology Conference, and he has been told Huang will discuss Yotta during his keynote on Monday.

One reason for the attention is a global imbalance in AI. If the technology has the potential to transform virtually every industry, as Huang and Microsoft CEO Satya Nadella argue, then countries like India, Indonesia or Turkey are at risk without access. In India, that could stymie scientific research, startup development and, more broadly, Modi's ambitions to create a technology superpower. "GPU disparity" is an increasingly popular term for the dilemma.

"Countries who don't have their own AI infrastructure and models will woefully lose the AI race," said Umakant Soni, co-founder of a nonprofit AI and robotics research park called ARTPARK.

Gupta sees a clear need to develop India-built AI models, trained with local languages and cultural diversity. "India needs sovereign AI, India needs sovereign models," he said.

Geopolitics is helping his case. Rising tensions between the U.S. and China have led the Biden administration to levy sweeping controls over the export of technologies to its geopolitical rival, including the very H100 Nvidia chips Yotta is buying. Cloud providers in the Middle East have also come under scrutiny after a key U.S. lawmaker urged the Commerce Department to probe the Chinese connections of Abu Dhabi-based AI firm G42.

Gupta figures he can supply Indian customers and others in Asia and the Middle East. Yotta already has half-dozen [data centers](#) in four Indian cities, and a new one opening in India's northeast. The entrepreneur named his company after the number eight in ancient Greek, representing one septillion.

"India is playing a bit of catchup," said Nruthya Madappa, a partner with the venture capital firm 3one4 Capital. "But because of the talent base, we see the catchup being very, very fast."

The seven-floor data center outside Mumbai is surrounded by electric fences, equipped with 850 cameras and includes seven layers of security. Mammoth diesel storage tanks hold enough fuel to run the facilities for 48 hours if the power goes out.

Gupta's partnership with Nvidia mandates such rigid protocols, along stringent specs for building the AI cloud business. He's sealed off the

facility's entire sixth floor for that purpose. An Nvidia team will arrive in the coming weeks to get the network up and running, with a target of starting operations in mid-May. Gupta calls the first H100 cloud service Shakti, the Hindi word for power.

He says he's sold out capacity for the day his network goes live, and has a waiting list of companies from India and beyond. Gupta is already looking forward to the next delivery of Nvidia chips, more than 16,000 scheduled for June. He'll do at least one thing differently though: hire guards since the value of the shipment could run into the hundreds of millions of dollars.

"Security?! I never thought of that!" said Gupta. "A lot of people want these."

2024 Bloomberg L.P.

Distributed by Tribune Content Agency, LLC.

Citation: Nvidia backs little-known upstart in India's biggest AI bet yet (2024, March 18)
retrieved 27 April 2024 from

<https://techxplore.com/news/2024-03-nvidia-upstart-india-biggest-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.