

New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel
Insights

[Try Now!](#) ×

THE CHANNEL CO.

CRN®



[HOME](#) ▶ [NEWS](#) ▶ [COMPONENTS PERIPHERALS](#) ▶ [2024](#) ▶ [INTEL REVEALS 8-CHIP GAUDI 3 PLATF...](#)

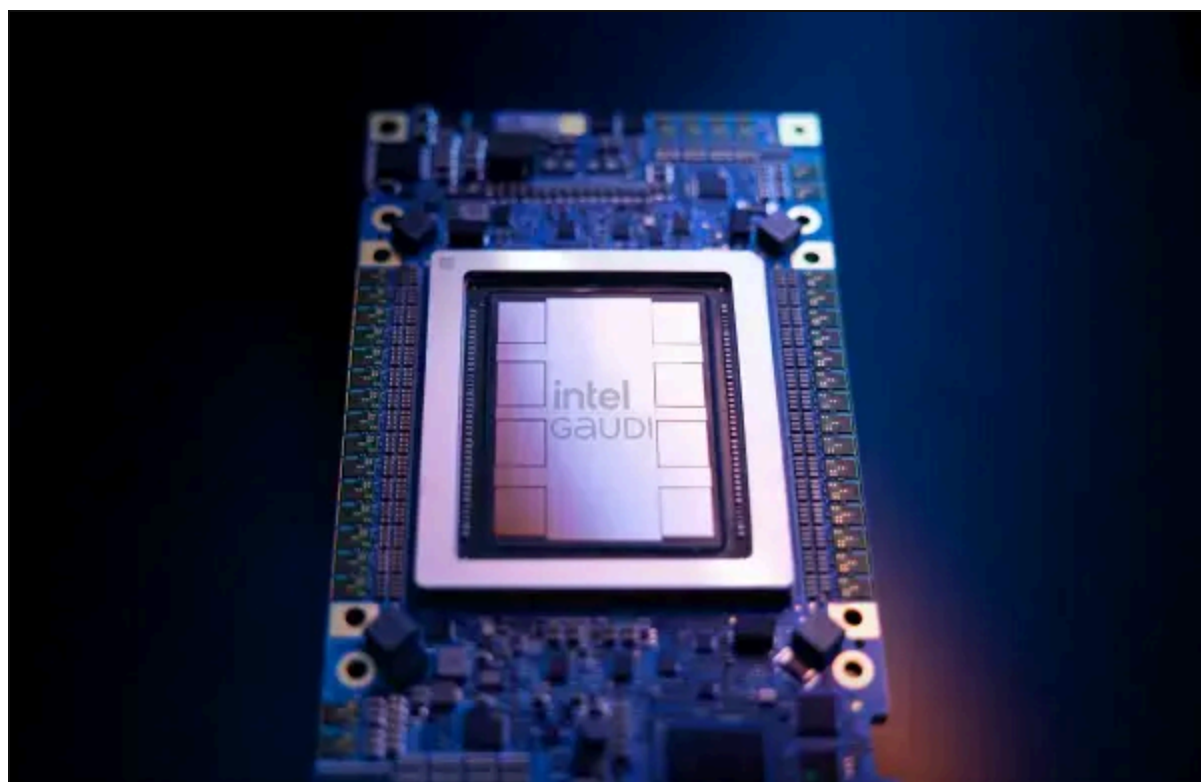
[Components & Peripherals](#) ▶

Intel Reveals 8-Chip Gaudi 3 Platform Price, Upending Industry Norm Of Secrecy

BY [DYLAN MARTIN](#)

JUNE 3, 2024, 11:00 PM EST

While public pricing for CPUs has been standard in the data center industry, it's been the opposite case for GPUs and other kinds of accelerator chips. An Intel exec explains to CRN why the company is changing its stance with the upcoming Gaudi 3 AI chips.



New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel
Insights

Try Now!

Intel said it's upending a norm of secrecy in the data center accelerator chip space by revealing the list price of the star server platform for its upcoming Gaudi 3 AI chip.

On Tuesday at Computex 2024 in Taiwan, the semiconductor giant said its universal baseboard with eight Gaudi 3 OAM modules will have a list price of \$125,000 for server vendors, including Dell Technologies, Hewlett Packard Enterprise, Lenovo and Supermicro, who plan to support the chip when it launches in the third quarter.

[Related: [Intel's Gaudi 3 AI Chip Targets Nvidia H100, H200; Scales To 8,000-Chip Clusters](#)]

Intel made the disclosure after rivals **Nvidia** and **AMD** discussed plans at Computex earlier in the week to release new and more powerful accelerator chips every year, forecasting greater challenges ahead for the Santa Clara, Calif.-based chipmaker.

The \$125,000 price puts the Gaudi 3 platform at two-thirds the estimated cost of Nvidia's eight-GPU HGX H100 platform, which serves as the basis for the rival's DGX H100 system and certain third-party servers, Anil Nanduri, vice president and head of Intel's AI acceleration office told CRN in an interview last Friday.

While Nvidia does not disclose pricing of its data center GPUs or platforms, Intel's calculation brings the estimated cost of an HGX H100 platform to roughly \$187,000. Nanduri noted that pricing can vary based on volume and how a server is configured.

Combining the price with its performance capabilities, the Gaudi 3 chip provides 2.3 times greater performance-per-dollar than Nvidia's H100 for inference throughput and 90 percent better performance-per-dollar for training throughput, according to testing by Intel.

The company also revealed a list price of \$65,000 for a universal baseboard outfitted with eight of its Gaudi 2 chips that debuted in 2022. This puts the Gaudi 2 platform at one-third the estimated cost of Nvidia's HGX H100 platform, it said.

CRN Answers

New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel
Insights

Try Now!

- [When are Intel Gaudi 3 systems available from vendors?](#)
- [How does Intel's Gaudi 3 pricing transparency help startups plan AI investments?](#)
- [What is the historical norm of pricing secrecy for data center accelerator chips?](#)
- [What are the benefits of an open ecosystem for AI solutions?](#)

What do you want to know?

Ask CRN

Intel did not provide pricing for individual Gaudi 3 chips. The chips will debut in the third quarter with a PCIe version and the OAM version, the latter of which is based the Open Compute Project's Accelerator Module design specification.

Why Intel Disclosed Gaudi 3's Platform Price

Nanduri said Intel decided to disclose the list prices of its Gaudi 3 and Gaudi 2 platforms because it will help customers, particularly startups who typically don't have visibility into behind-the-scenes pricing from vendors, plan investments in AI computing more easily.

"A lot of the innovation is happening from startups. They're not familiar with the ecosystems [and] how the purchasing, procurement [and request for quotes] work, so it helps them get a framing of what to expect," he said.

While public pricing for CPUs has been standard in the data center industry, it's been the opposite case for GPUs and other kinds of accelerator chips, with vendors like Nvidia and AMD failing to publicly disclose such information in the past. Until now, Intel also had kept pricing secret for its GPUs and accelerator chips.

"We've historically done that with CPUs. When it comes to GPUs, the norm was pricing was always kind of a secret sauce," Nanduri said.

Nanduri said the company decided to change its stance on pricing disclosures for accelerator chips as Gaudi 3 gained support not only from Dell, HPE, Lenovo and

New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel Insights

[Try Now!](#)

What was becoming clear to Intel is that customers are becoming increasingly concerned with customizing AI models in a way that is “computationally sustainable from a cost perspective,” according to the Intel executive. These customers are also sensitive to costs related to inferencing, where, for instance, the number of words—or tokens in AI parlance—produced at a certain rate for a large language model becomes important.

“When customers look at the tokens throughput per dollar or the cost of a million tokens, you see a huge disparity in terms of the cost, and a lot of it is fundamentally driven by either the model size [or] the cost of the infrastructure,” he said.



BY [DYLAN MARTIN](#) ▶

Dylan Martin is a senior editor at CRN covering the semiconductor, PC, mobile device, and IoT beats. He has distinguished his coverage of the semiconductor industry thanks to insightful interviews with CEOs and top executives; scoops and exclusives about product, strategy and personnel changes; and analyses that dig into the why behind the news. He can be reached at dmartin@thechannelcompany.com.



LEARN MORE: [AI](#) ▶ | [Generative AI](#) ▶ | [Accelerator Chips](#) ▶



RELATED CONTENT

AMD's OpenAI Deal A 'Major Validation Moment' For Chip Designer: Analyst

10 Hot MDR Vendors Making Moves In 2025

AMD Stock Skyrockets After Announcing 'Major' 6-Gigawatt OpenAI GPU Deal

Next-Gen Data Centers: Unleashing Performance With Dell PowerEdge And AMD EPYC

New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel
Insights

[Try Now!](#)

CRN AWARDS

Sign up to be notified about CRN awards ►



[Sign Up For Channel Chiefs 2026 ►](#)

CRN MAGAZINE

Browse ►



CRNtv

Get the scoop on top industry trends

CRNtv

ON LOCATION

New!

CRN Answers: AI-Powered Search for Trusted Tech & Channel Insights

[Try Now!](#)



NEWSLETTERS

Get the IT channel news you need, right to your inbox.

Subscribe ►



▲ [TO TOP](#)



© 2025 The Channel Company.
All rights reserved.

[SUBSCRIBE](#) [ABOUT](#) [CRN CONTACT](#) [CRN ETHICS](#) [POLICY](#) [THE CHANNEL COMPANY](#)

[Terms & Conditions](#) | [Privacy Policy](#) |
[Cookie Policy](#) | [Site Map](#)

Follow CRN

