December 1, 2020

AWS Leverages Habana Gaudi AI Processors

SANTA CLARA, Calif.--(BUSINESS WIRE)-- **What's New:** Today at AWS re:Invent 2020, AWS CEO Andy Jassy announced EC2 instances that will leverage up to eight Habana® Gaudi® accelerators and deliver up to 40% better price performance than current graphics processing unit-based EC2 instances for machine learning workloads[1]. Gaudi accelerators are specifically designed for training deep learning models for workloads that include natural language processing, object detection and machine learning training, classification, recommendation and personalization.

This press release features multimedia. View the full release here: https://www.businesswire.com/news/home/20201201006008/en/

A photo shows Habana Labs' HL-205 Gaudi Mezzanine Card. Gaudi-based EC2 instances deliver cost efficiency and high performance, while natively supporting common frameworks such as TensorFlow and PyTorch. (Credit: Habana Labs)

"We are proud that AWS has chosen Habana Gaudi processors for its forthcoming EC2 training instances. The Habana team looks forward to our continued collaboration with AWS to deliver on a roadmap that will provide customers with continuity and advances over time." –David Dahan, chief executive officer at Habana Labs, an Intel Company

Why It Matters: As the world's leading cloud provider, AWS is used by developers around the world to train their artificial intelligence (AI) models. However, the increase in complexity of machine learning models drives up both the time and cost to train, especially as more data becomes available and developers look to refine their models. Gaudi-based EC2 instances are designed to address these needs by delivering cost efficiency and high performance, while natively supporting common frameworks such as TensorFlow and PyTorch. And using Habana's SynapseAI Software Suite, developers will be able to easily build new or port existing training models

from graphics processing units to Gaudi accelerators.

How It Fits in Intel's AI and XPU Vision: Intel acquired Habana in 2019 to advance its AI strategy and strengthen its portfolio of AI accelerators for the cloud and data center. This includes a mix of products and technologies that power some of the most promising AI use cases in business, society and research. It also reflects the company's shift to delivering XPUs – a mix of architectures across CPUs, GPUs, FPGAs and more to help customers and the entire ecosystem unleash the potential of data.

"Our portfolio reflects the fact that artificial intelligence is not a one-size-fits-all computing challenge," said Remi El-Ouazzane, chief strategy officer of Intel's Data Platforms Group. "Cloud providers today are broadly using the built-in AI performance of our Intel® Xeon® processors to tackle AI inference workloads. With Habana, we can now also help them reduce the cost of training AI models at scale, providing a compelling, competitive alternative in this high-growth market opportunity."

More Context: Habana Gaudi AI Processors to bring lower cost-to-train to Amazon EC2 customers (Habana Labs Blog) | Customer Enablement of Habana® Gaudi® Amazon EC2 Instances (White Paper) | AWS and Intel | Gaudi AI Training

Intel Customer Stories: Intel Customer Spotlight on Intel.com | Customer Stories on Intel Newsroom

About Intel

Intel (Nasdaq: INTC) is an industry leader, creating world-changing technology that enables global progress and enriches lives. Inspired by Moore's Law, we continuously work to advance the design and manufacturing of semiconductors to help address our customers' greatest challenges. By embedding intelligence in the cloud, network, edge and every kind of computing device, we unleash the potential of data to transform business and society for the better. To learn more about Intel's innovations, go to newsroom.intel.com and intel.com.

[1]The price performance claim is made by AWS and based on AWS internal testing. Habana Labs does not control or audit third-party data; your costs and results may vary.

View source version on businesswire.com:
https://www.businesswire.com/news/home/20201201006008/en/

Liz Wu
1-503-696-2098
liz.wu@intel.com

Source: Intel Corporation