November 17, 2019

Intel Unveils New GPU Architecture with High-Performance Computing and AI Acceleration, and oneAPI Software Stack with Unified and Scalable Abstraction for Heterogeneous Architectures

NEWS HIGHLIGHTS

- Intel launches oneAPI, a unified and scalable programming model to harness the power of diverse computing architectures in the era of HPC/AI convergence.
- Intel introduces a general-purpose GPU optimized for HPC/AI acceleration based on the Xe architecture, code-named “Ponte Vecchio.”
- Intel unveils additional architectural details of the Aurora Supercomputer, delivering convergence at exascale at Argonne National Laboratory.

DENVER--(BUSINESS WIRE)-- At Supercomputing 2019, Intel unveiled its vision for extending its leadership in the convergence of high-performance computing (HPC) and artificial intelligence (AI) with new additions to its data-centric silicon portfolio and an ambitious new software initiative that represents a paradigm shift from today’s single-architecture, single-vendor programming models.

This press release features multimedia. View the full release here: https://www.businesswire.com/news/home/20191117005061/en/

The Aurora system at Argonne National Laboratory will be the first U.S. exascale system to leverage the full breadth of Intel’s data-centric technology portfolio, building upon the Intel Xeon Scalable platform and using Xe architecture-based GPUs, as well as Intel Optane DC persistent memory and connectivity technologies. (Credit: Intel Corporation)

Addressing the increasing use of heterogeneous architectures in high-performance computing, Intel expanded on its existing technology portfolio to move, store and process data more effectively by announcing a new category of discrete general-purpose GPUs optimized for AI and HPC convergence. Intel also launched the oneAPI industry initiative to deliver a unified and simplified programming model for application development across heterogenous processing architectures, including CPUs, GPUs, FPGAs and other accelerators. The launch of oneAPI represents millions of Intel engineering hours in software development and marks a game-changing evolution from today’s limiting, proprietary programming approaches to an open standards-based model for cross-architecture developer engagement and innovation.

More: oneAPI Fact Sheet | oneAPI Infographic | Intel Data Center News

“HPC and AI workloads demand diverse architectures, ranging from CPUs, general-purpose GPUs and FPGAs, to more specialized deep-learning NNPs, which Intel demonstrated earlier this month,” said Raja Koduri, senior vice president, chief architect, and general manager of architecture, graphics and software at Intel. “Simplifying our customers’ ability to harness the power of diverse computing environments is paramount, and Intel is committed to taking a software-first approach that delivers a unified and scalable abstraction for heterogeneous architectures.”

oneAPI: A Developer-Centric Approach to Heterogeneous Computing

The oneAPI initiative Intel launched today will define programming for an increasingly AI-infused, multi-architecture world. oneAPI delivers a unified and open programming experience to developers on the architecture of their choice without compromising performance and eliminating the complexity of separate code bases, multiple-programming languages, and different tools and workflows. oneAPI preserves existing software investments with support for existing languages while delivering flexibility for developers to create versatile applications.

oneAPI includes both an industry initiative based on open specifications and an Intel beta product. The oneAPI specification includes a direct programming language, powerful APIs and a low-level hardware interface. Intel’s oneAPI beta software provides developers a comprehensive portfolio of developer tools that include compilers, libraries and analyzers, packaged into domain-focused toolkits. The initial oneAPI beta release targets Intel® Xeon® Scalable processors, Intel® Core™ processors with integrated graphics, and Intel® FPGAs, with additional hardware support to follow in future releases. Developers can download the oneAPI tools, test drive them in the Intel oneAPI DevCloud, and learn more about oneAPI at software.intel.com/oneAPI.

Intel’s Data-Centric Strategy Delivers the Foundation for AI/HPC Convergence

Intel’s silicon portfolio is comprised of a diverse mix of architectures deployed in a range of silicon platforms. The foundation of Intel’s data centric strategy is the Intel® Xeon® Scalable processor, which today powers over 90 percent of the world’s Top500 supercomputers. Intel Xeon Scalable processors are the only x86 CPUs with built-in AI acceleration that are optimized to analyze the massive data sets in HPC workloads.

At Supercomputing 2019, Intel unveiled a new category of general-purpose GPUs based on Intel’s X^e architecture. Code-named “Ponte Vecchio,” this new high-performance, highly flexible discrete general-purpose GPU is architected for HPC modeling and simulation workloads and AI training. Ponte Vecchio will be manufactured on Intel’s 7nm technology and will be Intel’s first X^e-based GPU optimized for HPC and AI workloads. Ponte Vecchio will leverage Intel’s Foveros 3D and EMIB packaging innovations and feature multiple technologies in-package, including high-bandwidth memory, Compute Express Link interconnect and other intellectual property.

Building the Foundation for Exascale Computing

Intel’s data-centric silicon portfolio and oneAPI initiative lays the foundation for the

convergence of HPC and AI workloads at exascale within the Aurora system at Argonne National Laboratory. Aurora will be the first U.S. exascale system to leverage the full breadth of Intel's data-centric technology portfolio, building upon the Intel Xeon Scalable platform and using X^e architecture-based GPUs, as well as Intel Optane DC persistent memory and connectivity technologies. The compute node architecture of Aurora will feature two 10nm-based Intel Xeon Scalable processors (code-named "Sapphire Rapids") and six Ponte Vecchio GPUs. Aurora will support over 10 petabytes of memory and over 230 petabytes of storage. Aurora will leverage the Cray Slingshot fabric to connect nodes across more than 200 racks.

About Intel

Intel (NASDAQ: INTC), a leader in the semiconductor industry, is shaping the data-centric future with computing and communications technology that is the foundation of the world's innovations. The company's engineering expertise is helping address the world's greatest challenges as well as helping secure, power and connect billions of devices and the infrastructure of the smart, connected world – from the cloud to the network to the edge and everything in between. Find more information about Intel at newsroom.intel.com and intel.com.

© Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.

View source version on businesswire.com:
https://www.businesswire.com/news/home/20191117005061/en/

Steve Gabriel
408-655-5513
stephen.gabriel@intel.com

Source: Intel Corporation