CUSTOMER CASE STUDY

The Center for Computational Sciences (CCS) at Japan's University of Tsukuba is using a Cray CS300-AC™ to accelerate GPU computing — and fast track the race to exascale

Organization

University of Tsukuba
Center for Computational Sciences
Tsukuba, Japan
www.tsukuba.ac.jp

筑波大学
University of Tsukuba

Cray CS300-AC Cluster Supercomputer

The Cray CS300-AC™ cluster supercomputer is a highly scalable and air-cooled architecture that groups optimized, industry-standard building block server platforms into a unified, fully integrated system. Featuring the latest processor and network technologies and capable of scaling to 25 petaflops, the Cray CS300-AC is designed to meet the challenges of medium- to large-scale capacity and data-intensive workloads.

HA-PACS System Overview

- Peak performance: 802 teraflops
- Compute nodes: 268
- GPU: 1,072 NVIDIA® Tesla™ M2090 GPUs; 4 per node
- CPU: 536 eight-core Intel® Xeon® E5 processors; 2 per node
- Memory (TB): 40+
- Cluster size: 26 racks
- Energy consumption: 400 KW
- Software: HPC cluster software stack with Advanced Cluster Engine™

"This deployment will put us on the fast path towards exascale computing."

—Dr. Taisuke Boku
Professor and Deput Director, Center for Computational Sciences
University of Tsukuba

Customer

The Center for Computational Sciences (CCS) at the University of Tsukuba is an inter-university research facility aimed at enabling scientific discovery through the use of leading-edge computing technologies. Using a collaborative approach between scientific researchers from a variety of disciplines and computational and information scientists, CCS is working to advance computational science and technology.

Challenge: The HA-PACS Project

In pursuit of supporting meaningful scientific discovery, CCS undertakes research and development of high performance computing (HPC) systems as well as ongoing studies in computer and information science.

HA-PACS — the Highly Accelerated Parallel Advanced system for Computational Sciences — is CCS's current HPC project. Kicked off in 2011, the three-year effort is focused on maximizing the performance of GPU-accelerated computing with the end goal of moving HPC closer to exascale.

CCS is tackling three specific challenges:

1) porting codes to the graphics processing units (GPU) in areas such as life sciences, astrophysics and environmental science;

2) developing custom hardware called the Tightly Coupled Accelerator (TCA) that will speed performance by allowing direct communication between the GPUs — eliminating the need to go through the CPU; and

3) developing the new generation of algorithms and codes necessary to facilitate applications on accelerator technology.

"The largest issue [for] accelerated computing is how to fill the gap between its powerful internal computation performance and relatively poor external communication performance," says Taisuke Boku, leader of the High Performance Computing Systems group and CCS deputy director. "In some applications, we may need a paradigm shift toward a new generation of algorithms. HA-PACS will be the testbed for developing these algorithms."

Cray Inc.
901 Fifth Avenue, Suite 1000
Seattle, WA 98164
Tel: 206.701.2000
Fax: 206.701.2500
www.cray.com

Solution: Cray CS300-AC Cluster Supercomputer

To achieve their goal, CCS needed a configurable, energy-efficient, GPU-accelerated solution to serve as their demonstration system. They chose the Cray CS300-AC cluster supercomputer equipped with NVIDIA Tesla M2090 GPU processors and Intel Xeon E5 family processors.

The HA-PACS machine features 268 compute nodes, each with four GPUs and two 8-core Intel Xeon processors connected by the next-generation PCI Express for rich I/O bandwidth. Each node provides 2.99 teraflops of performance with 655 gigaflops coming from each GPU and 166 from each Intel Xeon processor for a total of 802 teraflops of peak performance.

The system is interconnected with dual-rail QDR InfiniBand and configured with Cray's HPC cluster software stack including Advanced Cluster Engine™ (ACE) remote management software. Higher density afforded by the GPUs means CCS is getting higher performance in a smaller footprint — the HA-PACS system consists of 26 racks and consumes just over 400 kilowatts. Additionally, the system gets a cooling efficiency boost from highly efficient power supplies and fans.

When the HA-PACS team completes development of their Tightly Coupled Accelerator (TCA) technology, they'll add the TCA cluster to the HA-PACS machine, giving it an additional 200-plus teraflops and pushing the integrated system to over a petaflop.

A single petaflop isn't anywhere close to exascale. But that's not the point. Dr. Boku says they don't think GPUs are the final solution to the exascale challenge; their performance and energy efficiency are simply an important part of the calculation. "This deployment will put us on the fast path towards exascale computing."