Argonne NATIONAL LABORATORY

ALCF Computing Resources

The Argonne Leadership Computing Facility operates supercomputing resources that support large-scale computing projects aimed at solving some of the world's most complex problems in science and engineering.

Aurora

Designed in collaboration with Intel and HPE, Aurora will be one of the nation's first exascale systems when it arrives in 2022. The supercomputer will be based on Intel's Xeon Scalable processors and high-performance Intel X^{e} GPU compute accelerators. It will rely on HPE's Cray EX supercomputer exascale-class architecture and Slingshot 11 interconnect technology, which can provide concurrent support for advanced simulation and modeling, AI, and analytics workflows. Aurora will leverage historical advances in software investments along with increased application portability via Intel's oneAPI. The system will also introduce a new I/O system called Distributed Asynchronous Object Storage (DAOS) to meet the needs of new exascale workloads.

- CY 2022 delivery
- 2 Intel Xeon scalable "Sapphire Rapids" processors; 6 X^{e} arch-based GPUs; Unified Memory Architecture; 8 fabric endpoints; RAMBO
- X^{e} arch-based "Ponte Vecchio" GPU archticture; Tile-based, chiplets, HBM stack, Foveros 3D integration
- CPU-GPU: PCIe; GPU-GPU: Xe Link; CPU-CPU: UPI
- > 10 PB aggregate system memory
- HPE Slingshot 11; Dragonfly topology with adaptive routing
- ≥ 1EF DP sustained performance
- 25.6 Tb/s per network switch, from 64–200 Gbs ports (25 GB/s per direction)
- ≥ 230 PB, ≥ 25 TB/s (DAOS) high-performance storage
- Intel oneAPI, MPI, OpenMP, C/C++, Fortran, SYCL/DPC++
- HPE Cray EX supercomputer software stack + Intel enhancements + data and learning
- HPE Cray EX supercomputer
- > 100 cabinets

Theta

Theta is an 11.7-petaflops supercomputer based on Intel processors and interconnect technology, an advanced memory architecture, and a Lustre-based parallel file system, all integrated by Cray's HPC software stack.

- Intel-Cray XC40 architecture
- 11.69 petaflops
- 64-core, 1.3-GHz Intel Xeon Phi 7230 processor per node
- 4,392 nodes
- 281,088 cores
- 843 TB of memory
- 70 TB of high-bandwidth memory
- Aries interconnect with Dragonfly configuration
- 24 racks

ThetaGPU

ThetaGPU is an NVIDIA DGX A100-based system. The DGX A100 comprises eight NVIDIA A100 GPUs that provide a total of 320 gigabytes of memory for training AI datasets, as well as high-speed NVIDIA Mellanox ConnectX-6 network interfaces.

- NVIDIA DGX A100 architecture
- 3.9 petaflops
- AMD EPYC 7742
- 24 nodes
- 24 TB of DDR4 memory
- 7, 680 GB of GPU memory
- 7 racks

U.S. DEPARTMENT OF ENERGY

Argonne National Laboratory is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC.

CONTACT
media@alcf.anl.gov
alcf.anl.gov

Data Storage

At the ALCF, disk storage provides intermediate-term storage for projects, offering a means to access, analyze, and share computational and experimental data. Tape storage is used to archive data from completed projects.

DISK STORAGE

The Theta storage system consists of 30 I/O nodes that connect to a storage array that controls 2,300 disk drives with a total useable capacity of 9 PB and a maximum aggregate transfer speed of 240 GB/s. Theta uses Lustre to access this storage. The ALCF also utilizes a 10 PB filesystem based on an IBM Elastic Storage Server (ESS) to host data for science running on the Theta and Cooley systems. The ESS system is a software defined storage system based on IBM's GPFS file system and consists of 60 I/O nodes controlling 7,260 disk drives.

The ALCF supports two 100 PB globally accessible Lustre filesystems named Grand and Eagle. Each storage array controls 8,480 disk drives with a sustained transfer speed of 650 GB/s in the current environment. Eagle storage allocations will be granted to projects focused primarily on data sharing and will be accessible by non-ALCF users via Globus. Grand storage will be provided with standard compute allocations and will only be accessible to ALCF users who are project members.

A new home filesystem, homefs, will reside on an all-flash storage array with a capacity of 220 TB and a transfer rate up to 24 GB/s. A new filesystem named Swift will reside on an all flash storage array with a capacity of 123 TB and transfer rate up to 48 GB/s. It is intended to be targeted by GPU-based workflows.

TAPE STORAGE

The ALCF has three 10,000-slot libraries. The tape technology is currently undergoing an upgrade to replace LTO-6 tape drives with LTO-8 tape drives. The upgrade should ultimately provide up to 300 PB of effective storage (approximately five times the amount provided by the LTO-6 tapes).

Networking

Theta has an internal proprietary network for communicating between nodes. InfiniBand enables communication between the I/O nodes and the storage system. Ethernet is used for external user access, and for maintenance and management of the systems.

The ALCF connects to other research institutions using up to 100 Gb/s of network connectivity. Scientists can transfer datasets to and from other institutions over fast research networks, such as ESnet and Internet2.

Testbeds

Argonne's Joint Laboratory for System Evaluation (JLSE) enables researchers to assess and improve next-generation computing platforms of interest to DOE. Established by Argonne's computing divisions and run by the ALCF, JLSE centralizes the laboratory's research activities aimed at evaluating future extreme-scale computing systems, technologies, and capabilities. JLSE users leverage existing infrastructure and next-generation hardware and software to explore low-level experimental computer and computational science, including operating systems, messaging, compilers, benchmarking, power measurements, I/O, and new file systems. Some notable JLSE testbeds include:and software to explore low-level experimental computer and computational science, including operating systems, messaging, compilers, benchmarking, power measurements, I/O, and new file systems. JLSE testbeds include:

Arcticus, Yarrow: Intel discrete GPU testbeds for ECP and ESP projects to develop, optimize, and scale applications and software for Aurora

Atos Quantum Learning Machine: Platform for testing and developing quantum algorithms and applications

Aurora Software Development Kit: Frequently updated version of the publicly available Intel oneAPI toolkit for Aurora development

HPE Comanche Prototype: ARM64 platform for exploring the Marvell Arm architecture

Intel Cooper Lake: Intel Xeon cluster for testing data types for AI and learning applications Cluster

Iris: Intel integrated Gen9 GPUs for ECP and ESP projects to develop, optimize, and scale applications and software for Aurora

NVIDIA GPUs: Clusters of P100, V100, A100 GPUs for preparing applications for heterogenous computing architectures

Presque: Intel DAOS nodes for testing the Aurora storage system

CONTACT
media@alcf.anl.gov
alcf.anl.gov