Next Stop: Exascale

Designed in collaboration with Intel and Cray, the ALCF's exascale system, Aurora, will help ensure continued U.S. leadership in high-end computing for scientific research.

With Aurora's 2021 arrival drawing closer, the ALCF continued to ramp up its efforts to prepare for the exascale system. In 2019, there was a flurry of activity that included a high-profile announcement event, the unveiling of new architectural details, collaborations to further develop an ecosystem that will enable science in the exascale era, and several training offerings designed to prepare researchers for Aurora.

On March 18, leaders from the U.S. Department of Energy (DOE), Argonne National Laboratory, Intel, and Cray (now part of Hewlett Packard Enterprise) came together to officially announce that the laboratory would be home to one of the nation's first exascale systems with the arrival of Aurora in 2021.

The event, hosted at Argonne, included remarks from former U.S. Secretary of Energy Rick Perry, Intel CEO Robert Swan, Cray CEO Pete Ungaro, University of Chicago President Robert Zimmer, U.S. Senator Dick Durbin, U.S. Representative Dan Lipinski, and Argonne Director Paul Kearns, who joined together to underscore the significance of reaching exascale.

From mapping the human brain to designing new functional materials to advancing our understanding of mysterious cosmological phenomena, the ALCF's forthcoming machine will enable researchers to pursue science and engineering breakthroughs that were not possible with previous supercomputers.

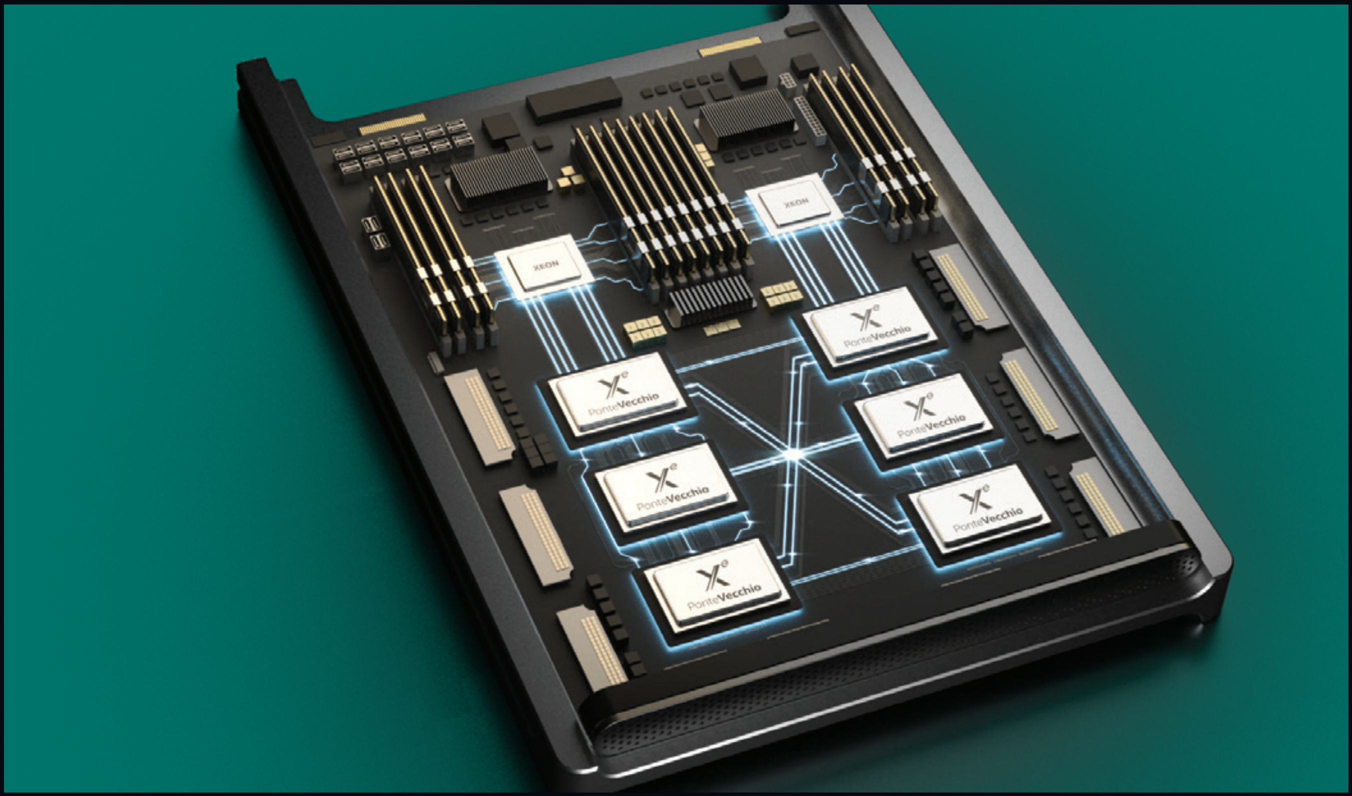
AURORA ARCHITECTURE

Aurora will be based on Intel's Xeon Scalable processors and high-performance Intel X^e GPU compute accelerators. The system will rely on Cray's Shasta exascale-class architecture and Slingshot 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 supercomputer will also introduce a new I/O system called Distributed Asynchronous Object Storage (DAOS) to meet the needs of new exascale workloads.

At the 2019 Supercomputing (SC19) conference in November, Intel unveiled several new architectural details and tools that have allowed researchers to ramp up their efforts to prepare for Aurora.

The system's compute nodes will feature two Intel Xeon Scalable processors (code-named "Sapphire Rapids") and six general-purpose GPUs based on Intel's X^e architecture (code-named "Ponte Vecchio") and optimized for HPC and AI workloads.

Intel also announced that Aurora will have over 10 petabytes of memory and over 230 petabytes of storage. It will leverage the Cray Slingshot fabric to connect nodes across the massive system.



Aurora's compute nodes will be equipped with two Intel Xeon Scalable processors and six general-purpose GPUs based on Intel's X^e architecture.
Image: Intel Corporation

Sustained Performance

≥1 Exaflop DP

Platform

Cray Shasta

Delivery

CY 2021

Aggregate System Memory

>10 PB

Compute Node
2 Intel Xeon scalable "Sapphire Rapids" processors; 6 X^e arch-based GPUs; Unified Memory Architecture; 8 fabric endpoints

Software Stack
Cray Shasta software stack + Intel enhancements + data and learning

Cabinets
>100

GPU Architecture
X^e arch-based "Ponte Vecchio" GPU; Tile-based chiplets, HBM stack, Foveros 3D integration, 7nm

On-Node Interconnect
CPU-GPU: PCIe
GPU-GPU: X^e Link

System Interconnect
Cray Slingshot; Dragonfly topology with adaptive routing

Network Switch
25.6 Tb/s per switch, from 64–200 Gbs ports (25 GB/s per direction)

High-Performance Storage
≥230 PB, ≥25 TB/s (DAOS)

Programming Environment
Intel oneAPI, MPI, OpenMP, C/C++, Fortran, SYCL/DPC++

The ALCF's Intel-Cray exascale system is scheduled to arrive in 2021.

In addition to the new architectural details, Intel launched its oneAPI initiative, which aims to provide a unified programming model that simplifies development for diverse workloads across different architectures.

TESTING AND DEVELOPMENT

The initial oneAPI beta software release provided a new avenue for researchers to prepare for the Aurora architecture. Available via the Intel DevCloud, users were able to gain free access to compilers, libraries, and tools for testing and development work in preparation for the exascale system.

Intel integrated GPUs, Gen9 and later, are also helping researchers gain a better understanding of Aurora's Ponte Vecchio GPU architecture. Available through Argonne's Joint Laboratory for System Evaluation (JLSE), Intel Gen9 integrated GPUs allow developers to test code performance and functionality using programming models that will be supported on Aurora.

EXASCALE PARTNERSHIPS

The ALCF is involved in numerous collaborative efforts to develop and deploy tools and technologies that are essential to enabling science in the exascale era. Staff researchers are engaged in several activities within DOE's Exascale Computing Project (ECP), a multi-lab initiative dedicated to delivering a capable exascale computing ecosystem for the nation. Argonne has strong presence on the ECP leadership team and in projects and working groups focused on application development, software technologies, and hardware and integration.

ALCF staff members continue to contribute to the advancement of software standards (e.g., C++, OpenMP), programming models (e.g., oneAPI, SYCL, Kokkos), and compilers (e.g., Clang/LLVM) that are critical to developing efficient and portable exascale applications.

The ALCF is partnering with Altair to leverage the company's PBS Professional software to create the scheduling system for Aurora. This effort will include working with the PBS Pro open source community to develop an effective solution for exascale job scheduling and workload management.

EARLY SCIENCE

The ALCF's Aurora Early Science Program (ESP) is designed to prepare key applications for the scale and architecture of the exascale machine, and field-test compilers and other software to pave the way for other production applications to run on the system.

Through open calls for proposals, the ESP has awarded pre-production computing time and resources to five simulation projects, five data projects, and five learning projects. The diverse set of projects reflects the ALCF's effort to create an environment that supports emerging data science and machine learning approaches alongside traditional modeling and simulation-based research.

By bringing together computational scientists, code developers, and computing hardware experts, the ESP creates a collaborative environment for optimizing applications and characterizing the behavior of the facility's future exascale system. In partnership with experts from Intel and Cray, ALCF staff members are helping train the ESP teams on the Aurora hardware design and how to develop code for it.

The Aurora Programming Workshop welcomed ECP and ESP researchers to the ALCF for guidance on preparing applications and software for the exascale system.

USER TRAINING

In 2019, the Argonne-Intel Center of Excellence (COE) kicked off a series of intensive, hands-on sessions called "hackathons" to help individual ESP teams advance efforts to port and optimize applications using the Aurora software development kit, early hardware, and other exascale programming tools.

The six 2019 hackathons covered a wide range of application areas and featured a variety of activities tailored to each project, including the implementation of kernels and mini-apps using various Aurora programming models; testing on Intel Gen9 integrated GPUs serving as development platforms; targeted presentations and deep dives; and lengthy Q&A sessions with Intel software and hardware experts.

COE staff coordinated with the ESP project teams before, during, and after the hackathons to prepare, facilitate, and follow up on activities. The hackathons also provided valuable input to Argonne and Intel on the merits and deficiencies of various programming approaches, and helped identify multiple bugs in the pre-alpha/alpha/beta versions of the software development kit.

In April, the Argonne-Intel COE held a three-day workshop to present a deep dive into the Aurora hardware and software environment for all ESP teams. The event covered a wide range of topics including the Aurora software stack; its DAOS I/O system; data science and analytics workloads; programming models; performance tools; and transitioning project workflows from Theta to Aurora.

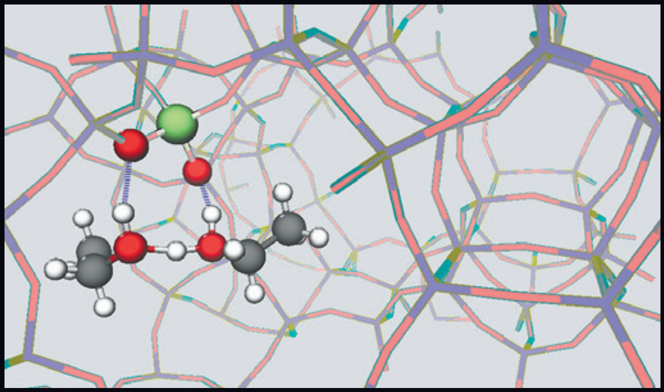
The ALCF hosted another exascale training event in September—the Aurora Programming Workshop—for both ESP and ECP research teams. The three-day workshop was focused on preparing applications and software technologies for Aurora. This event included presentations on Intel's Aurora programming models (e.g., OpenMP, SYCL/DPC++, OpenCL); open programming models (e.g., Kokkos, Raja); development hardware; and best practices and lessons learned thus far. Attendees also participated in hands-on sessions using the latest Aurora software development kit.

In addition to on-site events, the ESP continues to offer web-based tutorials to project teams on topics and tools relevant to leadership-scale computing resources, with an emphasis on data-intensive and machine learning subjects. The ECP and other DOE computing facilities also provide a multitude of workshops and webinars focused on exascale tools and topics.

Together, all of these efforts are preparing the research community to harness the immense computing power of Aurora and other future exascale systems to drive a new era of scientific discoveries and technological innovations.

Readying Aurora for Science on Day One

Simulation Projects



NWChemEx will provide the understanding needed to control molecular processes underlying the production of biomass. *Image: Thom H. Dunning Jr., University of Washington and Pacific Northwest National Laboratory*

Extending Moore's Law Computing with Quantum Monte Carlo

PI Anouar Benali
INST Argonne National Laboratory

High-Fidelity Simulation of Fusion Reactor Boundary Plasmas

PI C.S. Chang
INST Princeton Plasma Physics Laboratory

NWChemEx: Tackling Chemical, Materials, and Biochemical Challenges in the Exascale Era

PI Theresa Windus
INST Iowa State University and Ames Laboratory

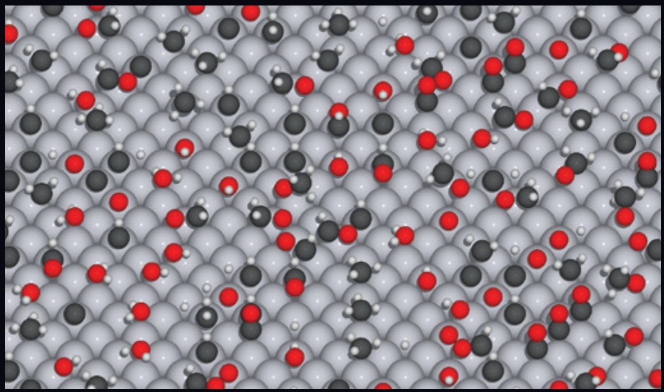
Extreme-Scale Cosmological Hydrodynamics

PI Katrin Heitmann
INST Argonne National Laboratory

Extreme-Scale Unstructured Adaptive CFD

PI Ken Jansen
INST University of Colorado Boulder

Data Projects



The catalysis project will combine data science techniques and quantum chemistry simulations to explore the otherwise intractable phase space resulting from gas phase molecules on catalyst surfaces to find relevant configurations and the lowest transition states between them. *Image: Eric Hermes, Sandia National Laboratories*

Exascale Computational Catalysis

PI David Bross
INST Argonne National Laboratory

Dark Sky Mining

PI Salman Habib
INST Argonne National Laboratory

Data Analytics and Machine Learning for Exascale Computational Fluid Dynamics

PI Ken Jansen
INST University of Colorado Boulder

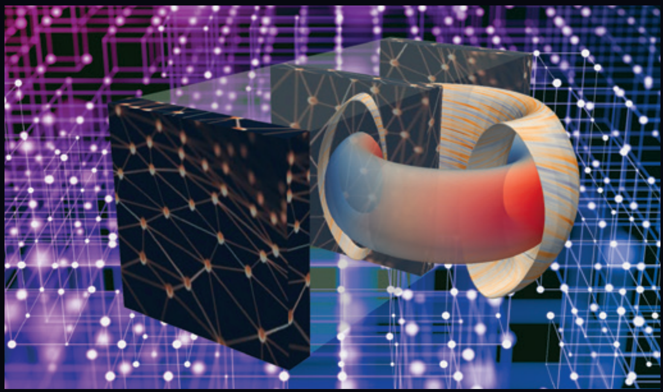
Simulating and Learning in the ATLAS Detector at the Exascale

PI Walter Hopkins*
INST Argonne National Laboratory
**The original PI Jimmy Proudfoot has retired.*

Extreme-Scale In-Situ Visualization and Analysis of Fluid-Structure-Interaction Simulations

PI Amanda Randles
INST Duke University and Oak Ridge National Laboratory

Learning Projects



The PPPL team's Fusion Recurrent Neural Network uses convolutional and recurrent neural network components to integrate both spatial and temporal information for predicting disruptions in tokamak plasmas. *Image: Julian Kates-Harbeck, Harvard University; Eliot Feibush, Princeton Plasma Physics Laboratory*

Machine Learning for Lattice Quantum Chromodynamics

PI William Detmold
INST Massachusetts Institute of Technology

Enabling Connectomics at Exascale to Facilitate Discoveries in Neuroscience

PI Nicola Ferrier
INST Argonne National Laboratory

Many-Body Perturbation Theory Meets Machine Learning to Discover Singlet Fission Materials

PI Noa Marom
INST Carnegie Mellon University

Virtual Drug Response Prediction

PI Rick Stevens
INST Argonne National Laboratory

Accelerated Deep Learning Discovery in Fusion Energy Science

PI William Tang
INST Princeton Plasma Physics Laboratory