Researchers are using the ALCF’s Theta supercomputer to advance our understanding of electronic stopping processes. This image is a snapshot from a non-equilibrium electron dynamics simulation of solvated DNA under proton irradiation. Image: Yosuke Kanai and Dillon C. Yost, University of North Carolina at Chapel Hill.
2019 Annual Report

02 Year in Review
04 ALCF Leadership
08 ALCF at a Glance

20 Preparing for Exascale
22 Next Stop: Exascale
26 Readying Aurora for Science on Day One

42 Expertise and Resources
46 ALCF Computing Systems
48 The ALCF Team

54 Science
56 Accessing ALCF Resources for Science
58 Advancing Design and Structure Prediction of Proteins and Peptides
   David Baker
59 FDTD Simulations to Facilitate Early-Stage Human Cancer Detection
   Allen Taflove
   Vadim Backman
60 Energy Exascale Earth System Model
   Mark Taylor
61 High-Fidelity Multiphysics Simulations to Improve Nuclear Reactor Safety and Economics
   Emily Shemon
62 Large-Eddy Simulation of Commercial Transport Aircraft Model
   Parviz Moin
   Sanjeeb Bose
63 Electronic Response to Particle Radiation in Condensed Matter
   André Schleife
64 Impact of Grain Boundary Defects on Hybrid Perovskite Solar Absorbers
   Wissam A. Saidi
65 Phase Transitions in Water-Ice-Vapor System
   Subramanian Sankaranarayanan
66 Predictive Simulations of Functional Materials
   Paul Kent
67 Astrophysical Particle Accelerators: Magnetic Reconnection and Turbulence
   Dmitri Uzdensky
68 Deep Learning at Scale for Multi-Messenger Astrophysics
   Eliu Huerta
69 Hadronic Light-by-Light Scattering Contribution to the Muon Anomalous Magnetic Moment
   Thomas Blum
70 High-Fidelity Gyrokinetic Simulation of Tokamak and ITER Edge Physics
   Chaongseok Chang
71 Kinetic Simulation of FRC Stability and Transport
   Sean Detrick
72 PLASM-IN-SILICO: HPC Modeling of High-Intensity Laser-Solid Interaction
   Jean-Luc Vay
74 ALCF Projects
79 ALCF Publications
The Argonne Leadership Computing Facility enables breakthroughs in science and engineering by providing supercomputing resources and expertise to the research community.
Microfluidics is a promising technique for detecting circulating tumor cells (CTCs) from patient blood samples. However, the transport and adhesion of CTCs in red blood cell (RBC) suspensions in microfluidics under flow is not fully understood. With a Director’s Discretionary allocation at the ALCF, researchers from Northern Illinois University are using high-performance computational models to study these processes. This image shows the interactions between CTCs (yellow), RBCs (red), microposts (blue cylinders), and hydrodynamics (streamlines). Image: Michael Hood and Jifu Tan, Northern Illinois University.
DIRECTOR’S MESSAGE

ALCF Leadership

MICHAEL E. PAPKA
ALCF Director
When the ALCF finally decommissioned its IBM Blue Gene/Q last December after seven years of reliable service, it had executed a total of 40 billion core-hours for scientists. Researchers had used Mira to make new advances in modeling everything from crystal plasticity to protein function to particle collision experiments. Once rated among the fastest and most energy efficient machines in the world, Mira had to make way for the ALCF’s incoming exascale system, Aurora.

Once it comes online in 2021, Aurora promises to be 100 times faster than both the now-retired Mira, and Theta, the ALCF’s current Intel/Cray resource. It will be a machine unlike any other in the world, an engineering marvel capable of concurrently supporting learning jobs and data analysis jobs, in addition to traditional simulation and modeling efforts. These new workloads hold the potential for an entirely new scale of discoveries.

The ongoing convergence of high-performance computing and artificial intelligence (AI) brings many exciting opportunities and challenges to the scientific community, and will continue to demand significant effort and coordination across our user community. I’m proud to say that both the ALCF and Argonne National Laboratory are already taking a lead role in this transformation, in part through the design and delivery of Aurora, a next-generation computing resource designed to support emerging AI and data workloads.

To that end, in 2019, the ALCF continued its many programs and preparations, some of them years underway, to ready a diversity of applications, and users, to take advantage of Aurora’s capabilities on day one. We have strengthened our services and frameworks, and implemented new ones. We have grown our training and outreach programs to reach a larger audience. And we have continued to share our world-class expertise with our colleagues in industry and the greater exascale computing community.

The ALCF continues to anticipate the requirements of its user community for producing breakthrough science in the coming decade by rethinking scientific workflows, optimizing large-scale simulation codes, and operating next-generation resources to fully exploit the increasing velocity of data.

We are also planning for the needs and work practices of an increasingly diverse user community. Our development of Petrel, in collaboration with the Globus team and the Data Science and Learning Division at Argonne, has led to a new way of handling data storage and sharing, and served as a prototype for the massive community filesystem we will deploy in 2021. In addition, the needs of our users to run unique and complex workflows led us to develop and deploy Balsam, a workflow manager that simplifies the execution of large collections of jobs with varying levels of concurrency and data dependencies. We have been working to leverage Balsam to integrate leadership computing for near-real-time experimental science at the Advanced Photon Source and at other laboratories within the DOE facility community.

With Aurora now fixed on the horizon, we remain as committed as ever to helping researchers deliver new scientific knowledge, like the game-changing findings you will read about in this Annual Report. The practical breakthroughs from the past year include a new “microscope in a computer” that can sense changes in large molecules in living cells—work that opens the door for more accurate and more efficient early-stage cancer screenings. Researchers also advanced a long-standing scientific challenge to control intrinsic magnetic properties in a single device—a capability that could be used to improve everything from computer memory to electric motors.

This report showcases the breadth of activities made possible only through the commitment of the ALCF’s highly talented staff. I am particularly eager to share how our people are delivering this transformation of computational science.
As we look toward the exascale era, we have placed a focus on identifying and cultivating the skills that our future workforce will require in a growing landscape of data science, machine learning, and artificial intelligence. Our participation in many outreach activities and events centered around these topics, as well as our engagement with broader conferences, is a testament to our effort in attracting future leaders in these areas.

With Argonne’s commitment to a diverse workforce as a backdrop, the ALCF continues to invest time and effort engaging with students ranging from high schoolers to postdoctoral candidates in order to build a pipeline for the next generation of computing experts. This past year, ALCF staff members organized high school coding and big data camps in collaboration with Argonne’s Educational Programs Office; co-chaired Argonne’s Introduce a Girl to Engineering Day for middle school girls; helped organize the lab’s Science Careers in Search of Women conference for high school girls; and organized and managed the Argonne Training Program for Extreme-Scale Computing, an annual event aimed at preparing early career researchers to use supercomputers for science.

Thinking about the future of computing provides the DOE national laboratories with a common interest in many areas, especially as we look to attract talent from a diverse pool of candidates. Collaborating with our sister labs to create a larger impact at conferences like the Richard Tapia Celebration of Diversity in Computing, the Grace Hopper Celebration, and Supercomputing (SC19) benefits computing across the DOE complex. As we plan out our strategic direction over the next several years, our engagements with students and historically underrepresented groups in the computing domain will help us get further in our commitment to creating a diverse, equitable, and inclusive work environment at Argonne and beyond.

Mira’s retirement on December 31, 2019, marked the end of an era at the ALCF. Our IBM Blue Gene/Q system was decommissioned after more than seven years and 40 billion core-hours of service to the scientific computing community. Over the course of its lifetime as a production supercomputer, Mira performed remarkably well, consistently exceeding key metrics for scheduled availability and utilization. Mira was also extremely reliable as evidenced by the number of full-machine runs carried out on the system. Our users routinely ran on 100 percent of Mira, performing more than 700 such runs over the past seven years.

In 2019, we launched a new ALCF user account and project management website, which brought many improvements, including a more user-friendly interface and enhanced security. We also began offering mobile tokens—a popular option that allows users to securely access ALCF systems via a smart phone app rather than logging in with a physical token.

We forged a partnership with Altair to adopt their PBS Professional workload manager for our next-generation systems. We will be working with the PBS Pro open source community to add some capabilities from our current Cobalt job scheduler, as well as some new features to support converged workloads in the exascale era.

We also selected a new global file system that will be deployed in 2020, providing an aggregate capacity of 200 PB and a transfer speed of 1.3 TB/s. The solution will be deployed as two filesystems, with one of them designed to allow our users to easily share data with collaborators outside of ALCF.
While Mira will be missed by ALCF users and staff alike, it was exciting to see the system close out its long and productive run on such a high note. For its final project, an Argonne-led team used the system for one of the most extensive cosmological simulations ever performed. With an allocation of nearly 800 million core-hours, the team, who helped launch Mira in 2012 with the massive Outer Rim simulation, harnessed the entire system to carry out a simulation that will form the basis for synthetic sky maps to be used by numerous cosmological surveys.

It is fitting that Mira—having been used to advance research in fields ranging from biochemistry and materials science to fluid dynamics and earth science—should continue to impact science through its final simulations and beyond. I cannot help but be amazed by the fact that, right before being powered down, this system was still capable of something so groundbreaking and expansive. The research community will be taking advantage of this work for a long time.

We also made major strides in our preparations for Aurora. The Aurora Early Science Program (ESP) continued to support work aimed at preparing key simulation, data, and learning applications for the scale and architecture of the system, ensuring that the research community can leverage its capabilities from the moment it enters production mode. Intel’s unveiling of several new Aurora architectural details, including a new category of general-purpose GPUs based on Intel’s Xe architecture, was extremely useful to the researchers participating in the ESP and DOE’s Exascale Computing Project. And the availability of development tools like Intel’s oneAPI programming environment and previous-generation Intel GPUs has allowed researchers to escalate their code development efforts for Aurora.

In 2019, we played a significant role in supporting various exascale projects, including the development of the Aurora hardware and software stack with Intel, and several activities within DOE’s Exascale Computing Project (ECP) and PathForward project.

With access to Intel Gen9 integrated GPUs and early versions of the Intel oneAPI software stack through Argonne’s Joint Laboratory for System Evaluation (JLSE), we worked closely with ECP application teams to port key kernels and project performance on Aurora’s Intel Xe architecture. We also held multiple hands-on workshops and webinars to help researchers prepare applications and software for our exascale system.

We continue to play an active role in software standards committees important to Aurora, including SYCL, OpenMP, and C++. Our team is engaged in porting programming models like Kokkos and Raja to Aurora, while adding new features to improve the usability of the Clang/LLVM compiler. In another effort to prepare for Aurora, and for the widening gap between compute and storage bandwidths, we have been working with ALCF users to enable in-situ visualization and analysis while the data is still in memory.

Our team also contributed several papers to key conferences this past year. For example, at SC19, we presented our work to enable near-real-time data analysis for DOE light sources using the ALCF-developed Balsam edge service; the development and deployment of DeepHyper, a scalable automated machine learning software package that has been leveraged by several ALCF projects; and a collaborative project with Cray and NERSC to develop a suite of portable benchmarks to quantitatively measure interconnect congestion and its management on HPC systems.
ALCF at a Glance

The Argonne Leadership Computing Facility (ALCF) is a U.S. Department of Energy (DOE) Office of Science User Facility that enables breakthroughs in science and engineering by providing supercomputing resources and expertise to the research community.

ALCF computing resources—available to researchers from academia, industry, and government agencies—support large-scale computing projects aimed at solving some of the world’s most complex and challenging scientific problems. Through awards of supercomputing time and support services, the ALCF enables its users to accelerate the pace of discovery and innovation across disciplines.

As a key player in the nation’s efforts to deliver future exascale computing capabilities, the ALCF is helping to advance scientific computing through a convergence of simulation, data science, and machine learning methods.

Supported by the DOE’s Advanced Scientific Computing Research (ASCR) program, the ALCF and its partner organization, the Oak Ridge Leadership Computing Facility, operate leadership-class supercomputers that are orders of magnitude more powerful than the systems typically used for open scientific research.

2019 BY THE NUMBERS

<table>
<thead>
<tr>
<th>Core-hours of compute time</th>
<th>8.7B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active projects*</td>
<td>413</td>
</tr>
<tr>
<td>Facility users*</td>
<td>1,246</td>
</tr>
</tbody>
</table>

Publications

| Publications | 275+ |

*Fiscal year 2019
2019 ALCF Users by Affiliation

- Academia: 535 users
- Government: 615 users
- Industry: 96 users

2019 U.S. ALCF Users by State

- 100+ Users: California, Illinois
- 01–10 Users: Alabama, Arkansas, Arizona, Connecticut, Delaware, Georgia, Idaho, Indiana, Kansas, Maryland, Minnesota, Missouri, Nevada, Ohio, Oregon, Rhode Island, South Carolina, Wisconsin
MIRA RETROSPECTIVE

The ALCF’s Mira supercomputer was decommissioned at the end of 2019, ending a long run of accelerating scientific discoveries and innovation.
Mira was the third fastest supercomputer in the world when it arrived in 2012.
Mira Retires After Years of Enabling Science

The IBM Blue Gene/Q system delivered 40 billion computing hours to more than 800 projects over the course of its lifetime.

Mira, the 10-petaflops IBM Blue Gene/Q supercomputer first booted up at the ALCF in 2012, was decommissioned at the end of 2019. Its work spanned seven-plus years and delivered 40 billion core-hours to more than 800 projects, solving nearly intractable problems in scientific fields ranging from biology to astrophysics.

Mira is the ALCF’s third and final system in the IBM Blue Gene architectural line, which began with the deployment of the 5.7-teraflops Blue Gene/L machine in 2005. That was followed by Intrepid, a 557-teraflops IBM Blue Gene/P system that served the scientific computing community from 2008 to 2013. When Mira came online in 2012, it was 20 times more powerful than Intrepid, giving researchers a tool that made it possible to perform simulations on unprecedented scales and create more accurate models of everything from combustion engines to blood flow. Mira was the pinnacle of the Blue Gene many-core architectural line, providing a combination of power and reliability that was unprecedented for its time.

When Mira debuted, it was ranked the third fastest system in the world on the TOP500 list and topped the Green500 list, which recognizes the most energy-efficient supercomputers. Seven years later, in 2019, Mira remained among the most powerful systems available for open science, sitting at number 22 on the most recent TOP500 list. Mira also remained the third ranked system on the Graph 500 list, a measure focused on a supercomputer’s ability to handle data-intensive applications.

One of the necessary advances that made Mira so energy-efficient involved directly cooling the machine with pipes carrying water instead of blowing air over the chips. This allowed the supercomputer to remove heat from the chips more quickly than a traditional air-cooling system.

To enable Mira to sink its teeth into the most challenging problems possible, its designers needed to rethink what a supercomputer should look like. Previous supercomputers were built with progressively more powerful processors, but eventually engineers hit a limit on how many transistors they could fit on an individual core. The answer came in the form of the IBM Blue Gene architecture that eventually resulted in Mira and fit 16 cores on a single node.

Each of Mira’s nearly 50,000 nodes functions like a nerve cell, relaying information at the speed of light, through fiber optic cables, to other parts of the machine. Getting these connections organized in the optimal configuration to reduce the time it takes for different parts of the computer to exchange information represents a fundamental challenge of designing a supercomputer.

A big part of what made Mira so remarkably effective at solving complex science challenges is how efficiently the machine was able to communicate across its nodes.
Even when different simulations were running simultaneously on different parts of the system, Mira was able to eliminate communication interference by isolating the traffic for each job.

Mira’s fiber optic network geometry is called the interconnect, routing like an interstate highway system the signals coming from each node. The complexity of the interconnect results from the introduction of extra dimensions that shrink the total space that signals need to cover. Previous versions of the Blue Gene architecture had simpler interconnects, but Mira’s was a defining accomplishment.

Another remarkable aspect of Mira is the number and variety of researchers who were able to take advantage of the system’s full computational horsepower with massive simulations that required using all of its nodes. Many supercomputers only do these full-machine runs once early in their lives and never again, but ALCF users carried them out routinely on Mira. Over the course of its lifespan, researchers performed more than 700 full-machine runs on Mira for studies ranging from cosmology to materials science.

Now that Mira has been decommissioned, the ALCF’s current leadership-class supercomputer, Theta, is serving as the facility’s primary system for open science until its forthcoming exascale machine, Aurora, arrives in 2021.
Core-hours of Compute Time

40B

Projects Supported

804

System Users

1,368

Full-Machine Runs

The number of jobs that utilized 100 percent of Mira’s nodes.

734

Scheduled Availability

The percentage of time a system is available to users (the target for leadership-class systems is 90 percent).

99%

Utilization

The percentage of time a system’s nodes run user jobs (80 percent is considered acceptable for leadership-class systems).

90%

Mean Time to Failure (MTTF)

MTTF is a measure of days, on average, to any unscheduled system outage. The time between failures increased over the course of Mira’s lifetime, illustrating one aspect of the system’s exceptional reliability.

Note: Mira numbers are from 2013-2019.
Mira’s Enduring Legacy of Scientific Computing Advances

To commemorate Mira’s lasting impact on scientific computing, we take a look back at some of the notable research advances made possible by the ALCF’s 10-petaflops IBM Blue Gene/Q supercomputer.

Improving the Design of Wind Turbines and Jet Engines

PI: Umesh Paliath
INST: GE Global Research

Scientists from GE Global Research used Mira to study the complex behavior of air as it passes through jet exhaust nozzles and over wind turbine blades. The team performed large eddy simulations to accurately characterize some of the key flow physics of multiscale turbulent mixing phenomena, measuring variables such as velocity, temperature, and pressure. GE uses information derived from these simulations to inform the design of quieter, more efficient wind turbines and jet engines.

Developing New Materials to Combat Friction

PI: Subramanian Sankaranarayanan
INST: Argonne National Laboratory

A multidisciplinary team from Argonne used Mira to identify and improve a new mechanism for eliminating friction. Their simulations led to the synthesis of a novel material that demonstrated superlubricity (near zero friction) at the macroscale for the first time. Limited to dry environments, potential applications for the material include wind turbine gears and magnetic disc drives in computers. In a follow-up study using the Mira simulations, the researchers also developed a self-generating dry lubricant that can re-adjust itself continuously, enabling it to last longer than traditional dry lubricant materials. The team’s research, which demonstrates how computation can advance the design and discovery of new materials, has resulted in multiple patents for the friction-reducing materials.
Enabling 3D Supernovae Simulations

PI  Sean Couch
INST  Michigan State University

A research team from Michigan State University has leveraged Mira to carry out some of the largest and most detailed 3D simulations ever performed of core-collapse supernovae. The emergence of petascale supercomputers like Mira made it possible to include high-fidelity treatments of rotation, magnetic fields, and other complex physics processes that were not feasible in previous supernova simulations. The team’s 3D simulations on Mira helped advance the study of supernovae, shedding new light on the physical mechanisms that drive these stellar explosions.

Improving the Performance and Production of Concrete

PI  William George
INST  National Institute of Standards and Technology

Concrete is the most widely used building material in the world, but the production of its ingredients, namely cement, is a significant contributor to the world’s greenhouse gas emissions. To enable the design of new, more sustainable mixtures of concrete, researchers from the National Institute of Standards and Technology (NIST) performed simulations on Mira to better understand concrete’s flow properties. The NIST team used their simulation results to create Standard Reference Materials for industrial researchers to calibrate concrete rheometers—instruments used to measure the flow of complex fluids—for materials development. Knowledge gained from the NIST simulations is critical to helping researchers expand the use of alternative materials and zero in on the optimal recipes for producing more sustainable concrete.

Advancing Microscopy Technique for Early Cancer Detection

PI  Allen Taflove and Vadim Backman
INST  Northwestern University

Researchers from Northwestern University tapped Mira to advance the development of an optical microscopy technique that can predict and quantify cancer risks at extremely early stages. The team’s technique, known as Partial Wave Spectroscopic (PWS) microscopy, is capable of detecting static intracellular nano-architectural alterations that are not accessible using conventional optical microscopy. Leveraging Mira, they performed simulations to inform the development of a new dynamic PWS technique that can detect macromolecular motions in living cells.

Shedding Light on the Dark Universe

PI  Salman Habib and Katrin Heitmann
INST  Argonne National Laboratory

An Argonne-led team used Mira to carry out some of the largest and most detailed simulations ever performed of the evolution of the universe. Their work resulted in the first simulations accurate enough to compare with large-scale cosmological surveys. The team’s simulations enabled the creation of cosmological maps that help improve the analysis of data from some of world’s largest astronomical surveys, such as the Dark Energy Survey. These synthetic sky maps allow scientists to advance studies of the gravitational forces at work on large-scale structures in the universe and, by extension, some of cosmology’s greatest mysteries, including dark energy and dark matter.
Simulating Particle Collision Events for the Large Hadron Collider

PI        Tom LeCompte and Taylor Childers
INST       Argonne National Laboratory

Particle collision experiments at CERN’s Large Hadron Collider (LHC) generate massive amounts of data that must be processed and analyzed to aid in the facility’s search for new physics discoveries. To help tackle this big data challenge, an Argonne-led team used Mira to simulate LHC collision events, marking the first time a leadership-class system was used for such simulations. Mira enabled the simulation of millions of particle collision events in parallel, while freeing the LHC Computing Grid to run other, less compute-intensive jobs. The team’s work demonstrated how DOE supercomputers can play a role in meeting the growing computing demands of the LHC and other data-intensive high energy physics experiments.

Informing the Design of Next-Generation Nuclear Reactors

PI        Emily Shemon and Aleksandr Obabko
INST       Argonne National Laboratory

Fast reactors are a promising nuclear energy technology that can substantially reduce radioactive waste and efficiently utilize natural nuclear resources. To help accelerate the design of safe and economical fast reactors, an Argonne-led research team has used Mira to model and analyze new reactor designs through high-fidelity multiphysics simulations. The end goal of nuclear modeling and simulation efforts on Mira and other DOE supercomputers is to remove some of the initial obstacles that the nuclear industry faces as it considers the design, licensing, and deployment of next-generation reactors.

Modeling Planetary and Stellar Magnetic Field Generation

PI        Jonathan Aurnou
INST       University of California, Los Angeles

Using Mira, a multi-institutional team led by the University of California, Los Angeles, developed high-resolution 3D models to study magnetic field generation on Earth, Jupiter, and the Sun at unprecedented levels of detail. To shed light on the dynamics that produce magnetic fields, the researchers performed massive simulations to model the interplay of magnetism, rotation, and turbulent convection occurring within the interiors of planets and stars. Mira allowed the team to resolve a range of spatial scales that were previously inaccessible with numerical simulations. A better understanding of the magnetic field generation process is critical to gaining new insights into the birth and evolution of the solar system.
Accelerating the Design of More Efficient Engines

PI       Sibendu Som
INST     Argonne National Laboratory

With access to Mira, Argonne researchers have worked with the automotive industry to develop powerful engine modeling and simulation capabilities that can help optimize complex fuel spray and combustion processes for a variety of fuels over a wide range of operating conditions. In one case, Argonne worked with a global petroleum and natural gas company to reduce engine design time from months to weeks by using Mira to simulate more than 2,000 engine-fuel combinations. Their work resulted in two optimized fuel-engine concepts with the potential to substantially improve fuel efficiency.

Advancing Protein and Peptide Design

PI       David Baker
INST     University of Washington

A research team from the University of Washington used Mira to develop and apply new computational methods aimed at improving protein structure prediction and design capabilities. Their work has included efforts to design mini-proteins called peptides for a wide range of therapeutic targets, including influenza, Ebola, HIV, and Alzheimer’s disease. After using the supercomputer to identify the most promising candidates for a given pathogen, the research team synthesized various proteins in a laboratory to experimentally validate their simulation results.

Taking QMC Simulations to New Heights

PI       Anouar Benali
INST     Argonne National Laboratory

Mira played an instrumental role in making the quantum Monte Carlo (QMC) method a more effective and accessible tool for the computational materials science community. Previously limited by a lack of sufficient computing power, the computationally demanding QMC method is capable of providing realistic predictions of materials properties that elude traditional methods like density functional theory. Mira’s massively parallel architecture was particularly well suited for the QMC approach, providing a launching pad for its application to a wide variety of materials. In one example, researchers from Argonne, Oak Ridge, Sandia, and Lawrence Livermore national laboratories used Mira to carry out QMC simulations that accurately captured the magnetic properties of the titanium oxide material, Ti$_4$O$_7$, for the first time.
PREPARING FOR EXASCALE

As the future home to Aurora, one of the nation’s first exascale systems, the ALCF is helping to lay the groundwork for the next generation of scientific computing.
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 Xe 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 Xe 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 Xe architecture.

Image: Intel Corporation

### Sustained Performance

≥1 Exaflop DP

### Delivery

CY 2021

### Platform

Cray Shasta

### Aggregate System Memory

>10 PB

#### Compute Node

2 Intel Xeon scalable “Sapphire Rapids” processors; 6 Xe arch-based GPUs; Unified Memory Architecture; 8 fabric endpoints

#### Software Stack

Cray Shasta software stack + Intel enhancements + data and learning

#### Cabinets

>100

#### GPU Architecture

Xe arch-based “Ponte Vecchio” GPU; Tile-based chiplets, HBM stack, Foveros 3D integration, 7nm

#### On-Node Interconnect

CPU-GPU: PCIe

GPU-GPU: Xe 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++
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

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

Simulation Projects

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 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

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 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

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
GROWING THE HPC COMMUNITY

As a leader in the HPC community, the ALCF is actively involved in efforts to broaden the impact of supercomputers and grow the body of researchers who can use them to advance science.
The ALCF's summer student program gives students an opportunity to tackle research projects that address issues at the forefront of scientific computing. Kavon Farvardin, a PhD student from the University of Chicago, spent his summer developing a real-time optimization tool called HALO (Wholly Adaptive LLVM Optimizer).
Driving Innovation for Industry

The ALCF’s industry partnerships help to strengthen the nation’s innovation infrastructure and expand the use of supercomputers for R&D efforts.

The ALCF’s Industry Partnerships Program is focused on growing the facility’s community of industry users by engaging prospective companies of all sizes, from start-ups to Fortune 500 corporations, that could benefit from leadership computing resources and collaborative opportunities with the ALCF and across Argonne.

Supercomputers like Theta and Mira enable companies to tackle problems that are too computationally demanding for traditional computing clusters. Access to ALCF systems gives them the ability to perform more complex simulations, achieve more accurate predictions, and create higher-fidelity models of everything from combustion engines to fusion energy devices.

The results allow companies to accelerate critical breakthroughs, verify uncertainties, and drastically reduce or eliminate the need to build multiple prototypes. Ultimately, ALCF-industry partnerships help to strengthen the nation’s innovation infrastructure and expand the use of HPC resources to advance research and development efforts.

In addition to gaining access to some of the world’s most powerful supercomputers, industry users also benefit from the ALCF’s expertise in simulation, data, and learning methods; application porting and optimization; performance tuning and scaling; and data analysis and visualization.

The ALCF has enhanced its industry outreach program by partnering with other Argonne user facilities and divisions, including the Technology Commercialization and Partnerships Division. This collaborative approach allows the ALCF to present a more complete picture of the laboratory’s resources and encourages additional collaborations. These efforts have resulted in broader engagements across the laboratory with a number of companies.

The ALCF is also actively involved in directing the DOE Exascale Computing Project’s Industry Council, an advisory group of senior executives from prominent U.S. companies interested in working with Argonne and other DOE laboratories to deploy exascale computing to improve their products and services.
The following project summaries illustrate how ALCF resources are helping companies to advance their R&D efforts.

### Advancing Fusion Energy Research

**PI**  Sean Detrick  
**INST**  TAE Technologies  
Researchers from TAE Technologies are performing simulations on Theta to accelerate their experimental research program aimed at developing a clean, commercially viable, fusion-based electricity generator. The team will use the simulation results to optimize a device for studying the confinement of energy with high plasma temperatures, and to inform the design of a future prototype reactor.

### Enabling the Design of More Efficient Engines

**PI**  Sibendu Som  
**INST**  Argonne National Laboratory  
Argonne researchers continue to work closely with the automotive industry to advance engine design through the use of ALCF supercomputing resources. In collaboration with companies like Convergent Science, Inc. and Aramco Research Center–Detroit, the research team is developing powerful engine modeling and simulation capabilities that can help optimize complex fuel spray and combustion processes for a variety of fuels over a wide range of operating conditions.

### Using Deep Learning for Industrial Design

**PI**  Rathakrishnan Bhaskaran  
**INST**  GE Global Research  
With an ALCF Data Science Program project, scientists from GE Global Research are leveraging machine learning and large datasets generated by wall-resolved large-eddy simulations to develop data-driven turbulence models with improved predictive accuracy. The researchers will apply this approach to turbomachinery, such as a wind turbine airfoil, demonstrating the impact that deep learning can have on industrial design processes for applications in power generation, aerospace, and other fields.

Visualization from a high-fidelity simulation of a heavy-duty engine fueled with a straight-run gasoline performed on ALCF computing resources by researchers from Argonne National Laboratory, Aramco Research Center–Detroit, and Convergent Science Inc. Image: Roberto Torelli, Argonne National Laboratory; Yuanjiang Pei, Aramco Research Center–Detroit
Shaping the Future of Supercomputing

ALCF researchers are at the forefront of numerous strategic activities that aim to push the boundaries of what’s possible with high-performance computing.

As home to some of the world’s most powerful computing resources, the ALCF breaks new ground with the development and deployment of each new supercomputer. ALCF staff members, collaborating with the researchers who use leadership-class systems to pursue scientific breakthroughs, are involved in the development and testing of new HPC hardware and software. This unique position the ALCF occupies affords the facility an important perspective on the trends, methods, and technologies that will define the future of supercomputing.

Leveraging this knowledge and expertise, ALCF researchers contribute to many forward-looking activities aimed at advancing the use of supercomputers for discovery and innovation.

These efforts include organizing workshops and meetings on topics like artificial intelligence and quantum computing; engagement in leading user groups and conferences; and contributions to the development of standards, benchmarks, and technologies that help propel continued improvements in supercomputing performance.
David Martin, manager of industry partnerships and outreach at the ALCF, provides an overview of ALCF computing resources at the 2019 HPC User Forum meeting hosted by Argonne National Laboratory.
Community Activities

**AI for Science Workshops**

This summer, Argonne put on two AI for Science workshops—one for laboratory researchers and another for students working at the lab. The workshops included a series of talks from Argonne AI experts highlighting the use of AI in their research and provided an opportunity for cross-disciplinary discussion and collaboration. Attendees were also able to participate in a full day of hands-on tutorial sessions that covered various AI tools and techniques. ALCF computer scientists served as presenters at both workshops, leading sessions on deep learning, reinforcement learning, autoencoders and generative adversarial models, and recurrent neural networks.

**Data Solution Task Force for DOE Light Sources**

With funding from DOE for a two-year pilot program, the five DOE light sources formed the Data Solution Task Force to demonstrate, build, and implement software, cyberinfrastructure, and algorithms that address universal needs for the facilities, ranging from enhanced data analysis capabilities to data storage and archival resources. As part of the task force, ALCF researchers are exploring the use of the Balsam edge service and workflow manager to connect light sources with DOE supercomputing facilities for real-time analysis of experimental data.

**DOE AI Town Hall Meetings**

ALCF staff members were among the many contributors to a series of four AI for Science town hall meetings convened by DOE laboratories. The town halls were aimed at collecting input on the opportunities and challenges facing the scientific community in the era of convergence of HPC and AI technologies, and the expected integration of large-scale simulation, advanced data analysis, data-driven predictive modeling, theory, and high-throughput experiments. An integrated, consolidated town hall report from the four town halls will highlight the most compelling problems where AI could have an impact and outline the requirements needed for research and facilities to realize these opportunities.

**Exascale Computing Project**

The ALCF is a key contributor to the DOE’s Exascale Computing Project (ECP), a multi-lab initiative aimed at accelerating the development of a capable exascale computing ecosystem. Several ALCF researchers are engaged in ECP-funded projects in application development, software development, and hardware technology. ALCF computing resources, particularly Theta, allow ECP research teams to pursue development work at a large scale. In the workforce development space, the ECP continues to fund the annual Argonne Training Program on Extreme-Scale Computing (ATPESC), which is organized and managed by ALCF staff.
HPC Standards, Benchmarks, and Technologies

ALCF staff members remain actively involved in the development of standards, benchmarks, and technologies that help drive continued improvements in supercomputing performance. Staff activities include contributions to the C++ Standards Committee, Cray User Group, HPC User Forum, Intel eXtreme Performance Users Group, MPI Forum, OpenMP Architecture Review Board, OpenMP Language Committee, and Open Scalable File Systems (OpenSFS) Board.

HPC User Forum Meeting

In September, Argonne National Laboratory hosted an HPC User Forum meeting that brought together users, vendors, and others interested in high-performance computing. The two-day event featured talks on the latest in HPC research, technologies, and initiatives. Argonne’s Rick Stevens, Valerie Taylor, and David Martin were among the featured speakers.

P3HPC Workshop

The annual Performance, Portability, and Productivity in HPC (P3HPC) workshop provides a forum for researchers and developers to discuss ideas and progress toward the goal of performance portability across current and future high performance computing platforms. ALCF staff members served on the program committee for the 2019 workshop in Denver, Colorado, and will organize the 2020 event.

Quantum Computing Training

In May, ALCF staff members organized a Quantum Computing Tutorial at Argonne National Laboratory. Designed to provide hands-on experience programming quantum simulators and computers, the tutorial gave participants instruction on how to program IBM quantum computers with Qiskit framework and how to use the Atos QLM35 quantum simulator housed in Argonne’s Joint Laboratory for System Evaluation.

SC19

Continuing Argonne’s long history of participation in the Supercomputing (SC) conference series, several researchers from the ALCF and the lab’s other computing divisions attended SC19 in Denver, Colorado. More than 90 Argonne staff members contributed to various conference activities, including technical paper presentations, invited talks, workshops, birds of a feather sessions, panel discussions, and tutorials, on topics ranging from exascale computing and big data analysis to AI and quantum computing. In addition, the ALCF organized an annual industry lunch at the conference, providing an opportunity for ALCF leadership to meet with industry users of HPC to discuss potential partnerships and strategies to improve their efficacy.
Engaging Current and Future HPC Users

The ALCF provides training and outreach opportunities that prepare researchers for efficient use of its leadership computing systems, while also cultivating a diverse and skilled HPC community for tomorrow.

Outreach to the facility’s user community is focused on providing expert-guided training for the growing array of computational resources, methods, and services available at the ALCF to support the advancement of scientific discovery.

In 2019, the ALCF hosted a number of events intended to optimize use of its leadership computing resources, including hands-on workshops aimed at improving the performance of simulation, data science, and machine learning applications; interactive videoconferences to connect users and developers from industry leaders such as Intel; and several training activities designed to prepare researchers for Aurora and science in the exascale era.

The ALCF also supports a wide variety of outreach activities directed at students, with staff members volunteering to engage participants. Multi-day camps centered around programming and big data visualization, as well as external events like Hour of Code, spark students’ interest in different aspects of scientific computing and introduce them to exciting career possibilities.

Additionally, the ALCF’s annual summer student program gives college students the opportunity to work side-by-side with staff members on real-world research projects and utilize some of the world’s most powerful supercomputers, collaborating in areas like computational science, system administration, and data science.
The ALCF hosts a variety of training events designed to educate users on the tools, systems, and frameworks that are available to accelerate their scientific computing projects.
**ALCF Computational Performance Workshop**

Held each spring, the annual ALCF Computational Performance Workshop is designed to help users boost code performance on ALCF systems. The three-day, hands-on workshop connects current and prospective users with the experts who know the systems best—ALCF computational scientists, performance engineers, data scientists, and visualization experts, as well as invited guests from Intel, ParaTools Inc. (TAU), and Rice University (HPCToolkit). With dedicated access to ALCF computing resources, the workshop provides an opportunity for attendees to work directly with these experts to test, debug, and optimize their applications on leadership-class supercomputers. One of the ALCF workshop’s primary goals is to help researchers demonstrate code scalability for INCITE, ALCC, and ADSP project proposals, which are required to convey both scientific merit and computational readiness.

**ALCF Developer Sessions**

The facility continued to host ALCF Developer Sessions, a monthly webinar series aimed at training researchers and increasing the dialogue between HPC users and the developers of leadership-class systems and software. Speakers in the series included developers from Intel, ARM, Globus, and Argonne, covering topics such as deep learning frameworks, cross-platform performance engineering, research data management, and in situ visualization and analysis.

**ATPESC 2019**

The annual Argonne Training Program on Extreme-Scale Computing (ATPESC) provides intensive hands-on training on the key skills, approaches, and tools needed to carry out research on current supercomputers and the exascale systems of the future. Organized by ALCF staff and funded by DOE’s Exascale Computing Project, ATPESC has a core curriculum that covers computer architectures; programming methodologies; data-intensive computing and I/O; numerical algorithms and mathematical software; performance and debugging tools; software productivity; data analysis and visualization; and machine learning and data science. Renowned scientists and HPC experts serve as lecturers and guide the hands-on laboratory sessions. The program, which was held at the Q Center in St. Charles, Illinois, this summer, welcomed 73 participants in 2019. With seven years now in the books, ATPESC has hosted nearly 500 participants since its inception. To extend ATPESC’s reach, videos of each lecture are captured and posted to Argonne’s YouTube channel.

**Aurora Hackathons**

The Argonne-Intel Center of Excellence began hosting a series of hackathons to help individual Aurora Early Science Program (ESP) teams port and optimize their applications for the forthcoming exascale system. The intensive, hands-on sessions pair the ESP teams with experts from the ALCF and Intel to advance code development efforts using the Aurora software development kit, early hardware, and other exascale programming tools.
Aurora Webinars

The ALCF continued to host a series of web-based training events exclusively for Aurora ESP project teams. The interactive webinars are designed to provide instruction and hands-on exercises on topics and tools relevant to current production machines, with an emphasis on data intensive and machine learning subjects. Webinar topics included the Intel Advisor and VTune Amplifier tools, OpenMP 4.5/5.0, machine learning frameworks, and SYCL/DPC++.

Aurora Workshops

In 2019, the ALCF hosted two hands-on workshops to help researchers ramp up their efforts to prepare for the forthcoming Aurora system. For the first event in April, the Argonne-Intel Center of Excellence hosted a three-day workshop to familiarize ESP projects teams with the Aurora hardware and software environment. In September, the ALCF held the Aurora Programming Workshop, a three-day event focused on application and software development for the exascale machine. Open to members of the ESP and DOE’s Exascale Computing Project (ECP), the workshop covered programming models, allowed attendees to share lessons learned and best practices, and provided hands-on sessions with the Aurora software development kit and other available tools.

Best Practices for HPC Software Developers

In 2019, the ALCF, OLCF, NERSC, and the Exascale Computing Project continued their collaboration with the Interoperable Design of Extreme-Scale Application Software (IDEAS) project to deliver a series of webinars—Best Practices for HPC Software Developers—to help users of HPC systems carry out their software development more productively. Webinar topics included containers in HPC, agile software practices, modern C++, and tools and techniques for floating-point analysis.

INCITE Proposal Writing Webinars

The Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program is one of the primary allocation programs providing access to DOE’s leadership computing facilities. Each year, the ALCF collaborates with the INCITE program and the OLCF to offer proposal writing webinars that provide attendees with an opportunity to ask questions and receive guidance on preparing an effective proposal.

Simulation, Data, and Learning Workshop

In October, the ALCF hosted its annual Simulation, Data, and Learning Workshop to help users improve the performance and productivity of simulation, data science, and machine learning applications on ALCF systems. Attendees worked directly with ALCF staff and industry experts, learning how to use data science tools and techniques at scale on ALCF systems. Topics covered included machine learning and deep learning frameworks (e.g., TensorFlow, Horovod, PyTorch, and DeepHyper), as well as workflow management services and containers.
Inspiring Students

**Big Data Camp**

In July, Argonne hosted a group of local high school students for its annual Big Data Camp. Led by ALCF Director Michael Papka and several ALCF colleagues and organized by Argonne’s Educational Programs Office, the five-day camp is designed to teach attendees to hone their analysis skills and give them a first-hand look at what it is like to be a data scientist. This year’s event welcomed a special guest, as U.S. Representative Bill Foster stopped by to kick off an activity in which the students had to use data analysis and visualization techniques to determine the cause of an 1854 cholera outbreak.

**CodeGirls@Argonne Camp**

The CodeGirls@Argonne Camp welcomes sixth- and seventh-grade girls to the laboratory each summer for a three-day event dedicated to teaching them the fundamentals of coding. The attendees get an opportunity to try out creative and computational thinking through activities that include programming robots. The camp also allows participants to meet and interview Argonne scientists and tour the ALCF’s machine room and visualization lab.

**CAPS High School Computing Workshop**

Over five days in July, ALCF staff members helped teach and mentor a group of local high school students at Argonne’s CAPS (Creative Approaches to Problems in Science) High School Computing Workshop. The camp curriculum promotes problem-solving and teamwork skills through interactive coding activities, including working with the Python language and programming a robot via a Raspberry Pi. The annual workshop is a joint initiative of the ALCF and Argonne’s Educational Programs Office.
Hour of Code
As part of the national Computer Science Education Week (CSEdWeek) in December, several ALCF staff members visited local classrooms to give talks and demos intended to spark interest in computer science. Working with students in classes from elementary to high school, the volunteers led a variety of activities designed to impart the basics of coding. CSEdWeek was established by Congress in 2009 to raise awareness about the need for greater computer science education.

Summer Student Program
Every summer, the ALCF opens its doors to a new class of student researchers who work alongside staff mentors to tackle research projects that address issues at the forefront of scientific computing. This year, the facility hosted 29 students ranging from undergraduates to PhD candidates. From exploring the potential of Kubernetes for data-intensive workloads to implementing real-time auto tuning for program optimization, the interns had the opportunity to gain hands-on experience with some of the most advanced computing technologies in the world. The summer program culminated with a series of seminars in which the students presented their project results to the ALCF community.

Women in STEM
Through participation in annual Argonne events like Introduce a Girl to Engineering Day and the Science Careers in Search of Women conference, ALCF staff members have the opportunity to connect with young women and introduce them to potential career paths in science, technology, engineering, and mathematics (STEM). The ALCF also promotes STEM careers to women through contributions to Argonne’s Women in Science and Technology group, AnitaB.org, and the Grace Hopper Celebration of Women in Computing.
EXPERTISE AND RESOURCES

The ALCF’s unique combination of supercomputing resources and expertise enables researchers to accelerate the pace of scientific discovery and innovation.
Yasaman Ghadar, ALCF assistant computational scientist, helps manage and organize the facility’s user training events.
In 2019, Mira and Theta were the engines that drove scientific discoveries and engineering breakthroughs at the ALCF. At around 10 petaflops each, the facility’s two production systems have ranked among the fastest supercomputers in the world for open science.

Supporting systems—Cetus, Vesta, and Iota—are used for debugging and test and development work.

Cooley, the facility’s visualization cluster, helps transform computational data into high-resolution images, videos, and animations, helping users to better analyze and understand simulations produced by ALCF supercomputers.

The ALCF’s supercomputing environment also includes advanced data storage systems and networking capabilities.

Additionally, Argonne’s Joint Laboratory for System Evaluation (JLSE) maintains a range of leading-edge hardware and software environments to enable researchers to evaluate and assess next-generation platforms and technologies.

**Expertise and Support**

The ALCF’s true value lies in its people. With a team comprised of HPC system and network administrators, computational scientists, computer scientists, data scientists, performance engineers, visualization experts, software developers, and support staff, the ALCF ensures facility users are able to get the most out of its supercomputers.

The ALCF’s exceptional staff and powerful supercomputing resources make the facility one of the world’s premier centers for scientific computing.
Theta is the ALCF’s 11.69 petaflops Intel-Cray supercomputer.
## ALCF Computing Systems

### Mira
10-petaflops IBM Blue Gene/Q supercomputer

<table>
<thead>
<tr>
<th>IBM Blue Gene/Q architecture</th>
<th>16-core, 1.6-GHz IBM PowerPC A2 processor per node</th>
<th>768 TB of memory</th>
<th>5D torus interconnect</th>
<th>48 racks</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 petaflops</td>
<td>49,152 nodes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>786,432 cores</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Theta
11.69-petaflops Intel-Cray supercomputer

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

### Cetus
IBM Blue Gene/Q system used to offload both debugging issues and alternative production workloads from Mira

<table>
<thead>
<tr>
<th>IBM Blue Gene/Q architecture</th>
<th>16-core, 1.6-GHz IBM PowerPC A2 processor per node</th>
<th>64 TB of memory</th>
<th>5D torus interconnect</th>
<th>4 racks</th>
</tr>
</thead>
<tbody>
<tr>
<td>838 teraflops</td>
<td>4,096 nodes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>65,536 cores</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Iota
Intel-Cray test and development platform

<table>
<thead>
<tr>
<th>Intel-Cray XC40 architecture</th>
<th>64-core, 1.3-GHz Intel Xeon Phi 7230 processor per node</th>
<th>44 nodes</th>
<th>2,816 cores</th>
<th>12.3 TB of memory</th>
<th>1 TB of high-bandwidth memory</th>
<th>Aries interconnect with Dragonfly configuration</th>
<th>1 rack</th>
</tr>
</thead>
<tbody>
<tr>
<td>117 teraflops</td>
<td>44 nodes</td>
<td>2,816 cores</td>
<td>12.3 TB of memory</td>
<td>1 TB of high-bandwidth memory</td>
<td>Aries interconnect with Dragonfly configuration</td>
<td>1 rack</td>
<td></td>
</tr>
<tr>
<td>64-core, 1.3-GHz Intel Xeon Phi 7230 processor per node</td>
<td>44 nodes</td>
<td>2,816 cores</td>
<td>12.3 TB of memory</td>
<td>1 TB of high-bandwidth memory</td>
<td>Aries interconnect with Dragonfly configuration</td>
<td>1 rack</td>
<td></td>
</tr>
</tbody>
</table>

### Vesta
IBM Blue Gene/Q test and development platform

<table>
<thead>
<tr>
<th>IBM Blue Gene/Q architecture</th>
<th>16-core, 1.6-GHz IBM PowerPC A2 processor per node</th>
<th>32 TB of memory</th>
<th>5D torus interconnect</th>
<th>2 racks</th>
</tr>
</thead>
<tbody>
<tr>
<td>419 teraflops</td>
<td>2,048 nodes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32,768 cores</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Cooley
Data analysis and visualization cluster

<table>
<thead>
<tr>
<th>Intel Haswell architecture</th>
<th>1 NVIDIA Tesla K80 GPU per node</th>
<th>47 TB of memory</th>
<th>3 TB of GPU memory</th>
<th>FDR InfiniBand interconnect</th>
<th>6 racks</th>
</tr>
</thead>
<tbody>
<tr>
<td>293 teraflops</td>
<td>126 nodes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two 6-core, 2.4-GHz Intel E5–2620 processors per node</td>
<td>1,512 cores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Supporting Resources

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

DISK STORAGE
The Mira system consisted of 384 I/O nodes that connected to 22 storage arrays that control 13,000 disk drives with a total useable capacity of 27 PB and a maximum aggregate transfer speed of 330 GB/s over two file systems. Mira used the General Parallel File System (GPFS) to access the storage. The Theta 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 file system 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.

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
The Mira and Theta systems each have 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
Through Argonne’s Joint Laboratory for System Evaluation (JLSE), the ALCF provides access to 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. JLSE testbeds include:

- Intel DAOS nodes (DCPMM and NVMe storage) with Intel DAOS file system v0.9
- Intel Xeon Phi Knights Landing Cluster
- NVIDIA DGX–1
- Intel Xeon Platinum Skylake Cluster
- Intel Iris Gen9 nodes (integrated GPUs) with Intel oneAPI (public beta)
- HPE Comanche Prototype ARM64 Cluster
- IBM Power System AC922
- Supermicro NVIDIA V100 and P100 cluster
- Atos Quantum Learning Machine
- Kubernetes Cluster with Rancher
- Petrel v3 (Ceph 3.2 PB filesystem)
The ALCF Team

**Operations**
The ALCF’s HPC systems administrators manage and support all ALCF computing systems, ensuring users have stable, secure, and highly available resources to pursue their scientific goals. This includes the ALCF’s production supercomputers, supporting system environments, storage systems, and network infrastructure. The team’s software developers create tools to support the ALCF computing environment, including software for user account and project management, job failure analysis, and job scheduling. User assistance specialists provide technical support to ALCF users and manage the workflows for user accounts and projects. In the business intelligence space, staff data architects assimilate and verify ALCF data to ensure accurate reporting of facility information.

**Science**
Computational scientists with multidisciplinary domain expertise work directly with ALCF users to maximize and accelerate their research efforts. In addition, the ALCF team applies broad expertise in data science, machine learning, data visualization and analysis, and mathematics to help application teams leverage ALCF resources to pursue data-driven discoveries.

With a deep knowledge of the ALCF computing environment and experience with a wide range of numerical methods, programming models, and computational approaches, staff scientists and performance engineers help researchers optimize the performance and productivity of simulation, data, and learning applications on ALCF systems.

**Technology**
The ALCF team plays a key role in designing and validating the facility’s next-generation supercomputers. By collaborating with compute vendors and the performance tools community, staff members ensure the requisite programming models, tools, debuggers, and libraries are available on ALCF platforms. The team also helps manage Argonne’s Joint Laboratory for System Evaluation, which houses next-generation testbeds that enable researchers to explore and prepare for emerging computing technologies.

ALCF computer scientists, performance engineers, and software engineers develop and optimize new tools and capabilities to facilitate science on the facility’s current and future computing resources. This includes the deployment of scalable machine learning frameworks, in-situ visualization and analysis capabilities, data management services, workflow packages, and container technologies. In addition, the ALCF team is actively involved in programming language standardization efforts and contributes to cross-platform libraries to further enable the portability of HPC applications.

**Outreach**
ALCF staff members organize and participate in training events that prepare researchers for efficient use of leadership computing systems. They also participate in a wide variety of educational activities aimed at cultivating a diverse and skilled HPC community for the future. In addition, staff outreach efforts include facilitating partnerships with industry and communicating the impactful research enabled by ALCF resources to external audiences.
Lenard Recognized as IBM Champion
Ben Lenard, ALCF HPC systems and database administrator, was selected to be part of the 2019 class of IBM Champions, a program that commends external experts who influence and mentor others to help them innovate with IBM technologies. Lenard was recognized for his efforts to amplify the use of IBM’s Data Server Manager and its Db2 database for researchers at the ALCF and beyond.

Salim Receives Best Workshop Paper
Misha Salim, ALCF computational scientist, received the Best Presentation Award at SC19’s XLOOP Workshop for “Balsam: Near Real-Time Experimental Data Analysis on Supercomputers.” The presentation covered how ALCF researchers used the Balsam edge service to enable near-real-time analysis of data collected at Argonne’s Advanced Photon Source (APS) via an x-ray photon correlation spectroscopy (XPCS) application running on the facility’s supercomputing resources.

Anisimov Co-Authors Book on Fast Multipole Method
Victor Anisimov, ALCF computational scientist, co-wrote a book titled, “Introduction to the Fast Multipole Method.” The book provides a user-friendly exploration of the theory and computer implementation of the fast multiple method (FMM), an algorithm that enables simulations in computational biophysics, chemistry, and materials science to deal with larger and more realistic models.

ALCF Staff Members Contribute to SC19 Technology Challenge Award
Several ALCF staff members contributed to an extensive collaboration that won SC19’s inaugural SCinet Technology Challenge for the “Real-Time Analysis of Streaming Synchrotron Data” demo. The Argonne-led team, which included members from Northwestern University, the StarLight Communications Exchange, Northern Illinois University, the University of Chicago, and the Metropolitan Research and Education Network, streamed experimental data from Argonne’s Advanced Photon Source to the ALCF for real-time image reconstruction. The researchers then used machine learning to enhance and de-noise the image data and streamed it back to the StarLight booth at SC19 for volume visualization. As part of the effort, the ALCF operations team played a key role in diagnosing and repairing an unexpected network failure that arose shortly before the scheduled demonstration.

Argonne Researchers Receive HPCwire Award
Researchers from Argonne National Laboratory, in collaboration with Convergent Science and Aramco Research Center-Detroit, were recognized with a 2019 HPCWire award for the “Best Use of HPC in Automotive.” With help from ALCF researchers, the team evaluated a new gasoline compression ignition by using Mira to analyze thousands of high-fidelity engine designs in days rather than months.
Staff Spotlights

The term “leadership” does not just apply to the ALCF’s leadership computing resources. It also applies to the people working to support facility users and maintain ALCF systems.

With a shared passion for research and innovation, ALCF staff members are helping to shape the future of supercomputing. The following pages highlight six staff members and some of their notable contributions in 2019.

Johannes is a postdoctoral appointee at the ALCF. An open-source developer in the LLVM compiler framework, he helps design and implement all OpenMP-related components in various runtimes and to design and implement optimizations for parallel programs. He also prototypes new features proposed for the OpenMP language so as to assess design choices and feasibility, and additionally helps mentor new contributors to the LLVM framework as well as any researchers who want to add their work to the production compiler. As a member of the OpenMP language committee, Johannes designs new language features and extensions, helps to generate outside proposals for extensions beneficial to HPC users, and prevents obtuse or excessively narrow features from standard integration.

On top of these tasks, Johannes remains engaged with greater student and research communities. He has twice acted as a mentor at the Afro-Academic, Cultural, Technological, and Scientific Olympics and last summer supervised three students for Google’s Summer of Code. In 2019, he organized the ISC High Performance conference’s first-ever LLVM tutorial and birds-of-a-feather session, as well as three consecutive LLVM Performance Workshops at the annual International Symposium on Code Generation and Optimization. For these workshops he also researches new applications for compilers so as to assist users’ efforts to better understand the theoretical performance limits of their algorithms and help them achieve speedups.
As co-lead for the catalyst team, Chris helps coordinate supporting and enabling science teams to complete high-impact, computationally intensive research campaigns on ALCF resources. Chris’ recent work supporting ALCF projects led by fellow Argonne collaborators resulted in a Nature Communications paper comparing experimental and simulated data to better understand the underlying processes of ultrafast imaging applications of intense x-ray free-electron laser pulses.

Chris is also co-lead for the Aurora Applications Working Group, helping to coordinate multiple teams from DOE’s Exascale Computing Project (ECP) and the ALCF’s Early Science Program to prepare their codes for Aurora. He also helped co-organize the first Aurora Programming Workshop this fall to introduce Aurora’s hardware and software to ECP application teams and help them begin their porting efforts. Chris’ interests include exploration of programming models and performance of code across different hardware and how best to provide guidance to software teams making their first venture into heterogeneous computing as they prepare for exascale.

JaeHyuk is a member of performance engineering group at the ALCF. Having used HPC extensively since 2008, he developed computational fluid dynamics codes for complex fluids during his PhD program at University of Illinois at Urbana-Champaign. In his postdoctoral training, he applied his HPC code to fluid-structure interactions, successfully resolving issues on an offshore oil platform. He subsequently joined the National Center for Supercomputing Applications as a research programmer and collaborated with scientists and engineers to tackle real-world problems for three years.

Since joining Argonne in 2018, JaeHyuk has worked on performance tools, math libraries, and programming models for Argonne’s upcoming exascale system, Aurora. Utilizing one of these programming models, he presented at SC19 a method for achieving a theoretical peak speedup on GPUs over CPUs with kernels from the quantum-chemistry software package GAMESS. In his day-to-day work, he collaborates with users and staff to measure the efficiency of their applications on multiple architectures via the Roofline performance model. He’s presented some of these results at such meetings as the 2019 Cray User Group Conference and Arm HPC User Group. JaeHyuk stays active in the OpenMP community by organizing and presenting webinars, tutorials, and birds-of-a-feather sessions for scientists and HPC vendors. Finally, he collaborates with the Exascale Computing Project’s ExaWind team to help maximize their scientific output as they leverage ALCF resources to predict the complex flow physics of wind farms.
Bethany joined the ALCF in November 2018. As a member of the data science team, she works with domain experts in structural dynamics, materials science, systems, and computational fluid dynamics to apply machine-learning-based solutions to their problems. Among these collaborations are efforts with researchers at the University of Washington to use machine learning to study dynamical systems; this work led to a paper in the *Proceedings of the National Academy of Sciences (PNAS)*, describing deep-learning method to learn governing equations from data. Bethany has also collaborated with Michigan Tech researchers to develop a deep-learning method for monitoring and detecting damage in infrastructure via video.

Bethany has helped to prepare the facility for Aurora by supporting engineering groups focused on the development of data-science software. She is also supervising the ALCF’s Margaret Butler postdoctoral fellow, Romit Maulik, in his studies of the use of machine learning in computational fluid dynamics, as well as the integration of simulation and AI. Over the summer she mentored a summer intern in visualizing Theta logs so as to better understand performance failures and improve reliability. Her outreach efforts have led her to deliver talks on AI at the Argonne Community Leaders Roundtable and on scientific machine learning at the Creative Approaches to Problems in Science Computing Workshop.

Arriving at the ALCF as a postdoc in 2013, Silvio has been an assistant computer scientist with the division since 2016. As a member of the visualization and analysis team, he supports ALCF users and helps them understand and extract scientific knowledge from large-scale simulation data. The visualizations his team produces using ALCF resources such as Theta and Cooley have been featured on the ALCF website and in its publications, displayed at the DOE’s booth at the SC conference (where they are regular finalists in the SC Scientific Visualization Showcase), and printed on the covers of prestigious scientific journals.

Silvio’s responsibilities at the ALCF also include serving as a member of the Scheduling Committee and assisting INCITE and ALCC users with their scientific visualization needs. As part of his outreach efforts, he teaches visualization methods to high school students at summer camps in coordination with Argonne’s Educational Programs Office, he also participates in ALCF training events and regularly provides tutorials on in-situ technologies at the SC and ISC conferences. He also serves in the IEEE Vis Conference Committee. In 2019, he chaired the SciVis Contest at the IEEE Vis Conference in Vancouver, Canada.

Silvio is intimately involved with in-situ visualization and exascale architectures. He collaborates with vendors to develop the highest quality visualization tools for Aurora, serves as the Argonne PI for ALPINE within the Exascale Computing Project to deliver in-situ algorithms and infrastructure for exascale machines, and is part of the SENSEI scalable in situ analysis and visualization multi-institutional team.
In his role as a systems integration administrator, Kyle is responsible for continuous integration (CI) at the ALCF. This means that he maintains and supports existing Jenkins and Gitlab resources for ALCF users and members of the Operations team. Since joining the facility, Kyle’s main focus has been on the Exascale Computing Project, working on a multi-site federated job run model and a custom Gitlab CI environment intended to support HPC workflows.

With deployment of CI set for ALCF in the next year, Kyle is currently preparing for productionization. As such, on a day-to-day basis he is occupied with solution design and the testing of custom development efforts to integrate CI across HPC resources, user accounts, and Gitlab Runner administration. He also collaborates with local science teams to help leverage the ECP CI solution in their work.
The ALCF is accelerating scientific discoveries in many disciplines, ranging from chemistry and engineering to physics and materials science.
A snapshot of silicene (shown in yellow), a 2D material made up of silicon atoms, as it grows on iridium substrate (shown in blue). The image was taken from a molecular dynamics simulation, which Argonne researchers used to predict the growth and evolution of silicene. Image: Henry Chan, Mathew Cherukara, Badri Narayanan, and Subramanian Sankaranarayanan, Argonne National Laboratory.
Accessing ALCF Resources for Science

As a national user facility dedicated to open science, any researcher in the world with a large-scale computing problem can apply for time on ALCF computing resources.

Researchers gain access to ALCF systems for computational science and engineering projects—typically with awards of millions of core-hours—through competitive, peer-reviewed allocations programs supported by the DOE and Argonne.

The ALCF also hosts competitive, peer-reviewed application programs designed to prepare key scientific applications and innovative computational methods for the architecture and scale of ALCF supercomputers.

Allocation Programs

INCITE
The Innovative Novel Computational Impact on Theory and Experiment (INCITE) program aims to accelerate scientific discoveries and technological innovations by awarding ALCF computing time and resources to large-scale, computationally intensive projects that address grand challenges in science and engineering.

ALCC
The ASCR Leadership Computing Challenge (ALCC) program allocates ALCF computing resources to projects that advance the DOE mission, help to broaden the community of researchers capable of using leadership computing resources, and serve the national interests for scientific discovery, technological innovation, and economic competitiveness.

DD
Director’s Discretionary (DD) projects are dedicated to leadership computing preparation, INCITE and ALCC scaling, and application performance to maximize scientific application efficiency and productivity on leadership computing platforms.

Application Programs

ADSP
The ALCF Data Science Program (ADSP) supports big data projects that require the scale and performance of leadership computing resources. ADSP projects focus on developing and improving data science techniques that will enable researchers to gain insights into very large datasets produced by experimental, simulation, or observational methods.

ESP
As part of the process of bringing a new supercomputer into production, the ALCF conducts its Early Science Program (ESP) to prepare applications for the architecture and scale of a new system. ESP projects represent a typical system workload at the ALCF and cover key scientific areas and numerical methods.
One Theta node-hour is equivalent to 13 Mira node-hours based on double-precision performance.

**2019 INCITE**
Million Node-Hours

**THETA**

- Biological Sciences: 12%
- Computer Science: 5%
- Earth Science: 5%
- Energy Technologies: 8%
- Engineering: 8%
- Materials Science: 22%
- Physics: 48%

**MIRA**

- Biological Sciences: 2%
- Computer Science: 11%
- Earth Science: 7%
- Energy Technologies: 8%
- Engineering: 33%
- Materials Science: 15%
- Physics: 24%

**2019 ALCC**
Million Node-Hours

**THETA**

- Biological Sciences: 4%
- Computer Science: 13%
- Earth Science: 25%
- Energy Technologies: 58%

**MIRA**

- Biological Sciences: 2%
- Computer Science: 11%
- Earth Science: 7%
- Energy Technologies: 8%
- Engineering: 33%
- Materials Science: 15%
- Physics: 24%

Note: ALCC data are from calendar year 2019.
Advancing Design and Structure Prediction of Proteins and Peptides

Proteins are complex molecules that drive virtually all cellular functions in living organisms. Using ALCF supercomputers, a team of researchers from the University of Washington is advancing protein structure modeling capabilities to enable the rational design of synthetic biomolecules, opening new frontiers for materials, catalysts, and therapeutics.

CHALLENGE
Protein design is an enormous combinatorial problem—an N-amino acid has $20^N$ possible sequences—that requires consideration of how amino acids interact with each other. This study was further complicated by the constraint that heterodimers (pairs of proteins formed from two identical molecules) bind exclusively to each other—that is, the pairs must be orthogonal.

APPROACH
Leveraging Mira to use the Rosetta software suite, the team efficiently sampled conformations to find 87 million working designs for four-helix backbones. The researchers then searched for unique hydrogen bonded networks that connected all four helices, finding 2,251 candidates and subsequently identifying low-energy sequences.

RESULTS
Of the 97 computationally selected designs that were stable and satisfied additional criteria, 39 were exclusive heterodimers. Four designs selected for x-ray crystal structure determination agreed well with the computational models and confirmed the designed hydrogen bond networks. The team also investigated permuting the hydrogen-bond networks in different helical repeat units in an attempt to expand the heterodimer set, generating 22 new constitutive heterodimers. Ultimately, six fully orthogonal heterodimer protein pairs were created. This work provides a path forward for one to computationally design specific, programmable binding—a property hitherto unique to DNA and RNA.

IMPACT
Engineered pairs of proteins that bind exclusively to one another would potentially carry large ramifications for medicine and biomaterial production. A general method for creating such pairs, as occurs in DNA, would be transformative: the encoding of specificity in DNA gave rise to technologies such as DNA origami and artificial circuits; the large set of interactions possible with heterodimers opens the door to innovations such as faster-responding protein-based control circuits.

PUBLICATIONS
Electrical and biomedical engineers at Northwestern University are using ALCF supercomputers to develop a low-cost, high-throughput microscopy technique capable of detecting macromolecular alterations to predict and quantify the risk of cancer at extremely early stages. The team intends to develop a technique whereby living cells can be observed to learn about their internal activity.

**CHALLENGE**
Understanding the link between intracellular motion and macromolecular structure is critical to addressing questions about macromolecular behavior in cells. The researchers have implemented finite-difference time-domain (FDTD) solution algorithms for Maxwell’s equations of classical electrodynamics. These algorithms comprise the physics basis of a “microscope in a computer” that synthesizes high-resolution microscope images. This capability facilitated development of a novel optical microscopy technique, dynamic partial wave spectroscopic (PWS) microscopy, capable of detecting the intracellular structure and macromolecular motion in living cells at length-scales below the diffraction limit.

**APPROACH**
The researchers are employing the FDTD software package Angora to solve Maxwell’s equations on nanoscale voxels within a simulated biological cell for an arbitrary lens-focused, apertured illumination. Integral transforms simulating refocusing lenses and apertures are applied to the FDTD-computed optical electromagnetic near field, culminating in the synthesis of full-color pixels at the image plane of the simulated microscope. These pixels are then analyzed for spectral content.

**RESULTS**
The researchers employed dynamic PWS to measure the intracellular nanoscale structure and macromolecular motion of living cells, sensitive to changes as small as 20 nanometers and achieving millisecond temporal resolution. FDTD computations performed in Angora were validated by comparison with experimental measurements involving nanosphere imaging phantoms. In-vitro application enabled the exploration of higher-order chromatin structure and dynamics changes attributable to cellular fixation, stem-cell differentiation, and ultraviolet (UV) irradiation. In doing so, the team discovered a new phenomenon: cellular paroxysm, a synchronous, nearly instantaneous burst of intracellular motion that occurs early during UV-induced cell death.

**IMPACT**
Early-stage cancer detection has been widely recognized as one of the most critical factors to successfully treat cancer and reduce mortality. Development of a low-cost, high-throughput microscopy technique that can sense macromolecular alterations in living cells is a key step towards this goal. The high spatial-temporal resolution capability of dynamic PWS combined with a parallel FDTD implementation opens the door for high-fidelity, high-throughput early-stage cancer screening. Using the dynamic PWS technique investigated during this study, seven different human cancers are now potential candidates for early detection with minimal false positives: lung, colon, ovarian, esophageal, pancreatic, thyroid, and prostate.

**PUBLICATIONS**
With the coming paradigm shift in computer architectures and programming models as capability moves to the exascale era, the Energy Exascale Earth System Model (E3SM) project aims to develop a cutting-edge climate and earth system that can tackle the most demanding climate research imperatives. Harnessing ALCF supercomputing resources, the E3SM project is addressing questions concerning the water cycle and cryosphere systems.

CHALLENGE

The research team is focusing on two grand challenge questions: (A) In the water cycle, how will more realistic portrayals of its important features (e.g., resolution, clouds, aerosols) affect river flow and associated freshwater supplies at the watershed scale? (B) In cryospheresystems, could a dynamic instability in the Antarctic Ice Sheet be triggered within the next 40 years? For (A), the team’s objective is to simulate changes in the hydrological cycle with a specific focus on precipitation and surface water in orographically complex regions, such as the western United States and Amazon headwaters. For (B), the team aims to examine the near-term risk of initiating the dynamic instability and onset of the Antarctic Ice Sheet’s collapse due to rapid melting by adjacent warming waters—the first fully coupled global simulation to include dynamic ocean-ice shelf interactions.

APPROACH

The team made extensive use of Theta to run the E3SM code, which comprises component models for atmosphere, ocean, sea ice, and land. Sixty-four tasks were assigned to every node, each with two OpenMP threads so as to establish intranodal parallelism.

The majority of the cores were allocated to the atmosphere model, a subset of which also ran the land and sea models; the remaining cores were allocated to the ocean model, which runs concurrently with the atmosphere model. Its examination of the risk of Antarctic Ice Sheet collapse represents the first fully coupled simulation to include dynamic ocean-ice shelf interactions.

RESULTS

Leveraging Theta, the researchers used initial conditions from the year 1950 to simulate a fifty-year period. High-resolution simulation runs with the E3SM code yielded more realistic renderings of strong tropical cyclones over lower-resolution runs, with wind speeds of up to 70 meters per second. One such storm followed a trajectory nearly identical to that of 1992’s Hurricane Andrew, one of the most destructive hurricanes in Atlantic history.

IMPACT

In addition to further advancing the predictive power of climate models and providing insight into the climatic effects of rapid changes in the earth’s ice content, the E3SM simulations have the potential to answer how water resources and the hydrological cycle interact with the climate system on both local and global scales. Hurricane hindcast simulations performed for this project demonstrated the high fidelity with which extreme weather can be modeled, while exposing parametric weaknesses that need improvement.

PUBLICATIONS

High-Fidelity Multiphysics Simulations to Improve Nuclear Reactor Safety and Economics

With the ability to recycle used nuclear fuels, fast reactors are a promising nuclear energy technology that can substantially reduce radioactive waste and efficiently utilize natural nuclear resources. It is a top priority of the DOE to accelerate the design of safe and economical fast reactors, which offer great potential as a sustainable, domestic energy source. In this ALCC project, the team leveraged ALCF computing resources to accurately assess the safety parameters of fast reactor designs through high-fidelity multiphysics modeling.

CHALLENGE
This project aimed to eliminate various modeling and simulation uncertainties in the calculation of a safety parameter called hot channel factors, which have historically been assessed using low-fidelity legacy physics solvers coupled with expensive mockup experiments. Today’s high-fidelity multiphysics tools offer improved accuracy over conventional techniques without the aid of expensive experiments, but they require the use of high-performance computing systems.

APPROACH
The researchers used the DOE-NEAMS (Nuclear Energy Advanced Modeling and Simulation)-developed SHARP Toolkit, which includes the PROTEUS neutronics code and the Nek5000 thermal hydraulics/computational fluid dynamics code, to carry out their high-fidelity multiphysics calculations. To reduce computational cost, this project also assessed a “zooming” tool for core simulations called SHARP Zoom, where local regions of interest are represented with full fidelity and other regions are represented with much coarser geometry. The use of ALCF supercomputers for these simulations allows for more accurate geometry and physics representations of the system than achieved with the legacy solvers.

RESULTS
The team successfully reduced or eliminated numerous modeling and simulation uncertainties, resulting in reduced hot channel factors compared to conservative legacy estimates. Smaller hot channel factors enlarge the safety margin to design limits or permit the reactor to operate at higher power, hence increasing economic output. The team also showed that SHARP Zoom has the potential to zoom anywhere in the core, making detailed information available at a fraction of the cost of a fully heterogeneous multiphysics core calculation and thereby greatly expanding the usability of high-fidelity tools in the field.

IMPACT
This project demonstrates high-fidelity modeling capabilities that can improve the economics and safety of advanced fast reactor designs, which supports DOE’s mission to facilitate the commercial deployment of such reactors as a clean and sustainable energy source.

PUBLICATIONS

This project, led by Stanford University researchers, aims to perform a large-eddy simulation (LES) of flow over a realistic model of an aircraft. Specifically, the researchers intend to demonstrate, for the first time, the predictive capability of low-dissipation LES methodologies for practical external aerodynamics configurations.

**CHALLENGE**
The major obstacle in accomplishing this goal is the presence of energetic near-wall eddies with diminishing length scales as one approaches a wall. Such eddies require a very large number of grid points in the inner-layer of the boundary layer; their resolution is infeasible even with the supercomputers to come in the next decade. The team is addressing this issue by modeling the effect of the small-scale near-wall turbulence on the large-scale resolvable fluid motion in the outer portion of the boundary layer. While the concept of wall modeling is now deemed indispensable for enabling predictive but affordable LES of practical aeronautical flows, application of wall-modeled LES (WMLES) to date largely has been limited to studies of canonical flows at low Reynolds numbers and 2D geometries. To this end, the team has conducted WMLES of the NASA Common Research Model (CRM) and the high-lift JAXA Standard Model (JSM), which are community research models for commercial transport aircraft. The high-fidelity simulation of the NASA CRM is considered a grand challenge problem for computational fluid dynamics.

**APPROACH**
To conduct its research, the team has leveraged ALCF supercomputing resources to deploy charLES, an unstructured, grid-based, multi-physics flow-solver. Designed to perform high-fidelity, large-eddy simulations, charLES combines novel mesh generation, low-dissipation numerical methods, advanced physical models, and massive parallel scalability. With charLES, the researchers have computed a variety of flows in realistic geometries that include physics such as hypersonic transitional and turbulent boundary layers, combustion instabilities in industrial gas turbines, and acoustic radiation from complex exhaust nozzels.

**RESULTS**
The team has carried out grid-converged calculations of the NASA CRM at the chord Reynolds number of 5 million at four angles of attack, and of the JSM model by capturing the flow features with a lower number of grid points. WMLES calculations have been computed for several geometries, including the Boeing Speed Bump, NASA Juncture Flow, and NACA 0012 configurations. Additional JSM simulations with wind tunnel geometry have been completed for angles of attack that span the lift curve.

**IMPACT**
Successful calculation of realistic engineering flows will facilitate the use of high-fidelity simulation techniques to help design future aircraft, particularly in take-off and landing configurations that lower-fidelity techniques have struggled to predict.

**PUBLICATIONS**

So as to develop our understanding of materials’ radiation tolerance, it is necessary that we create models of the early stages of radiation damage. Crucial to radiation damage is the notion of electronic stopping power, which refers to the transfer of energy from a charged particle travelling through a target to the electrons of the target compound. Researchers from the University of Illinois Urbana-Champaign used ALCF supercomputers to calculate the electronic stopping power in semiconductor materials and quantified the dynamics of proton projectiles.

**CHALLENGE**
Computational studies of radiation tolerance have typically relied upon semiclassical models that fail to predict the early stages of damage with consequences on the distribution of defects in target materials. Moreover, prior first-principles studies aimed at unraveling the effect of interfaces on radiation damage have been limited. Supercomputers on the scale of ALCF resources are necessary to model multiple trajectories in systems large enough to represent real-world applications.

**APPROACH**
Leveraging the ALCF’s Mira supercomputer, the researchers computed electronic stopping power using both excited-state and ground-state molecular dynamics, employing the Qb@ll code to model time-dependence and account for discrepancies that may arise from the effects of non-adiabatic and excited states. Simulating three different systems, each composed of 216 atoms—gallium phosphide (GaP), indium phosphide (InP), and a mixed material (In$_{0.5}$Ga$_{0.5}$P)—proton trajectories followed both two different crystalline axes and random directions through the crystals.

**RESULTS**
In all three materials studied, electronic stopping was found to depend strongly on trajectory and some differences could be explained by the spacing between atoms along different directions: Trajectories that approached closer to the centers of atoms with high electron density featured strong electronic stopping and were modeled by off-channel trajectories. The mixed material exhibited electronic stopping very close to the averages of the InP and GaP materials. Detailed analysis of electron density revealed reduced stopping near gallium atoms and enhanced stopping near indium atoms. Comparing results from excited-state and ground-state dynamics, the team concluded that excited electronic states must be considered when trying to understand radiation damage at the atomistic level and the emerging trajectory of the proton projectile. A recent paper used oxygen diffusion in proton-irradiated magnesium oxide to illustrate the influence of hot electrons on ion transport.

**IMPACT**
Phosphide-based semiconductors are promising materials for applications in extreme environments, due to their resistance to energetic, charged-particle radiation. These effects can be difficult to study directly, but first-principles calculations can provide atomic-level details to complement experiments and evaluate appropriateness of semi-classical methods. A better understanding of why these materials are tolerant to high levels of radiation will help to design new materials with improved resistance. This, in turn, can help improve a variety of applications across many disciplines.

---

**PUBLICATIONS**
Environmental effects and intrinsic energy-loss processes lead to fluctuations in the operational temperature of solar cells, which can profoundly influence their power-conversion efficiency. Researchers from the University of Pittsburgh and Hamad Bin Khalifa University leveraged ALCF supercomputers to determine how the electronic structure of methylammonium lead iodide changes with temperature by accounting for electron-phonon coupling and thermal expansion.

CHALLENGE
Temperature effects can be modeled by accounting for the interactions between the electrons and phonons in the system. Because the electron-phonon coupling is challenging to compute using first-principles methods, the commonly employed Allen-Heine-Cardona theory resorts to a low-order expansion in terms of vibrational normal modes. However, for some systems, such as methylammonium lead iodide, the higher-order terms are important, and thus the low-order expansion is insufficient to capture the temperature dependence of the electronic structure. For these cases, a real-space approach that accounts for interactions to all orders is required, although this necessitates the use of massively parallel, leadership-class computing resources.

APPROACH
The team computed the electron–phonon interactions for the orthorhombic and tetragonal phases of methylammonium lead iodide to all orders in electron-phonon coupling using a real-space Monte Carlo approach in conjunction with density functional theory calculations. These investigations took an approach similar to the one the group had previously used on the cubic phase of methylammonium lead iodide with the smaller unit cell. Spin-orbit coupling is included, as this is found to enhance the electron-phonon coupling strength. On the other hand, nonlocal correlations are not accounted for, as these had little effect on the coupling.

RESULTS
This project unambiguously confirmed for the first time the importance of high-order terms in the electron–phonon coupling by direct comparison with experiment. The researchers found that the band gap of methylammonium lead iodide increases with temperature, in excellent agreement with experimental results. They verified that anharmonic effects are only important near the tetragonal–cubic phase transition temperature. They also found that temperature has a significant effect on the effective masses and Rashba coupling in methylammonium lead iodide. At room temperature, electron–phonon coupling is found to enhance the band effective mass by a factor of two and to diminish the Rashba coupling by the same factor compared to $T=0\text{K}$ values. The team’s work is detailed in a paper published in The Journal of Physical Chemistry Letters.

IMPACT
This research, aligned with the DOE mission to advance clean energy, provides a comprehensive understanding of how temperature impacts the electronic structure of organic-inorganic perovskites that are of high interest for photovoltaic and optical applications.

PUBLICATIONS

While water is perceived to be one of the simplest substances in the world, modeling its behavior on the molecular level is a complex and challenging task. Leveraging ALCF supercomputers, a team from Argonne National Laboratory employed machine learning to develop a new, computationally inexpensive water model that more accurately represents the thermodynamic properties of water, including how water changes to ice.

**Challenge**
Accurate and computationally efficient molecular-level descriptions of mesoscopic behavior of ice-water systems remain a challenge despite continued advances in computing hardware. For example, a molecular-level picture of phase transformations in water is most desirable, but remains inaccessible to fully atomistic simulations due to system size and timescale limitations. Making use of a hierarchical optimization workflow and leadership-class computing resources, the Argonne research team was able to create machine-learned water models that correctly predict multiple physical properties of ice-water systems at a fraction of the computational cost of the atomistic water models currently available.

**Approach**
The researchers used coarse-grained modeling to treat individual water molecules as single particles in order to perform the simulation at a low computational cost, while machine learning helped ensure that their simpler model maintained the accuracy often sacrificed in these coarse-grained models. They utilized the ALCF’s Mira system to perform simulations of up to 8 million water molecules to study the growth and formation of interfaces in polycrystalline ice. The machine-learned coarse-grained models were trained against first-principles calculations, experimental results, and temperature-dependent properties from other simulations using a multilevel, hierarchical global optimization strategy.

**Results**
The models were able to successfully predict both the experimental melting point of ice and the density maximum of liquid water at a computational cost two to six orders of magnitude lower than existing water models. The reduced computational cost of these models relative to their atomistic counterparts makes them suitable for modeling mesoscopic phenomena involving polycrystalline ice such as mechanics of ice, ice grain growth, melting of ice crystals, and pollutant effects on ice nucleation. The team’s findings were published in *Nature Communications*.

**Impact**
The team’s machine learning-based model provides a tool that can accurately describe structural and thermodynamic properties of water and ice at a significantly lower computational cost than existing atomistic models. In addition, the machine learning workflow developed as part of this effort could be used to improve the performance and predictive capabilities of other atomistic and molecular models.
Magnetic damping (also known as Gilbert damping) affects the amount of energy necessary for a ferromagnetic device to operate and the speed at which it can do so. Using thin magnetic films typically employed in computer storage, researchers led by Oak Ridge National Laboratory identified an unexpected mechanism by which to manipulate, tune, and tailor Gilbert damping for novel applications in future devices.

CHALLENGE
To date, there has not been significant progress made in manipulating the intrinsic magnetic properties within a single device. In light of this, the observed phenomenon was completely unexpected. Leadership-class computing resources were required to optimize geometry and cell parameters of large cobalt-iron alloy simulations cells, which consist of up to 2,176 electrons.

APPROACH
In order to investigate the quantum-mechanical origin of the anisotropic damping, the researchers leveraged ALCF supercomputing resources to perform a series of density functional theory (DFT) calculations using Quantum Espresso, assisted by ALCF staff. The calculations were performed on three different cobalt-iron phases based on the cesium chloride, Zintl, and random alloy structures. To investigate effects resulting from magnetic interactions, non-collinear DFT calculations including electron spin-orbit coupling (SOC) effects were required.

RESULTS
By experimentally studying the angular dependence of spin-torque and inductive ferromagnetic resonance, the researchers observed magnetic Gilbert damping change by as much as a factor of four in epitaxial cobalt-iron film devices. The researchers discovered that the SOC effect on the damping anisotropy significantly amplified the local distortion in disordered cobalt-iron alloys, leading to SOC-induced changes in state-density at the Fermi level and magnetic anisotropy energy with respect to crystallographic axes. The computational results suggest that Gilbert damping anisotropy observed in the cobalt-iron films originates from the SOC's variation across different magnetization orientations in a disordered lattice.

IMPACT
Manipulation of a single device's intrinsic magnetic properties has largely gone unexplored. In identifying a key mechanism by which to manipulate, tune, and tailor Gilbert damping in magnetic materials for novel applications and energy-efficient devices, this work could help optimize the dynamic properties of future magnetic materials, leading to faster-access, higher-density computer storage, and improved motor and braking implements.
Astrophysical Particle Accelerators: Magnetic Reconnection and Turbulence

Magnetic reconnection and turbulence play significant roles in a wide variety of plasma environments, from tokamaks and Earth’s magnetosphere, to the solar corona and distant astrophysical sources where they can power intense gamma-ray flares. They are both fundamental plasma processes with important consequence, including the conversion of electromagnetic energy to particle kinetic energy. Besides heating the plasma, these processes can drive nonthermal particle acceleration (NTPA), resulting in a power-law particle energy distribution extending to very high energies. Understanding these mechanisms in the context of relativistic collisionless plasmas is especially important for high-energy astrophysical sources, where the existence of nonthermal particle populations is inferred from observations of nonthermal power-law radiation spectra. The acceleration mechanism, however, is a subject of continuing debate.

With this project, researchers from the University of Colorado Boulder are studying relativistic particle energization and resulting radiative signatures in reconnecting and turbulent plasmas, with the ultimate goal of determining the conditions and mechanisms operating in astrophysical sources.

CHALLENGE

Current research being performed for this project extends previous work on 3D collisionless magnetic reconnection and kinetic turbulence from electron-positron plasmas to electron-proton plasmas.

APPROACH

The team is using its open-source particle-in-cell (PIC) code Zeltron, which is capable of self-consistently incorporating the radiation reaction force to model cases where radiative losses significantly alter system dynamics and NTPA. The massively parallel, 3D simulations of reconnection and turbulence permit the researchers to examine the viability of these mechanisms as potential efficient astrophysical accelerators of large numbers of particles to ultrarelativistic energies sufficient to explain observed radiation.

RESULTS

The team generated a series of simulations, to show that the energetic particle distributions do converge with increased system size, which justifies extrapolating the simulation results to astrophysical scales. They also found that the time required to converge depended on system scale in a manner consistent with Fermi acceleration, and concluded that turbulence could explain nonthermal spectra observed in systems like the Crab Nebula. The researchers were able to find a correlation between the formation time of the electron energy spectrum and a previously discovered system-size dependence of NTPA. Further leveraging these simulations, the researchers are delving deeper into the microphysical mechanisms behind turbulent particle acceleration.

IMPACT

These results will lead to advances in the understanding of fundamental plasma physics processes and have important implications for modern high-energy astrophysics.

PUBLICATIONS


Deep Learning at Scale for Multi-Messenger Astrophysics

Multi-messenger astrophysics (MMA) refers to the observation of astrophysical phenomena using a multitude of cosmic messengers, including gravitational waves, electromagnetic waves, neutrinos, and cosmic rays. The astronomical facilities that probe the cosmos using these cosmic messengers produce datasets whose volume and velocity challenge the scalability and computational performance of standard data analytics algorithms. With this ADSP project, researchers are spearheading a science program at the interface of AI and HPC to address these challenges to maximize discovery with MMA discovery campaigns in the big-data era.

CHALLENGE
This project is developing AI algorithms to increase the depth and speed of gravitational wave searches, and to process in real-time terabyte-size datasets of telescope images to identify the electromagnetic counterparts of gravitational wave sources.

APPROACH
Leveraging ALCF supercomputing resources, the researchers are using TensorFlow and PyTorch to design neural network models for real-time detection and characterization of gravitational wave sources, as well as the classification of galaxy images in two electromagnetic surveys, namely the Sloan Digital Sky Survey (SDSS) and the Dark Energy Survey (DES). The main goal is to use gravitational wave observations of compact binary mergers to establish a redshift-distance relationship that may be used to measure the Hubble constant. Working together with ALCF researchers, this group has deployed optimized neural network models to be trained at scale with the open-source deep learning framework Horovod.

RESULTS
A paper published in Physical Review D demonstrates that deep learning can extract the complex gravitational wave signatures of eccentric binary black hole mergers from LIGO data. As shown in work in Physical Letters B, the team achieved the first application of deep transfer learning and distributed training to design neural network models for classifying galaxies in the SDSS and DES with state-of-the-art accuracy. This study also introduced a method to classify unlabeled galaxies in the DES. Other work demonstrated that deep learning algorithms can denoise gravitational wave signals embedded in non-Gaussian and non-stationary advanced LIGO noise. These new methodologies outperform traditional signal-processing techniques.

IMPACT
Future MMA observations will shed light on the matter composition of neutron stars, the origin of long gamma ray bursts, and the mass distribution of neutron stars. They also stand to elucidate the mechanisms that may trigger electromagnetic radiation when black holes and neutron stars collide, and will enable us to pinpoint the location of these events with unprecedented accuracy. Furthermore, the algorithms and computational frameworks designed and deployed in this project will benefit domain scientists and help promote deep learning as the default signal-processing method for big-data analytics.

PUBLICATIONS

Hadronic Light-by-Light Scattering Contribution to the Muon Anomalous Magnetic Moment

The Standard Model has been the definitive description of particle physics in the past 50 years, but it's being challenged by a discrepancy between the measured and predicted values of the magnetic moment of a particle called the muon. This project aims to quantify and reduce the largest uncertainties associated with this value in order to obtain the most precise calculation of the anomaly. These results, combined with data soon to come from Fermilab's Muon g-2 experiment, may lead to the discovery of new physics beyond the Standard Model.

**CHALLENGE**

The muon’s magnetic moment describes how this fundamental particle interacts with a magnetic field; it depends on all particles that can couple to the muon—including as-yet-undiscovered particles. The muon moment has been both measured in experiments and calculated theoretically, but those two values don’t quite match up—hinting at the existence of new physics. A higher-precision measurement could erase the discrepancy, so minimizing uncertainties in the theoretical calculation is necessary to reach a more definite conclusion. The largest uncertainty in the calculation comes from particles that interact through the strong force, known as hadronic contributions. Since these contributions can’t be solved with conventional perturbative methods at low-energy scales, researchers previously resorted to experimental data or approximations with large uncertainties.

**APPROACH**

The team used lattice quantum chromodynamics (QCD) to compute the hadronic contributions by allowing particles to exist only on a discrete lattice. They used Mira to solve the complex mathematical equations of QCD, which encode all possible strong interactions with the muon.

**RESULTS**

Culminating four years of running on Mira, the team recently submitted to *Physical Review Letters* a paper announcing the first ever complete calculation of the next-leading-order hadronic contribution to the muon anomaly from light-by-light scattering, including continuum and infinite volume limits. This followed an earlier paper published in *Physical Review Letters* detailing their successful determination of the most precise prediction of the leading-order hadronic vacuum polarization contribution to the muon’s magnetic moment. Both contributions contribute roughly the same to the total uncertainty in the Standard Model value.

**IMPACT**

These results suggest that the hadronic light-by-light contribution cannot explain the current discrepancy with experiment. Comparison with measurements from Fermilab’s Muon g-2 experiment could potentially lead to the discovery of new physics.

**PUBLICATIONS**

ITER, an international collaboration to design, construct, and assemble a burning plasma experiment that can demonstrate the scientific and technological feasibility of fusion energy, is a massive effort that relies on contributions from physicists around the world. This includes a team from the Princeton Plasma Physics Laboratory that is leading a three-year INCITE project aimed at furthering the understanding of the edge plasma physics in fusion reactors at the first-principles level.

**CHALLENGE**

There are two high-priority challenges in the advancement of ITER’s quest toward fusion energy: (1) the peak heat-flux density on the ITER divertor target plates in the high-confinement mode (H-mode) operation and (2) the achievability of the H-mode plasma condition (i.e., achieving the low-to-high confinement mode L-H transition at edge) that is essential for ITER to meet its tenfold energy gain goal. Understanding these nonlinear, multiscale, multiphysics problems has been an issue for over 35 years and is difficult due to the complexity of the tokamak edge geometry.

**APPROACH**

To address these challenging plasma physics questions, the PPPL team is using its 5D gyrokinetic particle-in-cell code, XGC1, to perform large-scale simulations on DOE leadership computing resources, including the ALCF’s Theta system. XGC1 has the unique capability to solve the boundary multiscale plasma problems across the magnetic separatrix (the boundary between the magnetically confined and unconfined plasmas) using first-principles-based kinetic equations. The magnetic field lines in the separatrix run through the x-point and continue outward, allowing waste heat and particles to escape and hit the divertor plate. If the energy deposited on the plate is concentrated in too narrow an area, it may physically damage it. By running XGC1 on Theta, the researchers are able to study the heat exhaust produced by fusion reactors.

**RESULTS**

The team’s simulations have revealed a surprising and critical effect of the x-point on the heat exhaust: a bump of electric field occurs at the x-point, causing plasma to circulate around it and preventing plasma particles from traveling directly between the upstream edge-plasma and downstream divertor areas. This shows that the present assumption used for the interaction between these edge-plasma components was inadequate, and more accurate equations and improved reduced models can now be developed. The team’s findings were published in *Physics of Plasma*.

**IMPACT**

Fusion energy has the potential to provide a safe, abundant, and carbon-free source of power. Achieving a tenfold energy gain in ITER experiments relies on a deeper understanding of edge plasma physics through large-scale, high-fidelity simulation. It has been known experimentally that the edge plasma behavior limits or enhances the core energy gain efficiency in steady-state fusion operation. DOE leadership computing resources, along with the team’s advanced XGC1 code, can allow for more accurate predictions of escaping waste heat and particles and will help inform the development of future, large-scale fusion experiments that are less vulnerable to internal damage.

**PUBLICATIONS**

Fusion energy offers the prospect of a carbon-neutral, environmentally responsible, inexhaustible energy source. Researchers from TAE Technologies, Inc. are using ALCF computing resources to accelerate their fusion plasma research program aimed at developing the world’s first commercially viable fusion-powered generator for electricity production.

CHALLENGE
The central challenge in generating fusion energy for commercial use is to keep plasma hot enough for long enough to sustain the fusion reaction. TAE Technologies has built “Norman,” an advanced beam-driven field-reversed configuration (FRC) plasma device, for experimental studies of the magnetic confinement of hot fusion plasmas. Heat can be lost from the FRC due to nonlinear wave-particle interactions, which start at the microscale and lead to kinetic turbulence. On the macroscale, nonlinear wave-particle interactions can also determine the global stability of the plasma.

APPROACH
To understand both the microscale and macroscale kinetic plasma physics in FRC devices, the research team has utilized Theta to develop the ANC kinetic micro-turbulence code and the FPIC kinetic macro-stability code to enable massively parallel 3D particle-in-cell (PIC) simulations of the Norman experiment. The ANC code, developed in collaboration with the University of California, Irvine, is being used to interpret experimental measurements of turbulence and to understand how heat loss (transport) scales with plasma temperature. The FPIC code is being used to help understand the global stability and self-organization of the plasma.

RESULTS
Developing these codes has allowed the TAE Technologies team to increase physics understanding by comparing experimentally measured and computationally simulated FRC turbulence for the first time. Using the ANC code on Theta, the researchers performed the first non-linear, global, cross-separatrix, turbulence simulations in FRC topology. The study revealed that short wavelength modes are linearly unstable outside of the FRC, but not on the interior. The simulations were validated against previous experimental observations and were found to agree well. These findings were published in *Nuclear Fusion*.

The team is using the FPIC code to study the global configuration stability of the FRC, simulating the impact of external actuators on plasma self-organization. These results will be used to inform the operating states and the feedback and control strategies that are being developed for Norman and future FRC devices.

IMPACT
This project will help accelerate TAE Technologies’ studies of the confinement of energy with high plasma temperatures and inform the design of a future prototype reactor. If the company’s technology can be extended to fusion-relevant temperatures, it will provide an economical path to clean, safe, and sustainable fusion energy.

PUBLICATIONS
The advent of high-power lasers amplified with the Chirped Pulse Amplification technique has laid the groundwork for a new branch of physics—ultra-high intensity (UHI) physics—for the study of light-matter interactions at extreme light intensities. Nowadays, PetaWatt (PW) lasers can deliver intensities exceeding $10^{22}$W/cm², at which point point matter turns into a strongly relativistic and out-of-equilibrium plasma. Beyond the fundamental interest presented by the study of these complex plasma regimes, laser-plasma interactions can produce very compact and ultra-short particle/light sources with promising applications for industry and medicine. This project, led by researchers from Lawrence Berkeley National Laboratory and Commissariat à l’Energie Atomique, uses massively parallel particle-in-cell (PIC) simulations on ALCF supercomputers to understand complex light-plasma interactions and overcome major challenges facing UHI physics.

**CHALLENGE**

The researchers intend to develop novel PW laser-based solutions for three major outstanding problems in UHI physics: (1) Can we produce high-charge compact electron accelerators with high-beam quality that will be essential to push forward the horizons of high-energy science? (2) Can we produce efficient and compact high-energy ion accelerators to democratize cancer hadron-therapy? (3) Can we overcome the current limitations of optical laser technology and reach extreme light intensities approaching the Schwinger limit, beyond which light in a vacuum self-focuses and electron-positron pairs are produced?

In particular, this project aims to show that so-called “relativistic plasma mirrors,” produced when a high-power laser hits a solid target, can provide simple and elegant paths to solving these three challenges.

**APPROACH**

The researchers’ work leverages recent transformative developments in the first-principles simulation of UHI laser-plasma interactions that enabled the 3D, high-fidelity modeling of plasma mirror sources. Close collaborations with teams at PW laser facilities help enable experimental validation of the devised solutions.

**RESULTS**

Combining correlated measurements with a detailed analysis of numerical simulations, it was demonstrated that the coupling mechanism between an intense laser field and a plasma (formed out of an initially solid target) depends on the steepness of the density gradient at the target surface, resulting from the plasma extension into vacuum. As this gradient is increased, the electron dynamics switch from temporally periodic to chaotic behavior due to a so-called stretching and folding mechanism. The chaotic behavior of the electrons eventually leads, via stochastic heating, to significant absorption of laser energy in the plasma.

**IMPACT**

The successful identification of laser-plasma parameters that maximize plasma heating should help optimize the efficiency of many particle and light sources generated during laser-plasma interactions. Such ion beams may find numerous applications, including radiography, fast ignition, warm dense matter studies, nuclear physics, radiation effects in electronics, and hadron therapy.
Diffusion Monte Carlo spin density difference between bulks of potassium-doped nickel oxide and pure nickel oxide, showing the effects of substituting a potassium atom (center atom) for a nickel atom on the spin density of the bulk.

Image: Argonne National Laboratory
### ALCF Projects

#### 2019 INCITE

<table>
<thead>
<tr>
<th>Field</th>
<th>Project Description</th>
<th>PI</th>
<th>Inst</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BIOLOGICAL SCIENCES</strong></td>
<td>Finite Difference Time Domain Simulations to Facilitate Early-Stage Human Cancer Detection</td>
<td>Allen Taflove</td>
<td>Northwestern University</td>
<td>Mira: 5,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>CHEMISTRY</strong></td>
<td>Advancing Design and Structure Prediction of Proteins and Peptides</td>
<td>David Baker</td>
<td>University of Washington</td>
<td>Mira: 5,000,000 Node-Hours, Theta: 715,000 Node-Hours</td>
</tr>
<tr>
<td></td>
<td>Dynamic Nanocluster Catalysis</td>
<td>Anastassia Alexandrova</td>
<td>University of California, Los Angeles</td>
<td>Mira: 20,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>EARTH SCIENCE</strong></td>
<td>Energy Exascale Earth System Model</td>
<td>Mark Taylor</td>
<td>Sandia National Laboratories</td>
<td>Theta: 2,354,000 Node-Hours, Summit: 700,000 Node-Hours</td>
</tr>
<tr>
<td></td>
<td>Extreme-Scale Simulations for Advanced Seismic Ground Motion and Hazard Modeling</td>
<td>Christine Goulet</td>
<td>University of Southern California</td>
<td>Mira: 1,620,000 Node-Hours, Summit: 134,000 Node-Hours</td>
</tr>
<tr>
<td></td>
<td>High-Resolution Climate Sensitivity and Prediction Simulations with the CESM</td>
<td>Gerald Meehl</td>
<td>NCAR</td>
<td>Mira: 13,250,000 Node-Hours</td>
</tr>
<tr>
<td><strong>ENERGY TECHNOLOGIES</strong></td>
<td>Bringing NE Regulatory Workflow on a Path to Exascale: LES Validation for PANDA</td>
<td>Aleksandr Obabko</td>
<td>Argonne National Laboratory</td>
<td>Mira: 17,700,000 Node-Hours</td>
</tr>
<tr>
<td><strong>ENGINEERING</strong></td>
<td>Crystal Plasticity from First Principles</td>
<td>Vasily Bulatov</td>
<td>Lawrence Livermore National Laboratory</td>
<td>Mira: 17,000,000 Node-Hours</td>
</tr>
</tbody>
</table>

#### DNS Reference Data for Turbulence Model Development on the Bachalo-Johnson ATB

<table>
<thead>
<tr>
<th>PI</th>
<th>Koen Hillewaert</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inst</td>
<td>Cenaeo</td>
<td>Mira: 7,875,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Laminar-Turbulent Transition in Swept-Wing Boundary Layers</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Lian Duang</td>
<td>Hours</td>
</tr>
<tr>
<td>Inst</td>
<td>Missouri University of Science and Technology</td>
<td>Mira: 6,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Large-Eddy Simulation of a Commercial Transport Aircraft Model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Parviz Moin</td>
<td>Hours</td>
</tr>
<tr>
<td>Inst</td>
<td>Stanford University</td>
<td>Mira: 24,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Shock-Induced Multi-Material Mixing</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Sanjiva Lele</td>
<td>Hours</td>
</tr>
<tr>
<td>Inst</td>
<td>Stanford University</td>
<td>Mira: 4,500,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Towards Ultimate Rayleigh-Benard Convection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PI</td>
<td>Janet Scheel</td>
<td>Hours</td>
</tr>
<tr>
<td>Inst</td>
<td>Occidental College</td>
<td>Mira: 12,500,000 Node-Hours</td>
</tr>
</tbody>
</table>

#### MATERIALS SCIENCE

<table>
<thead>
<tr>
<th>Field</th>
<th>Project Description</th>
<th>PI</th>
<th>Inst</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Advancing Electronic Stopping Simulation: From Solids to DNA</strong></td>
<td></td>
<td>Yasuke Kana</td>
<td>University of North Carolina at Chapel</td>
<td>Mira: 4,180,000 Node-Hours, Theta: 1,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Materials and Interfaces for Organic and Hybrid Photovoltaics</strong></td>
<td></td>
<td>Noa Marom</td>
<td>Carnegie Mellon University</td>
<td>Mira: 16,000,000 Node-Hours, Theta: 1,000,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Petascale Simulations for Layered Materials Genome</strong></td>
<td></td>
<td>Alichiro Nakano</td>
<td>University of Southern California</td>
<td>Mira: 12,500,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Predictive Simulations of Functional Materials</strong></td>
<td></td>
<td>Paul Kent</td>
<td>Oak Ridge National Laboratory</td>
<td>Theta: 1,750,000 Node-Hours, Titan: 1,000,000 Node-Hours, Summit: 100,000 Node-Hours</td>
</tr>
<tr>
<td><strong>Reactive Mesoscale Simulations of Tribological Interfaces</strong></td>
<td></td>
<td>Subramanion Senkaranarayanan</td>
<td>Argonne National Laboratory</td>
<td>Theta: 650,000 Node-Hours</td>
</tr>
</tbody>
</table>
PHYSICS
Ab-initio Nuclear Structure and Nuclear Reactions
PI  Gaute Hagen
INST Oak Ridge National Laboratory
HOURS Mira: 3,500,000 Node-Hours
Theta: 1,500,000 Node-Hours
Titan: 3,500,000 Node-Hours
Summit: 281,000 Node-Hours

Astrophysical Particle Accelerators: Magnetic Reconnection and Turbulence
PI  Dmitri Uzdensky
INST University of Colorado Boulder
HOURS Mira: 6,750,000 Node-Hours

Extreme-Scale Simulation of Supernovae and Magnetars from Realistic Progenitors
PI  Seon Couch
INST Michigan State University
HOURS Mira: 24,000,000 Node-Hours
Theta: 281,000 Node-Hours

High-Fidelity Gyrokinetic Simulation of Tokamak and ITER Edge Physics
PI  Chong-Seok Chang
INST Princeton Plasma Physics Laboratory
HOURS Mira: 1,500,000 Node-Hours
Theta: 1,500,000 Node-Hours
Titan: 1,050,000 Node-Hours

Kinetic Simulation of FRC Stability and Transport
PI  Seon Dettrick
INST TAE Technologies, Inc.
HOURS Mira: 750,000 Node-Hours

Lattice QCD
PI  Paul Mackenzie
INST Fermilab
HOURS Mira: 24,000,000 Node-Hours
Theta: 1,600,000 Node-Hours

N-Jettiness Subtraction for Precision Collider Phenomenology
PI  Radja Bougezel
INST Argonne National Laboratory
HOURS Mira: 1,250,000 Node-Hours
Theta: 400,000 Node-Hours

PETASCALE SIMULATIONS OF KINETIC EFFECTS IN IFE PLASMAS
PI  Frank Tsung
INST University of California, Los Angeles
HOURS Mira: 80,000,000 Node-Hours
Theta: 600,000 Node-Hours

PLASM-IN-SILICO: HPC Modeling of High-Intensity Laser-Solid Interaction
PI  Jean-Luc Vay
INST Lawrence Berkeley National Laboratory
HOURS Mira: 8,000,000 Node-Hours
Theta: 600,000 Node-Hours

Radiation Hydrodynamic Simulations of Massive Stars with Rotation
PI  Lars Bildsten
INST University of California, Santa Barbara
HOURS Theta: 1,300,000 Node-Hours

Towards a Definitive Model of Core-Collapse Supernova Explosions
PI  Adam Burrows
INST Princeton University
HOURS Theta: 1,750,000 Node-Hours

ALCC 2018–2019
The 2018–2019 ALCC award allocations are shown in core-hours. In 2019, the allocation programs transitioned to awarding DOE computing awards in node-hours.

CHEMISTRY
High-Fidelity Simulations of Flow and Heat Transfer During Motored Operation of an Internal Combustion Engine
PI  Paul Fischer
INST Argonne National Laboratory
HOURS 30,000,000 Core-Hours

COMPUTING
Demonstration of the Scalability of Programming Environments by Simulating Multiscale Applications
PI  Robert Voigt
INST Leidos, Inc.
HOURS 198,500,000 Core-Hours
ALCF: 10M, OLCF: 78.5M, NERSC: 20M

PORTABLE APPLICATION DEVELOPMENT FOR NEXT-GENERATION SUPERCOMPUTER ARCHITECTURES
PI  T.P. Straatsma
INST Oak Ridge National Laboratory
HOURS 60,000,000 Core-Hours
ALCF: 30M, OLCF: 30M

EARTH SCIENCE
Investigating the Impact of Improved Southern Ocean Processes in Antarctic-Focused Global Climate Simulations
PI  Mark Petersen
INST Los Alamos National Laboratory
HOURS 105,000,000 Core-Hours
ALCF: 35M, OLCF: 5M, NERSC: 65M

ENERGY TECHNOLOGIES
High-Fidelity Numerical Simulation of Wire-Wrapped Fuel Assemblies: Year 2
PI  Aleksandr Obabko
INST Argonne National Laboratory
HOURS 83,500,000 Core-Hours

High-Fidelity Simulation for Molten Salt Reactors: Enabling Innovation Through Petascale Computing
PI  Elio Merzari
INST Argonne National Laboratory
HOURS 140,000,000 Core-Hours

HPC4EnergyInnovation ALCC End-Station
PI  Peter Nugent
INST Lawrence Berkeley National Laboratory
HOURS 170,000,000 Core-Hours
ALCF: 20M, OLCF: 100M, NERSC: 50M

MULTIPHASE FLOW SIMULATIONS OF NUCLEAR REACTOR FLOWS
PI  Igor Bolotnov
INST North Carolina State University
HOURS 130,000,000 Core-Hours

ENGINEERING
Analysis and Mitigation of Dynamic Stall in Energy Machines
PI  Anupam Sharma
INST Iowa State University
HOURS 51,500,000 Core-Hours

MATERIALS SCIENCE
Imaging and Controlling Elemental Contrast of Nanocluster in Intense X-ray Pulses
PI  Phay Ho
INST University of Pittsburgh
HOURS 20,000,000 Core-Hours

Impact of Grain Boundary Defects on Hybrid Perovskite Solar Absorbers
PI  Wisam Saadi
INST University of Pittsburgh
HOURS 20,000,000 Core-Hours

Large-Scale Simulations of Heterogeneous Materials for Energy Conversion Applications
PI  Giulia Galli
INST The University of Chicago and Argonne National Laboratory
HOURS 100,000,000 Core-Hours

Modeling Fusion Plasma Facing Components
PI  Brian Wirth
INST Oak Ridge National Laboratory and University of Tennessee
HOURS 165,000,000 Core-Hours
ALCF: 80M, OLCF: 60M, NERSC: 25M

Predictive Modeling and Machine Learning for Functional Nanoporous Materials
PI  J. Ilja Siepmann
INST University of Minnesota
HOURS 58,000,000 Core-Hours
ALCF: 42M, NERSC: 36M

75
PHYSICS

Emulating the Universe
PI Katrin Heitmann
INST Argonne National Laboratory
HOURS 50,000,000 Core-Hours
ALCF: 10M OLCF: 40M

Hadronic Light-by-Light Scattering and Vacuum Polarization Contributions to the Muon Anomalous Magnetic Moment from Lattice QCD with Chiral Fermions
PI Thomas Blum
INST University of Connecticut
HOURS 162,000,000 Core-Hours

Nuclear Structure and Electric Dipole Moments with Physical Chiral-Symmetric Quarks
PI Sergey Syritsyn
INST RIKEN BNL Research Center
HOURS 50,000,000 Core-Hours

Scaling LHC Proton-Proton Collision Simulations in the ATLAS Detector
PI Eric Loncon
INST Brookhaven National Laboratory
HOURS 160,000,000 Core-Hours
ALCF: 80M OLCF: 80M

Semileptonic B- and D-meson Form Factors with High Precision
PI Ruth Van de Water
INST Fermi National Accelerator Laboratory
HOURS 247,000,000 Core-Hours

Simulations of Laser Experiments to Study MHD Turbulence and Non-Thermal Charged Particles
PI Petros Tzeferacos
INST The University of Chicago
HOURS 22,000,000 Core-Hours

ALCC 2019–2020

ENERGY TECHNOLOGIES
High-Fidelity Physics Simulations for DOE and Industry Fast Spectrum Nuclear Reactor Systems
PI Emily Sherman
INST Argonne National Laboratory
HOURS Theta: 880,000 Node-Hours

Nuclear Energy Industry Validation of Nek5000: ALAIN and HYMERES
PI Aleksander Obabko
INST Argonne National Laboratory
HOURS Theta: 340,000 Node-Hours

ENGINEERING
Towards Exascale Internal Combustion Engine Simulation with In-Situ Analysis
PI Muhsin Ameen
INST Argonne National Laboratory
HOURS Theta: 630,000 Node-Hours

MATERIALS SCIENCE
Accelerated Catalyst Discovery from First Principles Simulations and Machine Learning
PI Rajeev Surendran Assary
INST Argonne National Laboratory
HOURS Theta: 240,000 Node-Hours

Predictive Modeling and Machine Learning for Functionally Nanoporous Materials
PI J. Ilii Siepmann
INST University of Minnesota
HOURS Theta: 620,000 Node-Hours
Cori: 200,000 Node-Hours

Supercomputing for Automotive High-Temperature Alloy Design
PI Dongwon Shin
INST Oak Ridge National Laboratory
HOURS Theta: 350,000 Node-Hours
Summit: 391,000 Node-Hours

PHYSICS
Low Energy Neutrino-Nucleus Interactions
PI Soojo Postore
INST Washington University in St. Louis
HOURS Theta: 390,000 Node-Hours

Modeling the Response of Fusion Plasma Components
PI Brian Wirth
INST Oak Ridge National Laboratory
HOURS Theta: 200,000 Node-Hours
Summit: 250,000 Node-Hours

Neutrino Flux, Energy Deposition and Radiological Studies for the DUNE-LBNF Beamline
PI Igor Rakicho
INST Fermilab
HOURS Theta: 450,000 Node-Hours

The Next Leap Forward in LSST Sky Simulations
PI Katrin Heitmann
INST Argonne National Laboratory
HOURS Theta: 400,000 Node-Hours

Particle Acceleration in Plasma Jets: From Astrophysics to the Laboratory
PI Paulo E. Alves
INST SLAC National Accelerator Laboratory
HOURS Theta: 1,000,000 Node-Hours

Semileptonic B- and D-Meson Form Factors with High Precision
PI Aida El-Kahdra
INST University of Illinois at Urbana-Champaign
HOURS Theta: 400,000 Node-Hours
Cori: 330,000 Node-Hours

ALCF Data Science Program
Constructing and Navigating Polymorphic Landscapes of Molecular Crystals
PI Alexandre Tkatchenko
INST University of Luxembourg
HOURS Mira: 3,125,000 Node-Hours
Theta: 125,000 Node-Hours

Data-Driven Materials Discovery for Optoelectronic Applications
PI Jacqueline Cole
INST University of Cambridge
HOURS Theta: 468,750 Node-Hours

Machine Learning Magnetic Properties of van der Waals Heterostructures
PI Efthimios Kaxiras
INST Harvard University
HOURS Theta: 156,250 Node-Hours

Realistic Simulations of the LSST Survey at Scale
PI Katrin Heitmann
INST Argonne National Laboratory
HOURS Theta: 218,750 Node-Hours

X-ray Microscopy of Extended 3D Objects: Scaling Towards the Future
PI Chris Jacobsen
INST Argonne National Laboratory and Northwestern University
HOURS Theta: 187,500 Node-Hours

Aurora Early Science Program
Accelerated Deep Learning Discovery in Fusion Energy Science
PI William Tang
INST Princeton Plasma Physics Laboratory

Dark Sky Mining
PI Salman Habib
INST Argonne National Laboratory
Data Analytics and Machine Learning for Exascale Computational Fluid Dynamics
PI  Kenneth Jansen
INST  University of Colorado Boulder

Enabling Connectomics at Exascale to Facilitate Discoveries in Neuroscience
PI  Nicola Ferrier
INST  Argonne National Laboratory

Enabling Connectomics at Exascale to Facilitate Discoveries in Neuroscience
PI  Nicola Ferrier
INST  Argonne National Laboratory

Exascale Computational Catalysis
PI  David Bross
INST  Argonne National Laboratory

Extending Moore's Law Computing with Quantum Monte Carlo
PI  Anour Benali
INST  Argonne National Laboratory

Extreme-Scale Cosmological Hydrodynamics
PI  Katrin Heitmann
INST  Argonne National Laboratory

Extreme-Scale In-Situ Visualization and Analysis of Fluid-Structure-Interaction Simulations
PI  Amanda Randles
INST  Duke University and Oak Ridge National Laboratory

Extreme-Scale Unstructured Adaptive CFD
PI  Kenneth Jansen
INST  University of Colorado Boulder

High-Fidelity Simulation of Fusion Reactor Boundary Plasmas
PI  C.S. Chang
INST  Princeton Plasma Physics Laboratory

Machine Learning for Lattice Quantum Chromodynamics
PI  William Detmold
INST  Massachusetts Institute of Technology

Many-Body Perturbation Theory Meets Machine Learning to Discover Singlet Fission Materials
PI  Noa Maron
INST  Carnegie Mellon University

NWChemEx: Tackling Chemical, Materials, and Biochemical Challenges in the Exascale Era
PI  Thom Dunning
INST  Pacific Northwest National Laboratory

Simulating and Learning in the ATLAS Detector at the Exascale
PI  James Proudfoot
INST  Argonne National Laboratory

Virtual Drug Response Prediction
PI  Rick Stevens
INST  Argonne National Laboratory

2019 Director's Discretionary
The following list provides a sampling of the many Director's Discretionary projects at the ALCF.

BIOLOGICAL SCIENCES
Computational Analysis of Brain Connectomes for Alzheimer's Disease
PI  Joook Cha
INST  Columbia University
HOURS  Theta: 46,875 Node-Hours

Free Energy Landscapes of Membrane Proteins
PI  Benoit Roux
INST  The University of Chicago
HOURS  Theta: 78,125 Node-Hours

CHEMISTRY
Accelerated Catalyst Discovery from First Principles Simulations and Machine Learning
PI  Rajeev Suredran Assary
INST  Argonne National Laboratory
HOURS  Theta: 135,175 Node-Hours

Ionic Liquid as a Potential Electrolyte of High-Performance Lithium Ion Battery
PI  Zhengcheng Zhang
INST  Argonne National Laboratory
HOURS  Theta: 93,750 Node-Hours

COMPUTER SCIENCE
Simulating Realistic Quantum Computers
PI  Hal Finkel
INST  Argonne National Laboratory
HOURS  Theta: 251,000 Node-Hours

Workload Interference Analysis on Theta
PI  Kenneth Raffenetti
INST  Illinois Institute of Technology
HOURS  Theta: 65,630 Node-Hours

EARTH SCIENCE
Extreme-Scale Simulations for Advanced Seismic Ground Motion and Hazard Modeling
PI  Christine Goulet
INST  University of Southern California
HOURS  Theta: 111,250 Node-Hours

Terrestrial Ecosystem Carbon Cycle of the Conterminous U.S.
PI  Jinxun Liu
INST  United States Geological Survey
HOURS  Theta: 251,000 Node-Hours

ENERGY TECHNOLOGIES
Thermal Hydraulic Simulations for the Versatile Test Reactor
PI  Dillon Shaver
INST  Argonne National Laboratory
HOURS  Mira: 578,440 Node-Hours

PI  Surya Vegendla
INST  Argonne National Laboratory
HOURS  Theta: 437,500 Node-Hours

ENGINEERING
Data Analysis of Turbulent Channel Flow at High Reynolds Number
PI  Robert Moser
INST  University of Texas at Austin
HOURS  Mira: 937,500 Node-Hours

High-Fidelity Simulation of Supersonic Turbulent Flow-Structure Interaction and Mixing
PI  Ivan Bermejo Moreno
INST  University of Southern California
HOURS  Theta: 207,570 Node-Hours

Investigation of a Low Octane Gasoline Fuel for a Heavy-Duty Diesel Engine in a Low-Temperature Combustion Regime
PI  Sibendu Som
INST  Argonne National Laboratory
HOURS  Mira: 1,875,000 Node-Hours

Multiphase Simulations of Nuclear Reactor Thermal Hydraulics
PI  Igor A. Bolotnov
INST  North Carolina State University
HOURS  Mira: 625,000 Node-Hours

Parallel Hierarchic Adaptive Stabilized Transient Analysis
PI  Kenneth Jansen
INST  North Carolina State University
HOURS  Mira: 3,373,628 Node-Hours

Well-Resolved Simulations of Canonical Wall-Bounded Flows
PI  Ramesh Balakrishnan
INST  Argonne National Laboratory
HOURS  Mira: 2,531,250 Node-Hours

MATERIALS SCIENCE
Constructing and Navigating Polymorphic Landscapes of Molecular Crystals
PI  Alexandre Tkatchenko
INST  University of Luxembourg
HOURS  Mira: 3,125,000 Node-Hours

Large-Scale Real-Space Electronic Structure Calculations for Understanding Energetics of Complex Defects in Materials
PI  Phani Sudheer Motamarri
INST  University of Michigan
HOURS  Theta: 93,750 Node-Hours
### Phase Transitions in Water-Ice-Vapor System
- **PI**: Subramanian Sankaranarayanan
- **INST**: Argonne National Laboratory
- **HOURS**: Mira: 625,000 Node-Hours

### Rational Design of Ultrastrong Composites
- **PI**: Hendrik Heinz
- **INST**: University of Colorado Boulder
- **HOURS**: Mira: 206,250 Node-Hours

### Structure and Properties of Grain Boundaries in Materials for Energy Applications
- **PI**: Wissam Saidi
- **INST**: University of Pittsburgh
- **HOURS**: Mira: 750,000 Node-Hours

### PHYSICS

#### Extreme-Scale Cosmology
- **PI**: Katrin Heitmann
- **INST**: Argonne National Laboratory
- **HOURS**: Mira: 50,000,000 Node-Hours

#### MARS Energy Deposition and Neutrino Flux Simulations
- **PI**: Nikolai V. Mokhov
- **INST**: Fermilab
- **HOURS**: Theta: 65,630 Node-Hours

#### Pearl Necklace
- **PI**: Michael David Schneider
- **INST**: Imperial College London
- **HOURS**: Theta: 258,860 Node-Hours

#### Pion and Kaon Quark-Gluon Structure from Lattice QCD
- **PI**: Ian Cloet
- **INST**: Argonne National Laboratory
- **HOURS**: Theta: 65,925 Node-Hours

#### Unveiling the 3D Physics Behind Compact Ultrahigh Flux Neutron Sources
- **PI**: Frederico Fiuza
- **INST**: SLAC National Accelerator Laboratory
- **HOURS**: Mira: 3,125,000 Node-Hours
## ALCF Publications

Researchers who use ALCF resources are major contributors to numerous publications that document their breakthrough science and engineering. The refereed journal articles and conference proceedings represent research ventures undertaken at the ALCF through programs supported by the U.S. Department of Energy and Argonne National Laboratory.

The publications are listed by their publication dates. An asterisk after a name indicates an ALCF author. ALCF publications are listed online at alcf.anl.gov/science/publications.

<table>
<thead>
<tr>
<th>January</th>
</tr>
</thead>
</table>


February


Yost, D. C., and Y. Kanai. “Electronic Excitation Dynamics in DNA under Proton and α-Particle Irradiation,” Journal of the American Chemical Society (March 2019), ACS. doi: 10.1021/jacs.9b12148

April

Dzodzo, M. B. “Natural and Mixed Convection in Rectangular Enclosures and Channels Containing Liquid Metals and Partition Walls,” 4th Thermal and Fluids Engineering Conference (April 2019), Las Vegas, NV, ASTMFE.

do: 10.1615/TFEC2019.4.028998


May


July


**November**


December


About the Argonne Leadership Computing Facility

Argonne National Laboratory operates the Argonne Leadership Computing Facility (ALCF) as part of the U.S. Department of Energy's effort to provide leadership-class computing resources to the scientific community. The ALCF is supported by the DOE Office of Science, Advanced Scientific Computing Research (ASCR) program.

About Argonne National Laboratory

Argonne is a U.S. Department of Energy Laboratory managed by UChicago Argonne, LLC, under contract DE-AC02-06CH11357. The Laboratory's main facility is outside of Chicago, at 9700 South Cass Avenue, Lemont, Illinois 60439. For information about Argonne and its pioneering science and technology programs, visit anl.gov.

Availability of this Report (ANL/ALCF-20/01)

Online Access: U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available for free viaosti.gov.

Reports not in digital format may be purchased by the public from the National Technical Information Service (NTIS):

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd.
Alexandria, VA 22312
phone | 800.553.NTIS (6847) or 703.605.6000
fax | 703.605.6900
orders@ntis.gov
ntis.gov

Reports not in digital format are available to DOE and DOE contractors from the Office of Scientific and Technical Information (OSTI):

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
phone | 865.576.8401
fax | 865.576.5728
reports@adonis.osti.gov
osti.gov

ALCF Leadership: Michael E. Papka (Division Director), Jini Ramprakash (Deputy Division Director), Mark Fahey (Director of Operations), Katherine Riley (Director of Science), and Kalyan Kumaran (Director of Technology)

Editorial Team: Beth Cerny, Jim Collins, Nils Heinonen, and Laura Wolf

Design and production: Sandbox Studio, Chicago

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.