ORNL/SPR-2024/3352

US Department of Energy, Office of Science, High-Performance Computing Facility 2023 Operational Assessment Oak Ridge Leadership Computing Facility

April 2024

ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

ORNL/SPR-2024/3352

Oak Ridge Leadership Computing Facility

US DEPARTMENT OF ENERGY, OFFICE OF SCIENCE, HIGH-PERFORMANCE COMPUTING FACILITY 2023 OPERATIONAL ASSESSMENT OAK RIDGE LEADERSHIP COMPUTING FACILITY

Subil Abraham
J. Paul Abston
Ryan M. Adamson
Valentine Anantharaj
Ashley Barker
Aaron Barlow
Tom Beck
Katie L. Bethea
Josh Cunningham
Rafael Ferreira da Silva
Christopher Fuson
Olga Kuchar
Matt Lakin
Ryan Landfield
Jack Lange
Kellen Leland
Dustin Leverman
Hong Liu
Don Maxwell
Verónica Melesse Vergara
Bronson Messer
Ross Miller
Sheila Moore
James Morgan
Sarp Oral
Suzanne Parete-Koon
Ryan Prout
William Renaud
Tori Robinson
James Rogers
Michael Sandoval
Mallikarjun Shankar
Scott Simmerman
Betsy Sonewald
Christopher Stahl
In-Saeng Suh
Kevin G. Thach
Suzy Tichenor
Georgia Tourassi
Feiyi Wang
Spencer Ward
Christopher Zimmer

April 2024

Prepared by
OAK RIDGE NATIONAL LABORATORY
Oak Ridge, TN 37831
managed by
UT-BATTELLE LLC
for the
US DEPARTMENT OF ENERGY
under contract DE-AC05-00OR22725

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACE	Advanced Computing Ecosystem
ACCEL	Accelerating Competitiveness through Computational ExceLlence
ACM	Association of Computing Machinery
AD	Application Development
AFW	Air Force Weather
AI	artificial intelligence
ALCC	ASCR Leadership Computing Challenge
ALCF	Argonne Leadership Computing Facility
Argonne	Argonne National Laboratory
ARM	Atmospheric Radiation Measurement
ASCR	Advanced Scientific Computing Research
ATPESC	Argonne Training Program for Extreme-Scale Scientific Computing
BES	Basic Energy Sciences
CAAR	Center for Accelerated Application Readiness
CBI	Center for Bioenergy Innovation
CCSD	Computing and Computational Sciences Directorate
CRM	customer relationship management software
CNMS	Center for Nanophase Materials Science
CSEEN	Computational Scientists for Energy, the Environment, and National Security
DARTs	Days Away Restricted or Transferred
DD	Director’s Discretionary program
DFT	density functional theory
DOE	US Department of Energy
DVA	data visualization and analytics
ECP	Exascale Computing Project
ES&H	environment, safety, and health
ESGF	Earth Systems Grid Federation
Fermilab	Fermi National Accelerator Laboratory
FFT	fast Fourier transform
FL	federated learning
FrESCO	Framework for Exploring Scalable Computational Oncology
FRIB	Facility for Rare Isotope Beams
GPFS	General Parallel File System
GRETA	Gamma-Ray Energy Tracking Array
HPC	high-performance computing
HPL	High-Performance Linpack
HPSS	High-Performance Storage System
HR	ORNL Human Resources
INCITE	Innovative and Novel Computational Impact on Theory and Experiment
ISM	Integrated Safety Management
IRI	Integrated Research Infrastructure (IRI) Program
ITAR	International Traffic and Arms Regulations
JGI	Joint Genome Institute
JIF	journal impact factor

JLab	Thomas Jefferson National Accelerator Facility
LANL	Los Alamos National Laboratory
LBNL	Lawrence Berkeley National Laboratory
LCF	Leadership Computing Facility
LLM	large language model
LLNL	Lawrence Livermore National Laboratory
LQCD	Lattice Quantum Chromodynamics
LSMS	Locally Self-Consistent Multiple Scattering code
LVA	Live Visual Analytics
ML4NSE	Machine Learning for Neutron Scattering Experiment
MOSSAIC	Modeling Outcomes using Surveillance data and Scalable Artificial Intelligence for Cancer
MPI	Message Passing Interface
MTTF	mean time to failure
MTTI	mean time to interrupt
NAIRR	National Artificial Intelligence Research Resource
NAM	not a metric
NCCS	National Center for Computational Sciences
NCI	National Cancer Institute
NCRC	National Climate-Computing Research Center
NERSC	National Energy Research Scientific Computing Center
NGP	Next Generation Pathways Summer School
NIH	National Institutes of Health
NOAA	National Oceanic and Atmospheric Administration
NTCL	Nuclear Tensor Contraction Library
NuCCOR	Nuclear Coupled-Cluster Oak Ridge
NVFLARE	NVIDIA Federated Learning Application Runtime Environment
OA	overall availability
OAR	Operational Assessment Review
OCR	Operations Control Room
OLCF	Oak Ridge Leadership Computing Facility
ORNL	Oak Ridge National Laboratory
OUG	OLCF User Group
PAS	Personnel Access System
PEARC	Practice and Experience in Advanced Research Computing
PHI	protected health information
PI	principal investigator
PNNL	Pacific Northwest National Laboratory
PSDNS	Pseudo-Spectral Direct Numerical Simulation
QC	quantum computing
QCUP	Quantum Computing User Program
QRUC	Quantum Resource Utilization Council
RATS	Resource and Allocation Tracking System
RISE	Revolutionary Innovation for Sustainable Engines
RPS	Restricted Party Screening
RRTMGP	Rapid Radiative Transfer Model for GCMs in Parallel

RSS	research safety summary
SA	scheduled availability
Sandia	Sandia National Laboratories
SBMS	Standards Based Management System
SC	International Conference for High Performance Computing, Networking, Storage, and Analysis
SCREAM	Simple Cloud-Resolving E3SM Atmosphere Model
SEER	Surveillance, Epidemiology, and End Results program
SMCDC	Smoky Mountains Computational Sciences and Engineering Conference, Data Challenge Session
SNS	Spallation Neutron Source
SSH	Secure Shell
ST	Software Technology
SU	system utilization
TFIIH	transcription factor IIH
TTU	Tennessee Technological University
UIUC	University of Illinois at Urbana-Champaign
UT	University of Tennessee
VA	US Department of Veterans Affairs
WES	Workflow and Ecosystem Services
WDTS	Workforce Development for Teachers and Scientists
XNR1K	Experimental Nature Run at 1 km

HIGH-PERFORMANCE COMPUTING FACILITY
2023 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2024

EXECUTIVE SUMMARY

The Oak Ridge Leadership Computing Facility (OLCF) was established to accelerate scientific discovery by providing world-leading computational performance and advanced data infrastructure. As a US Department of Energy (DOE) Office of Science user facility, the OLCF has managed the successful deployment and operation of a succession of leadership-class resources dedicated to open science. In addition to these resources, the OLCF staff continually strive to develop innovative processes and technologies, improve security, and empower users through effective allocation management and comprehensive user support and training. These efforts support the advancement of science by the OLCF users and benefit high-performance computing (HPC) facilities around the world.

In calendar year (CY) 2023, the OLCF supported 1,676 users and 598 projects and exceeded all targets for user satisfaction. The facility received an average satisfaction score of 4.52 out of 5 on the annual user survey, and 94% of respondents reported a high satisfaction rate with the OLCF overall. Of the 3,619 user tickets submitted in CY 2023, OLCF staff resolved 97% within 3 business days. The facility opened Frontier to full scientific operations this year. Two projects conducted on Frontier received the Association for Computing Machinery (ACM) Gordon Bell Prize and the Gordon Bell Special Prize for Climate Modeling, and a third earned a nomination as a Gordon Bell Prize finalist. The facility's previous flagship machine, Summit, gained new life and was extended through 2024 in part to help provide resources to the Integrated Research Infrastructure (IRI) projects and the National Artificial Intelligence Research Resource (NAIRR) pilot program. The facility instantiated an Advanced Computing Ecosystem testbed in part to support IRI workflows. OLCF made interactivity easier and more accessible to users than ever through tools like Jupyter notebooks and workflows.

Throughout the year, OLCF staff ensured users had reliable access to the following HPC resources: the HPE Cray EX Frontier, IBM AC922 Summit, the General Parallel File System (Alpine), the Lustre file system Orion, and the archival storage system (High-Performance Storage System). Overall availability for both the compute resources and file systems exceeded 96% for CY 2023. The facility also successfully delivered on the allocation split of roughly 60%, 20%, and 20% of core-hours offered for the Innovative and Novel Computational Impact on Theory and Experiment (INCITE), the Advanced Scientific Computing Research Leadership Computing Challenge (ALCC), and Director's Discretionary (DD) programs, respectively (Section 2). Table ES-1 summarizes the 2023 OLCF metric targets and the associated results. More information can be found in Section 2 for each OLCF resource.

Table ES-1. 2023 OLCF metric summary.

Metric Description	CY 2023 Target	CY 2023 Actual
Overall OLCF score on the user survey will be 3.5 based on a statistically meaningful sample.	3.5	4.5
Time between Receipt of User Query (RT Ticket) and Center Response: 80% of OLCF problems will be addressed within 3 working days (72 hours) by either resolving the problem or informing the user how the problem will be resolved.	80%	94%

CAPABILITY JOBS: *For the CY following a new system/upgrade, at least 30% of the consumed node hours will be from jobs requesting 20% or more of the available nodes. In subsequent years, at least 35% of consumed core hours / node hours will be from jobs requiring 20% or more of cores/nodes available to the users.*		
Scientific and Technological Research and Innovation – Demonstrate Leadership Computing, Summit	35%	45.63%
Scientific and Technological Research and Innovation – Demonstrate Leadership Computing, Frontier	30%	79.08%
SCHEDULED AVAILABILITY (COMPUTE): *For the CY following a new system/upgrade, the scheduled availability (SA) target for an HPC compute resource is at least 85%. For year 2, the SA target for an HPC compute resource increases to at least 90%, and for year 3 through the end of life for the associated compute resource, the SA target for an HPC compute resource increases to 95%. Consequently, SA targets are described as 85/90/95%.*		
SA, Summit: Sustain scheduled availability to users, measured as a percentage of maximum possible scheduled.	95%	99.84%
SA, Frontier: Sustain scheduled availability to users, measured as a percentage of maximum possible scheduled.	85%	99.65%
OVERALL AVAILABILITY (COMPUTE): *For the CY following a new system/upgrade, the overall availability (OA) target for an HPC compute resource is at least 80%. For year 2, the OA target increases to at least 85%, and for year 3 through the end of life for the associated compute resource, the OA target increases to 90%. Consequently, OA targets are described as 80/85/90%.*		
OA, Summit: Sustain availability to users measured as a percentage of maximum possible.	90%	99.41%
OA, Frontier: Sustain availability to users measured as a percentage of maximum possible.	80%	97.64%
OVERALL AVAILABILITY (FILE SYSTEMS): *For the CY following a new system/upgrade, the OA target for an external file system is at least 85%. For year 2 through the end of life of the asset, the OA target for an external file system increases to at least 90%. OA targets are thus described as 85/90%.*		
OA, external file system Alpine: Sustain availability to users measured as a percentage of maximum possible.	90%	99.20%
OA, external file system Orion: Sustain availability to users measured as a percentage of maximum possible.	85%	98.42%
OA, archive storage: Sustain availability to users measured as a percentage of maximum possible.	90%	96.10%
SCHEDULED AVAILABILITY (FILE SYSTEMS): *For the CY following a new system/upgrade, the SA target for an external file system is at least 90%. For year 2 through the end of life of the asset, the SA target for an external file system increases to at least 95%. SA targets are thus described as 90/95%.*		
SA, Alpine: Sustain scheduled availability to users measured as a percentage of maximum possible scheduled.	95%	99.69%
SA, Orion: Sustain scheduled availability to users measured as a percentage of maximum possible scheduled.	90%	99.08%
SA, High Performance Storage System: Sustain scheduled availability to users measured as a percentage of maximum possible scheduled.	95%	99.65%

CY 2023 marked another year of sustained technological development and innovation for the OLCF. A team of OLCF researchers led an end-to-end examination of the effectiveness of large language models (LLMs) in scientific research that identified practical solutions and computational requirements for building and using LLM-based foundation models on Frontier. FORGE, the resulting suite of open foundation models, was trained on 257 billion tokens from more than 200 million publications and successfully demonstrated LLM use and effectiveness on scientific downstream tasks. The results will aid the development of the next generation of foundation models. Another initiative to explore detailed system operations for parameters such as energy efficiency developed a digital twin to simulate workloads on Frontier and replay telemetry data for visualization or validation of models. The OLCF accepted the Kronos Nearline Storage system, which strikes a balance between the high performance of scratch storage resources and the data integrity features of archival resources. The high performance, high durability, and expansive feature set empowers Kronos to support the IRI Architecture Blueprint's three science patterns. Enabling peering with the DOE High-Performance Data Facility may be achieved by supporting core services through Kronos and establishing staged data exchange and transfers as needed.

The OLCF also implemented and refined other technical innovations that will continually improve operations, including Advanced Computing Ecosystem Evaluations and testbeds to support and evaluate IRI patterns; leadership-scale capability benchmarks; scalable data hierarchy management; operational data workflow development; classical-quantum hybrid workflows; improved energy efficiency in Frontier; and continued efforts toward Federated Learning.

The scientific accomplishments of OLCF users and staff are a strong indication of long-term operational success and broad scientific impact, and this year's user projects and programs have continued to advance DOE's strategic objectives. Users published 50 papers in high-impact journals such as Nature Physics, Nano Letters, Trends in Plant Science, Nature Chemical Biology, and Advanced Materials.

The OLCF supported scientific accomplishments for a broad community of researchers in 2023, from traditional modeling and simulation projects to studies that exploit AI, ML, and big data analytics techniques. Emerging AI methods and mixed-precision approaches are becoming essential to the world's most significant computational campaigns. Frontier's leadership capability was critical to three finalist teams for the prestigious Gordon Bell award competition, including both winning teams who used leadership scale and innovative machine learning and low-precision computing in their campaigns. The Gordon Bell Prize is awarded each year at the International Conference for High Performance Computing, Networking, Storage and Analysis (SC) to recognize researchers who have made significant strides toward applying HPC systems to scientific applications.

One team led by University of Michigan researchers won the Gordon Bell Prize for their use of Frontier to simulate dislocations, or defects, in a magnesium system of nearly 75,000 atoms at near-quantum accuracy – one of the largest simulations of an alloy ever. The second team won the inaugural Gordon Bell Prize for Climate Modeling using Frontier to achieve record speeds in modeling worldwide cloud formations in 3D. Frontier's exascale speeds allowed the team to model a year's worth of global cloud formations in the span of a single day. The Gordon Bell Prize for Climate Modeling will be awarded annually through 2033 to recognize the contributions of climate scientists and software engineers toward improving climate modeling and the understanding of the Earth's climate system.

Besides the Gordon Bell honorees, OLCF systems supported a wide range of projects that tackled major scientific challenges. For example, a team from Georgia State University used Summit to conduct molecular dynamics simulations of the protein complexes that control human cell activity. Data from their study could help drive the quest to treat genetic disorders via such strategies as gene editing. A collaboration between ORNL's Center for Nanophase Materials Sciences and the OLCF, led by researchers from Stonybrook University, used Summit to conduct a coarse-grained molecular dynamics simulation of the million particles that make up a cicada's wing. The findings could lead to the use of nanopillars that replicate the self-cleaning, antibacterial qualities of cicada wings.

Another study led by ORNL and University of Tennessee (UT) researchers used Summit to perform large-scale 3D supernova simulations that offer the most complete picture yet of gravitational waves from exploding stars. The study's predictions of gravitational wave signatures emanating from these supernovae could serve as vital tools to identify these phenomena. A study led by Tokyo Institute of Technology scientists used Summit to simulate the rare isotope oxygen-28 in unprecedented detail. The findings could transform widely held theories of how atomic nuclei are structured.

To maintain a healthy management and workforce structure, the OLCF pursued a multipronged pipeline approach, starting with student programs and followed by on-the-job training programs. These efforts were supported by a focused recruitment campaign developed in collaboration with Human Resources (HR) and ORNL Communications.

The OLCF continued their strong engagement within the community to build the next generation HPC workforce by leading hands-on training sessions and participating in competitions and hackathons to reach audiences that are often underrepresented in the HPC community. OLCF staff were prominent participants at several conferences and events, such as SC, the Practice and Experience in Advanced Research Computing (PEARC) Student volunteer program, and the Tapia Diversity in Computing Conference, where they led workshops and introductory training for students new to HPC.

In CY 2023, the OLCF continued to support and strengthen several strategic engagements with other DOE Office of Science and applied office programs as well as sister federal agencies. External agency collaborating partners included the National Institutes of Health's (NIH) National Cancer Institute (NCI), the US Department of Veterans Affairs (VA), Air Force Weather (AFW), and National Oceanic and Atmospheric Administration (NOAA). These partnerships increased broader community engagement, contributed to the development of new software and capabilities, and supported overall operational innovation for OLCF and the agency partners.

Throughout the year, the OLCF maintained a strong culture of operational excellence, including risk management, workplace safety, and cybersecurity. The OLCF's rigorous risk management strategy anticipated and mitigated risks. Similarly, ORNL and the OLCF were committed to operating under the DOE's safety regulations that ensure a safe workplace. Technical staff tracked and monitored existing threats and vulnerabilities within the OLCF while continually developing tools and practices to enhance operations without increasing the facility's risk. The OLCF carried a strong safety culture through its day-to-day operations, and the facility met the zero accident performance criteria during large-scale workloads and multiple supercomputer installations.

The successful operation of the OLCF is the result of the extraordinary work by staff members in supporting the most capable HPC user facility in the world. The OLCF's staff are instrumental in identifying, developing, and deploying the innovative processes and technologies that support the advancement of science by the OLCF users and benefit the global HPC community.

ES.1 COMMUNICATIONS WITH KEY STAKEHOLDERS

ES.1.1 Communication with the Program Office

The OLCF communicates with the Advanced Scientific Computing Research (ASCR) Program Office through a series of regularly occurring events. These include weekly Integrated Project Team calls with the local DOE ORNL Site Office and the ASCR Program office, monthly highlight reports, quarterly reports, the annual Operational Assessment Report, and an annual “Budget Deep Dive.” Through a team of communications specialists and writers working with users and management, the OLCF produces a steady flow of reports and highlights for sponsors, current and potential users, and the public. See APPENDIX B for a list of science highlights submitted to ASCR.

ES.1.2 Communication with the User Community

The OLCF’s user-targeted communications are designed to relate science results to the HPC community and to help users more efficiently and effectively leverage OLCF systems. The OLCF offers many training and educational opportunities throughout the year for current facility users and the next generation of HPC users, as outlined in Section 1.

The impact of OLCF communications is assessed as part of an annual user survey. In the 2023 annual user survey, OLCF communications received a mean rating for users’ overall satisfaction of 4.5 on a scale of 5.0. The OLCF uses a variety of methods to communicate with users, including the following:

- Weekly email message
- General email announcements
- Automated notifications of system outages
- OLCF website
- Monthly conference calls
- OLCF User Council and Executive Board meetings
- One-on-one interactions with liaisons and analysts
- Office hours
- Social media
- Annual OLCF User Meeting
- Targeted training events (e.g., GPU hackathons or tutorials)

ES.2 SUMMARY OF 2023 METRICS

In consultation with the DOE program manager, a series of metrics and targets was identified to assess the operational performance of the OLCF in CY 2023. The 2023 metrics, target values, and actual results as of December 31, 2023, are noted throughout this report. The OLCF exceeded all agreed-upon metric targets.

ES.3 RESPONSES TO RECOMMENDATIONS FROM THE 2022 OPERATIONAL ASSESSMENT REVIEW

There was one recommendation from last year’s operational assessment.

> *The methodology for tracking research products that result from use of the facilities is not detailed or referenced in the report. Provide additional details about the methods used for tracking scientific accomplishments that resulted from staff support and facility use so that we can better assess its effectiveness and its potential adoption at other facilities.*

This information has been added to this year’s report in Section 8.1.1.2.

2023 1 – User Support Results

HIGH-PERFORMANCE COMPUTING FACILITY
2023 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2024

1. USER SUPPORT RESULTS

CHARGE QUESTION 1: Are the processes for supporting users, resolving users' problems, and conducting outreach to the user population effective?

OLCF RESPONSE: Yes. The OLCF supported 1,676 users and 598 projects (including Quantum) in CY 2023. The OLCF continued to leverage an established user support model for effectively supporting users based on continuous improvement, regular assessment, and a strong customer focus. One key element of internal assessment is the annual user survey. As part of the survey, users are asked to rate their overall satisfaction with the OLCF on a scale of 1 to 5, with a rating of 5 indicating "very satisfied." The mean rating for overall satisfaction with the OLCF in 2023 was 4.52. Overall ratings for the OLCF were positive: 94% of all survey respondents reported being "satisfied" or "very satisfied" with the OLCF.

The OLCF measures its performance by using a series of quantifiable metrics. The metric targets are structured to ensure that users are provided prompt and effective support, and that the organization responds quickly and effectively to improve its support process for any item that does not meet a minimum satisfactory score. The OLCF exceeded all metric targets for user satisfaction in 2023: 97% of tickets were resolved within 3 business days. The OLCF continued to enhance its technical support, collaboration, training, outreach, and communication and engaged in activities that promoted high-performance computing (HPC) and artificial intelligence (AI) to the next generation of researchers.

1.1 USER SUPPORT RESULTS SUMMARY

The OLCF's user support model comprises customer support interfaces, including user satisfaction surveys, formal problem-resolution mechanisms, user assistance engineers, and scientific and data liaisons; multiple channels for stakeholder communication, including the OLCF User Group Executive Board; and training programs, user workshops, and tools to reach and train both current facility users and the next generation of computer and computational scientists. The success of these activities and identification of areas for development are tracked by the annual OLCF user survey.

To promote continual improvement at the OLCF, users are sent surveys that solicit feedback on support services and user experiences at the facility. The 2023 survey was launched on October 9, 2023, and remained open for participation through November 27, 2023. The survey was sent to 1,508 users of the Innovative and Novel Computational Impact on Theory and Experiment (INCITE), Advanced Scientific Computing Research (ASCR) Leadership Computing Challenge (ALCC), and Director's Discretionary (DD) projects. These users also include members of the Exascale Computing Project (ECP) who logged into an OLCF system between January 1, 2023, and September 30, 2023. Although the ECP allocations come from the OLCF DD program, those responses were once again tracked separately from the DD responses. OLCF staff members were excluded from participation. A total of 786 users completed the survey for an overall response rate of 52.1%, as shown in Table 1-1. The results of the 2023 survey can be found on the OLCF website: https://www.olcf.ornl.gov/wp-content/uploads/OLCF-2023-User-Survey-Report.pdf.

Table 1-1. Annual survey response rate for 2023.

Survey response	2022 target	2022 actual	2023 target	2023 actual
Number of users the survey was sent to	N/A	1,396	N/A	1,508
Total number of respondents	N/A	645	N/A	786
Percent responding	N/A	46%	N/A	52.1%

Users were asked to rate their satisfaction on a 5-point scale, in which a score of 5 indicates a rating of "very satisfied," and a score of 1 indicates a rating of "very dissatisfied." The metrics were agreed on by the DOE and OLCF program manager, who defined 3.5 as "satisfactory." The effectiveness of the processes for supporting customers, resolving problems, and conducting outreach is measured, in part, by the key survey responses for User Support in Table 1-2.

Table 1-2. Key survey OLCF user responses for 2023.

Survey question	2022 target	2022 actual	2023 target	2023 actual
Overall OLCF Satisfaction Score on the User Survey	3.5/5.0	4.6/5.0	3.5/5.0	4.5/5.0
Overall Satisfaction with OLCF Support	3.5/5.0	4.6/5.0	3.5/5.0	4.5/5.0
Overall Satisfaction with OLCF Services	3.5/5.0	4.5/5.0	3.5/5.0	4.4/5.0
Overall Satisfaction with the Website	3.5/5.0	4.5/5.0	3.5/5.0	4.4/5.0
Overall Satisfaction with Communications	3.5/5.0	4.6/5.0	3.5/5.0	4.5/5.0
Overall Satisfaction with Problem Resolution	3.5/5.0	4.7/5.0	3.5/5.0	4.6/5.0
Show improvement on results that scored below satisfactory in the previous period	Results will show improvement in at least half of questions that scored below satisfactory (3.5) in the previous period.	No question scored below satisfactory (3.5/5.0) on the 2022 survey.	Results will show improvement in at least half of questions that scored below satisfactory (3.5) in the previous period.	No question scored below satisfactory (3.5/5.0) on the 2023 survey.

1.2 USER SUPPORT METRICS

The OLCF exceeded all user support metrics for 2023. The OLCF metric targets and actual results for user support by CY are shown in Table 1-3.

Table 1-3. User support metric targets and CY results.

Survey area	CY 2022 target	CY 2022 actual	CY 2023 target	CY 2023 actual
Overall OLCF satisfaction rating	3.5/5.0	4.6/5.0	3.5/5.0	4.5/5.0
Average of all user support services ratings	3.5/5.0	4.5/5.0	3.5/5.0	4.5/5.0

1.2.1 Average Rating across All User Support Questions

The calculated mean of answers to the user support services–specific questions on the 2023 survey was 4.5/5.0, thereby indicating that the OLCF exceeded the 2023 user support metric target and that users have a high degree of satisfaction with user support services. Users were asked to provide ratings of their satisfaction with support received from the wide variety of OLCF support and services. Respondents were highly satisfied with training (97%), communications (93%), documentation (93%), the projects and accounts team (96%), user assistance (95%), and the INCITE liaisons (94%).

Included below are select open-ended responses to "What are the best qualities of the OLCF?" that highlight various aspects of user support:

- "The OLCF clearly demonstrates that it cares about its users. Top-notch documentation, service, training, help. At every turn, the OLCF makes us feel heard and taken care of, combining cutting-edge compute with an even more impressive human element."
- "Top-flight computational power, reasonable queue times for the majority of the year, excellent documentation, excellent uptime."
- "The OLCF's leadership-class computing resources are crucial to conducting large-scale [density functional theory (DFT)] calculations on technologically important materials science problems. One key quality of OLCF is the good availability of the computing resources and high-quality technical support for troubleshooting and application performance related issues."
- "As incredible as the systems themselves are, the staff is far and away the most valuable resource at the OLCF. From leadership to liaisons to support folks to admin staff, everyone is always a great pleasure to work with and has enabled our group to take on problems we would not normally dream of attempting elsewhere."
- "OLCF enables leadership-scale high-fidelity simulations that are not possible on any other systems to improve our capabilities to analyze and design next generation aerospace vehicles. The hardware, software, and systems are the best and latest in the world, pushing the frontiers of HPC for the future."
- "OLCF remains by far the best computing facility I've ever used. Part of that is of course due to the hardware, which is without question world-leading. But what strikes me as its most valuable asset is the high quality of its staff, and as a result of that, all the documentation and training materials that are made available to users, the handling of issues, etc."

1.2.2 Improvement on Past Year Unsatisfactory Ratings

Each year, the OLCF works to show improvement on no less than half of any questions that scored below satisfactory (3.5) in the previous year's survey. However, all questions scored above 3.5 on both the 2022 and 2023 surveys. Although no results scored below satisfactory on the 2023 survey, a thorough review of the survey identified areas that could be improved or in which new services could be added to enhance the user experience at the OLCF. In some cases, the issue had already been addressed or a solution is in development and forthcoming.

- Data transfer: In the 2022 user survey, users identified data transfer at the OLCF as an area that could improve. In 2023, 13 new and faster data transfer nodes were purchased, and the process of upgrading Globus to Version 5 began to enable faster user data transfer. Data transfer documentation was clarified and updated, and data transfer examples were added in the *Frontier User Guide* to ensure that users would understand how to move data to the new Orion file system.

Additionally, the OLCF hosted two training events on the OLCF data ecosystem in 2023. The first took place on the July OLCF user call and was attended by 54 users. The second took place in August as part of the Frontier Training Workshop and was attended by 89 users. Both sessions described the OLCF data ecosystem and provided examples and best practices for using Globus, HTAR, and rsync to move data between resources using the center's data transfer nodes. These training sessions were recorded and posted in the training archive.

- Julia: Users requested Julia training in 2022 and 2023. Julia is an open-source general programming language that can be used for HPC and is interoperable with C, C++, and FORTRAN. The OLCF held a Julia for HPC tutorial in May 2023 that was open to OLCF, National Energy Research Scientific Computing Center (NERSC), and Argonne Leadership Computing Facility (ALCF) users. The 3-hour training session featured an introduction to Julia; guided exercises that allowed users to explore developing and running a Julia HPC mini-application with Threads, CUDA, and AMDGPU programming models; and parallel I/O using ADIOS2. The training utilized Jupyter and Pluto.jl notebooks for interactive computing. The event was popular, with 97 users out of the 110 registered in attendance.

- AI training: In 2023, the OLCF launched a new AI training series. Based on specific feedback from the 2022 user survey, the AI training series covered how to use specific machine learning and deep learning tools on OLCF systems. Additional details on the well-attended series are listed in Section 1.4.4.5.

- Data visualization: In 2023, the OLCF extended the Data Visualization and Analytics (DVA) training series by offering two events that highlight new tools on OLCF systems (Blender and Ascent). The DVA training series was created in 2022 and was expanded in 2023 due to feedback from the user survey. Series topics are listed in Section 1.4.4.7.

- Queue: Based on user feedback, the batch queue limits on the center's analysis cluster, Andes, were adjusted to better serve the system's workflows. The batch queues are monitored and adjusted where possible to better meet center workflow needs.

- Ability for quick video chat, office hours: Based on user feedback that requested the ability to discuss more expansive topics beyond Frontier, the bi-weekly office hours were expanded to cover all OLCF resources. Office hours are offered each Monday and Wednesday. Each office hour is limited to 5 different topics, and each one is assigned its own breakout room in Zoom. To request an office hours meeting, users simply sign up through the MyOLCF portal, which was also a new addition deployed in 2023, as described in https://docs.olcf.ornl.gov/support/index.html#olcf-office-hours. In the sign-up form, the project team must include information about the topic and details about the issues they would like to discuss. This helps OLCF staff coordinate with vendors to ensure someone knowledgeable in the subject can be available.

- Programming environment timeline: Changes to a system's programming environment can impact user workflows. To help document changes to Frontier's programming environment, changes are recorded and sorted by date within the Frontier user guide: https://docs.olcf.ornl.gov/systems/frontier_user_guide.html#system-updates. The timeline allows details of a programming environment change, including software versions and recompile requirements, to be collected in a single location. Placing this information in the Frontier user guide also makes it easier to locate the change timeline. Users are notified of changes to the programming environment through the weekly user email. The timeline allows user email notifications to be brief, thereby increasing the chance that the notice will be read.

1.3 PROBLEM RESOLUTION METRICS

The following OAR metrics were used for problem resolution.

- At least 80% of user problems are addressed (i.e., the problem is resolved or the user is told how the problem will be handled) within 3 business days.
- According to the user survey, average satisfaction ratings for problem resolution are "satisfactory" or better.

1.3.1 Problem Resolution Metric Summary

In most cases, the OLCF resolves reported problems directly by identifying and executing the necessary corrective actions. Occasionally, the facility receives problem reports that it may not be able to resolve because of factors beyond the facility's control. In such scenarios, addressing the problem requires OLCF staff to identify and carry out all corrective actions at their disposal for the given situation. For example, if a user reports a suspected bug in a commercial product, then prudent measures might be to recreate the issue; open a bug or ticket with the product vendor; provide the vendor with the necessary information about the issue; provide a workaround to the user, if possible; and track the issue to resolution with the product vendor, which may resolve the issue with a bug fix or workaround acknowledgment.

The OLCF uses Jira to track user-reported issues to ensure response goals are met or exceeded. Users can submit tickets in a variety of ways, including by email, telephone, and an online request form. Email remains the most common method. During CY 2023, the OLCF issued 3,619 tickets in response to user inquiries. The OLCF resolved 97% of issues within 3 business days, as shown in Table 1-4.

Nearly three-quarters (74.1%) of survey respondents submitted between one and five queries to the OLCF (via phone or email) in 2023. 95% of respondents were satisfied or very satisfied with OLCF's response to reported issues, with similarly high ratings for the *timeliness of responses to reported issues* (94% satisfied) and the *quality of technical advice given* (95% satisfied).

Table 1-4. Problem resolution metric summary.

Survey area	CY 2022		CY 2023	
	Target	Actual	Target	Actual
Percentage of problems addressed in 3 business days	80%	96%	80%	97%
Average of problem resolution ratings	3.5/5.0	4.7	3.5/5.0	4.6

1.4 USER SUPPORT AND ENGAGEMENT

The following sections discuss key activities and contributions in the areas that the OLCF recognizes as pillars of user support and engagement, including:

- A user support staff made up of account management liaisons, User Assistance engineers, Scientific Engagement Group liaisons, data liaisons, and visualization liaisons
- Multiple pathways to communicate with users, sponsors, and vendors
- Developing and delivering training to current and future users,
- Strong outreach to engage the next generation of HPC users, the external media, and the public.

1.4.1 User Support

The OLCF recognizes that users of HPC facilities have a wide range of needs that require diverse solutions—from immediate, short-term, trouble ticket–oriented support (e.g., assistance with debugging and optimizing code) to more in-depth support that requires total immersion and collaboration on projects.

The facility provides complementary user support vehicles that include user assistance and outreach staff; liaisons in scientific, data, and visualization areas; and computer scientists who assist with issues related to the programming environments and tools. The following sections detail some of the high-level support activities during CY 2023 and the specific OLCF staff resources available to assist users.

1.4.2 User Support Team

The OLCF addresses user queries; acts as user advocates; covers frontline ticket triage, resolution, and escalation; provides user communications; develops and delivers training and documentation; and installs third-party applications for use on the computational and data resources. The team also manages the OLCF's Resource and Allocation Tracking System (RATS), which is the authoritative source for most of the system, user, and project data at the OLCF. In addition to some of the initiatives already mentioned in the section above, some examples of OLCF initiatives in 2023 that helped improve the overall user experience are provided below, although some of them are very much behind the scenes.

1.4.2.1 MyOLCF

The OLCF provides users with a web-based self-service application called myOLCF (https://my.olcf.ornl.gov) that offers principal investigators (PIs) and project members timely and accurate data to empower decision-making for OLCF projects and self-service tools for project administration. MyOLCF makes relevant information about projects, users, project membership applications, project applications, resource allocations, help tickets, and project usage analytics available to PIs and project members.

MyOLCF also allows PIs and project members to perform administrative tasks without contacting the OLCF Accounts Team. Documentation on myOLCF can be found on the publicly available OLCF user documentation site. The software application is under continuous development, and new features and user experience improvements are deployed twice a month on average.

Highlights for myOLCF in 2023 include the following:

- Added office hour scheduling front end for users to request office hours with OLCF staff
- Enabled users to apply to *future* projects (projects defined but not yet enabled)
- Integrated Frontier information into application workflow
- Developed SummitPLUS proposal front end
- Added various other improvements to help with user interaction with the facility

1.4.2.2 RATS CRM

The OLCF's customer relationship management software, called RATS CRM, is under continuous development. RATS CRM provides the ability for the OLCF Accounts Team to process user and project applications, authorize user access to systems, track project allocations on systems, and manage certain configurations for HPC resources.

A big advancement this year was the migration of the RATS CRM application to OLCF's Kubernetes-based infrastructure, Slate. This has greatly improved the developer experience and agility for deploying new code.

Notable updates to RATS CRM in 2023 include:

- Integrated new file system policies for directory trees in support of Frontier's Orion file system
- Integrated Office Hour business logic to facilitate scheduling of Office Hours between users and OLCF staff
- Added new *health-check* API to improve RATS and MyOLCF integration
- Various other improvements to business processes and system integrations in support of the facility

1.4.2.3 RATS Report

RATS Report is vital to the OLCF's internal reporting and utilization tracking capabilities. It consists of various data pipelines, a data store, and a web-based front end for collecting information and providing dashboards for OLCF leadership. RATS Report also tracks system availability and allocation usage across programs for all OLCF compute resources and file systems.

Notable RATS Report highlights for 2023 include:

- Incorporated Frontier metrics
- Incorporated SummitPLUS metrics
- Implemented collection and display of usage metrics for new storage systems: Themis, the IBM Spectrum Archive system for the National Center for Computational Sciences (NCCS) open enclave; and Orion, the Lustre parallel file system for Frontier
- Made various improvements to data pipeline infrastructure

1.4.2.4 Advanced Plant Phenotyping Laboratory

The OLCF continues to work with data engineers from the Center for Bioenergy Innovation (CBI) in the Biosciences Division at ORNL to improve their data management practices and analysis capabilities. This work began in 2022 with the goal of creating an open data portal service to provide valuable images and data from the Advanced Plant Phenotyping Laboratory at ORNL—along with various data products from partner laboratories and collection sites—to their worldwide user base. These data portal services are now live within the ORNL private network, and the OLCF and CBI data management team are working to enable external access to these services. The OLCF has implemented a DevOps engine that provides automated continuous integration/continuous development processes and a GitLab-centric developer workflow for CBI data engineers. The project leverages the OLCF's OpenShift Kubernetes cluster, Slate, for the deployment of containerized web services, databases, and analysis scripts and uses the center's nearline storage system, Themis, for centralized data storage. The CBI data engineers are now working on new analysis scripts that will expand the number of data products they can provide to users around the world. Slate allows the project to leverage automated and on-demand compute resources to perform data analysis and processing. The centralized data storage and analysis platforms allow the CBI to share data with users in a faster and more efficient manner, enabling valuable science. The OLCF is uniquely positioned to provide the infrastructure, platforms, and support required to create the suite of integrated research services envisioned by the CBI.

1.4.2.5 OLCF Quantum Computing User Program

In 2023, the OLCF enhanced its strategic capability to offer users access to cutting-edge quantum computing (QC) resources to meet growing demand and support discovery and innovation in scientific computing. The program helps users understand QC's unique aspects and challenges, thereby enabling them to develop and test quantum algorithms on the increasingly capable systems available. The initiative, which includes research from all DOE Office of Science programs such as ASCR, Basic Energy Sciences (BES), Biological Environmental Research, High-Energy Physics, Fusion Energy Science, and Nuclear Physics, reflects a collaborative effort to expand the OLCF's Quantum Computing User Program (QCUP) with significant staff contributions from many areas inside and outside of the OLCF.

The following lists some of the notable 2023 QCUP efforts:

- The team renewed contracts with three QC hardware vendors (IBM, Rigetti Computing, Quantinuum/Honeywell) and added a fourth vendor, IonQ, to the roster of QC resources available to QCUP users. The team collaborated with vendors to maintain and improve the multistep processes for issuing user and project accounts, monitoring allocations, establishing and enforcing user agreements, tracking and reporting usage cases, updating hardware/software offerings, and providing user support.
- The external documentation site (https://docs.olcf.ornl.gov/quantum/index.html) was maintained and frequently updated to provide user support documentation.
- The fourth annual Quantum Computing User Forum was held in-person at ORNL in August and brought together over 100 attendees to discuss QC advancements and allow the opportunity for QCUP users to highlight their projects. The event featured 21 talks on QC-enabled research; panels; expert-led workshops from IBM Quantum, Quantinuum, and NVIDIA; and a poster session with international participation.
- The inaugural Quantum on the Quad event was held in November to facilitate lab-wide strategy discussions about ORNL's QC roadmap.
- In collaboration with NVIDIA, the OLCF hosted a training workshop during the Quantum Computing User Forum on how to use the CUDA Quantum tool on the OLCF Ascent and Summit HPC systems. This workshop demonstrated a pathway for how HPC users can utilize multiple CPUs, GPUs, and compute nodes to run a quantum simulator.
- OLCF staff members are working closely with the QCUP vendors to establish the infrastructure needed to further support hybrid QC+HPC computations at the center. By collaborating in software development of a given vendor's quantum library, the OLCF can guide how the vendor's API may interface with an HPC environment. Documentation on HPC quantum libraries was expanded to include multiple QCUP vendors (https://docs.olcf.ornl.gov/quantum/quantum_software/hybrid_hpc.html).
- The OLCF supported 80 total QCUP projects and 271 QCUP users in 2023 and tracked 52 publications that resulted from the use of these resources.

1.4.2.6 Jupyter at OLCF

The OLCF's Jupyter provides a critical function that allows OLCF users to perform data analysis and visualization tasks at a scale that does not require Summit's or Frontier's compute resources. Jupyter allows a user to simply browse to a URL and, by using their SSO token, launch a lab instance that grants them access to a prebuilt environment of various AI/machine learning/programming tools, a terminal for OLCF network access, and the ability to run data visualizations through a notebook. Jupyter exists on both the OLCF’s Open and Moderate enclaves, allowing users to access the file systems and computing resources associated with each one.

Jupyter launched in 2021 and saw significant use. The system has continued building a strong user base over 2022 and into 2023.

Notable Jupyter highlights include:

- Combined overall 2023 usage saw 424 unique users launch 15,436 Jupyter sessions
- Moderate enclave usage saw a 10% increase in unique users and a 43% increase in launched sessions, indicating a growing and active user base
- Open enclave usage saw a 40% increase in unique users and a 20% increase in launched sessions
- Open and Moderate combined to serve 351 training labs, which was a 37% increase over 2022 and provided a more seamless experience for HPC crash course students and users

Internal changes to Jupyter have also made it a more robust platform. Jupyter was updated to the RedHat Universal Basic Image 9, and the service proxy and database were updated to allow for a more consistent user experience. The updates also added multiple AI/machine learning packages and updated existing ones, notably PyTorch and TensorFlow, in response to user input.

1.4.2.7 Center-wide Scratch File System Transition

In 2023, the OLCF center-wide scratch file system migrated from Alpine (General Parallel File System [GPFS]) to Orion (Lustre). The transition required OLCF users to migrate data and workflows from Alpine to Orion prior to the decommissioning of Alpine in January 2024. Based on lessons learned from previous file system migrations, the OLCF created a data migration strategy and user notification campaign. To ensure all users were aware of the migration timeline, a web page was created to advertise the details of the timeline and migration procedures. The link was provided in concise email messages sent to the user community. Email messages were sent over time to all users through the weekly email communication. Direct emails were also sent to specific user groups (e.g., users with large amounts of data and users with purge protection). The notification campaign and procedures allowed the center and user community to successfully migrate workloads from Alpine to Orion before the January 1, 2024, deadline.

1.4.2.8 Scientific Liaisons

Scientific Liaisons—experts in scientific domains and computation—partner with facility users to obtain optimal scientific results from the OLCF’s computational resources and systems. The Scientific Liaisons include experts in chemical and materials sciences; nuclear physics such as nuclear structure and quantum chromodynamics; high-energy physics such as particle physics; astrophysics such as stellar evolution and cosmology; and climate science, geophysics, biology, biomedical sciences, and engineering. Some unsolicited user feedback on the Scientific Liaisons is provided below:

- “Our liaison, Mark Berrill, has been very helpful in answering questions, providing advice on portability/optimizations, and quickly diagnosing issues.”
- “Wayne Joubert’s assistance has been indispensable to augment CoMet with new features that are better suited for the climatype identification use case.”
- “Our INCITE Liaison (Antigoni Georgiadou) participates in our weekly calls and has been very helpful.”

- “The comprehensive simulations of the two burning plasmas ITER scenarios are the largest electromagnetic gyrokinetic simulations to date and thus required access to the Leadership Computing Facility (LCF) with the exascale speed, large memory, and fast I/O. Support by the INCITE liaison (Dr. Isaac Lyngaas) and ADIOS-2 group (Dr. Scott Klasky) has been extremely important. Dr. Isaac Lyngaas is also instrumental in porting GTC to the new exascale computer Frontier at OLCF.”
- “We are fortunate to have OLCF’s Matt Norman as both our INCITE project liaison and a team member under our ECP project. With Matt Norman’s assistance combined with the many resources of the ECP project, we have been able to explore a wide variety of techniques for getting maximal performance out of the Summit GPUs and Frontier GPUs.”
- “We are very grateful for the effectiveness and helpfulness of our [Center for Accelerated Application Readiness (CAAR)]/INCITE liaison Dr. Stephen Nichols. Dr. Nichols has been highly engaged every week for more than 3 years straight. He has played a vital role in helping us understand the details of HIP interfaces (for our Fortran codes), of rocfft work buffers, CEE and rocm library versions, and of memory monitoring on the GPUs. He also deserves much credit for the discovery that slab codes give better performance than pencils codes in our work. The quality of Dr Nichols’s work and the depth of his commitment as our project liaison demonstrates the importance of the OLCF and DOE’s investment in this type of PI-staff relationship for demanding projects.”
- “We would like to thank Balint Joo for his support with the developments related to HemeLB.”

Other liaison highlights include the following:

- Antigoni Georgiadou organized the "First Workshop on Enabling Predictive Science with Optimization and Uncertainty Quantification in HPC" as part of the SC23 conference.
- Balint Joo has engaged members of the computational biomedical community and has partnered with the team of Prof. Peter Coveney at University College London to work on simulation aspects related to human digital twins. In particular, he has assisted in porting the HemeLB code to model blood-flow to Frontier, getting it built, and assisted in initial benchmarking. The work enabled the team to not only use Frontier but assisted them in their migration to Aurora as well. Balint has also engaged with the group of Amanda Randles at Duke University and assisted in the running of benchmarks for the Harvey code on Crusher and Frontier to be used in a paper for the P3HPC workshop at SC23.
- Reuben Budiardja, a computational scientist in the Science Engagement Advanced Computing for Nuclear Particle and Astrophysics group, worked with a research group led by Spencer Bryngleson of Georgia Tech to port the group’s multicomponent flow code software to Frontier. Budiardja engaged with the group initially as they requested assistance through the OLCF Office Hours to overcome compiler issues with OpenACC. The group then attended an OLCF hackathon, which let them implement further optimization with the additional help from OLCF Center of Excellence liaisons Steve Abbots (HPE) and Brian Cornille (AMD). The final help came from Budiardja’s rewriting of the code I/O module, allowing the code to run at scale on over 9,000 Frontier nodes. With this success, the group is now preparing a proposal to compete for the next round of INCITE allocations. Reuben also assisted a team from NIST to port and run the Fire Dynamics Simulator code on Summit. Budiardja assisted the team in building the code, rectifying several compatibility issues, and providing the dependency library PETSc to enable their solver to use GPUs more effectively.

Frontier CAAR activities in Science Engagement that contributed to the successful Frontier CD4 project:

- Gustav Jansen: Nuclear Coupled-Cluster Oak Ridge (NuCCOR), a DOE NP-funded ab initio nuclear structure software suite developed primarily at ORNL, completed the OLCF5 CAAR project in 2023 and achieved a speedup of over 6× going from Summit to Frontier. Gustav R. Jansen, as the scientific liaison from the NAP group in NCCS, led the work by extracting performance-sensitive algorithms into a domain-specific library named the Nuclear Tensor Contraction Library (NTCL) that presents an architecture-independent interface to NuCCOR. NTCL contains hardware-specific plugins that provide a future-proof design that minimizes the necessary porting work for future resources at the OLCF.
- David Rogers: Particle-In-Cell on GPUs (nPIConGPU) uses a particle-in-cell model to simulate time-dependent electron motion in laser-driven plasmas. These simulations inform the design of next-generation, high intensity, focused lasers. At each timestep, macroparticles that represent electrons interact with a local electromagnetic field, and then the field within each cell is updated following Maxwell's equations. Algorithmic and hardware speedups combined to improve PIConGPU's throughput from 14.7e12 updates per second in a full-scale Summit run (late 2019) to 65.7e12 updates per second on Frontier (9,216 nodes in July 2022), thereby achieving 90% weak scaling efficiency and 4.5× higher simulation throughput.
- Arnold Tharrington: The OLCF5's NAMD, or Nanoscale Molecular Dynamics, CAAR effort was a collaboration between NCCS's Scientific Engagement Section, the University of Illinois at Urbana-Champaign's (UIUC's) Theoretical Biophysics Group and the Parallel Programming Laboratory with the goal of porting NAMD source code to Frontier's AMD GPUs. The OLCF liaisons were Josh Vermaas and Arnold Tharrington, and their liaison activities primarily consisted of the initial porting of NAMD to AMD GPUs by *Hipfying* NAMD source code, running benchmarks on Frontier and Summit, coordinating and leading the meetings between AMD and UIUC developers, and creating all final reports for CAAR documentation.
- Mark Berrill worked with the PI for Lattice Boltzmann Methods for Porous Media, or LBPM, to port the application to Frontier as part of the CAAR program. This entailed porting CUDA code to HIP, optimizing performance on Frontier, and performing scaling studies on Frontier to compare with the initial performance on Summit. Additionally, Mark made several improvements to the code as part of the porting task, including improvements to the communication, support for halo size >1, and improvements to the load balancing and analysis routines. The application port successfully met the target goal of a 4× speedup over Summit by achieving an actual speedup of 4.6× when running on 50% of both machines.
- Reuben Budiardja, a computational scientist in the Science Engagement section, was the OLCF liaison for the CAAR project Cholla. Budiardja's contribution to the project included reworking the fluid evolution of the code to be fully GPU resident, removing the need for data movement, and exploiting the unique Frontier architecture with its GPU-connected network interface for more efficient Message Passing Interface (MPI) communication. This work, combined with other algorithmic improvements by the team, enabled Cholla to achieve over a 20× speedup compared with its baseline run on Summit. The experience and lessons learned in porting and optimizing the suite of applications in the OLCF's CAAR program were documented and published in a paper for the 2023 ISC High Performance Computing conference.
- Stephen Nichols worked with the PIs on the CAAR-GESTS team to build two 3D fast Fourier transform (FFT) kernel codes (only performs the 3D FFT); two Pseudo-Spectral Direct Numerical Simulation (PSDNS) codes for use on Frontier; one set of codes for a 1D decomposition (the so-called "Slabs" code); and one set of codes for a 2D decomposition (the "Pencil" code). The two 3D FFT kernel codes were critical for developing an efficient algorithm without the overhead of the DNS-specific components. The two PSDNS codes were built on top of their respective 3D FFT kernel and have been tested on Frontier with problem sizes up to $32{,}768^3$ points (i.e., 32,768

points in all three coordinate directions) on 4,096 nodes in single precision and 8,192 nodes in double precision. Over the lifetime of the CAAR-GESTS project, Stephen worked closely with the Frontier Center of Excellence HPE and AMD staff members to identify and correct bugs in the compilers and the ROCM software and to improve performance in all aspects of the PSDNS algorithms by using sbcast and NVMe for reliable large-job submissions, OpenMP offloading for efficiency, fast and correct rocFFT execution, reliable and safe I/O, and efficient MPI all-to-all communications. CAAR-GESTS successfully completed an INCITE 2023 project and is starting its INCITE 2024 project.

- Markus Eisenbach: The Locally Self-Consistent Multiple Scattering (LSMS) code solves the Kohn-Sham equation for electrons in a solid by using the multiple scattering theory. By employing a real space formulation with a finite cutoff range for the maximal scattering distance, the code can achieve weak scaling in the number of atoms in the simulated system from single nodes to the full HPC system. The computations are dominated by dense linear algebra operations on double-precision, complex non-Hermitian matrices. As part of the CAAR effort, the CUDA version was ported to HIP to utilize the MI250x GPUs on Frontier. This port utilized routines provided by the rocBLAS and rocSolver libraries as well as HIP GPU kernels that transform the LSMS problem into a form suitable for the linear solver for operations that do not map readily onto standard library routines. By employing kernel profiling, the team was able to identify bottlenecks in the first implementation of these kernels. These bottlenecks were related to integer index and address calculations that interfered with the floating-point operations on the MI250x GPUs. Rearranging these operations significantly improved performance. The improvements to LSMS resulted in an ~7.5× improvement in the per-GPU performance when comparing one MI250x GCD on Frontier to a V100 on Summit. The team demonstrated weak scaling and maintained this performance up to calculations for a 1,048,576-atom system on 8,192 Frontier nodes. The CAAR effort has enabled the use of LSMS on Frontier for production calculations of electric conductivity and disorder in alloy systems as part of INCITE allocations.

1.4.3 OLCF User Group and Executive Board

1.4.3.1 OUG Monthly Webinar Series

PIs and users on approved OLCF projects are automatically members of the OLCF User Group (OUG) and remain members for 3 years following the conclusion of their OLCF project. The OUG meets 10 times a year via the OLCF User Conference Call webinar series. The OLCF User Call provides users with a forum to discuss issues and ask questions about OLCF resources and offers training on timely topics to start the discussion.

Table 1-5 contains a list of meetings that occurred in CY2023.

Table 1-5. OLCF User conference call webinar attendance.

Topic	Date	Speaker	Participants
Containers on Summit	January 25, 2023	Subil Abraham (OLCF)	112
Cybersecurity Best Practices	February 22, 2023	Ryan Adamson (OLCF)	77
Checkpointing Best Practices for Frontier	March 29, 2023	Scott Atchley (OLCF)	139
AI for HPC	April 26, 2023	Arjun Shankar, Junqi Yin, Wes Brewer (OLCF)	162
OLCF Storage and Orion Best Practices	May 31, 2023	Suzanne Parete-Koon (OLCF)	107
Blender on Frontier	June 28, 2023	Michael Sandoval (OLCF)	79
Data Transfer Overview	July 26, 2023	Suzanne Parete-Koon (OLCF)	54
Overview of Frontier Documentation	August 30, 2023	Subil Abraham (OLCF)	64
In Situ Visualization with Ascent	September 27, 2023	Cyrus Harrison, Nicole Marsaglia (LLNL)	73
Quantum Computing User Program	December 06, 2023	Michael Sandoval (OLCF)	84

1.4.3.2 OUG Executive Board

The OUG is represented by the OUG Executive Board. This board meets monthly for in-depth discussions with OLCF staff to provide feedback and guidance on training topics as well as the facility's resources and policies. OUG Executive Board terms are 3 years and are staggered so that 3 new members are elected each year. Additionally, an outgoing chair remains on the board for one year as an ex officio member if the term as chair is their third year on the board. Steven Gottlieb (Indiana University) replaced Eric Nielsen (NASA) as chair for the 2023–2024 term. Mike Zingale (State University of New York Stony Brook), Emily Belli (General Atomics), and Scott Callaghan (University of Southern California) began new 3-year terms that will conclude in 2026. The current Executive Board is listed at https://www.olcf.ornl.gov/about-olcf/oug/.

1.4.3.3 OLCF User Meeting

The 19th annual OLCF User Meeting was held on October 17–18, 2023. The meeting had in-person attendance for the first time in 3 years as well as a remote participation option for many events. The 2-day event featured presentations from OLCF staff and the OLCF user community and covered user experiences, Frontier updates and tips, an OLCF-6 update, data management talks, Summit updates, and many more topics. Day one concluded with a poster session, and each day included working lunch sessions that provided an opportunity for attendees to discuss multiple topics with OLCF staff. Topics during the lunch sessions included data storage, training, libraries, quantum, Jupyter, and many others. The meeting drew 118 participants in person and 42 virtually, including users and OLCF staff.

See https://www.olcf.ornl.gov/calendar/2023-olcf-user-meeting/ for more information.

1.4.4 Training

For 2023, the OLCF focused many of its training series and workshops on enabling users to become more productive on Frontier while still supporting the learning needs of those users doing work on Summit. The OLCF training program provides the user community with general HPC training as well as special topics needed to fully leverage the facility's cutting-edge HPC resources. In most cases, the training events are recorded, and the slides, recordings, and hands-on materials are made available to users through the OLCF training archive: https://docs.olcf.ornl.gov/training/training_archive.html. The OLCF continued to collaborate on training series and workshops with other HPC centers such as NERSC and the ALCF to maximize the benefit for the often-shared user bases and share the load of course development. Many training events were provided or enhanced by contributions from HPE, AMD, and NVIDIA staff.

In 2023, OLCF users had the choice of attending 33 training events—2 hackathons, 23 hosted by the OLCF and another 8 jointly hosted by the OLCF and either NERSC, the ALCF, or both.

The OLCF aimed to ensure that users onboarded throughout the year were never far from a training session that would help them get started on Frontier and at the OLCF. Eight events throughout the year focused on teaching users how to use the new Frontier system. Two of those events were three-day workshops that covered every aspect of using Frontier, two were hackathons, and the remaining four were user conference calls that covered timely updates about using the file systems, checkpointing, and documentation. In addition to these Frontier-specific events, the OLCF held regular Monday and Wednesday office hours throughout the year. These office hours were staffed by experts from the OLCF and the HPE/AMD Center of Excellence, and users could sign up to work on any problems they encountered while using any of the OLCF's systems. The OLCF also held a new user training workshop for Summit users.

The OLCF extended many of its series from 2022 into 2023—notably, the OUG Monthly Webinar Series and the Data and Visualization series were continued. New this year were the AI training series and performance portability series. The portability series also featured HIP and OpenMP target Offload mini-series that were well positioned to help Summit and other HPC system users modify their code bases for Frontier. The OLCF also had a Julia for HPC workshop and a DOE Cross-Facility Workflows training that were open to OLCF, ALCF, and NERSC users.

Survey respondents in 2023 rated the training at 4.5 out of 5.0, and, again, 97% were highly satisfied with the training overall. Usefulness of the online training archive and the quality of the training content received a high rating (96%), whereas the lowest rated aspect was the breadth of training events offered at 90%.

Notable highlights from training sessions are provided below.

1.4.4.1 Frontier Training Workshops

In 2023, the OLCF had two virtual 3-day workshops designed to train users how to use the Frontier exascale system. The virtual format was chosen to make the training easily accessible to as many users as possible. Sessions were hosted by AMD, HPE, and OLCF staff and covered important aspects of how to start running applications on Frontier and included overviews of the hardware, programing environment, programming models, job scheduler, and software and tools most commonly used for AI and HPC modeling and simulation on Frontier. The agenda, recordings, and slides are available in the links in Table 1-6 and at https://docs.olcf.ornl.gov/training/training_archive.html.

Table 1-6. Frontier training workshop attendance.

Topic	Date	Participants
Frontier Training Workshop	February 15–17, 2023	174
Frontier Training Workshop	August 23–25, 2023	89

1.4.4.2 Summit New User Training

The OLCF hosted a [Summit New User Training] that described best practices for using the OLCF, its clusters, and its data ecosystems. The training also covered examples and important aspects of using Summit, such as how to use the IBM Load Sharing Facility (LSF) resource sets to efficiently target Summit's CPUs and an GPUs. Summit users also were provided with an overview of the Summit Quick Start Guide, which covers documentation and recordings of past training sessions organized by topic. Thirty-seven users registered for this training and 24 attended. The training was recorded so that it could be viewed by users who joined later in the year.

1.4.4.3 Frontier Hackathons

OLCF hosted two Frontier hackathons in 2024—one in June and one in October. During the events, mentors drawn from the OLCF, AMD, HPE, and the ECP software and tools teams helped users solve a variety of challenges that they were having with running efficiently on Frontier. Those challenges included porting codes to HIP or OpenMP target offload for GPU acceleration, debugging, and profiling and optimization. Teams were selected by the pool of mentors based on how well their code development goals could be achieved during the event. Ten teams were selected for the June hackathon. These teams represented two INCITE projects, three DD projects, two ALCC projects, one CAAR, one ECP project, and one external collaboration project with LUMI. Eight teams were selected for the October hackathon, and these teams represented four INCITE projects, one ECP project, and three DD projects. The hackathons were held over Zoom to make them accessible to the remote users.

1.4.4.4 Performance Portability Training Series

The OLCF and NERSC collaborated to bring their users a performance-portability training series that consisted of nine events in 2023 (

Table 1-7). For large-scale simulation, the portability of a code base is becoming more important due to the variety of architectures being introduced in current and emerging HPC systems. For example, NERSC's Perlmutter, OLCF's Frontier, and ALCF's Aurora feature NVIDIA-, AMD-, and Intel-based GPUs, respectively. Performance-portable programming solutions have emerged to help ease developer transitions between such systems. Such solutions, for example, use a single interface to interact with multiple underlying programming models (e.g., CUDA, HIP, OpenMP). This training series informed users on currently available performance-portable programming solutions.

Within the performance-portability series, AMD presented a multipart HIP training intended to help new and existing GPU programmers understand the main concepts of the HIP programming model and the AMD GPU platform. Each part included a 1-hour presentation and example exercises. Events pages are linked in the table below.

Table 1-7. Performance-portability training series attendance.

Topic	Date	Primary Host	Participants
Advanced SYCL Techniques & Best Practices	May 23, 2023	NERSC	41
Intro to GPUs and HIP	August 14, 2023	OLCF	85
Porting Applications to HIP	August 28, 2023	OLCF	32
AMD Memory Hierarchy	September 18, 2023	OLCF	47
Introduction to OpenMP Offload Part 1: Basics of Offload	September 29, 2023	Both	69
GPU Profiling (Performance Timelines: Rocprof and Omnitrace)	October 2, 2023	OLCF	33
Introduction to OpenMP Offload Part 2: Optimization and Data Management 2023	October 6, 2023	Both	50
RAJA	October 10, 2023	OLCF	11
GPU Profiling (Performance Profile: Omniperf)	October 16, 2023	OLCF	33

1.4.4.5 AI Training Series

In 2023, the OLCF launched a new AI training series. Based on specific feedback from the 2022 user survey, the AI training series covered how to use specific machine learning and deep learning tools on OLCF systems. The series is ongoing (with tentative events planned for 2024), but the 2023 portion of the series spanned from April through October with four events. Although the focus was on OLCF systems, two AI training events were also open to non-OLCF users, thereby allowing NERSC users to participate in the "AI for Science at Scale – Introduction" and "SmartSim at OLCF" training events. Recordings and slides of the four events were made available on the OLCF Training Archive. Links to each of the event pages given in Table 1-8.

Table 1-8. AI training series attendance.

Topic	Date	Participants
AI for HPC Simulation Campaigns	April 26, 2023	162
AI for Science at Scale – Introduction	June 15, 2023	101
SmartSim at OLCF	July 13, 2023	78
AI for Science at Scale – Part 2	October 12, 2023	117

1.4.4.6 Quantum Training

In 2023, the OLCF started its QC training effort by offering two training events on how OLCF users can access QC resources and use them on the HPC systems. The first event highlighted the CUDA Quantum tool on the Ascent system during the 2023 Quantum Computing User Forum. The second event, an Overview of the QCUP, highlighted how existing OLCF users can apply for QC resources and routes for HPC+QC workflows.

1.4.4.7 Data Visualization and Analytics series

In 2023, the OLCF extended the DVA training series by offering two events (

Table 1-9) to highlight new tools on OLCF systems (Blender and Ascent). The DVA training series was created in 2022 and was expanded in 2023 due to feedback from the 2022 user survey.

Table 1-9. AI quantum training attendance.

Topic	Date	Participants
Blender on Frontier	June 28, 2023	79
In Situ Visualization with Ascent	September 27, 2023	74

1.4.4.8 DOE Cross-Facility Workflows Training

The DOE Cross-Facility Workflows Training was hosted jointly by the ALCF, NERSC, and the OLCF and featured an afternoon of training on tools that could enable researchers to manage their multicenter scientific workflows. Hands-on examples of GNU Parallel, Parsl, FireWorks, and Balsam were featured with explanations of how each tool could be accessed at the ALCF, NERSC, and the OLCF. The event drew a large registration of 196 users, 48 of whom identified as OLCF users.

1.4.5 Community Engagement

As STEM careers are projected to outpace non-STEM careers in growth over the next decade, there is a critical need to foster interest and diversity within these fields and within computing. OLCF staff organize seminars, workshops, and tutorials designed to build on existing HPC user training and to appeal to underrepresented groups in computing with the aim of broadening the HPC and STEM communities.

In 2023, the OLCF sustained several community engagement initiatives from 2022. These included providing challenge problems, HPC training, and resources for the Winter Invitational Classic Student Cluster Competition, organizing the Data Challenge session at the Smoky Mountains Computational Sciences and Engineering Conference, mentoring for the Faculty Hackathon, and running and improving the challenge-based hands-on HPC Crash Course.

The OLCF also began two new endeavors: the Next Generation Pathways Summer School (NGP) and the ECP Intro to HPC Boot Camp, both aimed at broadening participation in HPC. Notable highlights from these initiatives are outlined below.

1.4.5.1 ECP Introduction to HPC Boot Camp

The OLCF contributed training, panel lectures, and projects for the ECP's Introduction to HPC Bootcamp. This week-long event was held at NERSC and was aimed at students from underrepresented groups in computing and students from academic disciplines that do not typically utilize HPC. It centered on giving small groups of students a project that focused on energy justice that could be solved with HPC or the programming skills and algorithms needed as a foundation for HPC. The students spent the week learning programming skills and being guided by mentors. At the end of the week, each group of students presented their findings. NERSC provided a Jupyter Hub and HPC resources for the students, while ALCF helped plan and develop the format for the event. The students' travel and living costs for the week were paid for by the ECP, and the whole workshop was run and managed by the Sustainable Horizons Institute.

During the event, members of the OLCF worked with 20 students and 2 student peer-mentors to develop two projects that involved large power outage datasets provided by the DOE's EAGLE-I project. The first project asked the students to explore all the counties in the US during a heatwave and find the counties that had both a large number of power outages and many residents who were dependent on powered medical devices. The second project looked at the power outage data and the same heatwave but challenged the students to see if there were correlations between the number of power outages and a county's socioeconomics (based on census data). Both problems could utilize HPC because the datasets were very large, and the analysis could be done in parallel on each county for each time period explored. The OLCF also provided training for the whole camp on how to use Python for statistical analysis and organized a panel on HPC career paths. Of the 60 students who attended the Bootcamp, 26 applied for summer internships at national laboratories. ORNL's student groups also presented lightning talks on their experience at the 2023 Smoky Mountains Computational Sciences and Engineering Conference.

1.4.5.2 Next Generation Pathways to Computing Summer School

In the summer of 2023, ORNL launched NGP, a 5-week summer program that aims to inspire diverse participation, foster cultural awareness, and bridge the diversity gap in computing careers. The program brought together 10 high school students, including students from Title I high schools. NGP was selected as one of five Pathways Summer Schools that DOE's Workforce Development for Teachers and Scientists (WDTS) office awarded in 2023. The project team managing NGP and developing the curriculum included staff members from ORNL's Office of Research Education, the Spallation Neutron Source (SNS), and the OLCF.

Students began learning to program in Python using NVIDIA JetBot and Jupyter notebook-based coding workbooks designed by NVIDIA with contributions from SNS and OLCF staff. Pairs of students were also assigned ORNL mentors from the research and technical staff to work on small projects over the course of the program. Additionally, students took a modified version of the OLCF's 4-week Hands-On HPC Crash Course. OLCF staff taught extra sessions for each of the four modules of the Hands-On HPC Crash Course to ease students into the content. For modules that were presented virtually, an in-person instructor guided the students through the coursework.

OLCF staff members were instrumental in working with the Office of Research Education to secure the WDTS funding used for the course. They also led the curriculum design, set up and tested the NVIDIA Jetbots, taught many of the programming sessions, and mentored research projects for six of the students. By the end of the program, the 10 students were able to present posters about their research projects.

1.4.5.3 HPC Crash Course

In 2023, the OLCF continued to update and develop its Hands-On HPC Crash Course. A workbook guide was added for the C programming modules, a version of the course was updated to run on Frontier, and the version for Summit is still maintained. The OLCF continued to offer certificates for the course for college students who completed 7 of the course's 16 programming challenges and high school students who completed 2 challenges.

The course was again used to teach students how to use a supercomputer and about HPC-driven research for the 2023 Winter Invitational Classic Student Cluster Competition. The course was also given over four weeks in the summer and aimed at the ORNL summer interns and the NGP. Registration was also open to the students and professors from the 2023 Faculty Hackathon.

This version of the course was taught virtually for most students and in-person for the high school students. The summer course resulted in 65 certificates earned out of the 159 people registered. The OLCF also began collaborating with staff from Purdue's Rosen Center for Advanced Computing to modify the Frontier version of the course for Purdue's Anvil Supercomputer.

Purdue collaborated with the OLCF to conduct the SC23 student program's Hands-On HPC Crash Course, with the OLCF offering Frontier for the event at SC23 and Purdue offering Anvil for students to complete their homework and for 1 year after the course. Twenty-seven participants from SC23 earned course certificates.

OLCF staff members set up a LinkedIn page for course alumni to join so they could continue to receive training and workshop opportunities and so the OLCF team could follow their careers. This page replaced the course alumni mailing list.

1.4.5.4 Faculty Hackathon

For the second year in a row, the OLCF helped plan and participate in the Faculty Hackathon, a Science Gateways Community Initiative event designed to introduce faculty to HPC concepts, tools, and techniques applicable in their courses. This in turn exposes students to HPC, preparing them for further studies or careers in the field. This program broadens participation in HPC and creates an environment for collaboration among HPC educators from different institutions and industries.

In 2023, a total of six faculty teams, including seven faculty members from Kettering University, Texas A&M University – San Antonio, Jarvis Christian University, Southern Oregon University, Jackson State University, and the University of North Carolina at Pembroke, participated. ORNL mentored four of the faculty teams. See https://hackhpc.github.io/facultyhack-gateways23/teams.html.

Over the past year, OLCF staff worked to promote the OLCF's Pathways to the Supercomputing Initiative and other ORNL research experience opportunities among participating faculty and worked to connect faculty with the lab. OLCF mentor John Holmen plans to maintain the connection to his 2023 Kettering Faculty team by guest lecturing in Kettering's upcoming computing courses. Several students from the 2022 Faculty Hackathon teams have participated in the 2023 virtual OLCF Hands-On HPC Crash Courses as part of the Pathways to Supercomputing Initiative. OLCF will invite the 2023 professors to send their students to the 2024 Hands-On HPC Crash Course.

1.4.5.5 PEARC23 and SC23 Student Program

OLCF staff members participated in the PEARC23 Student program by serving as the lead organizers for the Resume Clinic workshop, the PitchIt! Perfecting your Elevator Speech workshop, and the Speed Networking workshop and by bringing a new negotiation workshop to the program. For the resume workshop, 10 professionals from national labs, industry, and universities reviewed student resumes and provided best practices and tips for how to tailor their resumes for different job fields. The event was well attended by students. For the PitchIt! Workshop, OLCF staff and university faculty showed students how to present a short sales pitch to introduce themselves and their research. Students then broke into teams to practice their pitches with volunteer mentors. In the Speed Networking workshop, students met with PEARC23 exhibitors in short one-on-one sessions to pitch their skills. In the Negotiation workshop, the students were taught how to navigate the job interview process, what goes into a salary negotiation, and all the factors to consider while negotiating so they can obtain the best offer possible.

OLCF staff also organized the Master the Art of Crafting and Delivering Your Elevator Pitch Workshop as part of the Students@SC program. The workshop taught them to present a short sales pitch to introduce themselves and their work to potential employers and collaborators. Additionally, OLCF staff delivered a workshop on Unlocking the Power of Negotiation that was similar to the one delivered at PEARC23 with the addition of providing students an opportunity to present to their peers and practice what they learned about negotiation. This provided students the ability to practice identifying what mattered to them most in a job.

Additionally, OLCF staff sat on panels for the SC23 Student Program, including the Navigating Career Waters: Finding Your Perfect Career Fit across Industries panel, in which staff provided insights on how to identify and grow the ideal career path.

OLCF staff also participated in multiple student mentorship programs organized at conferences. The individual staff member is assigned a student that they then mentor and maintain a professional relationship with well after the end of the conference, offering them continued guidance. OLCF staff members served as mentors in student programs in the PEARC23 and SC23 conferences.

1.4.5.6 Smoky Mountains Computational Sciences and Engineering Conference, Data Challenge session

In 2023, OLCF staff members organized the annual Smoky Mountains Computational Sciences and Engineering Conference, Data Challenge Session (SMCDC 2023). SMCDC provided an opportunity to tackle scientific data challenges that came from datasets contributed by ORNL, industry, and academia. These datasets were directly obtained from scientific simulations or instruments in physical and chemical sciences, electron microscopy, bioinformatics, engineering, materials science, neutron sources, urban development, and other areas. They were also accompanied by open-research questions and tasks for participants to solve. These challenges were intended to draw students, scientists, and researchers who might be at the beginning stages of incorporating data analytics into their work or research, but they also appealed to data analytics experts interested in applying novel techniques to datasets of national importance. SMCDC 2023 data sponsors provided seven challenges for participants to solve. In total, SMCDC 2023 received 37 registrations from 35 teams (2 teams registered for 2 challenges each). After peer reviews from SMCDC committee members, six solution papers were selected for publication by OSF DOI. The six teams were also invited to the Smoky Mountains Computational Sciences and Engineering Conference poster session and presented lightning talks about their work. The workshop also received positive feedback from students who expressed appreciation for the opportunity to work on real-world scientific data.

2. OPERATIONAL PERFORMANCE

CHARGE QUESTION 2: Did the facility's operational performance meet established targets?

OLCF RESPONSE: Yes. The OLCF provides highly capable and reliable systems for the user community. The 2023 reporting period covers production operations during the CY for the following HPC resources: the HPE Cray EX Frontier (April – December), the IBM AC922 Summit, the Alpine GPFS, the Lustre file system Orion (April – December), and the archival storage system (High-Performance Storage System [HPSS]). In 2023, the OLCF once again delivered all the compute hours committed to the three major allocation programs: INCITE, ALCC, and DD. The operational performance demonstrates that the OLCF delivered another prominent operational year of reliable and technically sufficient resources to the scientific research community.

2.1 RESOURCE AVAILABILITY

The OLCF exceeded all resource availability metrics in 2023 for the Frontier (Table 2-2) and Summit (Table 2-3) computational resources, HPSS data resources (Table 2-5), the Orion resources (Table 2-4) and the Alpine resources (Table 2-5). Supporting systems such as EVEREST, Andes, data transfer nodes, and Themis (nearline storage), were also offered. Metrics for these supporting systems are not provided. See APPENDIX C for more information about each of these systems. The following tables describe the availability of OLCF resources. More details on the definitions and formulae that describe the scheduled availability (SA), overall availability (OA), mean time to interrupt (MTTI), and mean time to failure (MTTF) are provided in APPENDIX D.

Operational performance metrics are provided for OLCF's Frontier, Summit, the HPSS archive system, and the Alpine external GPFS file system (Table 2-2–Table 2-7).

Table 2-1. OLCF Operational performance summary for Frontier.

	Measurement	2022 target	2022 actual	2023 target	2023 actual
HPE Cray EX Frontier	SA	N/A	N/A	85%	99.65%
	OA	N/A	N/A	80%	97.64%
	MTTI (hours)	NAM[a]	N/A	NAM	397
	MTTF (hours)	NAM	N/A	NAM	1,621

[a] NAM = Not a metric. No defined metric or target exists for this system. Data provided as reference only.

Table 2-2. OLCF Operational performance summary for Summit.

	Measurement	2022 target	2022 actual	2023 target	2023 actual
IBM AC922 Summit	SA	95%	98.52%	95%	99.84%
	OA	90%	97.87%	90%	99.41%
	MTTI (hours)	NAM[a]	857	NAM	1,396
	MTTF (hours)	NAM	1,726	NAM	4,205

[a] NAM = Not a metric. No defined metric or target exists for this system. Data provided as reference only.

Table 2-3. OLCF Operational performance summary for HPSS.

	Measurement	2022 target	2022 actual	2023 target	2023 actual
HPSS	SA	95%	99.88%	95%	99.65%
	OA	90%	98.58%	90%	96.10%
	MTTI (hours)	NAM	480	NAM	443
	MTTF (hours)	NAM	1,750	NAM	2,910

NAM data is provided as reference only.

Table 2-4. OLCF Operational performance summary for the Orion external Lustre file system.

	Measurement	2022 target	2022 actual	2023 target	2023 actual
Orion	SA	N/A	N/A	90%	99.08%
	OA	N/A	N/A	85%	98.42%
	MTTI (hours)	NAM	N/A	NAM	397
	MTTF (hours)	NAM	N/A	NAM	545

NAM data is provided as reference only.

Table 2-5. OLCF Operational performance summary for the Alpine external GPFS file system.

	Measurement	2022 target	2022 actual	2023 target	2023 actual
Alpine	SA	95%	98.47%	95%	99.69%
	OA	90%	97.57%	90%	99.20%
	MTTI (hours)	NAM	777	NAM	869
	MTTF (hours)	NAM	1,438	NAM	1,455

NAM data is provided as reference only.

In CY 2023, Summit saw significant increases in MTTI and MTTF. This is primarily due to the scheduled end of life for both the system and the DOE allocation programs which utilize this resource. The corresponding decrease in utilization had a positive impact on MTTI and MTTF. For a period of 1 year following either system acceptance or a major system upgrade, the SA target for an HPC compute resource is at least 85%, and the OA target is at least 80%. For year two, the SA target for an HPC compute resource increases to at least 90%, and the OA target increases to at least 85%. For year three through the end of life for the associated compute resource, the SA target increases to 95%, and the OA target increases to 90%. Consequently, for HPC compute resources, SA targets are described as 85%/90%/95%, and OA targets are described as 80%/85%/90%.

For a period of 1 year following either system acceptance or a major system upgrade, the SA target for an external file system is at least 90%, and the OA target is at least 85%. For year two through the end of life of the asset, the SA target for an external file system increases to at least 95%, and the OA target increases to at least 90%. Consequently, for an external file system, SA targets are described as 90%/95%, and OA targets are described as 85%/90%.

2.1.1 Scheduled Availability

SA is the percentage of time that a designated level of resource is available to users, excluding scheduled downtime for maintenance and upgrades. The user community must be notified of a maintenance event no less than 24 hours in advance of the outage (emergency fixes) for it to be considered scheduled downtime. Users will be notified of regularly scheduled maintenance in advance: no less than 72 hours prior to the event and preferably as many as seven calendar days prior. If that regularly scheduled maintenance is not needed, then users will be informed of the cancellation of that maintenance event in a timely manner. Any interruption of service that does not meet the minimum notification window is categorized as an *unscheduled* outage.

A significant event that delays the return to scheduled production by more than 4 hours will be counted as an adjacent unscheduled outage, as an unscheduled availability, and as an additional interrupt.

SA is described by Eq. (2.1). The OLCF has exceeded the SA targets for the facility's computational resources for 2022 and 2023 (Table 2-6).

$$SA = \left(\frac{\text{time in period} - \text{time unavailable due to outages in period}}{\text{time in period} - \text{time unavailable due to scheduled outages in period}}\right) * 100 \tag{2.1}$$

Table 2-6. OLCF operational performance summary: SA.

	System	2022 target	2022 actual	2023 target	2023 actual
SA	IBM AC922	95%	98.52%	95%	99.84%
		N/A	N/A	85%	99.65%
	HPSS	95%	99.88%	95%	99.65%
	Alpine	95%	98.47%	95%	99.69%
		N/A	N/A	90%	99.08%

2.1.1.1 OLCF Maintenance Procedures

Preventative maintenance is exercised only with the concurrence of the vendor hardware and software teams, the OLCF HPC Systems groups, and the OLCF Resource Utilization Council. Typical preventative maintenance activities include software updates, application of field notices, and hardware maintenance to replace failed components. Without concurrence, the systems remain in their respective normal operating conditions. Preventative maintenance is advertised to users a minimum of 2 weeks in advance if the maintenance activities include changing default software and a minimum of 1 week in advance if default software is not being changed.

2.1.2 Overall Availability

OA is the percentage of time that a system is available to users. Outage time reflects both scheduled and unscheduled outages. The OA of OLCF resources is derived by using Eq. (2.2).

$$OA = \left(\frac{\text{time in period} - \text{time unavailable due to outages in period}}{\text{time in period}}\right) * 100 \tag{2.2}$$

As shown in Table 2-7, the OLCF exceeded the OA targets of the facility's resources for 2022 and 2023.

Table 2-7. OLCF operational performance summary: OA.

	System	2022 target	2022 actual	2023 target	2023 actual
OA	IBM AC922	90%	97.87%	90%	99.41%
	Frontier	N/A	N/A	80%	97.64%
	HPSS	90%	98.58%	90%	96.10%
	Alpine	90%	97.57%	90%	99.20%
	Orion	N/A	N/A	85%	98.42%

2.1.3 Mean Time to Interrupt

MTTI is, on average, the time to any outage of the full system, whether unscheduled or scheduled. It is also known as MTBI. The MTTI for OLCF resources is derived by Eq. (2.3), and a summary is shown in Table 2-8.

$$\text{MTTI} = \left(\frac{\text{time in period} - (\text{ duration of scheduled outages} + \text{duration of unscheduled outages })}{\text{number of scheduled outages} + \text{number of unscheduled outages} + 1}\right) \tag{2.3}$$

Table 2-8. OLCF operational performance summary: MTTI

	System	2022 actual	2023 actual
MTTI[a] (hours)	IBM AC922	857	1,396
	Frontier	N/A	397
	HPSS	480	443
	Alpine	777	869
	Orion	N/A	397

[a] MTTI is not a metric. Data provided as reference only.

2.1.4 Mean Time to Failure

MTTF is the time, on average, to an unscheduled outage of the full system. The MTTF for OLCF resources is derived from Eq. (2.4), and a summary is provided in

Table 2-9.

$$MTTF = \frac{\text{time in period} - (\text{ duration of unscheduled outages })}{\text{number of unscheduled outages} + 1} \tag{2.4}$$

Table 2-9. OLCF operational performance summary: MTTF.

	System	2022 actual	2023 actual
MTTF[a] (hours)	IBM AC922	1,726	4,205
	Frontier	N/A	1,621
	HPSS	1,750	2,910
	Alpine	1,438	1,455
	Orion	N/A	545

[a] MTTI is not a metric. The data provided as reference only.

2.2 TOTAL SYSTEM UTILIZATION IN 2023

2.2.1 Resource Utilization Snapshot

During the operational assessment period (January 1, 2023–December 31, 2023), 33,923,873 Summit node hours were used outside of outage periods from an available 39,642,832 node hours. The total system utilization (SU) for the IBM AC922 Summit was 85.57%. On Frontier, 51,587,492 node hours were used outside of outage periods from an available 59,755,069 node hours. The total SU for the HPE Cray EX Frontier was 86.33%.

2.2.2 Total System Utilization

2.2.2.1 2023 Operational Assessment Guidance

SU is the percentage of time that the system's computational nodes run user jobs. The SU for OLCF resources is derived from Eq. (2.5). No adjustment is made to exclude any user group, including staff and vendors.

$$SU = \left(\frac{\text{core hours used in period}}{\text{core hours available in period}}\right) * 100 \qquad (2.5)$$

The measurement period is for 2023, regardless of the prescribed allocation period of any single program. As an example, the INCITE allocation period follows a CY schedule. The ALCC program follows an allocation cycle that runs for 12 months beginning July 1 of each year. The OLCF tracks the consumption of Summit and Frontier node hours by job. This method is extended to track the consumption of Summit and Frontier node hours by program, project, user, and system with high fidelity. Figure 2-1 shows the IBM AC922 Summit utilization by month and by program for CY 2023. Figure 2-2 shows the HPE Cray EX Frontier utilization by month and by program for CY2023. The three major OLCF user programs and usage by the ECP are represented, but the graph does not include consumed node hours from staff or vendor projects. For the 5th production year of Summit, utilization was 85.57%. For the 1st production year (as of April 5, 2023) of Frontier, utilization was 86.33%.

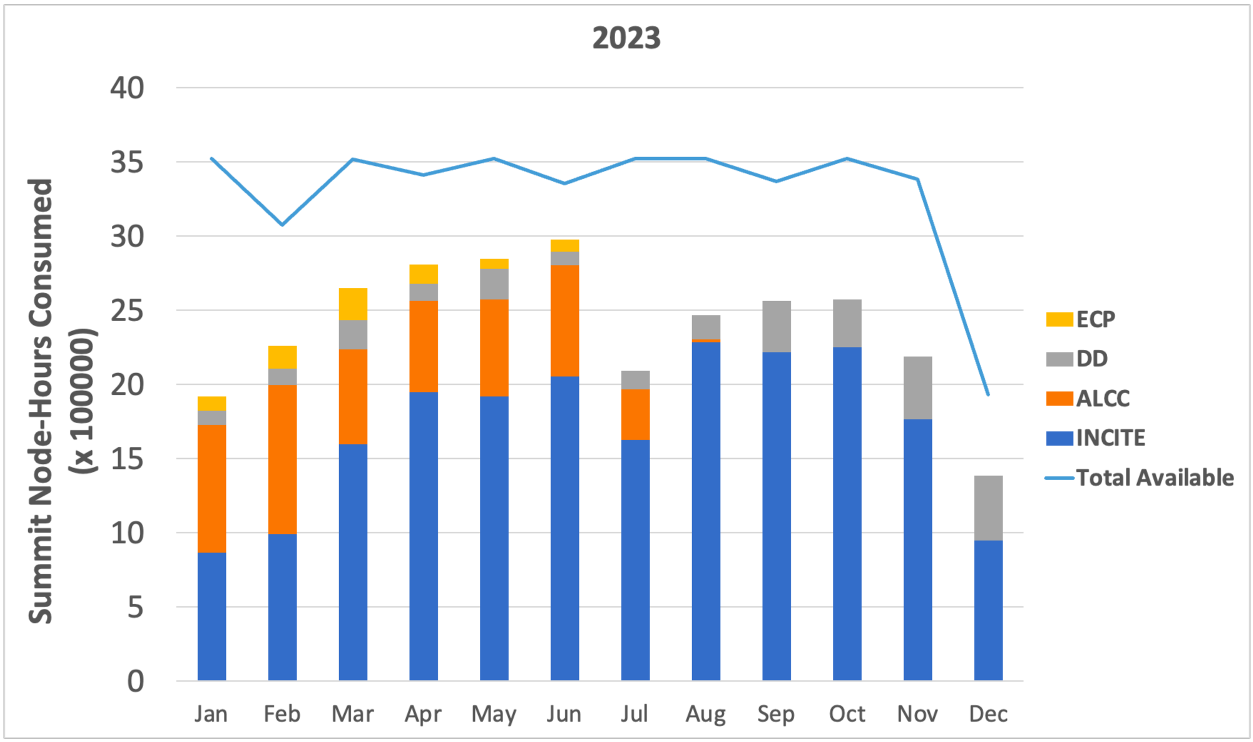

Figure 2-1. IBM AC922 resource utilization: Summit node hours by program for 2023.

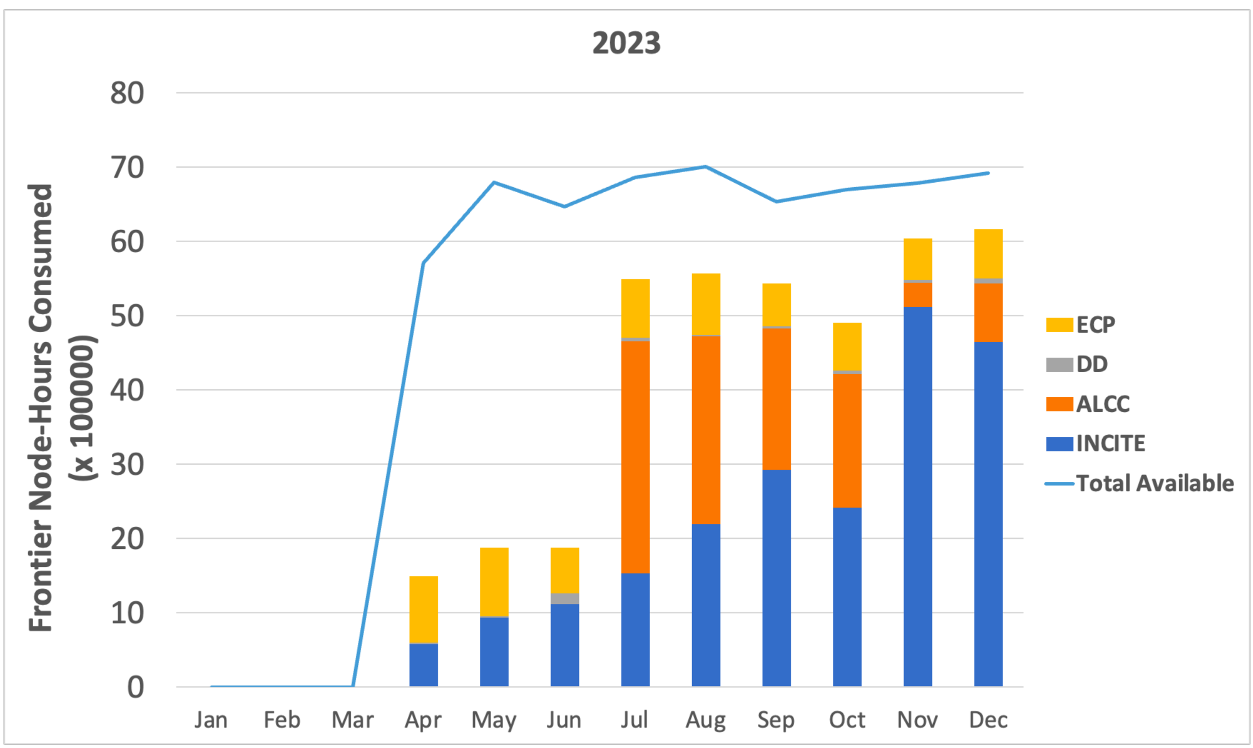

Figure 2-2. HPE Cray EX resource utilization: Frontier node hours by program for 2023.

2.2.2.2 Performance of the Allocation Programs

All allocation programs, including INCITE, ALCC, and DD, are aggressively monitored to ensure that projects within these allocation programs maintain appropriate consumption rates. The 2023 INCITE allocation program was the largest program in 2023, with a commitment of 18.8 million Summit node hours and 19.9 million Frontier node hours. The consumption of these allocation programs is shown in Table 2-10 (Summit) and Table 2-11 (Frontier). As shown, all commitments were exceeded for each allocation program on Summit and Frontier for 2023. This programmatic overachievement is due in part to the high uptime and diligent work of the OLCF staff.

Table 2-10. The 2023 allocated program performance on Summit.

Program	Allocation	Hours consumed	Percent of total
INCITE[a]	18,800,000	21,535,624	72%
ALCC[b]	Allocation spans multiple CY	4,947,986	17%
DD[c]	–	3,449,183	11%
Total[d]		29,932,793	100%

[a] Includes all INCITE program usage for CY 2023 (including the 13th bonus month usage in January 2023).
[b] Includes all ALCC program usage for CY 2023.
[c] Includes ECP.
[d] Does not include usage outside of the three primary allocation programs.

Table 2-11. The 2023 allocated program performance on Frontier.

Program	Allocation	Hours consumed	Percent of total
INCITE[a]	19,900,000	21,458,611	46%
ALCC[b]	Allocation spans multiple CY	17,388,893	38%
DD[c]	–	7,380,143	16%
Total[d]		46,227,647	100%

[a] Includes all 2023 INCITE program.
[b] Includes all ALCC program usage for CY 2023.
[c] Includes ECP.
[d] Does not include usage outside of the three primary allocation programs.

Non-renewed INCITE projects from 2023 continued running through January 2024 under the OLCF's 13th month policy. This policy permits an additional, final month for completion and was recognized as a best practice during a previous OAR. It also serves to maintain high utilization while new projects establish a more predictable consumption routine. ALCC projects from the 2023 allocation period (ending June 30, 2023) were also granted extensions as appropriate.

2.3 CAPABILITY UTILIZATION

To be classified as a *capability job*, any single job must use at least 20% of the leadership system's available nodes. For the CY following a new system/upgrade, at least 30% of the consumed node hours will be from jobs that request 20% or more of the available nodes. In subsequent years, at least 35% of the node hours consumed will be from jobs that require 20% or more of the nodes available to users. The metric for capability utilization describes the aggregate number of node hours delivered by capability jobs. The metric for CY 2023 was 35% for Summit based on years of service and 30% for Frontier, as described above. The OLCF continues to exceed expectations for capability usage of its HPC resources (Table 2-12 and Table 2-13 and Figure 2-3 and Figure 2-4).

Keys to successful demonstration of capability usage include the liaison roles provided by Science Engagement members who work hand-in-hand with users to port, tune, and scale code and the OLCF support of the application readiness efforts (i.e., CAAR) to actively engage with code developers to promote application portability, suitability for hybrid systems, and performance. The OLCF also aggressively prioritizes capability jobs in the scheduling system.

Table 2-12. OLCF capability usage on the IBM AC922 Summit system.

Leadership usage	CY 2022 target	CY 2022 actual	CY 2023 target	CY 2023 actual
INCITE	NAM[a]	57.38%	NAM	53.86%
ALCC	NAM	53.25%	NAM	26.51%
All projects	35%	56.51%	35%	45.63%

[a] NAM means no defined metric or target exists for this system. Data provided as reference only.

Table 2-13. OLCF capability usage on the HPE Cray EX Frontier system.

Leadership usage	CY 2022 target	CY 2022 actual	CY 2023 target	CY 2023 actual
INCITE	NAM[a]	N/A	NAM	77.87%
ALCC	NAM	N/A	NAM	90.50%
All projects	N/A	N/A	30%	79.08%

[a] NAM means no defined metric or target exists for this system. Data provided as reference only.

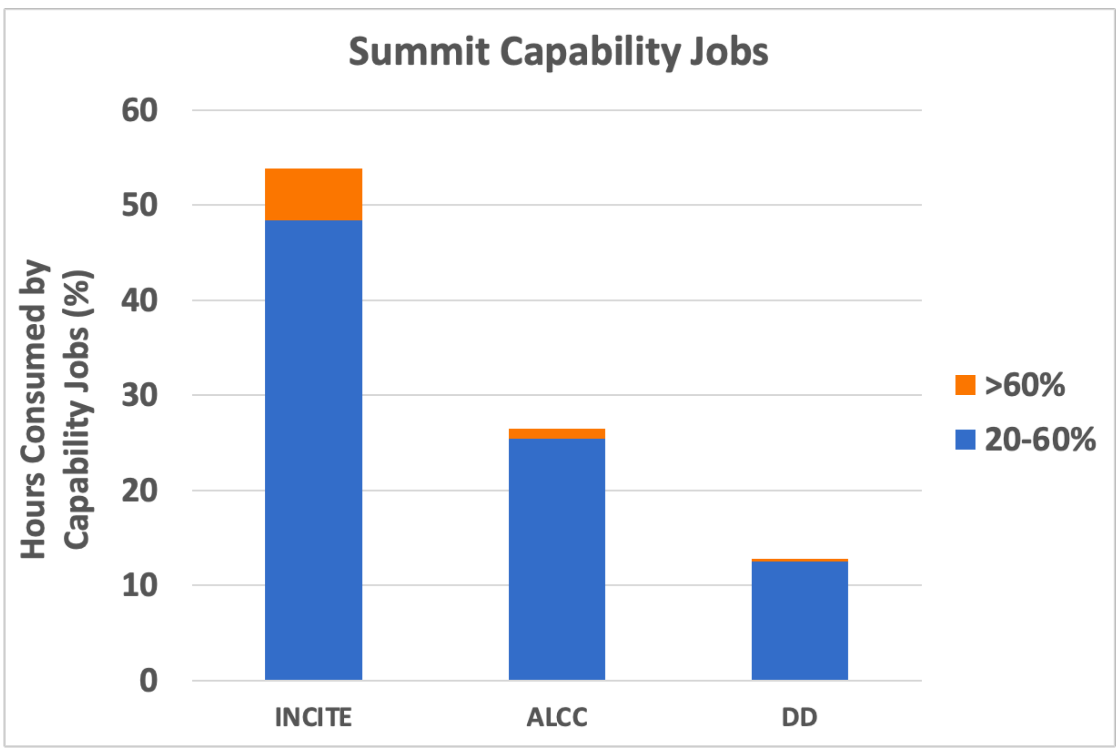

Figure 2-3. Summit capability usage by job size bins and project type.

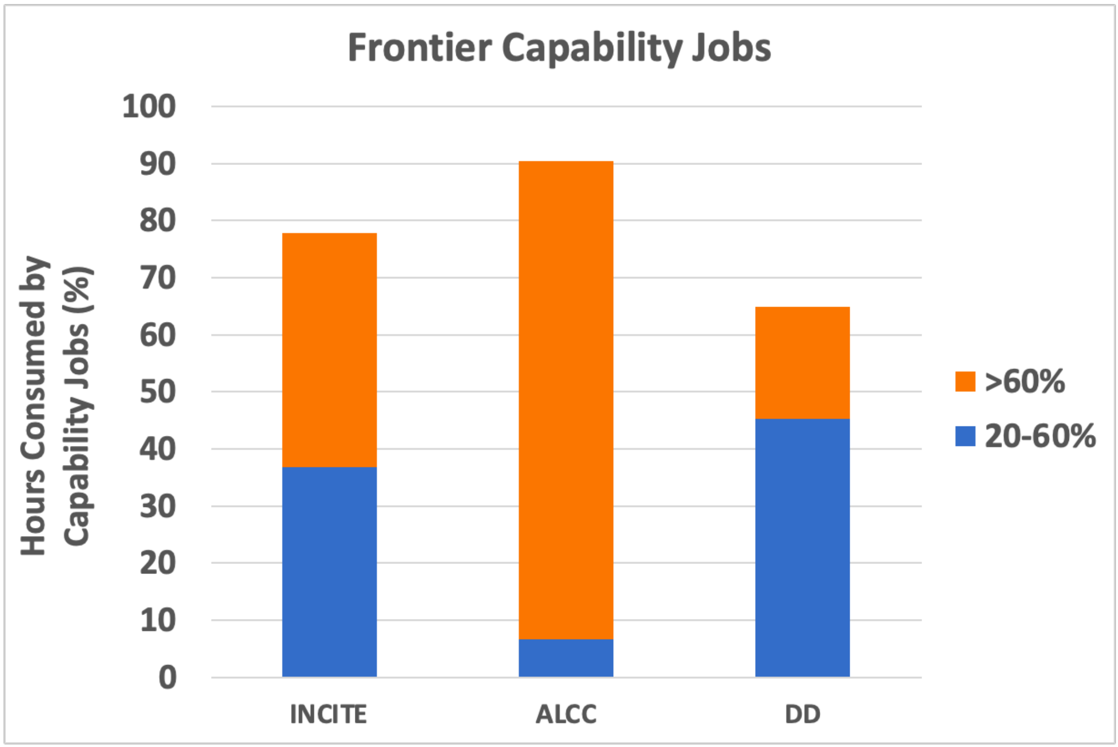

Figure 2-4. Frontier capability usage by job size bins and project type.

3. ALLOCATION OF RESOURCES

CHARGE QUESTION 3: (a) Did the allocation of computing resources conform with ASCR's published allocation policies (i.e., ratio of resources allocated between INCITE, ALCC, DD, and ECP)? (b) Was the allocation of DD computing resources reasonable and effective? (c) Did the Facility encounter issues with under- or over-utilization of user allocations? If so, was the Facility's management of these issues effective in maximizing productive use of resources while promoting equity across the user population?

OLCF RESPONSE: Yes. The OLCF continues to enable high-impact science results through access to its leadership-class systems and support resources. The allocation mechanisms are robust and effective. The OLCF enables compute and data projects through the DD user program. This program seeks to enable researchers through goals that are strategically aligned with ORNL and DOE, as described in Section 3.1. The allocation of resources among INCITE, ALCC, and DD programs matched the agreed-upon decomposition, with minor modifications on Frontier as the machine matured from early availability to full production throughout 2023. As of December 31, 2023, 119.5% of the total allocated hours for INCITE 2023 had been consumed on Summit and 107.9% of the allocation for INCITE 2023 had been consumed on Frontier. By extending a customary additional month of access to 2023 INCITE allocations that were not renewed for CY 2024, OLCF ultimately delivered 111% of the total allocated hours for INCITE 2023 on Frontier by January 31, 2023. No such extension of access was possible on Summit for January 2024 because the machine was taken out of service for the installation of a new parallel file system in support of the final year of Summit access in 2024.

3.1 SUMMARY OF ALLOCATION PROGRAMS

The primary allocation programs that provide access to OLCF resources are the INCITE program, the ALCC program, and the OLCF's DD program. The agreed-upon apportionment of resources for these programs is as follows: 60% of available resources were allocated to INCITE, 20% to ALCC, and 20% to DD, with up to half of the DD allocation (i.e., 10% of available resources) dedicated to ECP Application Development (AD) and Software Technology (ST) projects during each quarter. These ECP hours are reported as a separate category in the following discussion. These allocation percentages were different on Frontier when compared to a full machine-year for 2024 because the total allocation amounts were modified with ASCR concurrence to account for the move to full production on Frontier being offset from the start of the calendar year.

For CY 2023, as both AD and ST projects neared completion and Frontier moved to production, the allocation period for ECP projects was moved to bi-yearly (i.e., two allocation epochs for ECP, one beginning in January [H1] and the other beginning in July [H2]) for allocations on Frontier, which served as the primary proving ground for essentially all ECP projects. Individual allocation amounts for each ECP project for both half-years were determined in direct consultation with ECP to maximize availability and to ensure that individual projects had ample allocations to accomplish their goals when those exhibitions were possible and most effective. In CY 2023, the percentage of available compute time delivered to each of these programs is shown in Table 3-1. The ECP resources were generally under-allocated in each allocation period on both resources in CY 2023, and the resulting time available in each quarter was often used by other projects from the other allocation programs. It is also important to note that Summit was only allocated via ALCC for the first half of CY 2023 because the resource was not offered as part of the ALCC 2023–2024 call for proposals.

Table 3-1. Percentages of delivered time per allocation program for Summit and Frontier.

Resource	INCITE	ALCC	DD	ECP
Summit	72.0%	16.5%	8.8%	2.7%
Frontier	46.4%	37.6%	0.9%	15.1%

3.2 FACILITY DIRECTOR’S DISCRETIONARY RESERVE TIME

The OLCF primarily allocates time on leadership resources through the INCITE program and through the facility’s DD program. The OLCF seeks to enable scientific productivity via capability computing through both programs. Accordingly, a set of criteria is considered when making allocations, including the strategic impact of the expected scientific results and the degree to which awardees can effectively use leadership resources.

The goals of the DD program are threefold.

(1) To enable users to prepare for leadership computing competitions such as INCITE and ALCC (e.g., to improve and document application computational readiness)
(2) To broaden the community of researchers capable of using leadership computing by enabling new and nontraditional research topics
(3) To support R&D partnerships, both internal and external to ORNL, to advance the DOE and ORNL strategic agendas

These goals are aligned particularly well with three of the four OLCF missions.

(1) To enable high-impact, grand-challenge science and engineering that could not otherwise be performed without leadership-class computational and data resources
(2) To enable fundamentally new methods of scientific discovery by building stronger collaborations with experimental facilities as well as DOE offices that have large compute and data science challenges
(3) To educate and train the next-generation workforce in the application of leadership computing to solve the most challenging scientific and engineering problems

R&D partnerships are aligned with DOE and ORNL strategic agendas. These partnerships may be entirely new areas for HPC, or they may be areas that need nurturing. Examples include projects associated with the ORNL Laboratory Directed Research and Development program; programmatic science areas (e.g., fusion, materials, chemistry, climate, nuclear physics, nuclear engineering, bioenergy science and technology); and key academic partnerships (e.g., the University of Tennessee Oak Ridge Innovation Institute).

Also included in this broad category are projects that come to the OLCF through the Accelerating Competitiveness through Computational ExceLlence (ACCEL) Industrial HPC Partnerships outreach, which encourages opportunities for industrial researchers to access the leadership systems through the usual leadership computing user programs to conduct research that would not otherwise be possible.

The actual DD project lifetime is specified upon award: allocations are typically for 1 year or less. However, projects may request 3-month extensions or renewals up to an additional 12 months. The average size of a DD award on Summit in CY 2023 was roughly 28,025 Summit node hours, but awards can range from 1,000 to 200,000 node hours or more. The average size of a DD award on Frontier in

CY 2023 was roughly 24,853 Frontier node hours, and awards can range from 1,000 to 200,000 node hours or more. In 2023, the OLCF DD program participants (including ECP users) used approximately 11% of the total user resources on Summit (Table 3-1). See APPENDIX E for a full list of DD projects for CY 2023.

3.3 ALLOCATION PROGRAM CHALLENGES

The INCITE program suffered from under-utilization on Frontier throughout much of CY 2023, as teams began their first forays onto the new architecture. The OLCF modified job priorities for INCITE jobs relative to the usual policies to ensure that all allocated hours were delivered to the program in CY 2023. Specifically, starting in November 2023, the relative priority of INCITE jobs was elevated by increasing (by 5 days) the effective number of days they appeared to have been queued. Normally, INCITE and ALCC projects have a priority boost relative to DD and other jobs of a single day of aging. This increase in aging by 4 days, coupled with improved reliability and stability of the machine in the later parts of the CY, enabled INCITE projects to use (and exceed) their remaining time within the calendar year.

4. OPERATIONAL AND TECHNICAL INNOVATIONS

CHARGE QUESTION 4: (a) Have technical innovations been implemented that have improved the facility's operations? (b) Have management/workforce innovations been implemented that have improved the facility's operations? (c) Is the facility effectively utilizing their postdoctoral fellows?

OLCF RESPONSE:

(4a) Yes, the Facility has implemented technical innovations that improved the Facility's operations. Details can be found in Section 4.1.

(4b) Yes, the Facility has implemented management and workforce innovations that have improved the Facility's operations. Details can be found in Section 4.2.

(4c) Yes, the Facility has effectively utilized its postdoctoral fellows. Details can be found in Section 4.3.

4.1 TECHNICAL INNOVATIONS

OLCF has developed and implemented multiple technical innovations in 2023. These innovations and their operational impacts and improvements are described below.

4.1.1 OLCF-5 Best Practices in Large Scale HPC Procurement

In FY 2023, with the successful acceptance and CD-4 for OLCF-5 (Frontier), OLCF began to define OLCF-6. With the slowing of Moore's Law and the new reality that transistors are no longer decreasing in cost with newer generations, the easy increase in performance with each new generation is gone. To ensure that the next system meets the users' needs, members of NCCS worked to understand the usage of current systems (i.e., Summit and Frontier). The team then used these insights to help define requirements for OLCF-6. For example, the team identified that the compute system writes on average 1.5× the GPU memory capacity to storage each day. Users like to have their data available for at least 60 days. For OLCF-6, the team combined these two insights into a requirement that storage capacity is 90× the GPU memory (1.5 GPU memory/day × 60 days). The team is continuing to develop additional insights and will use options in the OLCF-6 request for proposals (RFP) to tailor the final system. This work was presented at SC23.

Additionally, members of the OLCF published an SC23 paper that describes the technology that underpins the Frontier supercomputer, the first exascale machine in the US. The paper evaluates the computational, network, and storage performance of Frontier. The paper was a finalist for the Best Paper award at SC23, a recognition of this work's significance in the field. Papers such as this one contribute operational insights for building and benchmarking the largest supercomputers, thereby shedding light on advancements in the domain to the wider HPC community.

Scott Atchley, Christopher Zimmer, John Lange, David Bernholdt, Veronica Melesse Vergara, Thomas Beck, Michael Brim, Reuben Budiardja, Sunita Chandrasekaran, Markus Eisenbach, Thomas Evans, Matthew Ezell, Nicholas Frontiere, Antigoni Georgiadou, Joe Glenski, Philipp Grete, Steven Hamilton, John Holmen, Axel Huebl, Daniel Jacobson, Wayne Joubert, Kim Mcmahon, Elia Merzari, Stan Moore, Andrew Myers, Stephen Nichols, Sarp Oral, Thomas Papatheodore, Danny Perez, David M. Rogers, Evan Schneider, Jean-Luc Vay, and P. K. Yeung. 2023. "Frontier: Exploring Exascale." In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC23)*. Association for Computing Machinery, New York, NY, USA, Article 52, 1–16. https://doi.org/10.1145/3581784.3607089.

4.1.2 OLCF-5 Frontier Node Screen Development

During the stabilization of Frontier, OLCF staff observed that when compute jobs failed, a cause for the failure was not always evident, and often no action could be taken to address the failure. An effort was undertaken to screen the compute and memory components of each node in the system using real application workloads such as LAMMPS, or the Large-Scale Atomic/Molecular Massively Parallel Simulator code. This node screen was designed to identify defective compute hardware that may have been contributing to the higher-than-expected node failure rate observed. Using the in-house OLCF Test Harness (https://olcf.github.io/olcf-test-harness/), a total of 13 tests were developed to efficiently isolate each node and execute tests independent of all other nodes in the system. The results were then reported to a database that was queried via Grafana dashboards. These tests were designed to be run on a single node up to the entire system at the same time. These tests include variants of LAMMPS, High-Performance Linpack (HPL), rocPRIM, HACC, AMG, and a vendor-provided memory stress application. To date, more than 4 million tests have been conducted across Frontier's 9,408 nodes. These tests successfully identified dozens of cases of defective hardware that were not detected by existing health checks.

By making the software available to the system integrators and through presentations at conferences OLCF is making this best practice available to our sister facilities.

4.1.3 OLCF-6 Workflow Benchmark

In 2023, the Advanced Technologies Section at the OLCF introduced an innovative end-to-end workflow benchmark (Figure 4-1) aimed at evaluating HPC system efficiency in processing dynamic, data-intensive workloads. Central to this initiative is the Machine Learning for Neutron Scattering Experiment (ML4NSE) application, which leverages a Temporal Fusion Transformer model to predict measurement times for various peaks in neutron scattering experiments, facilitating the integration of machine learning with HPC for improved experimental workflows and decision-making. This benchmark marks a significant step forward in the convergence of machine learning and HPC technologies by offering a powerful tool for enhancing scientific research efficiency and accuracy. By enabling near real-time predictive analysis in complex experiments such as neutron scattering, the ML4NSE application exemplifies the potential of machine learning-HPC integration to transform scientific discovery processes, paving the way for future advancements in computational science and its application across diverse research domains. Figure 4-1 shows the abstract overview of the OLCF-6 workflow benchmark. More information can be found at https://code.ornl.gov/workflow/olcf-6-benchmark.

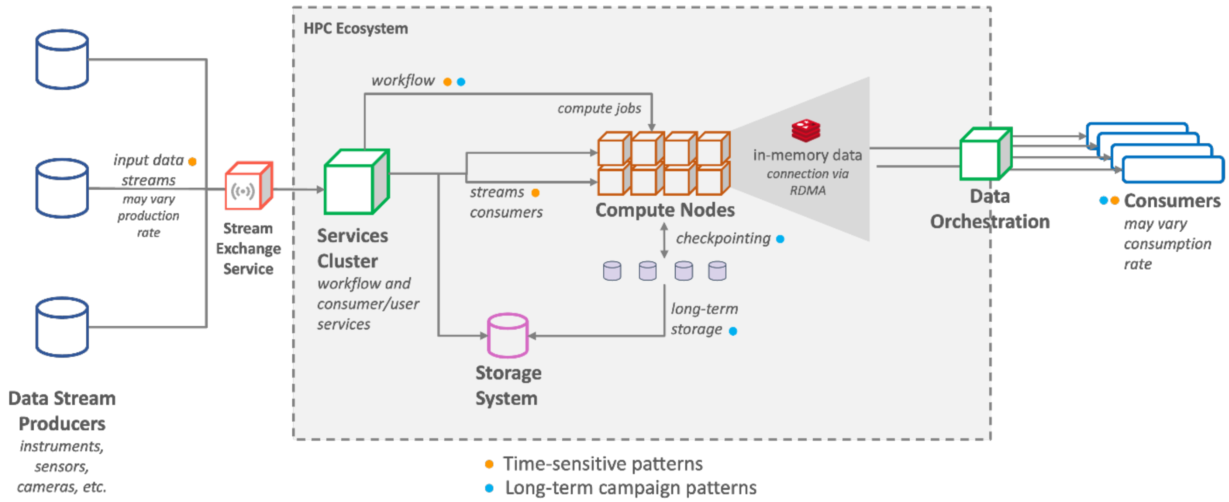

Figure 4-1. Abstract overview of a reference implementation of the workflow benchmark.

By making the software available as a public benchmark, the use of a true workflow benchmark is a new direction OLCF is making available to other supercomputing centers.

4.1.4 Data Management and Tiering Improvements

4.1.4.1 Kronos Nearline Storage System

Motivated by the imperative to enhance numerous subsystems within the prior-generation archival storage system, combined with the substantial costs involved, a competitive RFP was released in the summer of 2023 to address storage needs. In response to feedback from the OLCF user community and OLCF technical and scientific leadership, requirements surfaced to engineer a solution that strikes a balance between the high performance of scratch storage resources and the data integrity features of archival resources. This novel storage system class was coined *nearline*.

The Kronos Nearline Storage system consists of two tiers: warm and cold. The warm tier comprises 137 PB of IBM Storage Scale (GPFS) supporting streaming I/O and high disk bandwidth. The cold tier offers 400 PB of tape storage tailored for data with elevated durability requirements. Linking these storage tiers is the IBM Storage Archive software, which utilizes a flexible, high performance, and proven policy engine to migrate data based on policies tuned to meet the specific needs of programs and projects that use Kronos.

To enhance user interaction, the nearline storage resource offers a standard POSIX interface to mount volumes and append files, thereby enabling users to interact with Kronos from the command line using the same standard Linux tools employed with the scratch file systems. To facilitate user convenience, a new file listing command, *gls*, was developed. When listing files, *gls* color-codes the command line output to indicate whether files reside on tape, disk, or have been flagged for migration.

Kronos accessibility extends to three different classes of nodes:

(1) Data transfer nodes facilitate high-bandwidth data movement between OLCF resources and offsite locations via 400 GbE ESnet connectivity and leveraging the Globus toolkit.
(2) Open Shift clusters enable data workflows, portals, and related user services.
(3) Compute system login nodes are designed for lightweight interaction.

The high performance, high durability, and expansive feature set empowers Kronos to support the IRI Architecture Blueprint's three science patterns. Enabling peering with the DOE High-Performance Data Facility may be achieved by supporting core services through Kronos and establishing staged data exchange and transfers as needed. The transition to Kronos from existing archival collaborations (such as HPSS) has been discussed with HPDF partners for knowledge and best-practice exchange.

4.1.4.2 PoliMOR

Recent HPC systems increasingly deploy multi-tiered storage architectures to balance cost, performance, and capacity, requiring data management services to ensure optimal utilization and maintain performance and capacity levels. Policy engine Made to ORder (PoliMOR) (https://github.com/olcf/polimor), developed at the OLCF, is a scalable and customizable policy engine for automated data management in HPC storage systems. Utilizing a micro-service design with independently scalable agents, PoliMOR efficiently gathers file metadata and performs actions (e.g., purging or migration) based on site-defined policies. After completing functionality and performance testing on small-scale clusters, PoliMOR was installed on Orion, where large-scale testing on millions of files is currently underway. PoliMOR is slated for production by the end of Q1 2024, and testing results were recently presented at the 8th International Parallel Data Systems Workshop held in conjunction with SC23.

Anjus George, Christopher Brumgard, Rick Mohr, Ketan Maheshwari, James Simmons, Sarp Oral, and Jesse Hanley. 2023. "PoliMOR: A Policy Engine 'Made-to-Order' for Automated and Scalable Data Management in Lustre." *In Proceedings of the SC23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis (SC-W '23)*. Association for Computing Machinery, New York, NY, USA, 1202–1208. https://doi.org/10.1145/3624062.3624190

4.2 OPERATIONAL INNOVATION – MANAGEMENT/WORKFORCE

OLCF understands the significance of maintaining a healthy management and workforce structure for ensuring successful and continuous operation. To this end, OLCF is pursuing a multipronged pipeline approach. The first stage of OLCF's comprehensive pipeline starts with student programs followed by on-the-job training programs. To complement these efforts, the OLCF is also collaborating with HR and ORNL communications to develop a robust recruitment campaign.

4.2.1 ORNL's Next Generation Pathways to Computing Summer School

ORNL held the first NGP, one of five Pathways Summer Schools funded by DOE's WDTS Reaching a New Energy Sciences Workforce initiative. ORNL's NGP was the only school that focused on providing opportunities for high school students. In 2023, the summer school hosted 10 high school students from the East Tennessee region for five weeks. During the program, students were grouped into pairs and matched with an ORNL mentor. OLCF staff led the development of the curriculum for NGP, which included introductory programming lessons and provided an overview of HPC and AI topics. NGP participants were also invited to attend the HPC Crash Course offered by the OLCF and had in-person office hours to work through issues and explore topics more in-depth. In addition to learning about HPC and AI topics, the students utilized NVIDIA JetBot to gain hands-on experience and apply the concepts learned throughout the program.

Thanks to the success of the program in 2023, NGP was renewed by WDTS for summer 2024, and funding was increased to expand the program from 10 to 24 students.

4.2.2 Regional and On-Site Workforce Advancement

The OLCF also develops parallel outreach plans to multiple underrepresented schools across the US to bolster the OLCF's student pipeline. The OLCF also started a program with Tennessee Technological University (TTU) to form an academic alliance. In the spring of 2022, a delegation from ORNL, including the Computing and Computational Sciences Directorate's (CCSD's) Associate Laboratory Director, HR representatives, and technical leaders, met with the Office of the President at TTU to establish a Cooperative Education Program (co-op) between the two organizations. The goal was to provide a pipeline of students from TTU in the engineering disciplines needed to support NCCS. After several months of internal work at ORNL establishing the financial and organizational parameters around the co-op, it was launched in January of 2023. The inaugural CCSD co-op program accepted three TTU students in both Computer Engineering and Computer Science in the HPC Scalable Systems Group and the HPC Cybersecurity and Information Engineering Group. The students worked on projects that not only improved the performance, operations, and security of the current Frontier and Summit supercomputers, but they also contributed to research into tools and protocols that will impact future supercomputers and systems. Students were able to apply the theory they learned in the classroom to real-world technology and improve their collaboration and communication skills by working with their direct mentors and other staff members needed to help them accomplish their tasks. One student has extended his co-op term into the spring 2024 semester.

4.2.3 On-the-Job Training Opportunities

The other stage of the OLCF's pipeline focuses on on-the-job training. This stage has four distinct elements: job rotation, coaching, job instruction, and training through step-by-step assignments. The OLCF has instituted a staff development program for all new operators and currently has three early career system admins. As staff in the Operations Control Room (OCR) age out, the goal (just like the OLCF did in October 2021) is to promote operators to an early career admin position and replace them with a technician with an AS or a new college graduate with a BS in an IT-related field. The goal for the new operators is to learn from the ground up: start with NCCS facility operations, develop admin skills, and further education in IT (as needed). It is difficult to estimate their time in the OCR prior to promotion to early career admin because it is based on incoming skill set and degree. However, the OLCF has a 2–3-year goal for these candidates. As the early career staff members were hired, the OLCF looked at their skill sets and desires and placed them in section groups based on those criteria. The goal is for them to develop rapidly and to use their abilities with their embedded groups to transfer tasks that can be delegated to the operators. Again, how long they are in that role depends on effort, professional progress, and potentially education. The desire is for individuals to stay in the role for at least 1 year (unless advanced and other group leaders need their skills) but likely 2 years. Through the progression, the goal is to have a healthy mix of early career and seasoned OCR operators to increase the OCR abilities/skill set and to provide homegrown talent for the other groups. The OLCF is currently seeing success with their efforts and current staff. Because other groups across NCCS pick developed talent through this program, the OLCF is always looking for the next good entry-level person. Overall, this is a success for NCCS.

4.2.4 Frontier Recruitment Campaign

The OLCF, CCSD, and ORNL central communications teams collaborated with ORNL HR and recruiting to develop a targeted digital recruitment campaign focused on career opportunities provided by the launch of the Frontier supercomputer. The team held focus group meetings, developed messaging, designed materials, and explored paid placements and advertising to grow visibility of Frontier-related job opportunities and increase the number of qualified applicants for these positions. With the campaign, the OLCF learned that Google ads based on keyword search and Facebook ads provide the best return on investment. The OLCF spent $30–$33 per application received compared to the US industry average of

$260. Paid ads for general job categories provide 50%–90% more clicks than unpaid posts on social media sites. Trade publications provide less return on investment but are still valuable for increasing brand recognition. Posts with videos received 55% more clicks than static images, which aligns with expectations. Taglines that focused on diversity and solving big problems were most effective. Finally, a few additional observations must be tested to confirm their efficacy. For example, LinkedIn is more expensive than other social media platforms but may be valuable for difficult-to-hire job categories or specific positions, and posts by staff about specific positions may be the most effective use of unpaid ads on social media sites. This experiment was a helpful exercise in coordinating messages and building up recruitment efforts.

4.2.5 NCCS Virtual Career Fair 2023

Originally launched in 2021, the OLCF hosted the second iteration of the NCCS Virtual Career Fair on October 27, 2023. The event focused on recruiting for open positions in NCCS across all groups and was held via the GatherTown platform. The virtual nature of the event makes it easily accessible and enables the OLCF and NCCS to reach a broader audience. The event had 173 registrants from a wide range of universities and institutions. The candidates were interested in both internship and staff opportunities. More information about the NCCS Virtual Career Fair can be found at https://www.olcf.ornl.gov/calendar/virtual-career-fair-2/.

4.2.6 Open Science Data Hackathons

OLCF set up 4 PB of data from high-resolution climate simulations - Experimental Nature Run at 1 km (XNR1K) for training purposes. The OLCF facilitated direct access to the XNR1K data and provisioned access to OLCF computational resources via an Open Science Data Hackathon that concluded in August 2023. The 11 participating teams used Andes for data analytics, Summit for machine learning and AI workloads, and Jupyter notebook instances on Slate for interactive analytics. The hackathon participants were able to apply their own analytics tools to the XNR1K data for applications involving (1) Observing System Simulation Experiments for developing instruments for future satellites and for mission planning and (2) data-driven, sub-grid scale parameterization methods for earth system models. The XNR1K hackathon also offered the opportunity for mentoring and training the next generation of computational, data, and climate scientists and resulted in seven conference presentations (including a best paper award), two graduate student theses, and a value-added dataset.

4.3 POSTDOCTORAL FELLOWS

4.3.1 Program

DOE recognizes the need to train and retain computational scientists in a broad range of disciplines that support DOE and the nation's critical missions to maintain the US competitive advantage in high-performance and data-intensive scientific computing. When considering the ever-increasing capability of high-end computer architectures, there is a continuing and increasing need to ensure a well-trained computational science workforce in academia and industry and at the national laboratories. To address this need, DOE proposed that ASCR establish a postdoctoral training program at its user facilities, including the OLCF, the ALCF, and NERSC. This program, known as Computational Scientists for Energy, the Environment, and National Security (CSEEN) has the following objectives: (1) ensure an adequate supply of scientists and engineers who are appropriately trained to meet national workforce needs, including those of DOE, for high-end computational science and engineering with skills relevant to both exascale and data-intensive computing; (2) make ASCR facilities available through limited-term appointments for applied work on authentic problems with highly productive work teams and increasingly cross-disciplinary training; and (3) raise the visibility of careers in computational science and engineering to build the next generation of leaders in computational science. In CY 2019, the OLCF began to leverage additional funding from the ECP to augment the CSEEN program with additional postdoctoral fellows.

The OLCF CSEEN postdoctoral program seeks to provide opportunities to bridge the experience gap between the need to address domain science challenges and the need to cultivate high-performance software development expertise. One of the focus areas is to provide the skills required to port, develop, and use software suites on the leadership computing resources at the OLCF. The software development activities occur in conjunction with a CAAR project (both OLCF-5 funded and ECP funded). This model offers the greatest potential for scientific breakthroughs through computing and provides ample opportunity to publish in domain science literature. This approach will ensure that the postdoctoral trainees continue to build their reputations in their chosen science communities. Participants in the CSEEN postdoctoral program are encouraged to attend tutorials, training workshops, and training courses on select computer science topics. One of the most important outcomes for the postdoctoral trainee is the opportunity to publish and present research accomplishments.

In CY 2023, a total of nine postdocs were members of the OLCF workforce. Of those nine, six were fully supported by OLCF funds. Three of the postdocs were at least partially supported by sources outside the OLCF, including ECP AD and SciDAC projects. The background and current work of these postdocs in the Science Engagement section is described below.

4.3.1.1 Henry Monge-Camacho

Henry Monge-Camacho has been a member of the group since 2022. His research specializes in utilizing Lattice Quantum Chromodynamics (LQCD) for studies of nuclear physics. Mentored by Balint Joo, Henry was hired as part of the ECP to investigate performance portability in LQCD codes. During CY 2023 Henry worked on developing measurement codes for LQCD using the Kokkos performance-portability framework. Although Henry had no publications in CY 2023, he took on a variety of roles to prepare him for becoming a member of staff, including acting as a liaison to the CalLat group and participating in their ALCC projects, taking part in preparing an INCITE proposal, and submitting a successful SummitPLUS proposal as its PI. He also took part in OLCF-6 benchmark definition activities and attended Argonne Training Program for Extreme-Scale Scientific Computing (ATPESC). Henry was hired as a career staff member in the group at the end of the calendar year, and his new position began on Feb 1, 2024.

4.3.1.2 Tor Djaerv

Tor Djaerv joined the group as a post-doctoral associate in 2022. His research focuses on nuclear structure calculations using the NTCL tensor contraction code, and he is mentored by Gustav Jansen. Although he did not publish any journal publications in 2023, Tor made significant contributions to the NCTL tensor contraction code on the Frontier supercomputer, and this is anticipated to bear fruitful publications in the near future. Tor also attended the ATPESC.

4.3.1.3 Baishan Hu

Baishan Hu has been a member of the group since 2022. His activity is currently funded by the Physics Division through a SciDAC award, and he is mentored by Gustav Jansen. Baishan had three publications in CY 2023, two of which are in the prestigious *Physical Review Letters* journal.

Publications:

“Isomeric excitation energy for $^{99}In^{m}$ from mass spectrometry reveals constant trend next to doubly magic ^{100}Sn,” L. Nies, D. Atanasov, M. Athanasakis-Kaklamanakis, M. Au, K. Blaum, J. Dobaczewski, B. S. Hu, et al., *Phys. Rev. Lett.* 131, 022502 (2023). arXiv:2306.02033. DOI: https://doi.org/10.1103/PhysRevLett.131.022502.

“Observation of the Exotic 0+2 Cluster State in 8He,” Z. H. Yang, Y. L. Ye, B, Zhou, et al. *Phys. Rev. Lett.* 131, 242501 (2023). DOI: https://doi.org/10.1103/PhysRevLett.131.242501.

“Ab initio descriptions of A=16 mirror nuclei with resonance and continuum coupling,” S. Zhang, F. R. Xu, J. G. Li, B.S. Hu, et al., Phys. Rev. C 108, 064316 (2023). DOI: https://doi.org/10.1103/PhysRevC.108.064316.

4.3.1.4 Charles Stapleford

Charles Stapleford has been a postdoctoral associate in the group since 2020 and is mentored by Bronson Messer in the area of Nuclear Astrophysics, specifically neutrino interactions in supernovae. He worked on the development of the WeakLib and Thornado codes. His postdoctoral position finished December 31, 2023, and he has moved on to other opportunities. He had no publications in CY 2023.

4.3.1.5 Mariia Karabin

Mariia Karabin was a member of the group from 2020 until 2023. Her research focused on the investigation of alloy behavior using a combination of first principles DFT calculations, construction of surrogate models, and classical MD simulation. She was mentored by Markus Eisenbach. During CY 2023, she performed calculations to determine interactions that lead to ordering in multicomponent alloys, KKR calculations for disorder effects on the magnetism and electronic structure in CoSn and Ga-Bi-Te, and investigations of noise correction in QC calculations for the homogeneous electron gas.

Mariia presented part of her work at the 2023 APS March meeting in a talk titled “Material properties prediction using machine learning-based ab initio calculations” and was coauthor of two other talks at this conference. She also published a dataset of NiPt calculations (https://doi.org/10.13139/OLCF/1958172). Additionally, she coauthored to preprints that are currently under review for publication.

Publications:

M. Lupo Pasini, M. Karabin, and M. Eisenbach “Transferable prediction of formation energy across lattices of increasing size,” https://doi.org/10.26434/chemrxiv-2023-c14r3-v2.

W.-G. D. Ho, W. R. Mondal, H. Terletska, K.-M. Tam, M. Karabin, M. Eisenbach, Y. Wang, and V. Dobrosavljevic, “Mechanism of charge transfer and electrostatic field fluctuations in high entropy metallic alloys,” https://doi.org/10.48550/arXiv.2311.14463.

4.3.1.6 Tanvir Sohail

Tanvir Sohail has been a member of the group since September 2023. His research is in developing methods for multiscale modeling of materials using DOE leadership-class supercomputers. He is working on large-scale simulations of alloys, quantum materials, and polymer composites. He has one publication in CY 2023 in collaboration with Carbon and Composites Group at ORNL (https://doi.org/10.1002/advs.202305642). Tanvir is working on developing efficient workflow tools on

Frontier for ensemble runs of materials, performing distributed machine learning training of data, and will work on porting codes on GPUs.

4.3.1.7 Hector Hernandez Corzo

Hector Hernandez Corzo has been an integral part of the team since 2022. His research focuses on leveraging quantum mechanical methods to analyze the electronic structure of atoms and molecules. Under the mentorship of Dmytro Bykov, Hector joined the computational chemistry and advanced material modeling group. In CY 2022, Hector participated in ATPESC, enhancing his expertise in HPC. Throughout 2023, Hector focused on the development of code for highly parallel electronic structure calculations within the LS-DAlTON project. His work extended into the realm of AI, where he undertook projects aimed at advancing software development through AI applications. Collaborating closely with the operations team; Hector is involved in testing, evaluating, and incorporating AI libraries into the computing systems, thereby streamlining their accessibility for the center's future users. In the CY 2023, Hector played a mentorship role in one of the OLCF programs designed to assist professors from historically underrepresented colleges in integrating HPC into their curriculum. His scholarly output was impressive, with four published papers, including one for which he is the sole author. Hector was instrumental in assisting PIs with the preparation of INCITE, DD, and SummitPLUS proposals, serving as a co-PI on several occasions. In CY 2023, a notable achievement was Hector's successful proposal submission to the Office of Transition Technology concerning a sustainable energy and AI project for the Innovation Corporation program. This project, which was accepted, marked the first project from an OLCF team member to be granted approval, showcasing Hector's innovative approach and dedication to sustainable technology.

Publications:

Corzo, H. H., Hillers-Bendtsen, A. E., Barnes, A., Zamani, A. Y., Pawłowski, F., Olsen, J., Jørgensen, P., Mikkelsen, K. V., and Bykov, D. (2023). "Coupled cluster theory on modern heterogeneous supercomputers." *Frontiers in Chemistry*, 11, 1154526.

Hillers-Bendtsen, A. E., Bykov, D., Barnes, A., Liakh, D., Corzo, H. H., Olsen, J., Jørgensen, P., and Mikkelsen, K. V. (2023). "Massively parallel GPU enabled third order cluster perturbation excitation energies for cost-effective large scale excitation energy calculations." T*he Journal of Chemical Physics*, 158(14), 144111.

Corzo, H. H. (2023). "The importance of correlation in the molecular orbital picture." In *Chemical Reactivity: Theories and Principles* (Vol. 1, pp. 1-26). Springer Nature Singapore.

Abou Taka, A., Corzo, H. H., Pribram-Jones, A., and Hratchian, H. P. (2022). "Good Vibrations: Calculating Excited-State Frequencies Using Ground-State Self-Consistent Field Models." *Journal of Chemical Theory and Computation*, 18(12), 7286-7297.

4.3.1.8 Maria Batool

Maria Batool, who is jointly appointed by NCCS and CSED, is actively engaged in projects related to drug discovery and collaborative research efforts. Mentored by Jens Glaser from Science Engagement in NCCS and Debsindhu Bhowmik from Advanced Computing Methods for Health Sciences in CSED, Maria's primary focus is the development of HPC-enabled AI methods for identifying small molecule inhibitors targeting disease-related enzymes. Specifically, she contributes to ongoing collaborations aimed at discovering inhibitors for the serine hydroxymethyltransferase, a critical enzyme in various aggressive cancers. Maria's expertise in computational chemistry and drug discovery enhances her contributions to user projects at NCCS, where she utilizes AI tools to predict novel molecules for testing. Additionally, Maria’s involvement in projects that explore the effects of low-dose radiation exposure underscores her commitment to advancing understanding and mitigation strategies through molecule generation methods and computational simulations. This research holds promise for the development of novel radioprotective agents with applications across industries and medical contexts.

4.3.1.9 Khadiza Begam

Khadiza Begam joined the group as a postdoc in December 2023. Her expertise includes electronic structure-based spectral calculations, solvated NMR code implementation, molecular modeling to study condensed phase properties, and other experimental research. Currently, Khadiza is developing the bulk solvent model of the diffracted density mapping to study the proton contributions in biomolecular crystallography. Her work will expand fundamental understanding of biomedical science’s various chemical reactions and transfer processes.

5. RISK MANAGEMENT

CHARGE QUESTION 5: Does the Facility demonstrate effective risk management practices?

OLCF RESPONSE: Yes. The OLCF has a history of successfully anticipating, analyzing, rating, and retiring both project- and operations-based risks. The OLCF risk management approach is modeled after the Project Management Institute's best practices. Risks are tracked and retired, reclassified, or mitigated as appropriate. A change history is maintained for historical reference.

The major operational risks for the OLCF in CY 2023 are listed and described in this section. Planned mitigations and implementations are included in the subsequent descriptions. As of this writing, the OLCF has zero high-priority operational risks, but that could change because the risk management approach continuously reviews and assesses OLCF operations for new risks.

5.1 RISK MANAGEMENT SUMMARY

The OLCF's Risk Management Plan describes a regular, rigorous, proactive, and highly successful review process that is reviewed at least annually and updated as necessary. The plan covers both OLCF operations and its various projects (the conclusion of OLCF-5 and beginning of OLCF-6 in CY 2023). Each project execution plan refers to the main Risk Management Plan but may also incorporate project-specific adjustments. Risks are tracked in a risk registry database application that can track project and operational risks separately.

Operations risks are continually assessed and monitored by the risk owners, the facility management team, OLCF group leaders and section heads, and other stakeholders. When assessing risks, the OLCF management team focuses its attention on the high and moderate risks as well as any low risks within the impact horizons associated with the risk. Trigger conditions and impact dates are recorded in the risk notes narrative section of the register. Risk owners are proactive in tracking trigger conditions and impact horizons for their risks and bringing appropriate management attention to those risks, regardless of the risk rating level.

The OLCF reports a change summary of affected operations risks to the DOE program office as part of its monthly operations report. At the time of this writing, 34 active entries are in the OLCF operations risk register. In general, these fall into two categories: risks for the entire facility and risks for a specific portion of the facility. Facility-wide risks are concerned with safety, funding, expenditures, and staffing. The more focused risks are concerned with reliability, availability, and use of the system or its components (e.g., the computing platforms, power and cooling infrastructure, storage, networks, software, and user support).

The costs of handling risks are integrated in the budgeting exercises for the entire facility. For operations, the costs of risk mitigation are accepted, and residual risk values are estimated by expert opinion and are accommodated, as much as possible, in management reserves. This reserve is continually reevaluated throughout the year.

5.2 MAJOR RISKS TRACKED IN 2023

Table 5-1 contains the major risks tracked for OLCF operations in 2023. The selected risks are all rated medium or high in impact.

Table 5-1. 2023 OLCF major risks.

Risk ID/description	Probability/impact	Action	Status
723: Safety – personal injury	Low/medium	Mitigating	Reduce risk by monitoring worker compliance with existing safety requirements, having daily toolbox safety meetings, conducting periodic surveillance using independent safety professionals, having joint walk-downs by management and work supervisors, and emphasizing the stop-work authority of all personnel. Observations from safety walk-downs are evaluated for unsatisfactory trends (e.g., recurring unsafe acts). Unsatisfactory performance will receive prompt management intervention commensurate with the severity of the safety deficiencies.
1063: Programming environment and tools may be inadequate for future architectures	Medium/medium	Mitigating	OLCF has continued to work with vendors, standards committees, and users to ensure adequate tools are available for users on future architectures. Additionally, the OLCF tracks several risks related to programming environment (both compilers and accelerator libraries) and tools as part of the OLCF-5 risk register and routinely review the status of those risks and work to mitigate or possibly retire them.
1245: System unavailability due to mechanical/electrical system failure	Low/high	Mitigating	The system was designed with leak detection in mind. Mitigation involves performing all preventative maintenance, performing inspections, and monitoring where possible.

5.3 NEW OR RECHARACTERIZED RISKS SINCE LAST REVIEW

5.3.1 Recharacterized Risks

No risks in the OLCF Operations Risk Register were recharacterized during the review year.

5.3.2 New Risks in this Reporting Period

As a result of the OLCF-5 project's successful completion and Frontier's subsequent transition to operations, 10 remaining risks from the OLCF-5 register were transitioned from the OLCF-5 project risk register into Operations. These are shown in Table 5-2 below.

Table 5-2. Risks transferred from OLCF-5 project to operations.

Risk ID	Description	Probability/Impact	Action
1190	Environmental impacts of OLCF-5	Low/Very Low	Mitigate
1208	Programming environment and tools may be inadequate	Low/Low	Mitigate
1212	Vendor does not provide adequate internal management controls	Medium/Low	Mitigate
1276	Maturity and performance of OpenMP	High/High	Mitigate
1277	Quality of Fortran implementations	Medium/Medium	Mitigate
1279	Both Cray (C/C++ only) and AMD compilers are based on the LLVM compiler toolchain	Low/Medium	Accept
1281	Performance and correctness tools for AMD GPUs are immature	Medium/High	Mitigate
1284	Dependency on CUDA toolkit libraries	Medium/High	Mitigate
1285	Programming model interoperability problems	Medium/High	Mitigate
1342	HPE water quality on secondary cooling loop	High/Medium	Accept

5.4 RISKS RETIRED DURING THE CURRENT YEAR

No risks were retired from the OLCF Operations Risk Register during the review year.

5.5 MAJOR RISKS FOR NEXT YEAR

With Frontier now in operation and Summit continuing as part of the SummitPLUS program, both systems are critical to the success of OLCF in 2024. Risk ID 1245 (System unavailability due to mechanical/electrical system failure) directly impacts both systems. Although this is mitigated through preventative maintenance, inspections, and monitoring, it remains in place because a trigger of that risk could have significant effects.

Although major construction activities are not anticipated, routine maintenance is inevitable as are deliveries that may involve moving large equipment through the hallways and which can affect ingress/egress for staff and visitors. Work in and around the computer room and mechanical rooms requires vigilance for hazards such as electrical panels. Staff must also constantly be aware of situations that could create slip/trip/fall and other personal injury hazards. Staff should be reminded of their stop work authority and encouraged to exercise it as necessary (Risk 723).

Risks 406 (System cyber security failure) and 1240 (Failure to handle ECI properly) remain closely monitored as well. With ever-present cyber threats, constant vigilance on the part of the ORNL Cyber Security team and staff at large is crucial for reducing risks to the center caused by any such attack. Related is Risk 1240, especially as the center moves forward with the Scalable Protected Infrastructure program. Staff must remain vigilant to cybersecurity threats and of export control concerns addressed by these risks.

5.6 RISKS THAT OCCURRED DURING THE CURRENT YEAR AND THE EFFECTIVENESS OF THEIR MITIGATION

The risks listed in Table 5-3 were encountered and effectively mitigated in 2023. A short summary of the status and impact on the operations of the OLCF is included for each risk.

Table 5-3. Risks encountered and effectively mitigated in CY 2023.

Risk No. 1147	Data integrity issues in archival data in HPSS
Risk owner	Kevin Thach
Status	Mitigate
Probability	Low
Impact	Cost: Medium Schedule: Medium Scope/Tech: Low
Trigger Event	The risk was triggered in October when two files could not be recovered from HPSS due to a loss of two tapes in a 3+1 RAIT set. One was recovered from Alpine and the other from another source. It is statistically unlikely to hit this again, and we feel the existing stance with RAIT is adequately mitigating this risk.
Mitigations	Of the two files that could not be recovered from HPSS, one was recovered by the user from the Alpine file system and the other was recovered from another source. Upon further discussion with the risk owner, the OLCF thinks that the current mitigation measure (primarily RAIT) remains a sufficient/appropriate mitigation against data loss.
Triggers	Reported loss of data

6. ENVIRONMENT, SAFETY, AND HEALTH

CHARGE QUESTION 6: Does the facility exhibit a culture of continuous improvement in environment, safety, and health (ES&H) practices to benefit staff, users, the public, and the environment? Has the Facility implemented appropriate ES&H measures?

OLCF RESPONSE: Yes. The OLCF is committed to providing a safe workplace for ORNL employees, subcontractors, and visitors. The OLCF is committed to the zero-accident philosophy and believes that all accidents are preventable. The OLCF utilizes DOE safety regulations as specified in 10 CFR 851, ORNL's Standards Based Management System (SBMS), the Battelle Safe Conduct of Research Principles, and many other systems/work practices to obtain zero-accident performance.

The core functions of the Integrated Safety Management (ISM) process require defining the work scope, analyzing the scope for hazards, developing controls, performing work safely, and continuous improvement. These functions are the basis for safe operations in the OLCF data centers. ORNL SBMS contains several procedures for ISM that are implemented in the OLCF. ORNL's SBMS Subject Area Work Control is broken into smaller procedures that provide the requirements for research staff, maintenance staff, space management, and non-employees. The procedures for ISM in Research and Development, Implement Work Control for Operations and Maintenance, Maintain ISM in Laboratory Space, and Implement ISM for Nonemployees are utilized in the OLCF's safety management program. Each program is tailored for the target audience and provides different types of work control documents, research safety summary (RSS) for research staff and lab spaces, a work plan and job hazard evaluation for maintenance staff, and a subcontract with ES&H requirements, including subcontractor/vendor submittal of a hazard analysis. Each type of work control requires the identification of the hazards and the development of controls. Additionally, each type of work control is reviewed and approved. The approval level is a graded approach that depends on the hazard set. Approvals can include ES&H (ORNL and vendor), supervisor, company/vendor representative, technical project officer, group leader, division director, subject matter experts, or a combination. In addition to utilizing the ORNL subject matter experts, the OLCF center manager is a certified safety professional and is fully involved in work planning and safe execution of all center work.

Daily safety management in the facility also utilizes multiple tools. These tools include inspections of job sites or walk throughs, pre-task briefings, briefings on applicable lessons learned/safety snapshots, safety talks, hazard analysis briefings, stop-work issuance, post-task reviews, and management assessments. Management assessments are formal reviews and are recorded in the Assessment and Commitment Tracking System. The tracking system documents the completion of the ORNL ISM process and provides a means for analysis. The DOE ORNL Site Office participates in field implementation and documentation of all operational safety reviews and partners with the ORNL Offices of Institutional Planning and Integrated Performance Management and the Safety Services Division on independent safety management system assessments. The culture of safety at ORNL and the OLCF is reflected in these processes, which seek to reduce and prevent injuries to personnel and potential exposure to hazards associated with operation of the facility.

6.1 NORMAL DAY-TO-DAY OPERATIONS

Normal day-to-day operations of the OLCF include small system installations (same-rack installations), multiple-rack installations, un-racking and excessing equipment, infrastructure upgrades, routine facility maintenance, mechanical system maintenance, 24-7-365 operations support, and oversight of vendor/subcontractors performing on-site work.

6.1.1 Safety Performance

For CY 2023, the normal day-to-day operations met the zero accident performance criteria and remained safe, efficient, and effective. The ES&H performance for normal day-to-day operations included zero total recordable cases, zero Days Away Restricted or Transferred (DARTs), and zero first aid cases.

6.1.2 Normal Day-to Day ES&H Highlights

6.1.2.1 Center Support

The OLCF continued to operate 24-7-365 with operator support on each shift. However, the operator staff was short staffed due to illness and other personal impacts. To provide uninterrupted support, the OLCF resorted to implementing fill-in staff. Fill-in staff included the center manager and a previous operator. The center also identified the need for additional flexibility and succession planning. A job was posted, interviews conducted, and an additional operator selected.

6.1.2.2 Work Control/Monitoring

As in past years, the RSS was revised this year, and noise monitoring was conducted for the three data centers. The addition of new machines has impacted the noise levels in several areas. Noise monitoring results will be added to the RSS for staff to review. Additional postings have also been added to mark the new areas that require hearing protection.

6.1.2.3 Safety Talks

Safety talks are utilized to promote conversations about safety, safety culture, and the Safe Conduct of Research principles. The talks provide an opportunity to reinforce key safety culture principles by making them part of everyday conversations. Safety talk topics can be task-related safety items, facility safety items, or items related to safety away from work. Operations staff performed 114 safety talks in CY 2023.

6.1.2.4 Walk Throughs

Walk throughs are conducted at set intervals and in set room locations for the OLCF operators. Operators focus on facility conditions and observation of work practices. The frequency of walk throughs was modified in CY 2023—they are now conducted every 2 hours. The operator’s walk-through total for 2023 exceeds 4,000 events. Several issues such as alarms, unsafe conditions, and unsafe acts were identified during the walk throughs.

6.1.2.5 Center Hazard Analysis

All installations that require vendor support were conducted in accordance with the center’s hazard analysis. All individuals on site for site work received a pre-task hazard analysis briefing from the center manager or center staff.

6.1.2.6 Staffing or Succession Planning

In addition to adding operator staff, the center added a succession candidate for the center manager. The selected candidate is highly trained in electrical safety and working with support craft. The additional member adds oversight, electrical knowledge, and center management for facility preparation/installs.

6.1.2.7 Learning from Past Lessons Learned and Preparedness

The OLCF continues to learn from events and near misses. In December of 2022, the OLCF experienced a ruptured cooling hose under the floor of the 2nd floor data center. The rupture resulted in water dripping through the concrete floor and into the 1st-floor data center. In response to this event, the center realized that it was unprepared in terms of equipment to protect computing resources and to remediate the situation. In response, the facility initiated actions, including sealing the 2nd floor concrete slab and changing the old cooling hoses. The OLCF also purchased equipment for each floor that included protective materials and equipment to remediate water events. In the fall of 2023, another water event occurred on the 2nd floor and resulted in water dripping into the first floor. The event's impacts were minimized by using materials purchased after the 2022 event, clearly demonstrating a lesson learned.

6.2 LARGE-SCALE SUPERCOMPUTER INSTALLATION/ACTIVITIES

The CY 2023 large-scale workload included the completion of the Frontier start-up, delivery of a new NOAA computer (C6), removal of a NOAA computer (C3), and the installation of four new storage solutions.

For CY 2023, the large-scale installation/activities also met the zero accident performance criteria. All work was accomplished safely with zero total recordable cases, zero DARTs, and zero first aid cases. In addition, the work was performed without adverse impact to the facility, co-located ORNL employees, or the environment.

6.2.1 Large-Scale Supercomputer Installation/Activities ES&H Highlights

6.2.1.1 Vendor Hazard Analysis

All vendor personnel were briefed on the new hazard analysis by the OLCF center manager. In addition to the hazard analysis, (tasks, hazards, and controls), the applicable safe conduct of research principles, human performance, applicable lessons learned, and stop/suspend work authorization are covered during the briefing.

6.2.1.2 Hazardous Energy Control

During maintenance events that required hazardous energy control of the Frontier electrical system, normal practice was to "air gap" the electrical system prior to work. This practice allowed HPE personnel to perform hands-on maintenance that could subject an employee to an electrical hazard.

During the review cycle, it was noted that HPE personnel should be trained like ORNL personnel. The HPE site manager was tasked with identifying a set of on-site employees that would be depended on by HPE to perform tasks that could require lock/tag/verify as a control. The HPE manager provided the list, and a group was selected with the help of the OLCF manager. This group received the HPE electrical training as well as HPE's lock out/tag out training. The group was also briefed by the ORNL facility engineer and the OLCF manager to ensure that they understood the importance of the lock/tag/verify process and the importance of personal locks. Additionally, the OLCF manager discussed the process with HPE's ES&H staff, and ORNL will enroll all selected HPE employees in ORNL's lock/tag/verify training.

The OLCF also identified an inconsistency in arc flash safety gear for HPE personnel that open and close machine breakers. The OLCF manager suggested to HPE ES&H staff that HPE employees performing this task be provided the correct level of arc flash PPE. HPE then provided the correct PPE, which is similar to the clothing provided by ORNL for ORNL staff.

6.2.1.3 Stabilization of Frontier's Cooling Loops and Fluid Replacement

Stabilization of Frontier's cooling loops, which contained the original vendor-supplied water, as well as the replacement of the water with new coolant began during CY 2023. The effort included several meetings between the OLCF, HPE, and HPE's sub-tier water quality vendor. HPE analyzed the hazards and provided the hazard analysis(es), and the OLCF concurred with their findings.

6.2.1.4 Subcontractor ES&H Engagement

The engagement of the subcontractor's ES&H support is critical to ensuring that the vendor is adequately taking ownership of their staff members' safety and health and fulfilling subcontract ES&H requirements. This effort/engagement had not been past practice for the center or the vendor. In 2023, HPE's assigned ES&H support representative was requested on site at ORNL to spend time reviewing HPE's work at ORNL.

During this visit, the HPE ES&H professional was shown all HPE activities, including where staff members park, the walking path to the facility, how equipment is received, ORNL policies and programs, all center spaces in which HPE personnel work, and how the center operates. Both parties benefited from the visit, with HPE taking some ORNL ES&H program aspects back for incorporation into HPE's programs and HPE gaining a better understanding of HPE employee tasks and how ORNL/OLCF provides ES&H support for the work.

Additionally, the OLCF expressed their concern about HPE's lack of ES&H support for their employees and the lack of HPE-provided ES&H support for HPE sub-tier and temporary staff. HPE is in the process of developing a program that trains these temporary employees. The OLCF also benefited from the event. The OLCF received several suggestions for loading dock safety, ladder safety, and material handling for large hardware for future incorporation into the OLCF's practices.

HIGH-PERFORMANCE COMPUTING FACILITY
2023 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2024

7. SECURITY

CHARGE QUESTION 7: (a) Does the Facility exhibit a culture of continual improvement in cybersecurity practices? (b) Does the Facility have a valid cybersecurity plan and Authority to Operate? (c) Does the Facility have effective processes for compliance with applicable national security policies related to export controls and foreign visitor access?

OLCF RESPONSE: Yes. The OLCF maintains a strong culture of continuous operational improvement, especially in cybersecurity. The most recent OLCF Authority to Operate was granted on March 22, 2017, and is managed through an ongoing authorization process; no authorization termination date is set (Figure 7-1). Updates to the ORNL cybersecurity program plan as well as the supercomputing security zone system security plan have been made in CY 2023.

The technical staff members track and monitor existing threats and vulnerabilities to assess the risk profile of the OLCF operation. The Facility is committed to innovating in this area by developing open-source tools and employing cutting-edge practices that enhance the operation without increasing the OLCF’s risk profile.

The OLCF employs ORNL policies related to export control and foreign visitor access.

Department of Energy

ORNL Site Office
P.O. Box 2008
Oak Ridge, Tennessee 37831-6269

March 22, 2017

Mr. Kevin A. Kerr
Information Systems Security Manager
Oak Ridge National Laboratory
UT-Battelle, LLC
Post Office Box 2008
Oak Ridge, Tennessee 37831-6045

Dear Mr. Kerr:

AUTHORIZATION DECISION DOCUMENT FOR OAK RIDGE NATIONAL LABORATORY (ORNL) SUPERCOMPUTING ENCLAVE

Reference: Letter from Kevin A. Kerr to Johnny O. Moore, subject, *Contract DE-AC05-00OR22725, ORNL Supercomputing Enclave Approval to Operate*, dated February 21, 2017

As the Authorizing Official, I have reviewed the referenced request. The ORNL Supercomputing Enclave is authorized to operate. No additional conditions outside the substance of the request are required.

The information system is now being managed by an ongoing authorization process, thus an authorization termination date is not set. I accept the responsibility for performing all necessary activities associated with the ongoing authorization process.

If there are any questions or additional information is required, please contact John Young at (865) 576-7471 or youngjc1@ornl.gov.

Sincerely,

Johnny O. Moore, Manager
ORNL Site Office

Enclosure

cc w/enclosure:
Mike E. Bartell, ORNL
Amy D. Nuckols, ORNL
Neil Masincupp, SC-OR
Martha J. Kass, SC-OSO
John C. Young, SC-OSO

Document/Material Transmitted Contains Official Use Only Information.

When separated from enclosure, this document **does not** contain Official Use Only.

Figure 7-1. OLCF Authority to Operate.

7.1 SUMMARY

All IT systems that operate for the federal government must have authority to operate. This involves developing and obtaining approval for a policy and implementing a continuous monitoring program to confirm that the policy is effectively implemented. The ORNL accreditation package currently uses the NIST Special Publication 800-53, revision 5, Security and Privacy Controls for Federal Information Systems and Organizations, and the US Department of Commerce Joint Task Force Transformation Initiative (August 2009) as guidelines for security controls. The OLCF determined that the highest classification of data is moderate based on the guidelines for information classification in the Federal Information Processing Standards Publication 199, Standards for Security Categorization of Federal Information and Information Systems, Computer Security Division, Information Technology Laboratory, NIST. The OLCF is accredited at the moderate level for protecting the confidentiality and integrity of user and system information, and this accreditation authorizes the Facility to process sensitive, proprietary, and export-controlled data.

Security at the OLCF is built upon a strong configuration management baseline. Puppet, among other tools, is used to enumerate and deploy both security and operational configurations required by policy and best practices. HPC system images delivered from the vendor are augmented with a Puppet configuration to bring all nodes of a system into compliance. Other important controls include enforcement of multifactor authentication in the OLCF moderate enclave, lightweight and well-adopted configuration management procedures, adoption of DevSecOps principles (e.g., tight intergroup coordination and use of continuous integration) and strong incident response and triage capabilities (e.g., operational and security dashboards and frequent practice).

Year after year, cybersecurity planning is becoming more complex as the center continues its mission to enable world-class science. The center is very proactive and views its cybersecurity plans as dynamic documentation that it will preemptively respond with and modify as needed to provide an appropriately secure environment. The OLCF abides by the Health Insurance Portability and Accountability Act's Privacy and Security Rule to provide supercomputing resources to projects that contain protected health information (PHI). The OLCF also abides by the International Traffic and Arms Regulations (ITAR) and DOE I 471.7, Controlled Unclassified Information for projects that contain these levels of sensitive information.

7.2 SECURITY OPERATIONS

The security team performs a wide range of activities, including security policy development and assessment, event monitoring, incident handling and reporting, vulnerability scanning and triage, and security system engineering. The team maintains a DevSecOps mindset to automate security when possible and to integrate with the software development and systems/operations teams in the center to help design and secure new capabilities as they are developed. The security team has developed and enforced new procedures and policies to help formalize security related processes and to improve security posture. This includes a new security data review policy, which will help provide adequate information from data owners and allow the security team to properly analyze the data for any potential security issues. The team was involved in developing a new auditing procedure for data migration between systems, which included CITADEL`s migration from Summit to Frontier. The team has also been involved with internal procedures, and the OLCF has been involved in external documentation such as the NIST SP 800-23 draft. The SP 800-23 document outlines security architecture, threat analysis, and security posture for HPC. The team has also been involved in IRI discussions, in which the objective was to provide the PI with the OLCF's definition of Human Health Data and cover any edge case so that the application process can appropriately categorize their project into OLCF's Category 1, 2, or 3. This would ensure that Human Health Data that requires extra security controls via a Data Use Agreement or other Limited Data Set user agreements is categorized correctly and treated appropriately.

To facilitate continuous operational improvement, the security team has an open "Security Questions" channel by which anyone can ask security related questions about vulnerabilities, operational procedures, policy, and more. The team also participated in a Security of ASCR Facilities exercise, which was a collaboration with the other ASCR facilities and designed to share knowledge on how each facility handles incident response. The security team was also involved in SC23, where a team member hosted a talk during the International Workshop on Cyber Security in High Performance Computing. This talk discussed key security differences between enterprise systems and the common features of HPC environments.

In addition to project work, the security team also tracks significant cybersecurity events that are above and beyond normal baseline threats reported in the OLCF ticket tracking system. Baseline threats such as background SSH (Secure Shell) scanning, firewall probing, and software vulnerabilities patched within quarterly patching windows are expected. Examples of events above this baseline include the public disclosure of serious software vulnerabilities, detected suspicious user behavior, and observed but unexpected patterns of system event logs and metrics. In CY 2023, the OLCF security team tracked 30 above-baseline threats—none of which resulted in a compromise of OLCF systems by malicious attackers. Each of these events is treated as a live incident and used to practice the OLCF incident response plan several times throughout the year.

Because HPC systems are scientific instruments, care must be taken to determine the appropriate response for each vulnerability and potential threat. Industry-standard responses such as emergency patching and application allow-listing are difficult to implement quickly because scientific software stacks provided by the facility and users may need to be rebuilt and then tested to ensure the correctness of scientific results.

At its core, security has recently become a data science problem. Network intrusion detection, host-based intrusion detection, system and firewall logging, vulnerability scanning, netflow collection, and databases (e.g., hardware inventory, software inventory, user CRM) all have information to help an incident responder take appropriate action. The OLCF security team continues to develop Copacetic to help wrangle data from these different data sources and apply business rules to ensure that the actual state of the systems within the data center match the desired state. As zero-trust initiatives are completed in CY 2023, active monitoring of data sources to provide real-time access control decision-making will be required.

7.3 OLCF USER VETTING

The OLCF follows a set of rigorous controls for vetting user access, as defined by ORNL and DOE policy, to ensure compliance with export-control regulations and foreign visitor access policies.

7.3.1 OLCF Projects

Users must be added to an approved OLCF project to obtain access to OLCF resources. An ORNL export control officer reviews the scope of work for all OLCF user projects to determine whether there are any export-control restrictions to which the OLCF must adhere and to place an internal designation of category 1, category 2, or category 3 on each project. These categories then drive the business processes that are followed for each applicant and defined in Table 7-1.

Table 7-1. Export control review categories for projects.

Category designation	Category description	PI actions before project activation
Category 1	The category 1 rating is applied if the project is open fundamental research that does not involve proprietary input and/or output, sensitive data, and/or export-control restrictions above EAR99.	Sign OLCF PI agreement
Category 2	The category 2 rating is applied if the project involves proprietary input and/or output, sensitive subject areas, and/or export-control restrictions above EAR99 but below ITAR.	Sign OLCF PI agreement Participate in mandatory security call to review risks/restrictions associated with category 2 projects
Category 3	The category 3 rating is applied if the project involves proprietary input and/or output, sensitive subject areas, PHI, PII, and/or export-control restrictions above EAR99 to include ITAR.	Sign OLCF PI agreement Participate in mandatory security call to review risks/restrictions associated with category 3 projects

Sensitive information, including proprietary and export-controlled information, is segregated and protected in the specific project area to protect it from unauthorized access, and specific storage rules and requirements are relayed to the PI and individual project users to further prevent information mishandling. If a project is rated category 2 or above, then the project PI must participate in a mandatory security call with the Information Systems Security Officer (or designee) to review the risks and restrictions before the project is enabled.

Once the security call is complete and all other project requirements are met, the project is enabled in the OLCF RATS and labeled with the appropriate category, and any export control restrictions are added to the project.

7.3.2 OLCF Users

All users requesting access to OLCF resources are required to fill out the OLCF account application form and provide the project identification and PI for the project they are requesting to join. Based on the category of the project designated in RATS, the following requirements (Table 7-2) must be met before the user can be added to the project and provided access to OLCF resources.

Table 7-2. OLCF Project Category Requirements.

Project category	PI approval	ORNL Personnel Access System (PAS)[1] or Restricted Party Screening (RPS)[2]	UA[3]	Sensitive data rules[4]	Level 2 identity proofing[5]
Category 1	PIs must approve all user account requests to access to their project.	a. Approved PAS is required for all applicants that are not US citizens or lawful permanent residents. b. All US citizens or lawful permanent residents go through Restricted Party Screening (RPS).	Must have valid user agreement	N/A	Required
Category 2	PIs must approve all user account requests to access to their project.	c. Approved PAS is required for all applicants that are not US citizens or lawful permanent residents. d. All US citizens or lawful permanent residents go through RPS.	Must have valid user agreement	Must return signed sensitive data rules	Required
Category 3	PIs must approve all user account requests to access to their project.	e. Approved PAS is required for all applicants that are not US citizens or lawful permanent residents. f. All US citizens or lawful permanent residents go through RPS.	Must have valid user agreement	Must return signed sensitive data rules	Required

[1] PAS: The system that ORNL uses to process on-site and/or remote access for foreign nationals and non-employees.
[2] RPS: ORNL maintains a subscription to the Descartes Visual Compliance tool, which is used to look up applicants and their institutions that do not require PAS approval. If any hits are found on the user or the user's institution, then the information is turned over to the export control officer.
[3] UA: Serves as the master agreement that establishes the general terms and conditions, including the disposition of intellectual property, for work at any of ORNL's user facilities. A UA must be executed with each user's institution.
[4] Sensitive data rules: This form contains the user acknowledgment, which documents that users on a category 2 project are aware of the risks and rules for accessing the sensitive project.
[5] Level 2 identity proofing: The OLCF uses RSA SecurID tokens for authenticating to OLCF moderate resources. Level 2 identity proofing of all applicants is required as part of the OLCF moderate Certification and Accreditation. To achieve Level 2 identity proofing, applicants must have their identity and RSA SecurID token verified by a notary or an OLCF-designated registration authority. The token is not enabled until all the above steps are completed, including the return of the original notarized form.

8. STRATEGIC RESULTS

CHARGE QUESTION 8: (a) Are the methods and processes for monitoring scientific accomplishments effective? (b) Is the facility collaborating with technology vendors and/or advancing research that will impact next-generation HPC platforms? (c) Has the Facility demonstrated effective engagements with critical stakeholders (e.g., the Office of Science's Science Programs, DOE Programs, DOE national laboratories, User Facilities, and/or other critical US government stakeholders [if applicable]) to both enable mission priorities and gain insight into future user requirements?

OLCF RESPONSE: Yes. OLCF projects and user programs are advancing DOE's mission to ensure US security and prosperity by addressing its energy, environmental, and nuclear challenges through transformative science and technology solutions. The selected accomplishments described in this section highlight how the OLCF is advancing two strategic objectives of DOE's Strategic Plan Goal 1, "Science and Energy: Advance foundational science, innovate energy technologies, and inform data driven policies that enhance economic growth and job creation, energy security, and environmental quality...," as stated in the DOE Strategic Plan: 2014–2018 (March 2014):

- Strategic Objective 2: Support a more economically competitive, environmentally responsible, secure, and resilient US energy infrastructure
- Strategic Objective 3: Deliver the scientific discoveries and major scientific tools that transform the understanding of nature and strengthen the connection between advances in fundamental science and technology innovation

8.1 SCIENCE HIGHLIGHTS AND ACCOMPLISHMENTS

The Facility collects and reports annually the number of refereed publications that result (at least in part) from the use of the Facility's resources. For the Leadership Computing Facilities (LCFs), tracking is done for a period of 5 years following the project's use of the Facility. This number may include publications in press or accepted but not submitted or in preparation. This is a reported number, not a metric. Additionally, the Facility may report other publications when appropriate.

The Facility also regularly searches for and actively solicits information about possible scientific highlights produced by projects that use OLCF resources. Specific questions about the possible availability of such highlights are asked as part of each quarterly and annual report solicitation for projects. In addition, scientific liaisons to INCITE projects regularly report potential scientific highlights to the Facility management. These active avenues for discovering potential scientific highlights are augmented by a regular (i.e., roughly monthly) examination of the ongoing publication tracking described above.

Publications in high-impact journals or papers that garner a significant number of early citations (among other criteria) are deemed good candidates for features.

8.1.1.1 OLCF Publications Report

In 2023, 495 publications resulting from the use of OLCF resources were published based on a data collection completed on April 8, 2024. In this document, "year" refers to the calendar year unless it carries the prefix FY. In the 2022 OLCF OAR, 571 publications were reported (this number has been significantly revised upward; see below). A list of 2014–2023 publications is available on the OLCF website (https://www.olcf.ornl.gov/leadership-science/publications/). Guidance allows accepted and in-press publications to be reported, but the OLCF reports only publications appearing in print in the year under review. However, the OLCF continues to search for publications after the OAR is submitted to DOE each year, and the number of publications shown in previous OARs is updated in the current report. Table 8-1 provides the updated, verified, and validated publications count for the 2012–2023 period, showing overall consistent growth in both the total publications count and the number of publications in journals with high-impact factors. A small year-over-year decrease for 2023 was noted relative to the immediately preceding reporting years. Similar small decreases in total publication output were also observed in 2019 and 2014. Like 2023, these years marked the first full production availability of Summit and Titan, respectively.

Table 8-1. Summary of unique OLCF publications for 2012–2023.

Year	Unique, confirmed OLCF publications	High-impact publications with JIF* >10
2023	495	17
2022	571	17
2021	560	16
2020	525	19
2019	459	21
2018	496	20
2017	477	27
2016	467	33
2015	366	21
2014	315	16
2013	364	9
2012	342	20

*JIF = Journal impact factor

8.1.1.2 OLCF Publication Methodology

The Facility collects and reports annually the number of refereed publications resulting (at least in part) from use of the Facility's resources. This number may include publications in press or accepted but not submitted or in preparation. This is a reported number, not a metric. In addition, the Facility may report other publications where appropriate. The OLCF employs a multipronged approach to discovering, validating, and reporting publications, and this approach is explained in more detail below.

8.1.1.3 Discovery Methods for OLCF Publications

There are three primary methods of identifying OLCF publications: COBRA, Resolution, and self-reports.

COBRA is a system developed by ORNL to automate the discovery and management of publications related to scientific facilities and organizations. It uses various automatic methods to search for publications across multiple sources such as Web of Science, Google Scholar, Scopus, and more. It also verifies the publications by looking at the metadata, full text, acknowledgments, and funding agencies, among other information.

Resolution is ORNL's publication clearance and tracking system for publications with an ORNL author. To capture publications from Resolution, the COBRA administrator manually queries all publications in Resolution to collect any that contain OLCF keywords. If any are found, then the publications are imported into COBRA.

The final method of discovery is self-reporting from users. These reports are generally collected through quarterly and closeout reports. Upon receipt of reports, all publications listed in the report are manually verified first by querying COBRA to see if the publication has already been captured. If the publication is not already in COBRA, then OLCF staff will search online for the publication. Once found, a review of the acknowledgment and funding sections are conducted to locate the keywords and/or OLCF acknowledgment statement. If they meet the criteria, then they are added to COBRA.

8.1.1.4 Publication Validation

Once the publications are in COBRA, the next step is to validate each publication to ensure the work utilized OLCF resources. This is done by ORNL's library staff and/or OLCF staff. This is generally done by reviewing the publication and looking for the acknowledgment statement, associated keywords, and/or directly contacting the authors of the publication. Once validated, the publication is marked "confirmed" in COBRA.

8.1.1.5 How COBRA Results are Disseminated

Once the discovery process has been completed, COBRA has many features for publication management. COBRA provides the user with an interactive web-based GUI that allows for searching, filtering, and exporting of the data. The publication information data also comes from multiple sources and sources that currently have no openly available API (self-reports, internal systems, Google alerts); to accommodate these sources, COBRA also allows for publications to be entered manually. To increase efficiency in manual entry, COBRA can pull most of the metadata associated with a publication from its DOI. COBRA uses the Web of Science and Cross-Ref APIs to automatically populate many fields, including title, authors, and publication date. COBRA also calculates metrics based on the publication metadata such as high impact, citation count, and highly cited publications while keeping track of where publications are in their publication process. If greater detail is required, then a query can be executed on the data, and those records can be downloaded into a spreadsheet for further review.

The OLCF considers COBRA to be an invaluable addition to the publication workflow, and it greatly improves publication discovery and overall management beyond reliance on self-reports and ad-hoc tracking of publications (e.g., an Excel spreadsheet).

8.1.1.6 Scientific Accomplishments

The OLCF advances DOE's science and engineering enterprise by fostering robust scientific engagement with its users through the INCITE liaison program, the user assistance program, and the OLCF DD program. The following subsections provide brief summaries of select scientific and engineering accomplishments as well as resources for obtaining additional information. Although they cannot capture the full scope and scale of achievements enabled by the OLCF in 2023, these accomplishments advance the state of the art in science and engineering R&D across diverse disciplines and advance DOE's science programs toward mission goals. OLCF users published many breakthrough publications in high-impact journals in 2023, which is an additional indication of the breadth of these achievements (Table 8-2).

Table 8-2. Publications in high-impact journals in 2023.

Journal	Number of publications
Nature	1
Nature Materials	1
Nature Nanotechnology	1
Nature Physics	3
Advanced Materials	1
Nature Chemical Biology	2
ACS Nano	2
Nature Geoscience	1
Nano Letters	1
Nature Communications	2
Trends in Plant Science	1
Nucleic Acid Research	1

Altogether in 2023, OLCF users published 50 papers in journals with a journal impact factor (JIF) of greater than 7 and 17 papers in journals with a JIF greater than 10.

8.1.2 Science Highlights

8.1.2.1 Big Flex for Big Science

Simulations used exascale computing to approach quantum accuracy.

PI: Vikram Gavini, University of Michigan

Allocation Program: DD

The Science

Researchers used the world’s first exascale supercomputer to run one of the largest simulations of an alloy ever and achieve near-quantum accuracy. The study employed Frontier, the 1.14-exaflop HPE Cray EX supercomputer at DOE’s ORNL, to take a first-principles approach to simulation via the Schrödinger equation, which describes the evolution of microscopic systems, including their probabilistic nature.

The Impact

The team used an integrated computational framework on the Frontier and Summit supercomputers to simulate dislocations, or defects, in a magnesium system of nearly 75,000 atoms. Magnesium alloys make promising candidates for lighter alloys, but the lack of dislocations in magnesium’s atomic structure can lead to brittleness and cracking. Understanding dislocations in magnesium alloys could lead to lighter, more flexible alloys for industry. Results could be used to help design candidates for new alloys and to fuel other computational design efforts such as drug discovery.

Summary

A University of Michigan-led team won the Association for Computing Machinery's 2023 Gordon Bell Prize on November 16 for their study that used the Frontier supercomputer to run one of the largest simulations of an alloy ever and achieve near-quantum accuracy. The study led by Vikram Gavini employed Frontier to take a first-principles approach to simulation via the Schrödinger equation, which describes the evolution of microscopic systems, including their probabilistic nature. The team used DFT to simulate design and discovery of new materials. The model achieved the fastest sustained performance recorded for any materials simulation: 659.7 petaflops using Frontier. Results yielded the most detailed picture yet of such a structure at an accuracy approaching quantum accuracy (Figure 8-1). This information could be used to help design candidates for new alloys and to fuel other computational design efforts such as drug discovery.

This work brought together a novel use of AI for parameterizing a density functional trained on fully correlated quantum results. It achieved accurate energetics outside the small elements it was parameterized for (H, Li, N, and Ne). This transferability was enabled by targeting the underlying physics of electrons rather than fitting large datasets. A combination of algorithmic, communication, and GPU kernel optimizations led to reaching 659.7 petaflops on Frontier. OLCF Facility liaison and scientific collaborator David M. Rogers worked with the team to explore scientific and computational collaboration opportunities, prototypes of machine-specific optimizations, effective use of libraries, and reproducible code compilation and use. The software is directly available to run on Frontier via ORNL's user-managed module program, and the team is highly collaborative and supportive of groups with new materials applications and methodology needs.

Funding

This research was supported by the ECP, a collaborative effort of the DOE Office of Science and the National Nuclear Security Administration, the DOE BES program, and the DOE Office of Science's ASCR program.

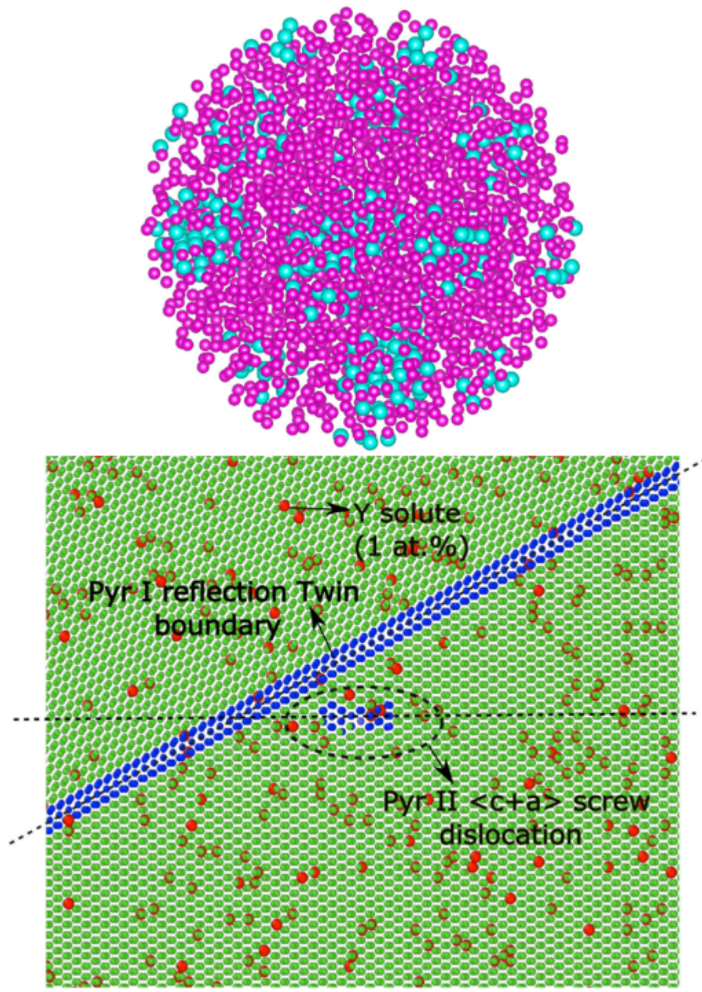

Figure 8-1. Top: A simulated quasicrystal structure in a ytterbium-cadmium alloy (Perlmutter). Bottom: Simulated magnesium system of nearly 75,000 atoms (Frontier). (Credit: Vikram Gavini.)

Publication: Sambit Das, et al., "Large-Scale Materials Modeling at Quantum Accuracy: Ab Initio Simulations of Quasicrystals and Interacting Extended Defects in Metallic Alloys." *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC23)*. https://doi.org/10.1145/3581784.3627037.

Related Link: "Big Flex for Big Science: Frontier search for lightweight, flexible alloys earns Gordon Bell Prize nomination," OLCF News (November 14, 2023)

8.1.2.2 Pulling Clouds into Focus

Frontier simulations bring long-range climate forecasts within reach.

PI: Mark Taylor, Sandia

Allocation Program: ECP

The Science

A research team used Frontier, the 1.14-exaflop HPE Cray EX supercomputer at ORNL, to achieve record speeds in modeling worldwide cloud formations in 3D. The computational power of Frontier, the fastest computer in the world, shrinks the work of years into days to bring detailed estimates of the long-range consequences of climate change and extreme weather within reach.

The Impact

The Energy Exascale Earth System Model (E3SM) project overcomes these obstacles by combining a new software approach with massive exascale throughput to enable climate simulations that run at unprecedented speed and scale. The Simple Cloud-Resolving E3SM Atmosphere Model (SCREAM) focuses on cloud formations as part of the overall project.

Summary

A team of scientists from across the national laboratory complex won the Association for Computing Machinery's 2023 Gordon Bell Special Prize for Climate Modeling for developing a model that uses Frontier to simulate decades' worth of cloud formations. The research team used Frontier to achieve record speeds in modeling worldwide cloud formations in 3D (Figure 8-2). The computational power of Frontier shrinks the work of years into days to bring detailed estimates of the long-range consequences of climate change and extreme weather within reach. Gauging the likely impact of a warming climate on global and regional water cycles poses one of the top challenges in climate change prediction. Scientists need climate models that span decades and include detailed atmospheric, oceanic, and ice conditions to make useful predictions. However, 3D models that can resolve the complicated interactions between these elements, particularly the churning, convective motion behind cloud formation, have remained computationally expensive, leaving climate-length simulations outside the reach of even the largest, most powerful supercomputers—until now. The E3SM project's Simple Cloud Resolving E3SM Atmosphere Model puts 40-year climate simulations, a longtime scientific goal, within reach.

OLCF liaison Dr. Matthew Norman was a member of the team whose contribution was performing most of the porting of the Rapid Radiative Transfer Model for GCMs in Parallel (RRTMGP) radiation code to portable C++ using the Yet Another Kernel Launcher library, of which he is the primary developer. The RRTMGP radiation code is one of the most expensive components of SCREAM used in the Gordon Bell submission with microphysics, dynamics, and subgrid-scale mixing being the other components dominating run time. The work was impressive not only because it is the first storm-resolving climate simulation with all physics included in the run time, but because it also ran at a throughput of around 500× real time, which enables simulation of actual climate timelines in a practical time frame when given enough node hours in an allocation.

Funding

This work was supported by the Energy Exascale Earth System Model project, funded by the DOE Office of Science's Office of Biological and Environmental Research, and by the ECP.

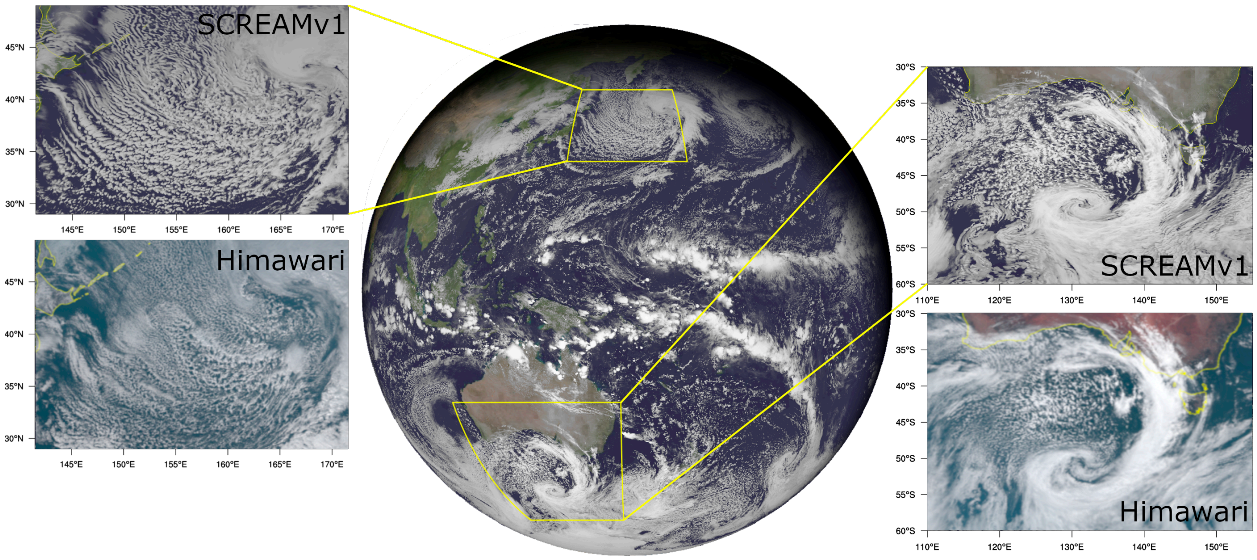

Figure 8-2. Cloud-resolving E3SM atmosphere model on Frontier. Years' worth of climate simulations at unprecedented speed and scale. (Credit: Bill Hillman/Sandia, DOE.)

Publication: Mark Taylor, et al. 2023. "The Simple Cloud-Resolving E3SM Atmosphere Model Running on the Frontier Exascale System." In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC23)*. Association for Computing Machinery, New York, NY, USA, Article 7, 1–11. https://dl.acm.org/doi/10.1145/3581784.3627044

Related Link: "Pulling Clouds into Focus: Frontier simulations bring long-range climate forecasts within reach," OLCF News (November 14, 2023)

8.1.2.3 New Insights into a Shapeshifting Protein Complex

Georgia State University researchers use Summit to chart the multitasking abilities of Transcription factor IIH (TFIIH) and gain new understanding of DNA repair.

PI: Ivaylo Ivanov, Georgia State University

Allocation Program: INCITE

The Science

Transcription factor IIH (TFIIH, pronounced "TF two H") is a veritable workhorse among the protein complexes that control human cell activity. It plays critical roles both in transcription—the highly regulated enzymatic synthesis of RNA from a DNA template—and in the repair of damaged DNA. But how can one protein assembly participate in two such vastly different genomic tasks? A team of Georgia State University researchers used the OLCF's Summit supercomputer to help answer that question. By conducting multiple molecular dynamics simulations of TFIIH in transcription and DNA repair-competent states and then contrasting the structural mechanisms at work, the team made an interesting discovery: TFIIH is a shapeshifter, reconfiguring itself to meet the demands of each task.

The Impact

This new, detailed picture of TFIIH's mechanical dynamics provides insights into the principal motions that allow TFIIH to remodel DNA in transcription initiation versus nucleotide excision repair. This can be useful information in the quest to treat genetic disorders. Learning how genetic mutations impair the function of TFIIH is the first step in designing therapeutic strategies such as gene editing.

Summary

A Georgia State University-led team used the Summit supercomputer to help understand how one protein assembly, TFIIH, can participate in two such vastly different and extremely important genomic tasks: playing critical roles both in transcription and in the repair of damaged DNA. By conducting multiple molecular dynamics simulations of TFIIH in transcription and DNA repair-competent states and then contrasting the structural mechanisms at work, the team made an interesting discovery: TFIIH is a shapeshifter, reconfiguring itself to meet the demands of each task (Figure 8-3). The team used graph algorithms to partition TFIIH's protein network into strongly connected components, thereby allowing them to identify dynamic modules—the pieces that move together. In turn, these models showed how the modules move with respect to other parts of the structure.

Funding

This study was funded by the National Science Foundation's Directorate for Biological Sciences, the National Institute of Environmental Health Sciences, and the National Cancer Institute (NCI).

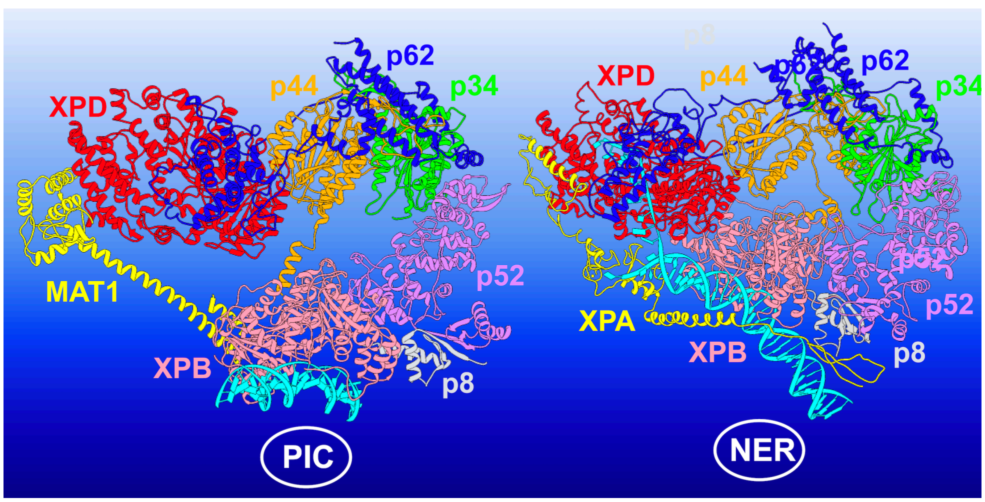

Figure 8-3. TFIIH's conformational switching between holo-PIC (transcription preinitiation) and NER (nucleotide excision repair). The protein complex changes its structure to execute each function. The TFIIH subunits are colored as follows: XPD red, p62 blue, p44 orange, p34 green, p52 purple, p8 light grey, and XPB pink. MAT1 and XPA are shown in yellow, and DNA is cyan. (Credit: Chunli Yan/Georgia State University.)

Publication: Yu, J., Yan, C., Dodd, T. et al. "Dynamic conformational switching underlies TFIIH function in transcription and DNA repair and impacts genetic diseases." *Nature Communications* 14, 2758 (2023). https://doi.org/10.1038/s41467-023-38416-6

Related Link: "New Insights Into a Shapeshifting Protein Complex: Georgia State University researchers use Summit to chart the multitasking abilities of TFIIH and gain new understanding of DNA repair," OLCF News (July 7, 2023)

8.1.2.4 Advancing Nanoscience Through Largescale MD Simulations

The OLCF teams with ORNL's Center for Nanophase Materials Sciences (CNMS) and Stony Brook University to learn how cicada wings kill bacteria.

PI: Jan-Michael Carrillo, ORNL

Allocation Program: DD

The Science

The wings of the cicada insect have the unusual ability to kill microbes on contact. Researchers at Stony Brook University's Department of Materials Science and Chemical Engineering developed a simple technique to duplicate the cicada wing's nanostructure and its nanopillars' microbe-killing ways. But it took a collaboration between ORNL's CNMS and the OLCF to discover how they work. Using the Summit supercomputer, a CNMS researcher ran a coarse-grained molecular dynamics simulation of about 1,000,000 particles (Figure 8-4). Those simulations showed how and where the bacteria's cell membrane stretched and collapsed within the local structure of the nanopillars—suggesting that membrane rupture occurs when the pillars generate sufficient tension within the lipid bilayer clamped at the edges of pillars.

The Impact

These findings could eventually lead to products that use such nanopillar designs for inherently antibacterial surfaces that are more effective than current chemical treatments. The Stony Brook team plans to continue using simulations to develop a more complete picture of the mechanisms at play, particularly the nanopillars' self-cleaning functionality, before applying the nanosurface to biomedical devices such as stents or catheters.

Summary

Teams of engineers, chemists, and biologists have analyzed the physical and chemical properties of cicada wings, hoping to unlock the secret of their ability to kill microbes on contact. When researchers at Stony Brook University developed a simple technique to duplicate the cicada wing's nanostructure, they were still missing a key piece of information. How do the nanopillars on its surface actually eliminate bacteria? Thankfully, they collaborated with an ORNL researcher who ran a coarse-grained molecular dynamics simulation of about 1,000,000 particles on the Summit supercomputer.

Funding

This study was partially funded by the National Science Foundation and the International Collaborative Research Program of the Institute for Chemical Research at Kyoto University. The work was performed at ORNL's CNMS, the OLCF, and the National Synchrotron Light Source-II at Brookhaven National Laboratory, all of which are DOE User Facilities.

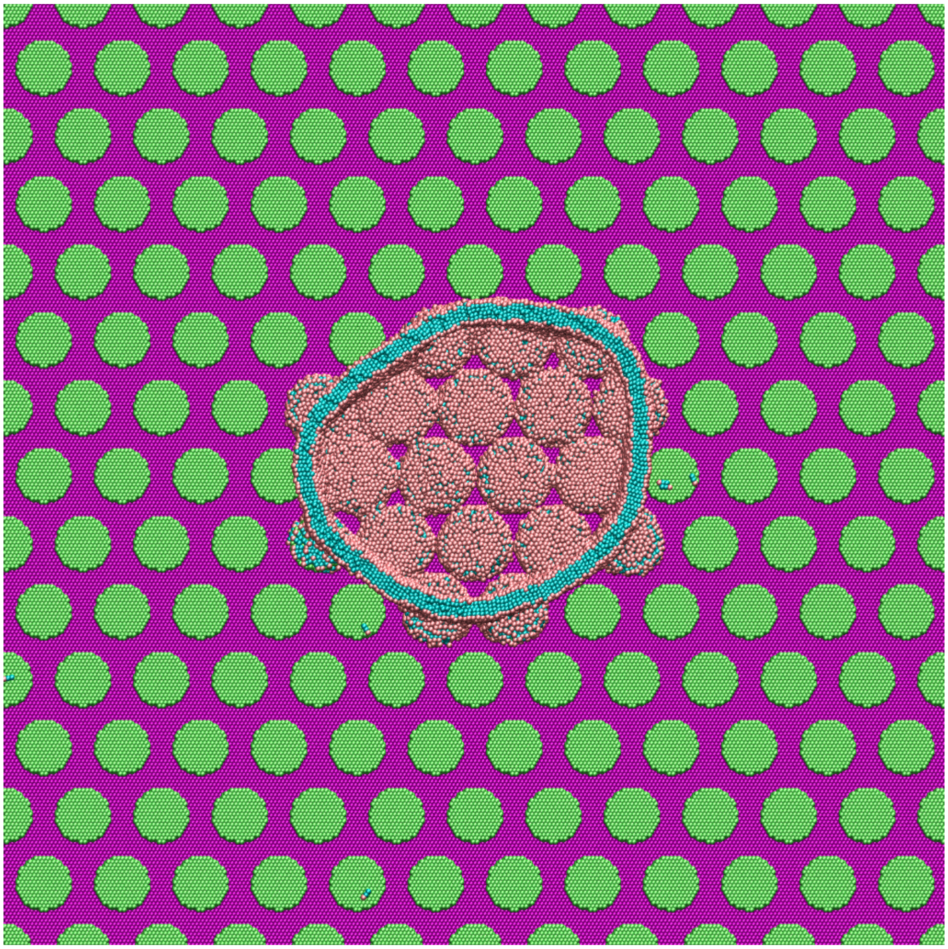

Figure 8-4. Simulated nanostructure of a cicada wing–like surface. The simulation was used to gain insight into the cicada wing's antibacterial properties. Top view cross-section: simulated lipid bilayer vesicles interact with nanopillars, showcasing the lipid arrangement and membrane rupture in high-curvature regions. (Credit: Jan-Michael Carrillo/ORNL.)

Publication: D. Salatto et al., "Structure-Based Design of Dual Bactericidal and Bacteria-Releasing Nanosurfaces," *ACS Applied Materials & Interfaces*, January 2023, https://doi.org/10.1021/acsami.2c18121.

Related Link: "Advancing Nanoscience through Largescale MD Simulations," OLCF News (July 14, 2023)

8.1.2.5 Reaching a New Summit for Supernova Simulations

University of Tennessee (UT), ORNL researchers use Summit to predict gravitational-wave signatures needed to detect distant supernovae

PI: Anthony Mezzacappa, University of Tennessee

Allocation Program: INCITE

The Science

As a result of large-scale 3D supernova simulations conducted on the OLCF's Summit supercomputer by researchers from UT and ORNL, astrophysicists now have the most complete picture yet of what gravitational waves from exploding stars look like (Figure 8-5). The team ran three supernova simulations with a supernova progenitor mass of 9.6, 15, and 25 solar masses, respectively. Their predictions of gravitational wave signatures emanating from these supernovae could serve as vital tools to help identify such infrequent, distant phenomena in the galaxy.

The Impact

The UT/ORNL team shares its gravitational wave predictions with observatories such the Laser Interferometer Gravitational-Wave Observatory. Gravitational-wave observatories use huge interferometers to detect minute flexing in spacetime due to gravitational waves caused by catastrophic events such as collisions between black holes or neutron stars. Astronomers compare these measurements with computer simulation data to determine their sources. So far, however, the observatory has yet to verify a core-collapse supernova. But that could change soon thanks to more sophisticated modeling data from this new study and from other supernova modeling efforts.

Summary

As a result of largescale 3D supernova simulations conducted on the OLCF's Summit supercomputer by researchers from UT and ORNL, astrophysicists now have the most complete picture yet of what gravitational waves from exploding stars look like. This is critical information in the quest to identify such phenomena in the universe.

Funding

This work was supported by the National Science Foundation Gravitational Physics Theory Program and by the DOE's Offices of Nuclear Physics and ASCR.

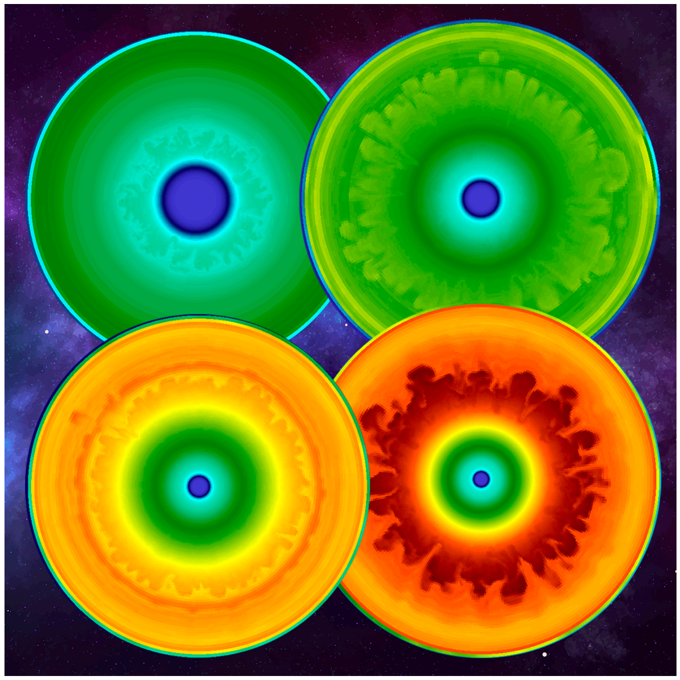

Figure 8-5. Simulated modes of convection deep within the stellar core and below the core collapse supernova shock wave. Such convection aids the explosion and generates gravitational waves at both low and high frequencies. (Credit: Anthony Mezzacappa, University of Tennessee.)

Publication: Mezzacappa, et. al., "Core collapse supernova gravitational wave emission for progenitors of 9.6, 15, and 25M," *Physical Review D* (2023), https://doi.org/10.1103/PhysRevD.107.043008.

Related Link: "Reaching a New Summit for Supernova Simulations," OLCF News (June 26, 2023)

8.1.2.6 Oxygen Unbound

Research teams find that rare isotope oxygen-28's "doubly magic" stability does not overcome its neutron-rich instability.

PI: Gaute Hagan, ORNL

Allocation Program: INCITE

The Science

Rare isotope oxygen-28 has been determined to be barely unbound by experiments led by researchers at the Tokyo Institute of Technology and by computer simulations conducted at ORNL. The findings from this first-ever observation of oxygen-28 answer a longstanding question in nuclear physics: can you get bound isotopes in a very neutron-rich region of the nuclear chart, where instability and radioactivity are the norm? Researchers found that despite its "doubly magic" configuration of 20 neutrons and 8 protons, which indicates inherent stability, oxygen-28 cannot overcome its neutron-rich instability and is not particle-bound.

The Impact

Isotopes—atoms of a particular element that have different numbers of neutrons—can be used for a variety of tasks, from tracking climate change to conducting medical research. Investigating rare isotopes such as oxygen-28 provides scientists with opportunities to test their theories of nuclear structure and to learn more about isotopes that have yet to be utilized in applications. According to the journal *Nature*, "If the results [of this project] can be replicated, physicists might need to update theories of how atomic nuclei are structured."

Summary

An ORNL team simulated oxygen-28 by using an emulator algorithm (Figure 8-6). These algorithms are trained to address selected sets of parameters and emulate first-principles calculations to solve the problem at lower fidelity, which reduces computational cost. Then, using a statistical approach called "history matching," the team crosschecked the models against other experiment data to iteratively produce an accurate model.

Funding

DOE's Office of Science supported the research.

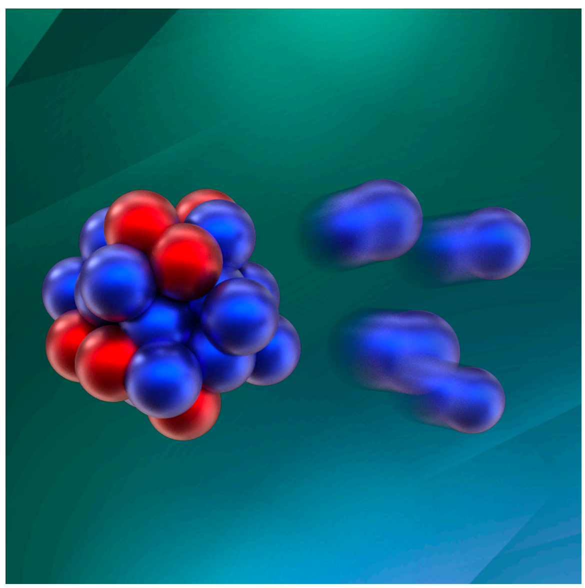

Figure 8-6. Oxygen-28 has been determined to be barely unbound through computer simulations conducted at ORNL. (Credit: Andy Sproles/ORNL.)

Publication: Y. Kondo et al. "First Observation of ^{28}O," *Nature*, 2023. https://doi.org/10.1038/s41586-023-06352-6.

Related Link: "Oxygen Unbound," OLCF News (August 30, 2023)

8.2 RESEARCH ACTIVITIES/VENDOR ENGAGEMENT FOR FUTURE OPERATIONS

8.2.1 Advanced Computing Ecosystem

Computing systems and the associated architectures required to support emerging applications and interconnected/integrated workflows are becoming more complex. To better understand these new technologies, evaluate their pros and cons, assess their operational impacts and capabilities, and develop tools to mitigate missing capabilities, the OLCF has aggregated related efforts under the Advanced Computing Ecosystem (ACE) umbrella. ACE is designed to help the OLCF with the upcoming OLCF-6 project and also with DOE's Integrated Research Infrastructure (IRI) Program. ACE consists of four fronts: testbeds, technology evaluations, research and development of foundational technologies, and IRI science pilots and workflows.

8.2.1.1 Testbeds

The ACE testbed is a unique OLCF capability that provides a centralized sandboxed area for deploying heterogeneous computing and data resources and facilitating the evaluation of diverse workloads across a broad spectrum of system architectures. ACE is designed to fuel productization of new HPC technologies as applicable to the OLCF and DOE missions. ACE is an open access environment consisting of HPC production-capable resources, and it allows researchers and HPC system architects to assess existing and emerging technologies more freely without the limitations of a production environment. Topics of interest include the following:

- IRI workflows and patterns (i.e., time-sensitive, data-intensive)
- Emerging compute architectures and techniques for HPC (e.g., Atmospheric Radiation Measurement (ARM), AI appliances, reconfigurable)
- Emerging storage architectures and techniques for HPC (e.g., object storage)
- Emerging network architectures and techniques (e.g., DPUs)
- Cloudification of traditional HPC architectures (e.g., multi-tenancy, preemptible queues)

The ACE testbed currently has the following computer resources:

- Defiant
  - 36 nodes: AMD Epyc CPU, 4 AMD MI100 GPUs, Slingshot 10 networking
  - Former Frontier early access system
- GraphCore
  - Dedicated AI appliance – BOWPod16 configuration
- Holly
  - Single Supermicro server with 8 NVIDIA H100 GPUs

The ACE testbed currently has the following storage resources deployed with close engagement with vendors:

(1) Polis – Lustre
  - ~1.6 PB
- Primarily spinning disk with some flash, connected to Defiant

(2) VastData – NFS
  - ~600 TB
  - NFS-over-RDMA storage appliance
  - Flash, connected to the IB fabric

(3) DAOS – Object Storage
   - 8 servers with ~30 TB flash each and IB connections (currently being upgraded from HDR to NDR)

The ACE testbed also contains OLCF's Wombat cluster, which was established in 2018 to evaluate the maturity and compatibility of the ARM AArch64 architecture for scientific HPC applications. Over time, the testbed has been updated as needed. During 2022, a large-scale study of typical HPC software on AArch64 and AArch64+GPU hardware was conducted. Those results were presented in March 2023 at the International Workshop on Arm-based HPC: Practice and Experience (IWAHPCE-2023), which is part of the HPC Asia 2023 conference.

Following the publication of that study, research emphasis on Wombat shifted away from general software ecosystem evaluation and toward more specific fields. In particular, more researchers started working with the NVIDIA BlueField DPUs available on some nodes. In February, the OLCF hosted a one-day workshop attended by about two dozen ORNL researchers and led by a few NVIDIA employees. The purpose of the workshop was to allow the attendees to try out an early, alpha-quality version of OpenMPI that used the DPUs to offload collective operations. The attendees were able to run this code on the actual DPU hardware in the Wombat nodes. During the summer months, researchers from the Modeling Outcomes using Surveillance data and Scalable Artificial Intelligence for Cancer (MOSSAIC) project at the NCI looked at using the encryption capabilities of the DPUs to securely distribute the AI model data. The model data contained private health information from cancer patients and required additional security and encryption. Most recently, another group of researchers (mostly from the Georgia Institute of Technology and led by Oscar Hernandez of ORNL) has started looking at using the DPUs for OpenMP offloading. More details can be found at https://www.olcf.ornl.gov/olcf-resources/compute-systems/testbeds/.

8.2.1.2 OLCF 6 Cloud Requirements Evaluation

The capability of cloud computing services to meet the computing needs of HPC applications has been a growing question within the DOE. To better understand the capabilities offered by the cloud, the OLCF has been evaluating the potential of cloud environments to support OLCF workloads. This has been a multiyear effort and has consisted of an initial phase of basic capability evaluations and an extended pilot project to identify the potential benefits and shortcomings of selecting a cloud-based option as the system architecture for OLCF-6. The first evaluation phase consisted of an initial study that looked at the capability of standard commodity cloud environments to support the leadership-class HPC workloads that comprise the OLCF mission set. The study’s findings concluded that existing cloud infrastructures were not capable of meeting the OLCF mission requirements. However, fundamentally, nothing would prevent a cloud provider from designing and building a custom solution that would be competitive with on-premises HPC solutions.

Based on those initial findings, the OLCF began exploring the potential of a cloud-based offering as an option for the upcoming OLCF-6 procurement. Although the OLCF believes that a cloud-based solution is technically feasible, it also presents several novel risks and unknown challenges that must be considered during the OLCF-6 selection process. To identify as many of the risks and challenges as possible, a second evaluation phase was initiated; this phase consisted of deploying a pilot cloud testbed environment as an experimental dry run for a possible future OLCF-6 system. The intention of this pilot is to deploy a production-ready system on Microsoft Azure in an off-site data center. The deployment of the testbed is meant to replicate (as near as is feasible) the processes that would be followed for an OLCF-6 scale system. These processes include the negotiation of a procurement contract and service agreement directly between Microsoft and the OLCF, the integration of the testbed with existing NCCS environmental infrastructure, development of new and reconciliation of existing security policies, and the adaptation of existing system operations and management procedures to map to a cloud-based environment. The ultimate purpose of the pilot project is to identify unknown challenges and barriers and to mitigate as many risks as possible for a possible OLCF-6 deployment.

8.2.1.3 FORGE: Large Language Model and OLCF-6 AI Benchmark

Large language models (LLMs) are poised to revolutionize the way scientists conduct scientific research in many profound ways. However, both model complexity and pretraining cost are impeding effective adoption for the wider science community. Identifying suitable scientific use cases, finding the optimal balance between model and data sizes, and scaling up the model training are among the most pressing issues that must be addressed. A team of researchers at the OLCF conducted a study to provide practical solutions for building and using LLM-based foundation models to target scientific research and use cases. They present an end-to-end examination of LLMs' effectiveness in scientific research, including their scaling behavior and computational requirements on Frontier.

The team also developed a suite of open foundation models called FORGE. These models have about 26 billion parameters using 257 billion tokens from over 200 million scientific articles and achieve performance either on par or superior to comparable state-of-the-art models. The team has demonstrated the use and effectiveness of FORGE on scientific downstream tasks. Overall, the team's approach and findings have significant implications for developing language models that can better understand and interpret scientific language and thus support scientific research and discovery. While these approaches primarily involve using model embeddings as a knowledge representation (e.g., phase classification and energy regression) or model-generated sequences (e.g., chemical or genomic sequences) to navigate a large search space, the team is experimenting with scientific applications that will include combining LLMs with data of different modalities such as instrumental parameters and images. These diverse applications would highlight the versatility of LLMs and their potential to enhance scientific research and analysis.

Based on this study, the OLCF developed a benchmark for the OLCF-6 RFP. It aims to stress the system in terms of low-precision computation, communication bandwidth, and energy usage. Correspondingly, the evaluation also considers computational performance (i.e., teraflops), scaling efficiency, and energy efficiency (i.e., teraflops/Watt). This is the first large-scale AI benchmark used for OLCF system acquisition.

J. Yin, S. Dash, F. Wang, and M. Shankar, FORGE: Pre-Training Open Foundation Models for Science, proceedings of SC23, 2023.

8.2.1.4 Digital Twin for HPC

Over the past year, the OLCF embarked on an ambitious initiative to develop a comprehensive digital twin of the Frontier supercomputer. This twin includes 3D asset modeling with virtual and augmented reality capabilities, telemetry data assimilation, AI/machine learning integration, simulations, and reinforcement learning for optimization. Key simulations under development include (1) a transient simulation of the thermo-fluid cooling system from cooling tower to cold plate, (2) a rectifier loss model predicting heat generation and rectification losses, (3) a job scheduling simulator, and (4) a parallel discrete-event simulator to study network congestion. This digital twin offers insights into operational strategies and what-if scenarios and elucidates complex, cross-disciplinary transient behaviors. The digital twin also serves as a design tool for future system prototyping. Built on an open software stack (Modelica, SST Macro, Unreal Engine) with an aim to foster community-driven development, the OLCF formed a partnership with supercomputer centers around the world as well as HPC suppliers to develop an open framework for modeling supercomputers. Figure 87**Error! Reference source not found.** illustrates the OLCF Frontier digital twin project.

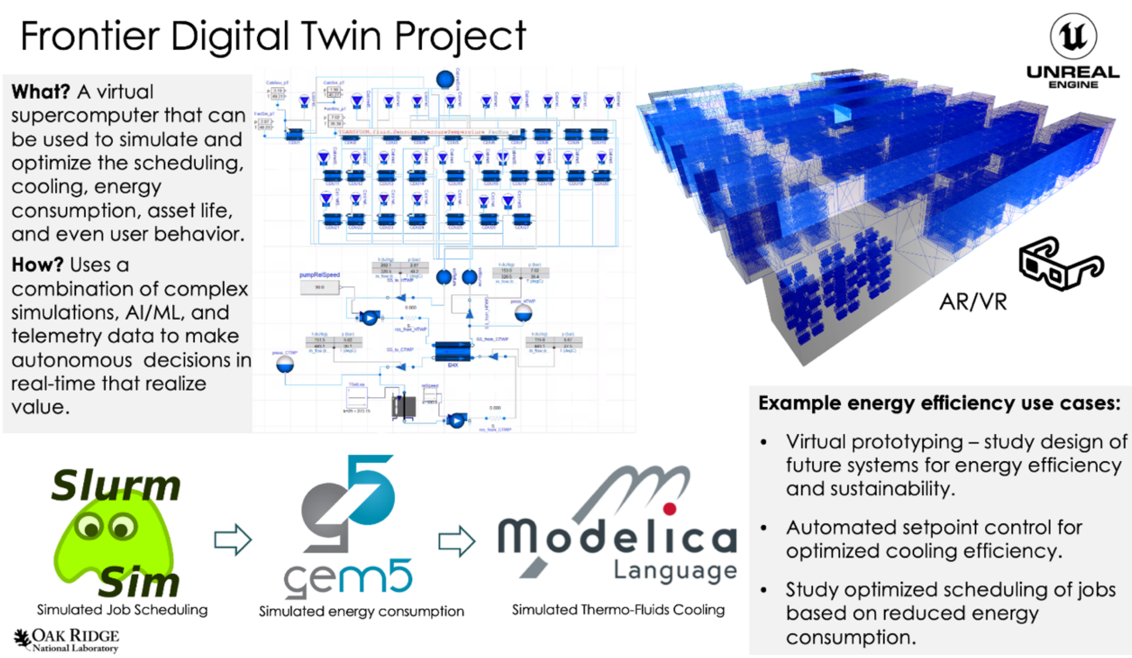

Figure 8-7. The OLCF Frontier digital twin project.

To date, the team has finished initial developments of (1) the scheduling simulator, (2) the power consumption model, (3) the cooling model, and (4) the virtual reality representation of the supercomputer and central energy plant. The team has also been actively working on developing a model of the Slingshot network based on SST Macro parallel discrete event simulator. This has involved profiling several applications on Crusher and Frontier to understand communication profiles, which will then need to be integrated with the simulation engine. The OLCF can use the current version of the digital twin to either simulate workloads on the digital twin or replay telemetry data for visualization or validation of the models. With the current state of the model, the OLCF can investigate strategies for more efficient power consumption, such as implementing a load-sharing strategy for dynamically turning on/off rectifiers as needed.

Harnessing Visual Analytics and Machine Learning on Large-Scale Telemetry for Enhanced Operational Insights at OLCF

The OLCF is expanding its power and energy analytics, enhancing access to detailed component-level power and energy time series data, and adding contextualized job allocation details. This initiative builds on years of effort to enable large-scale, real-time operational data analytics. By improving access to this comprehensive data, the OLCF is paving the way for a future of data-driven energy efficiency strategies.

Interactive Visual Analytics: A new visual analytics web application, called Live Visual Analytics (LVA), was created by OLCF staff (**Error! Reference source not found.**). LVA offers an interactive platform for exploring high-dimensional power, energy, thermal, and job allocation data with ease. Designed for interactive detailed analysis, it allows users to navigate vast data volumes quickly (response times under three seconds) thanks to extensive data reduction and preprocessing enabled by a real-time stream data processing infrastructure. **Error! Reference source not found.** illustrates the OLCF LVA dashboard for the Summit supercomputer.

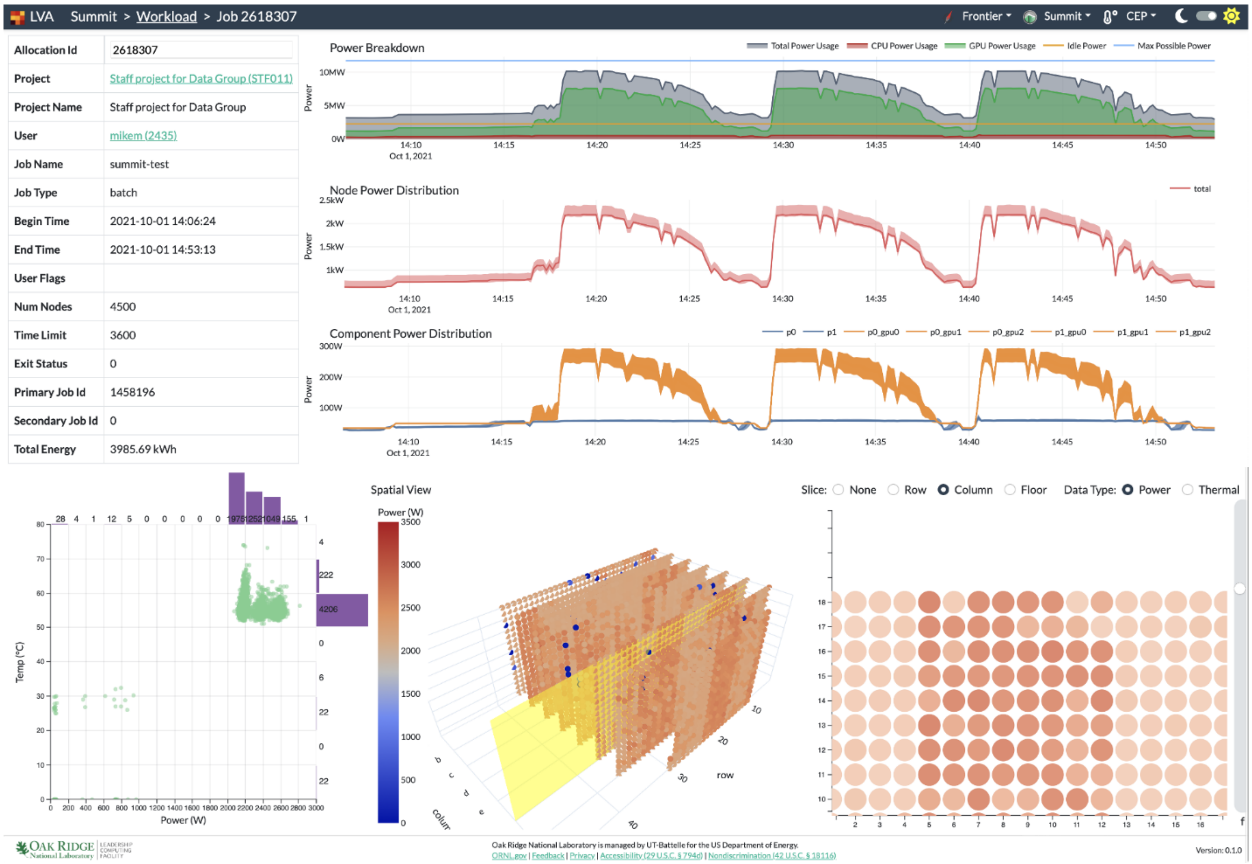

Figure 8-8. Live Visual Analytics portal. Enables users to interactively examine GPU-level power consumption allocated to a large computation.*Machine Learning for Power Profile Classification:* The OLCF team has developed an innovative machine learning pipeline for real-time clustering of job-level power consumption time series. This method overcomes the computational challenges and limitations of existing clustering techniques by using generative adversarial network clustering and a pretrained neural network open-set classifier. With this approach, which considers over 180 features per job time series, the OLCF aims to deploy a real-time power profile prediction tool that will enable tailored automated responses based on application power consumption types.

8.2.1.5 Federated Learning

Federated learning (FL) is a technique for sharing machine learning models without sharing raw data. FL-trained models generally converge faster and fit more closely to the data than independently trained models on smaller individual datasets. Depending on the FL algorithm, privacy-preserving properties can be proven, which may enable much better models to be trained in healthcare, genomics, and other privacy-sensitive fields.

In 2023, OLCF staff members continued to work with collaborators at NVIDIA to make advancements toward a production HPC-FL deployment on OLCF systems. This work centered around making NVIDIA's Federated Learning Application Runtime Environment (NVFLARE) more compatible with HPC environments, specifically integrating FL clients and job launching mechanisms for HPC resource managers (e.g., Slurm). At first, the NVFLARE client software was not capable of being launched via a job launcher (i.e., Slurm's *srun*) and therefore could not run across multiple nodes on Frontier or Summit. Being able to launch an NVFLARE client in a multi-tenant and multi-node environment required code changes to NVFLARE, and this is now remedied owing to the OLCF's collaboration with NVIDIA in

2023. **Error! Reference source not found.** depicts the NVFLARE integration with the Framework for Exploring Scalable Computational Oncology (FrESCO) code base on OLCF's Frontier.

Additional advancements were made during various hackathons with a pilot research project at ORNL. This pilot project focuses on developing deep learning capabilities within FrESCO. This project is currently partnered with ORNL's Knowledge Discovery Infrastructure and the OLCF to utilize the secure computing platform called CITADEL. The effort is helping both groups explore how FL and HPC can work together for scientific discovery. The FL application will keep the data siloed but enable the development of a common model, thereby creating a new avenue for collaborations and advancements. By working with this project, members of the OLCF have been able to integrate their custom library into NVFLARE, allowing the OLCF to launch their code through an NVFLARE client across Frontier and Summit (Figure 89**Error! Reference source not found.**). Before this, the OLCF had primarily focused on the provided examples from NVLFARE and not custom code from ORNL researchers. This pivot led to a deeper understanding of how to customize workflows within NVFLARE and will feed future FL use cases in the OLCF.

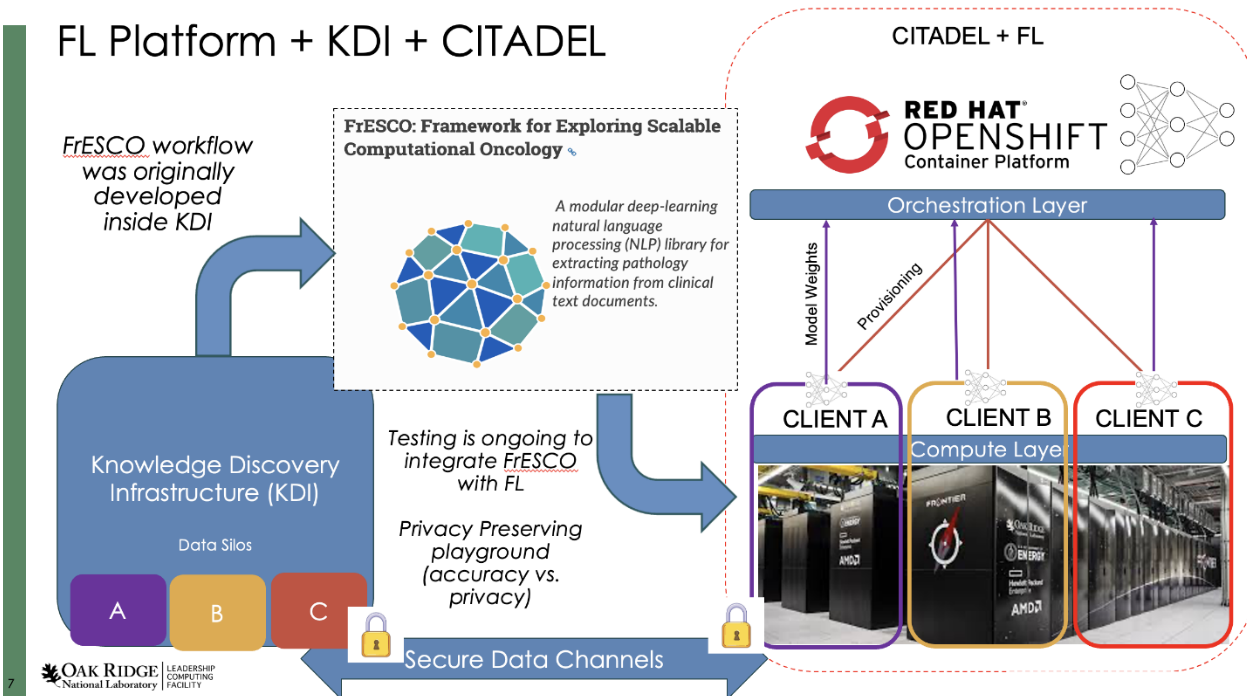

Figure 8-9. Depicts the integration of the FrESCO code base with NVFLARE on Frontier.

8.2.2 Industry Engagement

In addition to vendor engagement in the testbeds and tools R&D for future systems, OLCF continued to strengthen ties with industry collaborators. ACCEL, OLCF's industrial partnership program, had a strong year in 2023 with 39 projects underway during the year—22 new projects launched and 17 that began prior to 2022 continued throughout the year. New projects used both Frontier and Summit.

These 39 industrial projects represented 11% of the total projects provided to external user programs, INCITE, ALCC, and DD (including the ECP).

Industrial projects ran on Summit, Frontier, and/or Crusher, with some projects running on more than one system: 23 had Summit allocations, 9 had Frontier allocations, 6 had allocations on more than one system, and 1 had an allocation solely on Crusher.

These projects used 4,598,739 Summit node hours, which represents approximately 15% of the total Summit hours, and 16,798,981 Frontier node hours, which represents approximately 36% of the total Frontier hours.

8.2.2.1 Industrial Project Allocations Across External User Programs

In 2023, of the total allocated industrial project hours on Frontier and Summit combined,

- 67% were allocated through INCITE
- 23% via ALCC and DD
- 10% through the OLCF DD program

On Summit,

- 81% of the industrial project hours were allocated through INCITE
- 10% via ALCC
- 9% through the OLCF DD program

On Frontier,

- 54% of the industrial project hours were allocated through INCITE
- 36% via ALCC
- 10% through the OLCF DD program

8.2.2.2 Observations about the Industrial Projects

The 22 new projects in 2023 received their awards via INCITE (5 projects), ALCC (3 projects), and DD (14 projects).

INCITE

Industry had five new INCITE awards, four of which had access to both Frontier and Summit. Industry partnered with universities or national laboratories in four of the projects, thereby ensuring broader community engagement.

ALCC

Industry fared well in the ALCC program with three new awards of time on Frontier. GE Aviation continues as the lead industrial user of Frontier with its ALCC award. In addition, Westinghouse and Cadence collaborated with university PIs. Commercial software vendor Cadence is the only commercial software vendor to date that has ported and scaled its software to run on Frontier.

OLCF DD

Summit Awards

Four new industrial users received DD awards for Summit: Nissan Technical Center North America, Dridam Flight, Vayuh Inc., and Veracity Nuclear. The latter three are all small businesses.

Dridam Flight is a small business engaged in the development of supersonic and hypersonic vehicles.

Vayuh, Inc is a small business that uses physics and AI to prepare sub-seasonal predictions and climate risk modelling to better enable climate-based decision-making (e.g., selection, pricing, portfolio management) in energy markets, insurance, and supply chains. They focus particularly on extreme or anomalous events.

Veracity Nuclear is a small business that specializes in developing and applying integrated nuclear core and plant simulation capabilities to manage complex reactor constraints, achieve peak operational efficiency, and increase the overall economic benefit of operation.

These small businesses reveal again that large-scale problems that require leadership computing are not the purview of only large companies. Small businesses also have complex problems that benefit from access to OLCF resources.

Frontier Awards

Two new industrial users received DD awards on Frontier: global industrial turbine manufacturer Baker Hughes as well as Vayuh Inc. which received a Frontier award after completing its Summit project.

The nuclear industry continues to turn to the OLCF for access to the leadership-class systems. Veracity Nuclear, a new industrial user, was granted a DD allocation on Summit Both General Fusion and Westinghouse, in collaboration with Penn State University, were also awarded DD allocations on Frontier. Additionally, Zap Energy and Vayuh, Inc., another new industrial user each received DD allocations on both Summit and Frontier. Furthermore, General Atomics received an INCITE award on both Summit and Frontier.

8.2.2.3 Industrial Project Highlight

GE Aerospace Makes a Bold Move Toward Sustainable Energy in Flight

GE Aerospace and long-time partner Safran Aircraft Engines are pursuing a revolutionary new open fan engine architecture for GE’s jet engines through their Revolutionary Innovation for Sustainable Engines (RISE) technology demonstration program. Their goal is to deliver over 20% lower fuel consumption and CO_2 emissions compared to today’s most efficient aircraft engines.

To evaluate the potential efficiency of this new architecture while simultaneously reducing engine noise levels, engine designers need to understand airflow behavior around the blades of the engine, including the complex physics of turbulence, mixing, and potential flow for separation. GE engineers used Frontier to perform first-ever 3D large eddy simulations at realistic flight-scale conditions with unprecedented detail, revealing breakthrough insights into improving aerodynamic and acoustic performance in this next-generation aircraft engine technology (Figure 8-10).

The high-fidelity simulations performed on Frontier provide unique insights to better understand important factors in novel propulsion design, such as how complex turbulent flow physics measured in smaller scale prototype rig tests scale up to flight performance of the full-scale engine. The prior state of the art was limited to a reduced scale by the available computational resources. Those simplifications introduce uncertainty when translating engineering knowledge to the real thing. The processing power and scale of Frontier is enabling GE to move beyond those previous limitations.

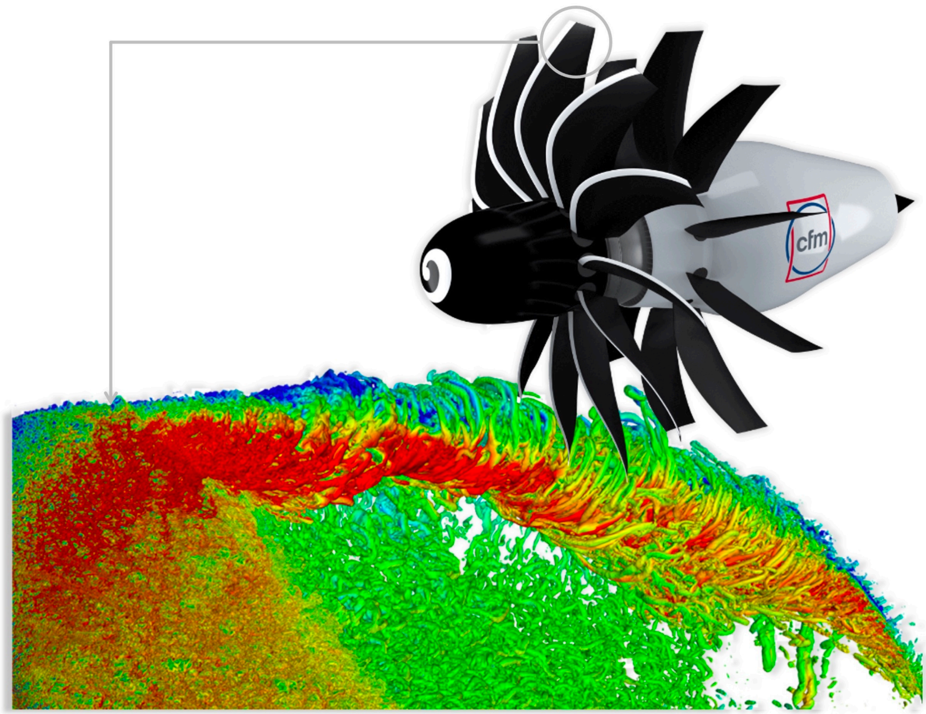

Figure 8-10. Top: A rendering of the CFM RISE program's open fan architecture. Bottom: A visualization of turbulent flow in the tip region of an open fan blade using the Frontier supercomputer at ORNL. (Credit: CFM, GE Research [CFM is a 50–50 joint company between GE and Safran Aircraft Engines].)

Publication: In development

Related Link: "GE Aerospace runs one of the world's largest supercomputer simulations to test revolutionary new open fan engine architecture," OLCF News (June 17, 2023)

The Impact

Access to Frontier is lending crucial support to GE's pursuit of reducing and eventually eliminating fossil fuel use for aircraft propulsion and achieving sustainable energy in flight.

Fluid dynamics simulations on Frontier are enabling GE to extend insights from prior research to realistic flight conditions and discover new ways of controlling turbulence, improving fan performance, and more efficiently guiding future physical testing. The study also offers a rich high-fidelity dataset from which machine learning algorithms can train reduced-order models useful in exploring design trade-offs.

Dr. Stephan Priebe, senior engineer in Computational Fluid Dynamics & Methods at GE Research summed it up: “Frontier has unlocked the ability for us to numerically fly our fans during the design phase, years in advance of actual flight testing, with a level of detail previously unattainable.”

Through diverse partnerships, the OLCF cultivates strategic engagements with other DOE Office of Science programs, DOE Applied Programs, the ECP, and other federal agencies to increase the OLCF user community, better understand current and future DOE mission needs related to the Facility, and leverage opportunities for additional hardware, software, application, and operational innovations.

8.3 DOE PROGRAM ENGAGEMENTS/REQUIREMENTS GATHERING

Through diverse partnerships, the OLCF cultivates strategic engagements with other DOE Office of Science programs, DOE Applied Programs, the ECP, and other federal agencies to broaden the OLCF user community, better understand current and future DOE mission needs related to the Facility, and leverage opportunities for additional hardware, software, application, and operational innovations. Below are examples of some of the collaborations and outcomes that benefitted the OLCF and the broader DOE community.

8.3.1 Ongoing Engagements with Office of Science Observational and Experimental Facilities Investigators

Although PIs funded by DOE Office of Science program offices commonly use the OLCF, in recent years many PIs have explored discretionary allocations to establish workflows from their observational facilities to the OLCF resources. Using institutional resources, Office of Science Biological and Environmental Research–sponsored investigators in the ARM program use operational data workflows that originate from their sensors in their data management facility and have set up exploratory efforts with OLCF data and analytics (JupyterHub) portals. BES-sponsored scientists who use facilities such as ORNL’s CNMS and SNS have successfully won discretionary time on Summit to perform AI-driven electron microscopy and design-of-experiment campaigns (porting their codes to GPUs), respectively. The movement of data from the facilities is supported by traditional data transfers and new tools to simplify the user experience (such as the DataFlow tool discussed in Section 4.1.4). Additionally, scientists from the LCLS femtosecond x-ray laser facility at SLAC have obtained three SummitPLUS allocations that will facilitate new workflows for rapid processing of data and training of AI models on massive datasets. This work will help analyze and compress emergent raw data and steer ongoing experiments. The partnership will enhance the coupling of a major BES experimental facility with OLCF exascale resources. The lessons learned from workflows for High Energy Physics programs (e.g., ATLAS-supporting BigPANDA workflows), NP programs (e.g., ALICE), and FES’s deep learning for real-time prevention of disruption have informed workflows deployed on Slate for Summit and the design of the Frontier login nodes.

8.3.1.1 Ongoing Engagements with Atmospheric Radiation Measurement

In 2023, NCCS partnered with the ARM program to enhance their data transformation workflows. NCCS deployed over 5 PB of disk storage and 28 PB of tape storage into the NCCS Open nearline system (Themis) to provide ARM with a highly durable storage resource. The OLCF offers data management support and expertise in these resources, thereby ensuring safe, long-term storage that also performs well enough to meet researcher demands.

The hosting of ARM’s HPC resources and data at the OLCF offers several benefits to both the OLCF and ARM. The OLCF leveraged the Multi-Category Security feature of Slurm to provide a limited node sharing capability to ARM. This improved cluster throughput and reduced the time from job submission to end data result, better enabling these workflows in an HPC space. Supporting a project such as ARM showcases the unique service capabilities and expertise that the OLCF provides.

8.3.1.2 Science Pilots and Workflows Driving Future IRI Requirements

SLAC/OLCF IRI project

Connecting world-leading experimental facilities such as nuclear science labs and light sources to exascale computing facilities is a major focus of the IRI effort at the DOE. Experimental facilities such as the new LCLS-II beamline at SLAC can produce data at a rate 10,000× that of the previous LCLS source. This holds great potential for accelerating discovery in biology, chemistry, and materials but presents major problems related to the data deluge that requires data transfer and coupled HPC. In the spring of 2023, Thomas Beck, Dilip Asthagiri, Patrick Widener, and Rafael Ferriera da Silva from the OLCF collaborated with a team from SLAC on a Brave bio-preparedness proposal. This led to further discussions of approaches for tying SLAC and ORNL efforts to IRI goals. This work will enable accelerated discovery in complex problems in biology related to photosynthesis and RNA structure. Thomas Beck attended the SLAC Users Group Meeting in September 2023 and had extensive discussions with a range of domain, data, and computer scientists at SLAC. He informed them of the emerging SummitPLUS program, which led to the submission of three successful proposals for allocations (led by Jana Thayer, Fred Poitevin, and Ryan Coffee). Two of these are related to real-time data analysis computational tools (including data reduction) that can help guide ongoing experiments as the data emerges. A third project will develop a large-scale foundational model for a massive image dataset produced at SLAC over the last several years (5 PB in size). The training weights obtained in the AI/machine learning modeling will then be shipped back to SLAC for use in data analysis during ongoing experiments. Teams from SLAC and ORNL are meeting on a biweekly cadence to plot a course for the many data workflow and computational aspects of the collaborative projects.

Gamma-Ray Energy Tracking Array IRI project

Gamma-Ray Energy Tracking Array (GRETA) is a state-of-the-art gamma-ray detector under construction at Lawrence Berkeley National Laboratory (LBNL) for the DOE, Office of Science (Nuclear Physics), in collaboration with ORNL, Argonne, and the Facility for Rare Isotope Beams (FRIB) at Michigan State University. GRETA will be an essential instrument for FRIB and will support a broad nuclear physics program after delivery in 2025. As a first of its kind, GRETA will cover an entire sphere and have a fully networked streaming readout at a rate of 4 GB/s going into an online data pipeline responsible for buffering, real-time analysis, data reduction, event building, and storage. This workflow is an example of a time-sensitive IRI pattern because of the time constraint on the real-time analysis.

Gustav R. Jansen from the Science Engagement Advanced Computing for Nuclear Particle and Astrophysics group in NCCS is part of the GRETA data pipeline team and is responsible for developing the real-time data analysis component. He is also integral to the team that developed Deleria, a generalization of the GRETA data pipeline that enables cross-facility workflows. As a proof of concept, Deleria was deployed successfully in the IRI testbed in the fall of 2023, using the ESnet testbed and a network loop corresponding to an 80 ms latency between the simulated experiment and the real-time analysis component of the data pipeline. Currently, Gustav is leading the work to deploy Deleria on the IRI testbed resources at the OLCF. The goal is to demonstrate that leveraging ASCR resources can increase the performance and capabilities of GRETA and show that nuclear physics experiments can perform real-time analysis at ASCR facilities by providing reusable concepts, patterns, and software.

Earth Systems Grid Federation IRI project

The Earth Systems Grid Federation (ESGF) is a globally distributed data infrastructure that has been in operation for over 10 years. The core strengths of ESGF are federated search and access plus robust data management. Currently, the ESGF project is re-architecting the underlying system to improve user experience, advance integration with computing facilities, ease platform administration, and expand data management capabilities. In terms of advancing integration with computing facilities, this is where the ASCR IRI can provide support—specifically in enabling automated and trusted integration with ASCSR HPC centers for offloading computational workflows. Currently, members of the Software Services Development and Workflow and Ecosystem Services (WES) groups are working to enable users to execute workflows that leverage ESGF data with unprecedented ease, flexibility, and scale. Enabling the seamless orchestration of cross-facility workflows across ORNL, Argonne, and JLab would allow users to quickly and even automatically redirect their computations to optimal environments. This IRI will promote ease of processing for users by handling many of the authentication, workflow management, and data transfer barriers currently present in executing cross-facility jobs. Key OLCF projects that will support this effort include the OLCF Facility API, Frontier, and INTERSECT.

Joint Genome Institute IRI project

The Joint Genome Institute (JGI) is a multi-project user facility located at LBNL. These projects are focused on large-scale genomic sciences that address questions of relevance to DOE's Office of Biological and Environmental Research missions in sustainable biofuel and bioproducts production, global carbon and nutrient cycling, and biogeochemistry, among others. Several projects involve producing and processing microbial, fungal, algal, and plant genome data. As a direct consequence of these activities, challenges arise from the massive amounts of data transmission, processing, and storage required. A cross-facility approach to computational and data processing may alleviate some of these challenges. Currently, members of the WES group are working to support the JGI produced workflows and service artifacts such as JGI Analysis Workflows Service on the IRI ACE testbed, which is a precursor to an infrastructure to facilitate the data transfer, storage, and workflow processing platform. The testbed can be seamlessly integrated as part of the existing JGI infrastructure, thus augmenting the compute and storage capabilities of the overall system.

8.3.2 Engagement with the National Institutes of Health and the National Cancer Institute

The collaboration between ORNL and NIH and NCI began as a strategic partnership in 2016 with the aim of advancing specific areas of cancer research and HPC development by applying advanced computing, predictive machine learning/deep learning models, and large-scale computational simulations.

Collaborative research is performed under the project known as MOSSAIC and consists of (1) developing large-scale, state-of-the-art Transformer language models for clinical information extraction; (2) building new capabilities for biomarker and recurrence detection; (3) pushing novel research in abstention and uncertainty quantification so that models can be effectively deployed in clinical practice; (4) conducting lab studies to evaluate the performance of the models in real-world cancer registry settings; and (5) enabling large-scale Transformer training on LCF systems. Collectively, these efforts aim to modernize the national cancer surveillance program and enable near real-time cancer surveillance in the US.

This research poses several significant computational challenges. Transformer models require pretraining on very large datasets with billions of elements, which itself requires access to large HPC resources. For example, these models require hundreds of Summit nodes for training using data parallelism.

Additionally, Transformer models must be trained on the raw text of cancer pathology reports, and this text includes PHI.

To better support these efforts, the OLCF launched a security framework called CITADEL, which allows researchers to use supercomputers for research that contains protected health data. ORNL's unmatched combination of a secure PHI enclave and a protected computing environment—through CITADEL—allows the secure use of Summit for MOSSAIC research and is crucial for the application of large-scale Transformer models on the cancer pathology report corpus. CITADEL is a good example of an innovation achieved in one program that can benefit multiple programs, such as the US Department of Veterans Affairs (VA), and other OLCF user communities.

NIH/NCI's collaboration with the OLCF has led to several notable outcomes, both for the broader HPC community and for the MOSSAIC research team. The computing capabilities of the OLCF enable efficient iterative development of deep learning models, driven by the challenges of extracting information from the NCI data. Deploying in this way will allow integration with the new CITADEL capability and support running the Transformer training pipeline with sensitive PHI data on the OLCF systems. The Transformer training modules will also be made available to the OLCF user community, allowing domain scientists in other fields to train their own domain-specific Transformer models on these systems.

Other major accomplishments of this collaboration have been the development of a new, scalable, and more efficient way to develop and deploy APIs to the Surveillance, Epidemiology, and End-Results (SEER) registries and other relevant stakeholders. The MOSSAIC team developed a new, modular PyTorch API designed to enable quick and easy swapping between (1) deep learning models; (2) different data preprocessing techniques, uncertainty quantification methods, and other pre- and post-processing methods; and (3) different cancer surveillance tasks, including but not limited to path-coding, identifying reportable path reports, identifying report type, detection of biomarkers, and identifying cases of recurrent metastatic disease.

In 2023, the MOSSAIC team continued to grow the Case-Level Multi-Task Hierarchical Self Attention Network, which is a natural language processing algorithm that autocodes cancer pathology reports at participating SEER registries and does so 18× faster than traditional methods. The algorithm saves 46,000 person hours per year and is a significant step toward near real–time population cancer surveillance. As the algorithm is adopted by more registries, the team is working on privacy-preserving methods that would enable data- and model-centric approaches to train the algorithm in a distributed environment. These advances demonstrate the power of AI for improving population health surveillance.

Building on this foundation, the team recently used the MIMIC III dataset to evaluate various attention mechanisms used to train AI models to focus on specific information in clinical documents. They then used these findings to improve the models that extract cancer data from electronic cancer pathology reports.

The MOSSAIC team also used Summit and CITADEL to develop a specialized transformer model called Path-BigBird, which they used to process 2.7 million cancer pathology reports from six SEER registries. Path-BigBird has the potential to streamline extraction of pathology data and could outperform traditional deep learning approaches for gathering cancer data such as sites, histology, and incidences.

8.3.3 Engagement with the US Department of Veterans Affairs

In 2016, the VA partnered with DOE and ORNL to revolutionize the health care of veterans and, by extension, all Americans via advanced data analytics and HPC. The DOE-VA collaboration advances the missions of both the VA and DOE and leverages each agency's unique resources to support efforts that would not otherwise be possible. The VA's Office of Research and Development uses DOE's supercomputing facilities, expertise in big data, and HPC to advance veterans' treatment and the VA's medical and genomic research. The Million Veteran Program is using Summit to identify genetic markers for early onset prostate cancer (before the age of 55). Supported by the security framework CITADEL, which protects sensitive personal health data so it can be studied directly on the supercomputer, the project is using computationally efficient transformer models on Summit to capture interactions within and between genes, a capability not viable without supercomputers.

These partnerships have resulted in unprecedented developments in security that enable the NIH and VA to use supercomputing for research that contains sensitive data and have strengthened the HPC ecosystem at ORNL.

8.3.4 Engagement with Air Force Weather and National Oceanic and Atmospheric Administration

ORNL and Air Force Weather (AFW) launched two HPE Cray EX supercomputers, named Miller and Fawbush, to support the AFW's numerical weather modeling at much higher speed and fidelity. Thanks to the powerful systems, which reached Certification of Operational Readiness in February 2021, AFW is developing new specialized models such as full physics cloud forecasting and a global hydrology model. ORNL is also providing a unique system of safeguards, including separate, dedicated power sources for each machine, failsafe features through the Slurm resource scheduler, and dynamic load balancing. The early receipt of these systems, in advance of OLCF-5, provided staff with critical decision points related to the full software stack, the tools for managing these systems, and invaluable experience delivering production-capable systems based on the new EX platform. These lessons significantly benefitted the OLCF-5 Frontier project by supporting key decisions early in that project in advance of the delivery of the initial test and development systems.

In 2023, an AFW-funded team tested a new hierarchical flood modeling framework that can simulate an entire flood event and translate precipitation to inundation. The study identified precipitation forecasts as one of the most important sources of error in flood models while demonstrating the accuracy of the framework for predicting future floods.

The OLCF also supported an AFW-funded study on the influence of natural climate variability on monsoons and other extreme weather events in Pakistan. The study found that natural climate variability, which includes factors such as sea surface temperature and jet stream anomalies, could account for over 70% of observed extremes, with climate change potentially adding to the severity.

ORNL and NOAA established the National Climate-Computing Research Center (NCRC) through a strategic partnership in 2009 with a goal to leverage leading HPC and information technologies to develop, test, and apply state-of-the science computer-based global climate simulation models based upon a strong scientific foundation. In 2021, based on success achieved over the program's lifespan, the program was extended another 5 years. The NCCS hosts the program's primary computational resource, named Gaea, which consists of an HPE-Cray EX 3000 HPC system providing over 10 petaflops of peak computation. In 2023, work began to procure and install a new climate modeling system with a peak performance of over 10 petaflops. Combined with the program's existing HPC resource, the new system will increase the program's peak performance to over 20 petaflops. The NCRC supports and complements NOAA's climate mission to understand climate variability and change to improve society's ability to plan and respond. The partnership enhances the capacity of NOAA's existing climate research centers by supporting the development of next-generation models, building computational capacity, cultivating a highly trained computational workforce, and engaging the global user community.

The partnership allows not only the opportunity for collaborative R&D but also the opportunity to develop and harden tools, methods, and best practices that improve operation and efficient use of HPC resources within the NCCS. For example, a project that began in 2021 and continued in 2023 has been developing the ability to run portable workloads within containers on NOAA's HPC resources. The NCCS project members are also undertaking similar efforts to develop containers that will run on Summit and Frontier. Lessons learned from each effort benefit both NOAA and the OLCF, and the broadened perspective helps to harden and improve the end result for all programs in the NCCS.

8.3.5 Exascale Computing Project Engagement Summary

A primary and natural place for community engagement has been DOE's ECP, the goal of which is to develop software and applications and influence the development of hardware technology—all to facilitate the successful deployment and operation of capable exascale systems. In 2023, the ECP continued to fund efforts in national labs, academia, and industry; these timely investments significantly aided the OLCF in delivering a capable exascale system, Frontier, with robust system and application software that can address the science gaps.

The OLCF made Frontier available to all ECP teams in April of 2023. All Application Development (AD) and Software Technology (ST) teams ran on Frontier, and all 70 ST products demonstrated critical functionality on Frontier. All ECP KPP-1, KPP-2, and KPP-3 performance metrics were met using Frontier.

The OLCF's engagement with the ECP included three primary thrust areas described below and many other interactions (e.g., staff involvement in ECP AD and ST projects).

8.3.5.1 OLCF-ECP Application Development/Software Technology Engagements

The OLCF has a long history of readying applications for its forthcoming architectures, dating to before the delivery of the OLCF-3 system (Titan), then with OLCF-4 (Summit), and now with OLCF-5 (Frontier). CAAR has served as a successful collaboration point for application teams, vendors, and tool developers to exploit hierarchical parallelism within applications in preparation for next-generation architectures. The OLCF partnered with the ECP to augment the OLCF CAAR portfolio for Frontier with an additional 12 ECP AD teams. These teams were selected to diversify OLCF application readiness efforts funded through the OLCF-5 CAAR to ensure that the OLCF had a broad suite of applications ready to use Frontier. These projects were also selected with an eye toward matching the architectural strengths of Frontier with the appropriate computational motifs and methods employed by these ECP applications. The teams that continue to partner with the OLCF and OLCF Scientific Computing liaisons are listed in Table 8-3.

Table 8-3. Listing of OLCF ECP engagement applications, the ECP AD PI, and the OLCF Scientific Engagement liaison.

WBS/ECP AD Project	Project PI	OLCF Liaison
2.2.1.01 LatticeQCD	Andreas Kronfeld (Fermilab)	Balint Joo
2.2.1.02 NWChemEx	Theresa Windus (Ames Laboratory)	Dmitry Liakh
2.2.1.03 GAMESS	Mark Gordon (Iowa State University)	Dmytro Bykov
2.2.1.05 ExaAM	John Turner (ORNL)	Stephen Nichols
2.2.2.02 Combustion-PELE	Jackie Chen (Sandia)	Kalyana Gottiparthi
2.2.2.03 ExaSMR	Steven Hamilton (ORNL)	Mark Berrill
2.2.2.05 WDMApp	Amitava Bhattacharjee (PPPL)	Antigoni Georgiadou
2.2.3.01 ExaStar	Dan Kasen (LBNL)	Austin Harris
2.2.3.02 ExaSky	Salman Habib (Argonne)	Antigoni Georgiadou
2.2.3.05 E3SM-MMF	Mark Taylor (Sandia)	Matt Norman
2.2.4.02 ExaSGD	Slaven Peles (PNNL)	Philip Roth
2.2.4.04 ExaBiome	Kathy Yelick (LBNL)	Philip Roth

PPPL = Princeton Plasma Physics Laboratory

With ECP funding, these ECP AD teams were provided with staff expertise from the OLCF Scientific Engagement section, access to the system vendor's COE from both Cray and AMD, access to early testbed hardware provided by the Frontier vendor, and support from postdoctoral researchers through the CSEEN program in 2023. In addition, because of the dependence of AD projects on components of the ECP ST portfolio, the ECP ST projects also had access to the COE resources and access to early testbed hardware.

This mutually beneficial partnership enabled ECP to benefit from the application readiness lessons learned and best practices developed during the two prior instantiations of the CAAR program. Additionally, the expected OLCF application and software portfolio ready for Frontier was further diversified by including these additional projects in the CAAR program.

The OLCF also contributed directly to the ECP via the Hardware Integration area as well by leading or participating in efforts in application support, software integration, facility resource utilization, and training and productivity.

APPENDIX A. 2023 OLCF Operational Assessment

There was one recommendation from last year's operational assessment.

> *The methodology for tracking research products that result from use of the facilities is not detailed or referenced in the report. Provide additional details about the methods used for tracking scientific accomplishments that resulted from staff support and facility use so that we can better assess its effectiveness and its potential adoption at other facilities.*

This information has been added to this year's report in Section 8.1.1.2.

APPENDIX B. 2023 OLCF Operational Assessment

Table B-1. Science highlights.

Date	Science Highlights Submitted to DOE Title	Writer	URL
1/30/23	Computational Study Finds Genetic Links, Therapy Targets for Varicose Veins	Stephanie Seay	https://www.olcf.ornl.gov/2023/01/30/computational-study-finds-genetic-links-therapy-targets-for-varicose-veins/
2/8/23	The Most Advanced Bay Area Earthquake Simulations Will be Publicly Available	Aliyah Kovner	https://www.olcf.ornl.gov/2023/02/08/the-most-advanced-bay-area-earthquake-simulations-will-be-publicly-available/
2/10/23	Autocoding Cancer	Betsy Sonewald	https://www.olcf.ornl.gov/2023/02/10/autocoding-cancer/
2/22/23	ORNL Supercomputing Resources Support Simulation of Influenza Virus Showing Universal Vaccine Promise	Quinn Burkhart	https://www.olcf.ornl.gov/2023/02/22/ornl-supercomputing-resources-support-simulation-of-influenza-virus-showing-universal-vaccine-promise/
3/3/23	UTK-ORNL Course is Teaching Next Generation of Data Center Designers	Betsy Sonewald	https://www.olcf.ornl.gov/2023/03/02/utk-ornl-course-is-teaching-next-generation-of-data-center-designers/
3/30/23	Pioneering Frontier: Forging a File System	Matt Lakin	https://www.olcf.ornl.gov/2023/01/17/forging-a-file-system/
4/10/23	U.S. Department of Energy's INCITE Program Seeks Proposals for 2024 to Advance Science and Engineering at U.S. Leadership Computing Facilities	Katie Bethea	https://www.olcf.ornl.gov/2023/04/10/us-department-of-energys-incite-program-seeks-proposals-for-2024-to-advance-science-and-engineering-at-us-leadership-computing-facilities/
4/12/23	ORNL, NOAA Launch New Supercomputer for Climate Science Research	Betsy Sonewald	https://www.ornl.gov/news/ornl-noaa-launch-new-supercomputer-climate-science-research
4/14/23	NREL, GE Research Team Find Critical Adjustments to Improve Wind Turbine Design	Brooke Van Zandt	https://www.olcf.ornl.gov/2023/04/14/nrel-ge-research-team-find-critical-adjustments-to-improve-wind-turbine-design/
4/26/23	Simulating a More Detailed Universe with Frontier	Coury Turczyn	https://www.olcf.ornl.gov/2023/04/26/simulating-a-more-detailed-universe-with-frontier/
5/10/23	Best of 2023: Abston, keeper of Frontier and Summit data centers, honored by peers	Matt Lakin	https://www.olcf.ornl.gov/2023/05/10/best-of-2023/

Date	Science Highlights Submitted to DOE Title	Writer	URL
5/17/23	Putting Quantum circuits to the test	Matt Lakin	https://www.olcf.ornl.gov/2023/05/17/putting-quantum-circuits-to-the-test/
5/19/23	Learning with the flow: GE study on Summit could lead to cleaner, greener jet flights	Matt Lakin	https://www.olcf.ornl.gov/2023/05/19/learning-with-the-flow/
5/22/23	Early Frontier users seize exascale advantage, grapple with grand scientific challenges	Matt Lakin	https://www.olcf.ornl.gov/2023/05/22/early-frontier-users-seize-exascale-advantage-grapple-with-grand-scientific-challenges/
5/22/23	Predicting the Future of Fission Power	Matt Lakin	https://www.olcf.ornl.gov/2023/05/22/predicting-the-future-of-fission-power/
5/24/23	The OLCF's Matt Sieger Selected to Lead OLCF-6	Coury Turczyn	https://www.olcf.ornl.gov/2023/05/24/the-olcfs-matt-sieger-selected-to-lead-olcf-6/
5/24/23	The NCCS's operational orchestrator of high-performance computing: Ashley Barker	Coury Turczyn	https://www.olcf.ornl.gov/2023/05/24/the-nccss-operational-orchestrator/
5/30/23	Grid Modeling Tool Successfully Launches on World's Fastest Supercomputer	Elsie Puig-Santana	https://www.olcf.ornl.gov/2023/05/30/grid-modeling-tool-successfully-launches-on-worlds-fastest-supercomputer/
6/1/23	LLNL's Diachin takes helm of ECP: ORNL's Ashley Barker will serves as Diachin's Deputy	Betsy Sonewald	https://www.olcf.ornl.gov/2023/06/01/llnls-diachin-takes-helm-of-does-exascale-computing-project/
6/13/23	Visionary report unveils ambitious roadmap to harness the power of AI in scientific discovery	Scott Jones	https://www.olcf.ornl.gov/2023/06/13/visionary-report-unveils-ambitious-roadmap-to-harness-the-power-of-ai-in-scientific-discovery/
6/13/23	Summit study fathoms troubled waters of ocean	Matt Lakin	https://www.olcf.ornl.gov/2023/06/13/summit-study-fathoms-troubled-waters-of-ocean-turbulence/
6/17/23	GE Aerospace Runs One of the world's Largest Supercomputer Simulations to Test Revolutionary New Open Fan Engine Architecture	Katie Bethea	https://www.olcf.ornl.gov/2023/06/16/ge-aerospace-runs-one-of-the-worlds-largest-supercomputer-simulations-to-test-revolutionary-new-open-fan-engine-architecture/

Date	Science Highlights Submitted to DOE Title	Writer	URL
6/18/23	An Oak Ridge-led group combines algorithms and supercomputers to reveal information that scientists missed	ASCR Discovery	https://ascr-discovery.org/2023/06/not-so-plain-sight/
6/27/23	Reaching a New Summit for Supernova Simulations	Coury Turczyn	https://www.olcf.ornl.gov/2023/06/26/reaching-a-new-summit-for-supernova-simulations/
6/28/23	Exascale blastoff: Early users helped launch Frontier to success	Matt Lakin	https://www.olcf.ornl.gov/2023/06/28/exascale-blastoff/
6/30/23	UnifyFS team wins IPDPS award for open-source software	Betsy Sonewald	https://www.olcf.ornl.gov/2023/06/30/unifyfs-team-wins-ipdps-award-for-open-source-software/
7/5/23	Peering into the Universe's Dark Matter	Coury Turczyn	https://www.olcf.ornl.gov/2023/06/26/peering-into-the-universes-dark-matter/
7/7/23	Georgia State researchers use ORNL supercomputer to gain new insights into DNA repair	Coury Turczyn	https://www.olcf.ornl.gov/2023/07/07/new-insights-into-a-shapeshifting-protein-complex/
7/11/23	Air Force Weather-funded research aims to improve predictability of extreme weather	Betsy Sonewald	https://www.olcf.ornl.gov/2023/07/11/air-force-weather-funded-research-aims-to-improve-predictability-of-extreme-weather/
7/14/23	Advancing nanoscience through largescale MD simulations: Scientists use ORNL's Summit supercomputer to learn how cicada wings kill bacteria	Coury Turczyn	https://www.olcf.ornl.gov/2023/07/13/advancing-nanoscience-through-largescale-md-simulations/
7/18/23	Exascale's New Frontier: WarpX, Exascale Modeling of Advanced Particle Accelerators	Coury Turczyn	https://www.olcf.ornl.gov/2023/07/18/exascales-new-frontier-warpx/
7/25/23	College Students have a blast with ORNL during virtual high-performance computing competition	Betsy Sonewald	https://www.olcf.ornl.gov/2023/07/24/students-have-a-blast-with-ornl-in-2023-virtual-cluster-competition/
7/28/23	Improving wildfire predictions with Earth-scale climate models	Stephanie Seay	https://www.ornl.gov/news/improving-wildfire-predictions-earth-scale-climate-models

Date	Science Highlights Submitted to DOE Title	Writer	URL
7/28/23	Shedding Light on Singlet Fission Materials	Coury Turczyn	https://www.olcf.ornl.gov/2023/07/27/shedding-light-on-singlet-fission-materials/
8/7/23	Quantum Experts Gather for Fourth Annual Quantum Computing User Forum at ORNL	Betsy Sonewald	https://www.olcf.ornl.gov/2023/08/07/quantum-experts-gather-for-fourth-annual-quantum-computing-user-forum-at-ornl/
8/7/23	LLNL team reaches milestone in power grid optimization on world's first exascale supercomputer	Jeremy Thomas	https://www.llnl.gov/article/50066/llnl-team-reaches-milestone-power-grid-optimization-worlds-first-exascale-supercomputer
8/8/23	Exascale's New Frontier: WDMApp	Coury Turczyn	https://www.olcf.ornl.gov/2023/08/08/exascales-new-frontier-wdmapp/
8/23/23	ORNL wins six R&D 100 research awards	Sara Shoemaker	https://www.ornl.gov/news/ornl-wins-six-rd-100-research-awards#:~:text=Technologies%20developed%20by%20researchers%20at,of%20finalists%20announced%20last%20week.
8/23/23	Exascale's New Frontier: ExaSMR	Coury Turczyn	https://www.olcf.ornl.gov/2023/08/23/exascales-new-frontier-exasmr/
8/30/23	Oxygen-28 Unbound: Research teams find that rare isotope oxygen-28's 'doubly magic' stability does not overcome its neutron-rich instability	Coury Turczyn	https://www.olcf.ornl.gov/2023/08/30/oxygen-28-unbound/
8/30/23	Exascale Drives Industry Innovation for a Better Future	Caryn Meissner	https://www.olcf.ornl.gov/2023/08/31/exascale-drives-industry-innovation-for-a-better-future/
9/6/23	OLCF to host 19th annual user meeting in October	Betsy Sonewald	https://www.olcf.ornl.gov/2023/09/06/olcf-to-host-19th-annual-user-meeting-in-october/
9/11/23	ExaSMR Nominated for 2023 ACM Gordon Bell Prize	Coury Turczyn	https://www.olcf.ornl.gov/2023/09/11/exasmr-nominated-for-2023-acm-gordon-bell-prize/
9/22/23	ORNL Researchers Develop Open-Source Mixed-Precision Benchmark Tool	Coury Turczyn	https://www.olcf.ornl.gov/2023/09/22/ornl-researchers-develop-open-source-mixed-precision-benchmark-tool/
9/22/23	OLCF Launches New Allocation Program, SummitPLUS	Betsy Sonewald	https://www.olcf.ornl.gov/2023/09/21/olcf-launches-new-allocation-program-summitplus/

Date	Science Highlights Submitted to DOE Title	Writer	URL
9/25/23	DOE's First 'Intro to HPC' Bootcamp Focuses on Energy Justice and a New Model for Workforce Development	Betsy Sonewald	https://www.olcf.ornl.gov/2023/09/22/does-first-intro-to-hpc-bootcamp-focuses-on-energy-justice-and-a-new-model-for-workforce-development/
10/13/23	New Study Reveals the Influence of Natural Climate Drivers on Extreme Monsoons in Pakistan	Betsy Sonewald	https://www.olcf.ornl.gov/2023/10/13/new-study-reveals-the-influence-of-natural-climate-drivers-on-extreme-monsoons-in-pakistan/
10/16/23	Exascale Day 2023	Betsy Sonewald	https://www.olcf.ornl.gov/2023/09/26/exascale-day-2023/
10/19/23	Exascale's New Frontier: MFIX-Exa	Coury Turczyn	https://www.olcf.ornl.gov/2023/10/19/exascales-new-frontier-mfix-exa/
10/27/23	New flood modeling research aims to produce more accurate predictions	Betsy Sonewald	https://www.ornl.gov/news/new-flood-modeling-research-aims-produce-more-accurate-predictions
10/30/23	ESnet Turns On 400G Circuits to Four DOE National Labs, Supercharging Multi-Site Scientific Research	Betsy Sonewald	https://www.olcf.ornl.gov/2023/10/30/esnet-turns-on-400g-circuits-to-four-doe-national-labs-supercharging-multi-site-scientific-research/
11/10/23	Exascale's New Frontier: EXAALT, Predicting how structural materials in fission and fusion reactors will hold up	Coury Turczyn	https://www.olcf.ornl.gov/2023/11/10/exascales-new-frontier-exaalt/
11/13/23	OLCF Hosts First In-Person User Meeting Since Breaking Exascale Barrier	Quinn Burkhart	https://www.olcf.ornl.gov/2023/11/13/olcf-hosts-first-in-person-user-meeting-since-breaking-exascale-barrier/
11/14/23	Big Flex for Big Science	Matt Lakin	https://www.olcf.ornl.gov/2023/11/14/big-flex-for-big-science-frontier-search-for-lightweight-flexible-alloys-earns-gordon-bell-prize-nomination/
11/14/23	Pulling Clouds into Focus: Frontier simulations bring long-range climate forecasts within reach (Gordon Bell Special Prize Finalist)	Matt Lakin	https://www.olcf.ornl.gov/2023/11/14/pulling-clouds-into-focus/
11/14/23	The Journey to Frontier	Matt Lakin	https://www.ornl.gov/journeytofrontier
11/14/23	Super speeds for super AI: Frontier sets new pace for artificial intelligence	Matt Lakin	https://www.ornl.gov/news/super-speeds-super-ai-frontier-sets-new-pace-artificial-intelligence

Date	Science Highlights Submitted to DOE Title	Writer	URL
11/14/23	Super construction at super scale: How ORNL built new home for Frontier	Matt Lakin	https://www.ornl.gov/news/super-construction-super-scale-how-ornl-built-new-home-frontier
11/16/23	Cloud Simulations on Frontier Awarded Gordon Bell Special Prize for Climate Modeling	Matt Lakin	https://www.olcf.ornl.gov/2023/11/16/cloud-simulations-on-frontier-awarded-gordon-bell-special-prize-for-climate-modeling/
11/16/23	Frontier Search for Lightweight, Flexible Alloys Wins Gordon Bell Prize	Matt Lakin	https://www.olcf.ornl.gov/2023/11/16/frontier-search-for-lightweight-flexible-alloys-wins-gordon-bell-prize/
11/27/23	Exascale's New Frontier: SOLLVE, Advancing OpenMP programming interface into the exascale era	Coury Turczyn	https://www.olcf.ornl.gov/2023/11/27/exascales-new-frontier-sollve/
12/11/23	Data Curation for the Exascale Era	Coury Turczyn	https://www.olcf.ornl.gov/2023/12/11/data-curation-for-the-exascale-era/
12/13/23	ORNL Scientists Generate Molecular Datasets at Extreme Scale	Coury Turczyn	https://www.olcf.ornl.gov/2023/12/12/ornl-scientists-generate-molecular-datasets-at-extreme-scale/
12/14/23	Exascale’s New Frontier: Kokkos	Coury Turczyn	https://www.olcf.ornl.gov/2023/12/13/exascales-new-frontier-kokkos/
12/18/23	Custom Software Speeds Up, Stabilizes High-Profile Ocean Model	Elizabeth Rosenthal	https://www.olcf.ornl.gov/2023/12/18/custom-software-speeds-up-stabilizes-high-profile-ocean-model/
12/19/23	OLCF Announces SummitPLUS Allocations	Coury Turczyn	https://www.olcf.ornl.gov/2023/12/19/olcf-announces-summitplus-allocations/

APPENDIX C. 2023 OLCF Operational Assessment

The Oak Ridge Leadership Computing Facility (OLCF) provided the Frontier computational resource (Table C-1) and the Alpine and HPSS data resources for production use in 2023. Supporting systems such as EVEREST, Andes, and data transfer nodes were also offered. Metrics for these supporting systems are not provided. The following systems were operational during this reporting period.

HPE Cray EX (FRONTIER) RESOURCE SUMMARY

Debuting in 2022 at 1.102 exaflops, Oak Ridge National Laboratory's (ORNL's) Frontier supercomputer is the world's first exascale system. ORNL's long history of supercomputing excellence enables scientists to expand the scale and scope of their research, solve complex problems in less time, and fill critical gaps in knowledge. The OLCF team tuned the system and improved the Linpack score in May 2023, adding 0.92 exaflops to the result for a new score of 1.194 exaflops. Frontier is an HPE Cray EX supercomputer comprising 74 Olympus rack HPE cabinets, each with 128 AMD compute nodes, for a total of 9,408 AMD compute nodes. Each Frontier compute node consists of one 64-core AMD Optimized 3rd Gen EPYC CPU (with 2 hardware threads per physical core) with access to 512 GB of DDR4 memory. Each node also contains four AMD MI250X GPUs, each with 2 Graphics Compute Dies (GCDs) for a total of 8 GCDs per node. The programmer can think of the 8 GCDs as 8 separate GPUs, each having 64 GB of high-bandwidth memory. The CPU is connected to each GCD via an Infinity Fabric CPU-GPU link, allowing a peak host-to-device and device-to-host bandwidth of 36 + 36 GB/s. The 2 GCDs on the same MI250X are connected with an Infinity Fabric GPU-GPU link with a peak bandwidth of 200 GB/s. The GCDs on different MI250X GPUs are also connected with an Infinity Fabric GPU-GPU link, and the peak bandwidth ranges from 50 to 100 GB/s based on the number of Infinity Fabric connections between individual GCDs.

IBM AC922 (SUMMIT) RESOURCE SUMMARY

The OLCF installed and deployed an IBM AC922 system, Summit, which became available for full production on January 1, 2019. Summit comprises 4,608 high-density compute nodes, each equipped with two IBM POWER9 CPUs and six NVIDIA Volta GPUs (Table C-2). In total, the Summit system is capable of 200 petaflops of peak computational performance and was recognized as the most powerful system in the world for its performance on both the High-Performance Linpack and the Conjugate Gradient benchmark applications from June 2018 until June 2020. Three new cabinets with a higher memory footprint were added to the Summit system in July 2020 to support COVID-19 research. Additionally, three more cabinets were added to the system in March 2021. These cabinets will provide spare parts for Summit in its 6th year of operation in 2024.

Table C-1. Frontier 2023.

System	Access	Type	CPU	GPU	Computational description			Interconnect
					Nodes	Node configuration	Memory configuration	
Frontier	Full production	HPE Cray EX	AMD Optimized 3rd Gen EPYC CPU (64 core)	AMD Instinct MI250X GPUs, each feature 2 Graphics Compute Dies (GCDs) for a total of 8 GCDs per node	9,408	64-core AMD Optimized 3rd Gen EPYC CPU (1/node) + AMD Instinct MI250X GPUs (4/node)	700 PB HDD + 11 PB Flash Performance Tier 9.4 TB/s and 10 PB Metadata Flash Lustre	4-port HPE Slingshot 200 Gbps (25 GB/s) NICs providing a node-injection bandwidth of 800 Gbps (100 GB/s)

Table C-2. Summit 2023.

System	Access	Type	CPU	GPU	Computational description			Interconnect
					Nodes	Node configuration	Memory configuration	
Summit	Full production	IBM AC922	3.45 GHz IBM POWER9 (22 core)	1,530 MHz NVIDIA V100 (Volta)	4,608	IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node)	(4,608) 512 GB DDR4 and 96 GB HBM2 per node; (54) 2TB DDR4 and 192 GB HBM2 per node; >10 PB DDR4 + HBM + Non-volatile aggregate	Mellanox EDR 100 Gbps InfiniBand (Non-blocking Fat Tree)

GPFS FILE SYSTEMS (ALPINE AND WOLF) RESOURCE SUMMARY

In January 2019, the OLCF deployed Alpine as its next-generation global file system to support the computational resources in the OLCF. Alpine is a single GPFS (General Parallel File System) namespace with a usable capacity of 250 PB and a file system–level performance of 2.5 TB/s. The Alpine file system is the default high-performance parallel file system for all of the OLCF's moderate compute resources.

In March 2017, the OLCF procured, installed, and deployed the Wolf GPFS, which serves as the center-wide file system for the computational resources in the Open Production enclave. Wolf provides a total storage capacity of 8 PB and up to 120 GB/s of performance.

LUSTRE FILE SYSTEM (ORION) RESOURCE SUMMARY

In April 2023, the OLCF deployed Orion as its next-generation global file system to support the computational resources in the OLCF. Orion is a multi-tiered (flash and HDD), single Lustre namespace with a usable capacity of 700 PB. The demonstrated file system–level performance from Fronter is 11 TB/s write and 14 TB/s read. The Orion file system is the default high-performance parallel file system for all the OLCF's moderate compute resources.

DATA ANALYSIS AND VISUALIZATION CLUSTER (ANDES) RESOURCE SUMMARY

A new data analytics cluster named Andes went into production on December 9, 2020. The primary purpose of the data analytics cluster is to provide a conduit for large-scale scientific discovery through pre- and post-processing of simulation data generated on Summit. Users with accounts on Innovative and Novel Computational Impact on Theory and Experiment (INCITE) or ASCR Leadership Computing Challenge (ALCC) supported projects are automatically given accounts on the data analytics cluster. Director's Discretionary (DD) projects may also request access to this cluster. Andes is a 704-node cluster, and each node contains two 16-core 3.0 GHz AMD EPYC processors and 256 GB of main memory. Andes offers nine additional heterogeneous nodes, each of which provides 1 TB of main memory and two NVIDIA Tesla K80 (Kepler GK210) GPUs. The data analytics cluster is currently connected to the OLCF's 250 PB high-performance GPFS file system, Alpine.

HIGH-PERFORMANCE STORAGE SYSTEM RESOURCE SUMMARY

The OLCF provides a long-term storage archive system based on the High-Performance Storage System (HPSS) software product co-developed by IBM, Los Alamos National Laboratory, Sandia National Laboratories, Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and ORNL. The ORNL HPSS instance is currently over 136 PB in size and provides ingestion rates of up to 30 GB/s via a 22 PB front-end disk cache backed by a 21-frame Spectra Logic TFinity tape library that houses 81 IBM TS1155 tape drives and over 17,000 tape media slots, thereby providing ORNL a current capacity of 180 PB, which is expandable well into hundreds of PBs. The archive's average ingestion rate ranges between 100 and 150 TB/day. The archive environment consists of hardware from Dell, Brocade, DataDirect Networks, and Spectra Logic.

VISUALIZATION RESOURCE SUMMARY

The EVEREST facility has three computing systems and two separate state-of-the-art visualization display walls. The primary display wall spans 30.5 × 8.5 ft. and consists of eighteen 1920 × 1080 Barco projection displays arranged in a 6 × 3 configuration for a 32:9 aspect ratio at 11,520 × 3,240. The secondary display wall is being upgraded to eighteen 1920 × 1080 BARCO LCD displays arranged in a 6 × 3 configuration to provide a secondary 32:9 aspect ratio. There are several 86 in. mobile interactive touch displays for easy and fast collaboration. Multiple augmented reality systems provide an interactive scalable room space equipped for mixed-reality data exploration and analysis. These combined technologies, along with OLCF staff expertise, allow scientists to analyze complex scientific datasets in an immersive environment and communicate abstract concepts in an intuitive visual format.

NEARLINE STORAGE (THEMIS) SUMMARY

Themis entered production in late CY 2022 to bridge the retention requirement gap between scratch (immediate term) and archival (permanent) data storage use cases while streamlining resource access requirements. In response to user needs, the system is also intended to facilitate gateways and other data-sharing mechanisms within the National Center for Computational Sciences, ORNL, and the broader research community. Themis consists of a 14 PB Spectrum Scale file system backed by a 37 PB IBM TS4500 tape library. The Spectrum Scale layer will be capable of an aggregate bandwidth of 45 GB/s, moving to 65 GB/s after network upgrades. The tape component offers dual-copy to provide resiliency with an aggregate bandwidth of 7.68 GB/s. The system was expanded by 11 PB of Spectrum Scale and 7 PB of tape in late Q1 of CY 2023, with additional expansions to both disk and tape components later in the year.

PROTECTED DATA INFRASTRUCTURE SUMMARY

The OLCF now provides a production enclave to support the processing of data subject to the Health Insurance Portability and Accountability Act and International Traffic and Arms Regulations. Using the infrastructure in this enclave, users can submit and run Protected Data workloads on Summit. The enclave provides a dedicated login node for submission of Protected Data jobs and access to the Protected Data GPFS file system, Arx. Arx provides a total storage capacity of ~3.3 PB and up to 30 GB/s of performance.

OLCF HPC RESOURCE PRODUCTION SCHEDULE

The OLCF computational systems entered production according to the schedule listed in Table C-3. This list includes historical data associated with the Cray XT5, the very small overlap in December 2011 beginning with the introduction of the Cray XK6, and the series of Cray XK systems first available in 2012 and 2013.

Table C-3. OLCF HPC system production dates, 2008–present.

System	Type	Production date[a]	Performance end date[b]	Notes
Frontier	HPE Cray EX	April 5, 2023	Present	
Summit	IBM AC922	March 19, 2021	Present	Production with 4,608 hybrid CPU-GPU nodes: IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node). Three cabinets added as a spare parts cache for an optional 6th year.
Summit	IBM AC922	July 1, 2020	Present	Production with 4,608 hybrid CPU-GPU nodes: IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node). Three cabinets added for COVID-19 research.

System	Type	Production date[a]	Performance end date[b]	Notes
Summit	IBM AC922	January 1, 2019	Present	Production with 4,608 hybrid CPU-GPU nodes: IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node).
Spider III (Alpine)	GPFS	January 1, 2019	Present	250 PB GPFS single namespace file system.
Spider II	Lustre parallel file system	October 3, 2013	August 2, 2019	Delivered as two separate file systems: /atlas1 and /atlas2, 30+ PB capacity.
Orion	HPE ClusterStor E1000 v.2	April 5, 2023	Present	700 PB Lustre single namespace file system.
Eos	Cray XC30	October 3, 2013	August 2, 2019	Production with 736 Intel E5, 2,670 nodes.
Titan	Cray XK7	May 31, 2013	August 2, 2019	Production with 18,688 hybrid CPU-GPU nodes: AMD Opteron 6274/NVIDIA K20X.
JaguarPF	Cray XK6	September 18, 2012	October 7, 2012	Production at 240,000 cores until September 18, when partition size was reduced to 120,000 AMD Opteron cores. Additional Kepler installation. TitanDev access terminated.
JaguarPF	Cray XK6	February 13, 2012	September 12, 2012	Full production until September 12, when partition size was reduced to 240,000 AMD Opteron cores. Beginning of Kepler installation.
JaguarPF	Cray XK6	February 2, 2012	February 13,2012	Stability test. Restricted user access. 299,008 AMD Opteron 6274 cores. Includes 960-node Fermi-equipped partition.
JaguarPF	Cray XK6	January 5, 2012	February 1, 2012	Acceptance. No general access. 299,008 AMD Opteron cores.
JaguarPF	Cray XK6	December 12, 2011	January 4, 2012	142,848 AMD Opteron cores.
JaguarPF	Cray XT5	October 17, 2011	December 11, 2011	117,120 AMD Opteron cores.
JaguarPF	Cray XT5	October 10, 2011	October 16, 2011	162,240 AMD Opteron cores.
JaguarPF	Cray XT5	September 25, 2009	October 9, 2011	224,256 AMD Opteron cores.
JaguarPF	Cray XT5	August 19, 2008	July 28, 2009	151,000 AMD Opteron cores.

[a] The production date used for computing statistics is either the initial production date or the production date of the last substantive upgrade to the computational resource.
[b] The performance end date is the last calendar day that user jobs were allowed to execute on that partition.

APPENDIX D. 2023 OLCF Operational Assessment

2023 OPERATIONAL ASSESSMENT GUIDANCE

Scheduled Availability

Scheduled availability (SA) (Eq. [D.1]) in high-performance computing facilities is the percentage of time that a designated level of resource is available to users, excluding scheduled downtime for maintenance and upgrades. The user community must be notified of the need for a maintenance event window no less than 24 hours in advance of the outage (emergency fixes) for it to be considered scheduled downtime. Users will be notified of regularly scheduled maintenance in advance on a schedule that provides sufficient notification, no less than 72 hours before the event and preferably as many as 7 calendar days prior. If that regularly scheduled maintenance is not needed, then users will be informed of the maintenance event cancellation in a timely manner. Any service interruption that does not meet the minimum notification window is categorized as an unscheduled outage.

$$SA = \left(\frac{\text{time in period} - \text{time unavailable due to outages in period}}{\text{time in period} - \text{time unavailable due to scheduled outages in period}}\right) * 100 \qquad \text{(D.1)}$$

A significant event that delays a return to scheduled production by more than 4 hours will be counted as an adjacent unscheduled outage, unscheduled availability, and additional interrupt.

Overall Availability

Overall availability (OA) (Eq. [D.2]) is the percentage of time that a system is available to users. Outage time reflects both scheduled and unscheduled outages.

$$OA = \left(\frac{\text{time in period} - \text{time unavailable due to outages in period}}{\text{time in period}}\right) * 100 \qquad \text{(D.2)}$$

Mean Time to Interrupt

Mean time to interrupt (MTTI) is the time, on average, to any outage of the full system, whether unscheduled or scheduled (Eq. [D.3]).

$$MTTI = \left(\frac{\text{time in period} - (\text{ duration of scheduled outages} + \text{duration of unscheduled outages })}{\text{number of scheduled outages} + \text{number of unscheduled outages} + 1}\right) \qquad \text{(D.3)}$$

Mean Time to Failure

Mean time to failure (MTTF) is the time, on average, to an unscheduled outage of the full system (Eq. [D.4]).

$$MTTF = \frac{\text{time in period} - (\text{ duration of unscheduled outages })}{\text{number of unscheduled outages} + 1} \qquad \text{(D.4)}$$

System Utilization

System utilization (SU) is the percentage of time that the system's computational nodes run user jobs. No adjustment is made to exclude any user group, including staff and vendors (Eq. [D.5]).

$$SU = \left(\frac{\text{core hours used in period}}{\text{core hours available in period}}\right) * 100 \qquad \text{(D.5)}$$

APPENDIX E. 2023 OLCF Operational Assessment

DIRECTOR'S DISCRETIONARY PROJECTS UTILIZING SUMMIT (ENHANCED ENCLAVE) ENABLED (AT ANY POINT) IN CY 2023

Table E-1. Director's Discretionary projects utilizing Summit (Enhanced Enclave) in CY 2023.

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
ARD144_MDE	Sriram Shankaran	GE Aviation	20,000	0	High Pressure-Low Pressure Turbine Interaction Dynamics
CSC487_MDE	Joseph Lake	ORNL	30,000	27	Population Health Research with Sensitive Data
GEO141_MDE	Dalton Lunga	ORNL	7,000	16,176	Global Scale Gravitational Modeling with AI
LRN029_MDE	Nathan Hariharan	Bowhead Total Enterprise Solutions LLC	20,000	0	Development of Machine Learning Based Surrogate Modeling Methods to Support Aerospace Applications
MED116_MDE	Heidi Hanson	ORNL	30,000	40,205	Privacy-Preserving Transformer Models for Clinical Natural Language Processing
MED118_MDE	Silvia Crivelli	LBNL	820,000	49,332	Develop Language Models for EHR Data
MED122_MDE	Mayanka Chandra Shekar	ORNL	10,000	47	Cincinnati Children's Hospital Medical Center (CCHMC) Mental Health Trajectories
SYB109_MDE	Michael Garvin	ORNL	20,000	0	Genomic Structural Variation
SYB110_MDE	Daniel Jacobson	UT, ORNL	20,000	0	Combinatorial Associations: From EHR or Epistasis

DIRECTOR'S DISCRETIONARY PROJECTS UTILIZING SUMMIT (MODERATE ENCLAVE) ENABLED (AT ANY POINT) IN CY 2023

Table E-2. Director's Discretionary projects utilizing Summit (Moderate Enclave) in CY 2023.

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
ARD147	Wesley Harris	MIT	20,000	202,037	Direct Numerical Simulations of Hypersonic Boundary Layer Transition on a Flat Plate

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
ARD153	Zhi Jian Wang	University of Kansas	20,000	3,012	Wall-Modeled Large Eddy Simulation of High Lift Configuration
ARD155	Vineet Ahuja	Whisper Aero Inc.	20,000	2,199	CFD Simulations of Novel Airframe Configurations with Distributed Electric Propulsors
ARD161	Jeffrey Slotnick	Boeing	60,000	79,834	Fifth High Lift Prediction Workshop CFD Technology Assessment
AST031	Pierre Ocvirk	Universite de Strasbourg, Strasbourg Astronomical Observatory	10,000	2,828	Reionization And Its Impact on The Local Universe: Witnessing Our Own Cosmic Dawn
AST146	Brian O'Shea	Michigan State University	20,000	0	Measuring Performance and Scaling of Kokkos-accelerated Athena++ on Summit
AST154	Philipp Moesta	University of Amsterdam	20,000	21,619	Dynamical Space-time GRMHD Simulations of Neutron-star Mergers and Remnants
AST163	Christian Cardall	ORNL	20,000	5,131	INCITE Preparation for 3D+1D Core-Collapse Supernova Simulations with GenASiS
AST166	Gaurav Khanna	UMass Dartmouth	5,000	0	Mixed-Precision WENO Method for Hyperbolic PDE Solutions
AST176	Eliu Huerta Escudero	UIUC, Argonne	35,000	22,730	Geometric Deep Learning for Gravitational Wave Astrophysics
AST183	Brian O’Shea	Michigan State University	10,000	1,981	Feedback and Energetics from Magnetized AGN Jets in Galaxy Groups and Clusters
ATM112	Wei Zhang	ORNL	81,000	106,104	Aim High: Air Force R&D Collaboration
ATM121	Xin-Zhong Liang	U. of Maryland	15,000	0	Dashboard for Agricultural Water use and Nutrient management (DAWN)
ATM122	Branko Kosovic	NCAR, UCAR	20,000	0	GPU-Resident Real-Time Large-Eddy Simulations of Stably Stratified Atmospheric Boundary Layers
ATM127	Ngoc-Cuong Nguyen	MIT	20,000	6,820	High-Fidelity Space Weather Modeling

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
ATM129	Veerabhadra Kotamarthi	Argonne	20,000	8,618	Distributed Wind Obstacle Model
ATM135	Mayur Mudigonda	Vayuh Inc.	10,000	3,399	vayuh-deep-s2s
ATM137	Ngoc-Cuong Nguyen	MIT	20,000	0	A Discontinuous Galerkin Approach for the Physics-Based Modeling of the Coupled Ionosphere-Thermosphere System
BIE119	Abhishek Singharoy	Arizona State University	40,000	0	Membrane Models of Biological Energy Transfer
BIF139	Chongle Pan	OU	20,000	129	Biolearning
BIF140	Dali Wang	ORNL	20,000	0	Efficient Prediction of Lineage Tree and Interactions with Graph Neural Networks and Deep Reinforcement Learning
BIF141	Thomas Beck	ORNL	15,000	8,913	A Multimodal Atlas of Cells and Circuits in the Mouse Brain
BIP167	Philip Kurian	Howard University	20,000	4	Computing Many-Body Van Der Waals Dispersion Effects in Biomacromolecules
BIP215	Julie Mitchell	ORNL	75,000	59,549	Training An Advanced Model for Structure-Based Proteomics
BIP218	Michael Kiebish	Berg	20,000	2,705	Supercomputing-Guided Drug Discovery and Artificial Intelligence in the Pharmaceutical Sciences
BIP222	Ruth Nussinov	Leidos, Inc.	15,000	0	Modeling Microbial Proteins and Their Interactions with The Human Proteins
BIP227	Giulia Palermo	University of California – Riverside	20,000	0	Dynamics and Mechanism of Transposon-Encoded CRISPR-Cas Systems
BIP230	Yinglong Miao	University of Kansas	19,800	9,563	Accelerated Molecular Simulations of Protein Interactions with SARS-CoV-2 Spike
BIP232	Pin-Kuang Lai	Stevens Institute of Technology	20,000	11,120	Integrating Molecular Dynamics Simulations and Machine Learning to Accelerate Antibody Development

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
BIP233	Dilnoza Amirkulova	Proctor & Gamble Company	20,000	10,140	Molecular Modeling of Surfactant Interactions with Phospholipids Bilayers Mimicking Corneal Epithelium
CFD124	Bamin Khomami	UT	200,000	610	Elucidating the Molecular Rheology of Entangled Polymeric Fluids via Direct Comparison of NEMD Simulations and Model Predictions
CFD142	Balaji Jayaraman	GE	20,000	13,804	Characterizing Coastal Low-Level Jets and their Impact on Offshore Wind Farms
CFD154	Spencer Bryngelson	Georgia Tech	5,000	10,337	Accelerated Sub-Grid Multi-Component Flow Physics
CFD159	Vittorio Badalassi	ORNL	5,000	4,353	Fusion Energy Reactor Models Integrator (FERMI)
CFD161	Mark Kostuk	General Atomics	18,000	13,133	ALMA for SNS
CFD163	Ioannis Nompelis	University of Minnesota	20,000	0	Porting of an implicit finite volume solver for compressible flow to GPUs
CFD165	Federico Municchi	Colorado School of Mines	20,000	13,050	CFD Simulations of Heat and Mass Transport in Solar Powered Membrane Distillation Systems
CHM155	Peter Coveney	University College London	80,000	42,882	COMPBIO
CHM156	Remco Havenith	U. of Groningen	7,500	24,681	TURTLE
CHM160	Andre Severo Pereira Gomes	CNRS	20,000	2,265	PRECISE: Predictive Electronic Structure Modeling of Heavy Elements
CHM174	Monojoy Goswami	UT, ORNL	20,000	36,094	Multi-Scale/Multi-Physics Molecular Simulations at the Chemical Sciences Division
CHM181	Kurt Mikkelsen	University of Copenhagen	10,000	32,360	Massively Parallel GPU-Enabled Cluster Perturbation Methods
CHM184	Darrin York	Rutgers	20,000	0	AMBER Drug Discovery
CHM186	Anna Gudmundsdottir	University of Cincinnati	10,000	5,392	Photopopping Crystals
CHM187	Vyacheslav Bryantsev	ORNL	20,000	118,183	Molten Salts Modeling with Ab Initio Accuracy

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
CHM188	Thanh Do	UT	20,000	10,206	Seeking Novel N-Methylated Macrocyclic Peptide Conformers While Optimizing MD Methodology
CHM189	Alex Ivanov	ORNL	20,000	27,273	Pm Complex Chemistry Relativistic Simulations
CHP121	Jonathan Owens	GE	10,000	10,863	Materials Discovery and Characterization for CO_2/H_2O Co-Adsorption
CHP124	Richard Friesner	Columbia University	5,000	5,978	Quantum Simulations of Photosystem II and Organometallic Enzymes
CLI131	Samuel Hatfield	European Centre for Medium-Range Weather Forecasts	8,000	2	A Baseline for Global Weather and Climate Simulations at 1 KM Resolution
CLI137	Forrest Hoffman	ORNL	10,000	24	Earth System Grid Federation 2 (ESGF2)
CLI138	Moetasim Ashfaq	ORNL	20,000	48,550	Analytical Frameworks for Sub-Seasonal to Multi-Decadal Climate Predictions and Impact Assessments
CLI143	Nathaniel Collier	ORNL	5,000	0	Identifying Ecosystems Vulnerable to Climate Change
CLI144	Peter Thornton	ORNL	20,000	607	Ultra-High Resolution Land Process Simulation Using GPUs And OpenACC on Summit
CLI146	Jiafu Mao	ORNL	5,000	5,225	Ecosystem Resilience to Thermal Extremes: Urbanization Impacts
CLI900	Valentine Anantharaj	ORNL	5,000	33,177	Provisioning of Climate Data
CMB148	Aditya Konduri	IISc, IISc Bangalore	15,000	6	Scalable Mathematically Asynchronous Algorithms for Flow Solvers
CMB151	Venkat Raman	U. of Michigan	15,000	0	Detonation Simulations
CPH005	Dario Alfe	University College London	5,000	6,013	New Frontiers for Material Modeling via Machine Learning Techniques with Quantum
CPH138	Sumit Sharma	OU	20,000	2,969	Computational Studies of Interactions of Small Molecules with Biological and Polar Interfaces
CPH141	Ivan Stich	Slovak Academy of Sciences	20,000	7,580	QMCPACK Performance Comparison and Tests

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
CPH149	Alberto Nocera	University of British Columbia	20,000	27,560	Accelerating Tensor Network Methods for Establishing Quantum Advantage using Quantum Annealing
CSC143	Norbert Podhorszki	ORNL	30,000	53,864	ADIOS – The Adaptable IO System
CSC357	Laxmikant Kale	UIUC	15,000	0	CharmRTS
CSC369	Dmitry Pekurovsky	UC San Diego	15,000	36	Scalable Software Framework for Multidimensional Fourier Transforms
CSC377	Allan Grosvenor	Microsurgeonbot	20,000	94,131	Expertise-as-a-Service via Scalable Hybrid Learning: R&D Supporting Improvements to Hybrid Intelligence Agent Tooling
CSC380	Bronson Messer	ORNL	150,000	125,479	CAAR for Frontier
CSC382	Catherine Schuman	UT	5,000	16,670	Scalable Neuromorphic Simulation and Training
CSC401	Barbara Chapman	SUNY Stony Brook	10,000	1	Support for the Deployment of OpenMP in Large-Scale Scientific Applications
CSC427	Michela Taufer	UT	15,000	3	A4NN: Accelerating Scientific AI Leveraging Open Data and Open Models
CSC452	Abhinav Bhatele	LLNL, U. of Maryland	20,000	19,137	Performance Analysis and Tuning of HPC and AI Applications
CSC470	Christopher Stanley	ORNL	20,000	349	Flexible Privacy-Enabled Platform for Sensitive Applications
CSC491	Terry Jones	ORNL	15,000	0	WEAVABLE
CSC492	John Cohn	IBM	10,000	0	Structure Transformer for Code Generation
CSC505	Mark Coletti	ORNL	20,000	44,872	Increasing Search Effectiveness for MENNDL via Variable-Length Representation
CSC506	Johannes Blaschke	LBNL	10,000	0	NESAP
CSC511	Geoffrey Lentner	Purdue University	10,000	0	HyperShell Development and Scaling

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
CSC519	Maciej Cytowski	Pawsey Supercomputing Centre	15,000	0	Pawsey SC Access for Early Testing Purposes
CSC524	Van Ngo	ORNL	20,000	2	Benchmarking and Development of the MiMiC Framework for Highly Efficient QM/MM Simulations on Leadership Supercomputers
CSC528	Ang Li	PNNL	19,000	10,014	nwqsim
CSC529	Pieter Abbeel	UC Berkeley	20,000	15,174	Learning Common Sense Reasoning through Large Scale Video Generation
ENG123	Ramesh Balakrishnan	Argonne	10,000	20	Direct Numerical Simulation of Separated Flow over a Speed Bump at Higher Reynolds Numbers
ENG130	Zhiting Tian	Cornell	18,000	3,003	Understanding Heat Transport in Complex, Ultra-Low Conductivity Materials
ENG131	Pablo Seleson	ORNL	20,000	23,108	A Scalable Cabana-Based Peridynamics Fracture Simulator
FUS137	William Fox	Princeton	20,000	5,572	Energetics of Collisionless Plasmas in the Laboratory and Space
FUS147	Noah Reddell	HPE	20,000	0	Study of High Energy Density Z-Pinch Plasma Stability by Kinetic Model on GPU
FUS150	Yury Osetskiy	ORNL	30,000	13,435	Ab Initio-Based Machine Learning of Surface Damage by Fusion Plasma
GEN026	Bran Radovanovic	ORNL	0	0	DC-DataCenterPowerResearch
GEO112	Philip Maechling	USC	40,000	338,234	Extreme-Scale Simulations for Advanced Seismic Ground Motion and Hazard Modeling
GEO142	Dalton Lunga	ORNL	18,000	13,837	Scalable GeoAI Workflows for Satellite Image Analytics
GEO146	Kohei Fujita	University of Tokyo	20,000	0	Scalable Dynamic Stochastic Earthquake Simulation with Low-Ordered Unstructured Finite Elements on Frontier
GEO149	Barry Schneider	NIST	50,000	91,389	Physics-Based Modeling for Forest Fuel Management

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
HEP134	Dirk Hufnagel	Fermilab	2,000	0	HEPCloud-FNAL
LGT100	Andre Walker-Loud	LBNL	60,000	169,406	The Structure and Interactions of Nucleons from the Standard Model
LGT114	Amy Nicholson	University of North Carolina at Chapel Hill	10,000	9,323	Electromagnetic Corrections to the Nucleon Axial Charge
LRN018	Vikram Jadhao	Indiana University	20,000	0	Training Machine Learning Surrogates for Simulations using Adaptive Methods
LRN023	Maria Martin Aguirre	Dridam Flight LLC	20,000	0	Data Assimilation for Digital Twins of Supersonic and Hypersonic Vehicles
LRN025	Steven Young	ORNL	20,000	6,389	Model Parallel Neural Architecture Search
LRN026	Massimiliano Lupo Pasini	ORNL	20,000	33,573	Scalable Accelerated Training of Physics Informed Graph Convolutional Neural Networks for Material Science and Chemistry
LRN030	Charles Cao	UT	10,000	0	Formal Verification of Neural Networks in AI-based Decision Making
MAT198	Ivan Oleynik	University of South Florida	50,000	19,632	Predictive Simulations of Phase Transitions in Dynamically Compressed Materials
MAT201	David Lingerfelt	ORNL	100,000	100,503	Center for Nanophase Materials Sciences
MAT226	Jan Michael Carrillo	UT, ORNL	5,000	24,000	Molecular Dynamic Simulations of Amphiphilic Oligomer Membranes: Design Rules Toward Stable Membranes Capable of Learning and Memory
MAT242	Yangyang Wang	ORNL	10,000	5,672	Resolving Flow-Induced Mesoscopic Structures in Polymeric Materials
MAT243	Aliya Tychengulova	Satbayev University	5,000	8	Study on the Catalytic Mechanism and Properties of Materials for Water-Splitting in Natural and Artificial Photosynthesis
MAT250	Massimiliano Lupo Pasini	ORNL	20,000	32,126	Scalable Accelerated Training of Physics-Informed Deep Learning Models for Material Science

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
MAT253	Jonathan Schwartz	U. of Michigan	15,000	599	Fused Multi-Modal Electron Tomography
MAT255	Ankit Roy	PNNL	15,000	0	Radiation Response in Lithium Ceramics Using Cascade Simulations
MAT256	Zongtang Fang	Nissan Technical Center North America	20,000	19,082	Computational Studies of Spent NMC Cathode Materials Toward Direct Recycling
MED112	Sumitra Muralidhar	VA, Veterans Administration, BVARI	30,000	506	Genome-Wide Phenome-Wide Association Study in the Million Veteran Program
MED121	Hong-Jun Yoon	ORNL	20,000	2,692	Skin Cancer Classification Systems with Mobile Devices
MED122	Greeshma Agasthya	ORNL	0	0	CCHMC Mental Health Trajectories
NFI123	Benjamin Collins	ORNL	20,000	138	VERA-Excore
NFI124	Scott Beckman	Washington State University	4,000	0	Molten Salt INCITE Benchmarking Tests
NME100	Ram Mohan	North Carolina Agriculture and Technology State University	1,000	0	Nano to Engineering Scale Modeling – Materials, Mechanics, and Manufacturing
NPH126	Aurel Bulgac	U. of Washington	30,000	46,514	Nuclear Matter Dynamics in Real Time and the Heaviest Elements in Nature
NPH143	Kenneth McElvain	UC Berkeley	20,000	16,419	Neutron Star Interiors: Nuclear-matter Calculations from Chiral Effective Field Theory
NRO107	Shinjae Yoo	Brookhaven	10,000	5	Spatiotemporal Learning
PHY149	Markus Eisenbach	ORNL	20,000	4,658	Pushing the Limits of Classical Simulation of Hard Quantum Circuits via Novel Tensor Network Algorithms and Accelerated High Performance Computing
PHY166	Deborah Levin	UIUC	20,000	0	High Fidelity Kinetic Simulations of Plasma-Based Flows Using Heterogeneous Computational Strategies

Project ID	PI	Institution	Most recent summit allocation	Summit usage	Project name
PHY171	Michael Wilking	SUNY Stony Brook	20,000	20,077	Markov Chain Monte Carlo Development for Neutrino Physics
STF050	Olga Kuchar	ORNL	1,000	0	DOI Data Archive for Non-OLCF Users
TUR140	Shanti Bhushan	Mississippi State University	10,000	266	Neural Network Framework to Enhance Unsteady Turbulent Flow Predictions

DIRECTOR'S DISCRETIONARY PROJECTS UTILIZING FRONTIER (MODERATE ENCLAVE) ENABLED (AT ANY POINT) IN CY 2023

Table E-3. Director's Discretionary projects utilizing Frontier (Moderate Enclave) in CY 2023.

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
ARD154	Frank Ham	Cascade Technologies	300,000	0	Predicting Broadband Noise Signatures of Urban Air Mobility Vehicles
AST031	Pierre Ocvirk	Universite de Strasbourg, Strasbourg Astronomical Observatory	10,000	0	Reionization and Its Impact on the Local Universe: Witnessing Our Own Cosmic Dawn
AST163	Christian Cardall	ORNL	20,000	3,905	INCITE Preparation for 3D+1D Core-Collapse Supernova Simulations with GenASiS
AST181	Evan Schneider	University of Pittsburgh	500,000	11,416	Cholla
AST182	Jay Kalinani	RIT	20,000	133	Benchmarking for the New GPU-Accelerated Open-Source GRMHD Code AsterX
AST185	Benjamin Prather	LANL	20,000	695	Horizon-Scale Variability Modeling for the EHT
AST186	Carl Fields	LANL	16,000	13,310	Exascale Efforts for Characterizing Turbulent Convection in Core-Collapse Supernovae
AST187	Gregory Werner	U. of Colorado	20,000	31	PIC Simulation of Astrophysical Plasmas

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
AST190	Pierre Ocvirk	Universite de Strasbourg, Strasbourg Astronomical Observatory	10,000	0	Understanding the Epoch of Reionization Through Analysis of Cosmic Dawn III
AST191	Philipp Moesta	University of Amsterdam	40,000	11,374	Exascale Simulations of Binary Neutron-Star Mergers
AST192	David Vartanyan	Carnegie Institution for Science	10,000	0	Core-Collapse Supernovae to Shock Breakout
AST193	Ka Ho Yuen	LANL	15,000	0	Exascale Multiphase Magnetized Turbulence Simulations with Cosmic Ray Support and Chemistry Network
ATM112	Wei Zhang	ORNL	60,000	1,985	Aim High: Air Force R&D Collaboration
ATM136	Wei Zhang	ORNL	20,000	0	Quantifying Feedbacks of Climate Intervention Under Climate Change
BIF147	Kathy Yelick	UC Berkeley, LBNL	20,000	0	Exabiome
BIP167	Philip Kurian	Howard University	20,000	0	Computing Many-Body Van Der Waals Dispersion Effects in Biomacromolecules
BIP215	Julie Mitchell	ORNL	10,000	7,289	Training An Advanced Model for Structure-Based Proteomics
BIP232	Pin-Kuang Lai	Stevens Institute of Technology	20,000	0	Integrating Molecular Dynamics Simulations and Machine Learning to Accelerate Antibody Development
BIP234	Harel Weinstein	Weill Cornell Medicine	20,000	14	Project BIP109 Stepping up to Frontier
BIP235	Ron Dror	Stanford	30,000	3,783	Advancing the Rational Design of Functionally Selective GPCR-Targeted Drugs
BIP250	Jerry Parks	ORNL	18,000	0	BPGbio
BIP251	Josh Vermaas	NREL, Michigan State University	20,000	0	Bacterial Microcompartment Permeability Assays in the Computational Microscope
CFD154	Spencer Bryngelson	Georgia Tech	10,000	1,197	Accelerated Sub-Grid Multi-Component Flow Physics
CFD165	Federico Municchi	Colorado School of Mines	20,000	0	CFD Simulations of Heat and Mass Transport in Solar Powered Membrane Distillation Systems

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
CFD172	Jonathan MacArt	Notre Dame	15,000	2,747	Frontier: Data Assimilation for Predictive Flow Simulations
CFD173	Yuuichi Asahi	JAERI	30,000	12	Optimization of City Wind Flow Simulation Code on Frontier
CFD174	Eric Johnsen	U. of Michigan	15,000	9,027	Numerical Simulations of Shocks and Interfaces
CFD175	Peter Vincent	Imperial College London	16,550	9	Application of PyFR to Methane Super-Emitter Detection using Frontier
CFD183	Vittorio Michelassi	GE Global Research	20,000	0	Fluent GPU Performance Assessment for High-Efficiency Waste Heat Recovery Axial Expanders
CHM174	Monojoy Goswami	UT, ORNL	5,000	0	Multi-Scale/Multi-Physics Molecular Simulations at the Chemical Sciences Division
CHM181	Kurt Mikkelsen	University of Copenhagen	15,000	185	Massively Parallel GPU-Enabled Cluster Perturbation Methods
CHM191	Andre Severo Pereira Gomes	CNRS	15,000	183	PRECISE
CHM193	Taisung Lee	Rutgers	20,000	0	A Next-Generation Extendable Simulation Environment for Affordable, Accurate, and Efficient Free Energy Simulations
CHM194	Kedan He	Eastern Connecticut State University	1,000	0	Computational Chemistry and Drug Design: Molecular Recognition and Binding Prediction Powered by Machine Learning and Deep Learning Approaches
CHP124	Richard Friesner	Columbia University	15,000	14,377	Quantum Simulations of Photosystem II and Organometallic Enzymes
CLI131	Samuel Hatfield	European Centre for Medium-Range Weather Forecasts	10,000	808	A Baseline for Global Weather and Climate Simulations at 1 KM Resolution
CLI137	Forrest Hoffman	ORNL	10,000	0	Earth System Grid Federation 2 (ESGF2)
CLI138	Moetasim Ashfaq	ORNL	20,000	0	Analytical Frameworks for Sub-Seasonal to Multi-Decadal Climate Predictions and Impact Assessments

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
CLI181	Samuel Hatfield	European Centre for Medium-Range Weather Forecasts	20,000	0	Preparing for INCITE 2024
CMB147	Joseph Oefelein	Georgia Tech	15,000	0	Analysis of Combustion and Wave Dynamics in Rotating Detonation Engines
CMB150	Olivier Desjardins	Cornell	20,000	0	High Fidelity Modeling of Spray Formation and Dispersion
CPH005	Dario Alfe	University College London	20,000	68	New Frontiers for Material Modeling via Machine Learning Techniques with Quantum
CPH154	Alberto Nocera	University of British Columbia	20,000	15	Quantum Annealing Simulation of Spin Glasses Using Large Scale Tensor Network Methods
CSC143	Norbert Podhorszki	ORNL	200,000	0	ADIOS – The Adaptable IO System
CSC380	Bronson Messer	ORNL	355,000	289,488	CAAR for Frontier
CSC524	Van Ngo	ORNL	20,000	1,477	Benchmarking and Development of the MiMiC Framework for Highly Efficient QM/MM Simulations on Leadership Supercomputers
CSC528	Ang Li	PNNL	18,000	4	nwqsim
CSC533	William Severa	Sandia	20,000	0	COINFLIPS
CSC536	Chathika Gunaratne	ORNL	20,000	0	SAGESim – Scalable Agent-based GPU Enabled Simulator
CSC538	Irina Rish	University of Montreal	20,000	1,017	Large-Scale AI Models
CSC546	Guojing Cong	ORNL	20,000	4	Generative Pretraining with Graph Neural Networks For Materials for a Diverse Set of Property Predictions
CSC547	Abhinav Bhatele	LLNL, U. of Maryland	20,000	50	Performance Analysis and Tuning of HPC and AI Applications
CSC549	Dhabaleswar Panda	Ohio State University	20,000	0	MVAPICH
CSC550	Ramakrishnan Kannan	ORNL	15,000	1	gnn1e12
CSC581	Mahantesh Halappanavar	PNNL	10,000	0	ExaGraph

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
ENG123	Ramesh Balakrishnan	Argonne	15,000	6,108	Direct Numerical Simulation of Separated Flow over a Speed Bump at Higher Reynolds Numbers
FUS147	Noah Reddell	HPE	40,000	27	Study of High Energy Density Z-Pinch Plasma Stability by Kinetic Model on GPU
FUS151	Chang Liu	Princeton Plasma Physics Laboratory	15,000	1,188	Hybrid Simulation of Energetic Particles Instabilities and Transport in Fusion Plasmas
FUS152	Colin McNally	General Fusion	20,000	0	Beyond Neoclassical Closures via Kinetic Monte-Carlo Calculations
GEN189	Veronica Melesse Vergara	ORNL	15,000	0	Frontier External Collaboration Project
GEO112	Philip Maechling	USC	10,000	5,214	Extreme-Scale Simulations for Advanced Seismic Ground Motion and Hazard Modeling
GEO146	Kohei Fujita	University of Tokyo	20,000	12,828	Scalable Dynamic Stochastic Earthquake Simulation with Low-Ordered Unstructured Finite Elements on Frontier
GEO147	James McClure	VA Tech	500,000	0	LBPM
GEO148	Jeroen Tromp	Princeton	20,000	528	Global Adjoint Tomography on Frontier
GEO151	Kim Olsen	San Diego State University	35,000	0	Preparing AWP-ODC for Exascale Earthquake Ground Motion Research on Frontier
GEO152	Dalton Lunga	ORNL	20,000	413	Large Vision Models for Geospatial Applications
GEO157	Kohei Fujita	University of Tokyo	40,000	0	Finite-Element Seismic Simulation with Model Optimization on Frontier
GEO158	Kaushik Bhattacharya	Caltech	20,000	0	A Learning-Based Multiscale Model for Reactive Flow in Porous Media
HEP114	Salman Habib	Argonne	100,000	107,012	2.2.3.02 ADSE01-ExaSky: Computing the Sky at Extreme Scales
HEP133	Rajan Gupta	LANL	2,000	2	High Impact Nucleon Matrix Elements Affected by N-Excited States

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
LGT100	Andre Walker-Loud	LBNL	5,000	12,759	The Structure and Interactions of Nucleons from the Standard Model
LGT114	Amy Nicholson	University of North Carolina at Chapel Hill	15,000	3,523	Electromagnetic Corrections to the Nucleon Axial Charge
LGT123	Amy Nicholson	University of North Carolina at Chapel Hill	15,000	0	Electromagnetic Corrections to the Nucleon Axial Charge
LRN031	Pilsun Yoo	ORNL	20,000	39	Inverse design of UV-Vis Property of Molecules Using a Workflow with Machine Learning Methods
LRN032	Jessica Inman	Georgia Tech	12,500	2	Large Scale Formal Verification of Neural Networks
LRN033	Mayur Mudigonda	Vayuh Inc.	20,000	4,471	Toward Foundation Models for Weather Forecasting Using Physics and Deep Learning
LRN036	Dan Lu	ORNL	20,000	2,564	AI Foundational Model for Weather and Climate Project
LRN037	Pei Zhang	ORNL	20,000	0	Multiscale Foundation Models for Physical Systems
LRN039	Allan Grosvenor	Microsurgeonbot	20,000	712	Autonomously Driven Software – Frontier Pursuit of Level 4 Autonomy
MAT226	Jan Michael Carrillo	UT, ORNL	20,000	11,320	Molecular Dynamic Simulations of Amphiphilic Oligomer Membranes: Design Rules Towards Stable Membranes Capable of Learning and Memory
MAT258	Corey Oses	Johns Hopkins University	15,000	687	Startup
MAT259	Rodrigo Freitas	MIT	3,000	93	Phase-Stability of Complex Solid-Solutions During Nonequilibrium Processes
MAT263	Wenzhe Yu	Argonne	20,000	0	GPU Acceleration of Large-Scale Many-Body Perturbation Theory Using OpenMP 5
MAT265	Victor Fung	Georgia Tech	20,000	4	Foundational Graph Neural Network Models for Materials Chemistry

Project ID	PI	Institution	Most recent frontier allocation	Frontier usage	Project name
MAT268	Patxi Fernandez-Zelaia	ORNL	10,000	0	Physics Informed Generative AI for Materials Design
NFI125	Aurel Bulgac	U. of Washington	50,000	11,669	Microscopic Simulations of Real-Time Nuclear Dynamics in the Exascale
PHY129	Alexander Tchekhovskoy	Northwestern University	20,000	0	Simulating Neutron Star Binary Merger Remnant Disks and Tilted Thin Disk
PHY180	Mark van Schilfgaarde	NREL	20,000	0	Green's Function Description of the Electron Phonon Interaction
PHY184	Nicholas Sauter	LBNL	20,000	0	ExaFEL
TUR141	Niclas Jansson	Kungliga Tekniska Högskolan (KTH Royal Institute of Technology)	20,000	0	Extreme-Scale High-Fidelity Turbulence Simulations of Convection
TUR143	Mahendra Verma	IIT Kanpur	15,000	0	Anisotropic Magnetohydrodyamic Turbulence with Applications to Solar Physics