ORNL/SPR-2023/2849

US Department of Energy, Office of Science High Performance Computing Facility Operational Assessment 2021
Oak Ridge Leadership Computing Facility

April 2022

OAK RIDGE
National Laboratory

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

ORNL/SPR-2023/2849

Oak Ridge Leadership Computing Facility

US DEPARTMENT OF ENERGY, OFFICE OF SCIENCE HIGH PERFORMANCE COMPUTING FACILITY OPERATIONAL ASSESSMENT 2021 OAK RIDGE LEADERSHIP COMPUTING FACILITY

Subil Abraham
J. Paul Abston
Ryan M. Adamson
Valentine Anantharaj
Aaron Barlow
Ashley Barker
Tom Beck
Katie L. Bethea
Adam Carlyle
Daniel Dietz
Christopher Fuson
Benjamin Hernandez Arreguin
Jason J. Hill
Ryan Landfield
Don Maxwell
Rachel McDowell
Veronica Melesse Vergara
Bronson Messer
Ross Miller
Sheila Moore
Sarp Oral
Thomas Papatheodore
Suzanne Parete-Koon
Ryan Prout
Sherry Ray
William Renaud
James Rogers
Mallikarjun Shankar
Woong Shin
Scott Simmerman
Betsy Sonewald
Kevin G. Thach
Aristeidis Tsaris
Georgia Tourassi
Justin L. Whitt
Junqi Yin

April 2022

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 v
LIST OF TABLES v
EXECUTIVE SUMMARY vii
1. USER SUPPORT RESULTS 1
1.1 USER SUPPORT METRICS 2
1.1.1 Average Rating across All User Support Questions 3
1.1.2 Improvement on Past Year Unsatisfactory Ratings 3
1.2 PROBLEM RESOLUTION METRICS 5
1.2.1 Problem Resolution Metric Summary 5
1.3 USER SUPPORT AND ENGAGEMENT 6
1.3.1 User Support 6
1.3.2 User Support Roles 7
1.3.3 OLCF User Group and Executive Board 10
1.3.4 Training 11
1.3.5 Community Engagement 15
2. OPERATIONAL PERFORMANCE 17
2.1 RESOURCE AVAILABILITY 17
2.1.1 Scheduled Availability 19
2.1.2 Overall Availability 19
2.1.3 Mean Time to Interrupt 20
2.1.4 Mean Time to Failure 20
2.2 TOTAL SYSTEM UTIIZATION IN 2021 21
2.2.1 Resource Utilization Snapshot 21
2.2.2 Total System Utilization 21
2.3 CAPABILITY UTILIZATION 23
3. ALLOCATION OF RESOURCES 25
3.1 SUMMARY OF ALLOCATION PROGRAMS 25
3.2 FACILITY DIRECTOR'S DISCRETIONARY RESERVE TIME 25
3.3 ALLOCATION PROGRAM CHALLENGES 26
4. OPERATIONAL INNOVATION 29
4.1 OPERATIONAL INNOVATION – TECHNICAL 29
4.1.1 Software Deployment, CI, and E4S 29
4.1.2 Revealing Power, Energy, and Thermal Dynamics of the Summit Supercomputer through Analysis of High-Volume Operational Data 30
4.1.3 Supporting Integrated Data and Workflow Infrastructures 31
4.1.4 User-Managed Software 32
4.1.5 NCI/NIH Whole-Slide Imaging Service on Slate 32
4.2 OPERATIONAL INNOVATION – MANAGEMENT/WORKFORCE 32
4.2.1 Crash Courses for HPC and ML/DL Ecosystem Community Training 32
4.2.2 Winter Classic Invitational Student Cluster Competition 33
4.2.3 Pathways to Supercomputing Initiative 34
4.2.4 OLCF Recruiting, Mentoring, and Workforce Development Efforts 34
4.2.5 Development of a Cross-Facility ASCR Cybersecurity Working Group 37
4.3 POSTDOCTORAL FELLOWS 37
4.3.1 Computational Scientists for Energy, the Environment, and National Security Postdoctoral Program 37
5. RISK MANAGEMENT 45
5.1 RISK MANAGEMENT SUMMARY 45

5.2 MAJOR RISKS TRACKED IN 2021 46
5.3 NEW OR RECHARACTERIZED RISKS SINCE LAST REVIEW 46
5.3.1 Recharacterized Risks 46
5.3.2 New Risks in This Reporting Period 46
5.4 RISKS RETIRED DURING THE CURRENT YEAR 46
5.5 MAJOR RISKS FOR NEXT YEAR 47
5.6 RISKS THAT OCCURRED DURING THE CURRENT YEAR AND THE EFFECTIVENESS OF THEIR MITIGATION 47
6. ENVIRONMENT SAFETY AND HEALTH 49
6.1 FRONTIER INSTALLATION 50
6.2 FRONTIER INJURY 51
7. SECURITY 53
7.1 SUMMARY 54
7.2 SECURITY OPERATIONS 55
7.3 OLCF USER VETTING 56
7.3.1 OLCF Projects 56
7.3.2 OLCF Users 57
8. STRATEGIC RESULTS 59
8.1 SCIENCE HIGHLIGHTS AND ACCOMPLISHMENTS 59
8.1.1 OLCF Publications Report 60
8.1.2 Scientific Accomplishments 60
8.1.3 Science Highlights 61
8.2 RESEARCH ACTIVITIES / VENDOR ENGAGEMENT FOR FUTURE OPERATIONS 69
8.2.1 Summit on Summit 69
8.2.2 Accelerated Data Analytics and Computing (ADAC) Institute 70
8.2.3 Enabling Novel Capability on OLCF Machines 70
8.2.4 ML/DL Acceptance, Comparative analysis 70
8.2.5 Exploring Power, Energy, and Thermal Dynamics of the Summit Supercomputer through Analysis of High-Volume Operational Data 71
8.2.6 Industry Engagement 72
8.3 DOE PROGRAM ENGAGEMENTS / REQUIREMENTS GATHERING 76
8.3.1 Ongoing Engagements with SC Observational Facilities Investigators 76
8.3.2 Ongoing Engagements with Atmospheric Radiation Measurement (ARM) 76
8.3.3 Engagement with the National Institutes of Health (NIH) and the National Cancer Institute (NCI) 76
8.3.4 Engagement with the US Department of Veterans Affairs (VA) 77
8.3.5 Engagement with Air Force Weather (AFW) and National Oceanic and Atmospheric Administration (NOAA) 78
8.3.6 Exascale Computing Project Engagement Summary 80
APPENDIX A. RESPONSES TO RECOMMENDATIONS FROM THE 2020 OPERATIONAL ASSESSMENT REVIEW A-1
APPENDIX B. TRAINING, WORKSHOPS, AND SEMINARS B-1
APPENDIX C. OUTREACH PRODUCTS C-1
APPENDIX D. OPERATIONAL SYSTEMS SUMMARY D-1
APPENDIX E. BUSINESS RESULTS FORMULAS E-1
APPENDIX F. DD PROJECTS ENABLED AT ANY POINT IN CY 2021 ....F-1

LIST OF FIGURES

Figure 2.1. 2021 IBM AC922 resource utilization—Summit node hours by program. 22
Figure 2.2. Summit capability usage by job size bins and project type. 24
Figure 4.1. Scientific software stream framework from incubation to installation. 30
Figure 7.1. OLCF Authority to Operate. 54
Figure 8.1. Optimized geometry of GeSe monolayer. 62
Figure 8.2. Crystal structure of a representative cuprate superconductor and simulated photoemission spectrum. 64
Figure 8.3. The Oleynik team simulated a split elastic-inelastic shock wave moving through a single crystal diamond. 66
Figure 8.4. An ORNL study used a DL language model to sort through billions of chemical sequences to find molecules that might block two of the primary protein components of the coronavirus. 67
Figure 8.5. A visualization of deuterium-tritium density fluctuations in a tokamak driven by turbulence. 69
Figure 8.6. Summit power and energy trends (CY 2020). 71
Figure 8.8. A row of upstream bars produces highly turbulent flow that gets accelerated through a high-pressure turbine blade row and interacts with the blade surface, causing significant temperature variations. 75
Figure D.1. 2021 GPU-enabled usage by program. D-4

LIST OF TABLES

Table 1.1. Annual survey response rate for 2021. 2
Table 1.2. Key survey OLCF user responses for 2021. 2
Table 1.3. User support metrics targets and CY results. 3
Table 1.4. Problem resolution metric summary 6
Table 1.5. OLCF user rating for MyOLCF for 2021. 7
Table 2.1. OLCF Operational performance summary for Summit. 17
Table 2.2. OLCF Operational performance summary for HPSS. 18
Table 2.3. OLCF Operational performance summary for Alpine, the external GPFS file system. 18
Table 2.4 OLCF operational performance summary: Scheduled availability. 19
Table 2.5 OLCF operational performance summary: Overall availability. 20
Table 2.6. OLCF operational performance summary: Mean time to interrupt. 20
Table 2.7. OLCF operational performance summary: Mean time to failure. 20
Table 2.8. The 2021 allocated program performance on Summit. 22
Table 2.9. OLCF capability usage on the IBM AC922 Summit system. 23
Table 5.1. 2021 OLCF major risks. 46
Table 5.2. Risks encountered and effectively mitigated in CY 2021. 47
Table 7.1. Export control review categories for projects. 56
Table 7.2. OLCF Project Category Requirements. 57
Table 8.1 Summary of unique OLCF publications for 2012–2020. 60
Table 8.2. Publications in high-impact journals in 2021. 61
Table 8.3. Listing of OLCF ECP engagement applications, the ECP AD PI, and the OLCF Scientific Engagement liaison. 80
Table B.1. 2021 OLCF training and outreach events. B-1

Table D.1. Summit 2021. .... D-1
Table D.2. OLCF HPC system production dates, 2008–present. D-3
Table F.1. DD projects enabled at any point in CY 2021. .... F-1

Executive Summary

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2022

EXECUTIVE SUMMARY

Oak Ridge National Laboratory's (ORNL's) Leadership Computing Facility (OLCF) continues to surpass its operational target goals of supporting users; delivering fast, reliable computational ecosystems; creating innovative solutions for high-performance computing (HPC) needs; contributing to the community to build the next generation HPC workforce, and managing risks, safety, and security associated with operating some of the most powerful computers in the world. The results can be seen in the cutting-edge science conducted by users and the praise from the research community.

Calendar year (CY) 2021 saw continued excellence in research supported by the OLCF's leadership-class computing resources, including Summit (the nation's most powerful supercomputer), the global scratch file system Alpine, the Scalable Protected Infrastructure (SPI), the Exploratory Visualization Environment for Research in Science and Technology (EVEREST), and the archival mass-storage resource High-Performance Storage System (HPSS). While maintaining access and exceptional user support for Summit, the OLCF continued to make progress on the installation and deployment of Frontier, which will be the nation's first exascale system when it comes online at the start of CY 2023. Users have already begun running and optimizing scientific codes on Crusher, the OLCF test and development system equipped with Frontier's architecture.

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, and at this time there are no high-priority operational risks. Similarly, ORNL and the OLCF were committed to operating under the US Department of Energy's (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.

CY 2021 was filled with outstanding results and accomplishments, including a very high rating from users on overall satisfaction for the eighth consecutive year; a tremendous number of node hours delivered to 1,671 researchers on Summit; and the successful delivery of 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), Advanced Scientific Computing Research Leadership Computing Challenge (ALCC), and Director's Discretionary (DD) programs, respectively (Section 2). COVID-19 research remained a focus in 2021, and the ALCC and DD programs allocated over 1 million Summit hours to the COVID-19 High Performance Computing Consortium. These accomplishments, coupled with the high utilization rates (i.e., overall and capability usage), represent the fulfillment of the promise of leadership-class machines: efficient facilitation of leadership-class computational applications. Table ES.1

summarizes the 2021 OLCF metric targets and the associated results. More information can be found in Section 2 for each OCLF resource.

Table ES.1. 2020 OLCF metric summary.

Metric Description	CY 2021 Target	CY 2021 Actual
Overall OLCF score on the user survey will be 3.5 based on a statistically meaningful sample.	3.5	4.6
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%	96%
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%	50.68%
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%	99.71%
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%	97.88%
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%	99.07%
OA, archive storage: Sustain availability to users measured as a percentage of maximum possible.	90%	97.56%
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%	99.88%
SA, High Performance Storage System: Sustain scheduled availability to users measured as a percentage of maximum possible scheduled.	95%	99.96%

The scientific accomplishments of OLCF users are a strong indication of long-term operational success, with publications this year in such high-impact journals as *Nature, Nature Communications, Nature Physics, Advanced Materials, ACS Nano, Accounts of Chemical Research,* and *Chemical Reviews.* Crucial domain-specific discoveries facilitated by resources at the OLCF are described in the *High Performance Computing Facility Operational Assessment 2021 Oak Ridge Leadership Computing Facility* (OAR) Section 8. For example, researchers used Summit to simulate the optical properties of the nanomaterial germanium selenide. These simulations allow the team to make more confident predictions about similar or more complex materials that might have applications in photodetectors, gas sensors, and lithium-ion batteries.

The OLCF supported scientific accomplishments for a broad community of researchers in 2021, from traditional modeling and simulation projects to studies exploiting artificial intelligence (AI), machine learning (ML), and big data techniques. Twelve finalist teams, including the four winning teams, used Summit to conduct their research in the competition for the prestigious Gordon Bell Prize. The Gordon Bell Prize is awarded each year at the International Conference for High Performance Computing, Networking, Storage, and Analysis to recognize researchers who have made significant strides toward applying HPC systems to scientific applications. Another four teams, including the winning team, used Summit for research that was nominated for the Gordon Bell Prize for HPC-based COVID-19 research.

OLCF systems continue to support data-intensive science including deep learning (DL). For example, an ORNL-led team used a DL language model called *Bidirectional Encoder Representations from Transformers*, or BERT, to streamline the search for potential COVID-19 treatments. Researchers used Summit to pretrain BERT on a dataset of 9.6 billion molecules, which allowed the model to sort through billions of chemical sequences to find molecules that might block two of the primary protein components of SARS-CoV-2 in a matter of hours. Previously, this work might have taken years. This study was nominated for the Association of Computing Machinery Gordon Bell Special Prize for High Performance Computing–Based COVID-19 Research.

Another team, led by scientists at the University of South Florida, used machine-learned descriptions of interatomic interactions on Summit to model more than 1 billion carbon atoms at quantum accuracy and observe how diamonds behave under extreme pressure and temperature. The study will help scientists better understand how carbon behaves under extreme conditions, which is crucial for inertial confinement fusion and could help uncover the internal structure of carbon-rich planets like Uranus.

Finally, physicists from Lawrence Berkeley National Laboratory modeled the interactions between negatively charged electron particles in a material and the interactions between electrons and phonons, the smallest unit of vibrational energy in a material. The researchers used Summit to model millions of particle states to produce one of the team's largest calculations of copper-based superconductors to date, which could lead to a better understanding of the mechanisms of a unique family of more efficient copper-based superconductors.
OLCF operations remained a focus in CY21, and the facility has been a leader in the creation and development of tools and policies dedicated to advancing computational science and

disseminating knowledge through collaborations, workshops, and other engagement activities. In collaboration with ALCF and NERSC, the OLCF software teams standardized software packaging, testing, and deployment frameworks, continued innovative work on HPC data center cooling using energy-efficient cooling methods, and developed new software such as DataFlow, a software stack that provides secure access to remote data services for air-gapped instruments.

The OLCF continued their strong engagement within the community to build the next generation HPC workforce and pioneered the "Hands-on with Summit" training to reach audiences that are often underrepresented in the HPC community, both within and outside ORNL. OLCF staff were prominent at several conferences and events, such as the International Conference for High Performance Computing, Networking, Storage and Analysis, the PEARC Student volunteer program, and the Tapia Diversity in Computing Conference, where they led workshops and introductory trainings for students new to HPC.

In CY21, the OLCF continued to support and strengthen several strategic partnerships with other DOE Office of Science programs and Federal agencies. These partnerships, which include ongoing collaboration and system support for Atmospheric Radiation Measurement, the National Cancer Institute, the US Department of Veterans Affairs, Air Force Weather, and the National Oceanic and Atmospheric Association, increase community engagement, contribute to the development of new software and capabilities, and support overall operational innovation.

The successful deployment and operation of a succession of leadership-class resources is the result of the extraordinary work by the OLCF staff in supporting the most capable HPC user facility in the world. The OLCF staff are pivotal to identifying, developing, and deploying the innovative processes and technologies that support the advancement of science by the OLCF users and benefit other HPC facilities around the world.

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 US Department of Energy (DOE) ORNL Site Office and the ASCR Program office, monthly highlight reports, quarterly reports, the annual OAR, an annual "Budget Deep Dive," an annual independent project review, and the OLCF annual report. 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 C for a list of science highlights submitted to ASCR.

As a result of ongoing restrictions caused by the COVID-19 pandemic, on-site tours, meetings, conferences, and seminars were limited in 2021. Workshops, meetings, and events were held virtually or with a hybrid model following local and facility pandemic guidelines. Regular facility tours continued to be offered through the virtual facility tour via the Matterport software system.

ES.1.2 Communication with the User Community

The OCLF's communications with users are tailored to the objectives of relating science results to the large community and helping users more efficiently and effectively use OLCF systems. The OLCF offers many training and educational opportunities throughout the year for current facility users and the next generation of HPC users outlined in Section 1.

The impact of OLCF communications is assessed as part of an annual user survey. In the 2021 annual user survey, OLCF communications received a mean rating for users' overall satisfaction of 4.4. The OLCF uses a variety of methods to communicate with users, including the following:

- Weekly email message
- Welcome packet
- 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
- Social networking
- Annual OLCF User Meeting
- Targeted training events (i.e., GPU hackathons or tutorials)

ES.2 Summary of 2021 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 2021. The 2021 metrics, target values, and actual results as of December 31, 2021, are noted throughout this report. The OLCF exceeded all agreed-upon metric targets.

ES.3 RESPONSES TO RECOMMENDATIONS FROM THE 2020 OPERATIONAL ASSESSMENT REVIEW

In April 2021, the operational activities of the OLCF for CY 2020 were presented to the DOE sponsor. The review committee of that report identified no recommendations.

User Support Results

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2022

1. USER SUPPORT RESULTS

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

OLCF RESPONSE: Yes. The Oak Ridge Leadership Computing Facility (OLCF) at the US Department of Energy's (DOE's) Oak Ridge National Laboratory (ORNL) supported 1,671 users and 535 projects (including Quantum) in CY 2021. 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 2021 was 4.6. Overall ratings for the OLCF were positive: 96% of all survey respondents reported being "satisfied" or "very satisfied" with the OLCF.

The OLCF measures its performance 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 2021: 96% 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) to the next generation of researchers.

USER SUPPORT RESULTS SUMMARY

The OLCF's user support model comprises customer support interfaces, including user satisfaction surveys, formal problem-resolution mechanisms, user assistance analysts, and scientific and data liaisons; multiple channels for stakeholder communication, including the OLCF User Council; 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 using the annual OLCF user survey.

To promote continual improvement at the OLCF, users are sent surveys soliciting their feedback regarding support services and their experience as users of the facility. The 2021 survey was launched on October 07, 2021 and remained open for participation through November 16, 2021. The survey was sent to 1,443 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, which includes the Exascale Computing Project (ECP) projects, who logged into an OLCF system between January 01, 2021, and September 30, 2021. Even though the ECP allocations come from the OLCF DD program, we once again tracked those responses separately from the DD responses. OLCF staff members were excluded from participation. A total of 706 users completed the survey, for an overall response rate of 49% as shown in Table 1.1. The results of the 2021 survey can be found on the OLCF website at https://www.olcf.ornl.gov/wp-content/uploads/2022-olcf-user-survey-report.pdf.

Table 1.1. Annual survey response rate for 2021.

Survey Response	2020 target	2020 actual	2021 target	2021 actual
Number of users the survey was sent to	N/A	1,260	N/A	1,443
Total number of respondents	N/A	688	N/A	706
Percent responding	N/A	55%	N/A	49%

Users were asked to rate their satisfaction on a 5-point scale, where 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 3.5/5.0 as satisfactory. The effectiveness of the processes for supporting customers, resolving problems, and conducting outreach is in part measured by the key survey responses for User Support in Table 1.2.

Table 1.2. Key survey OLCF user responses for 2021.

Survey question	2020 target	2020 actual	2021 target	2021 actual
Overall OLCF satisfaction score on the user survey	3.5/5.0	4.6/5.0	3.5/5.0	4.6/5.0
Overall Satisfaction with OLCF Support	3.5/5.0	4.4/5.0	3.5/5.0	4.6/5.0
Overall Satisfaction with OLCF Services	3.5/5.0	4.4/5.0	3.5/5.0	4.4/5.0
Overall Satisfaction with the Website	3.5/5.0	4.4/5.0	3.5/5.0	4.6/5.0
Overall Satisfaction with Communications	3.5/5.0	4.6/5.0	3.5/5.0	4.4/5.0
Overall Satisfaction with Problem Resolution	3.5/5.0	4.5/5.0	3.5/5.0	4.5/5.0
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 2020 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 2021 survey.

1.1 USER SUPPORT METRICS

The OLCF exceeded all user support metrics for 2021. The OLCF metric targets and actual results by CY for user support are shown in Table 1.3.

Table 1.3. User support metrics targets and CY results.

Survey area	CY 2020 target	CY 2020 actual	CY 2021 target	CY 2021 actual
Overall OLCF satisfaction rating	3.5/5.0	4.6/5.0	3.5/5.0	4.6/5.0
Average of all user support services ratings	3.5/5.0	4.4/5.0	3.5/5.0	4.5/5.0

1.1.1 Average Rating across All User Support Questions

The calculated mean of answers to the user support services–specific questions on the 2021 survey was 4.5/5.0, indicating that the OLCF exceeded the 2021 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 available. Respondents were highly satisfied with training (97%), the projects and accounts team (96%), user assistance (96%), issue response (95%), and documentation (95%). Included below are select open-ended responses to "What are the best qualities of the OLCF?" that highlight various aspects of user support.

"In my experience the best qualities of OLCF are:

- "Great computing resources coupled with well-written documentation for proper usage, and wonderful user training events throughout the year."

- "OLCF's user guides and documentations are among the best I have used. The machines themselves are of course also amazing resources."

- "The OLCF facilities and support staff are some of the best that I have had the pleasure of interacting with. It is a tremendous asset to have access to state-of-the-art systems, and I cannot commend OLCF enough."

- "OLCF is a perfect service, and I am grateful for the privilege of access. A few words about the best qualities that come to my mind: high performance, easy use, perfect documentation, excellent support, training slides and videos, wide range of the latest software."

- "The OLCF is a vital facility for the fostering and support of the highest performance computing available to scientists in many fields. From our perspective, the successful and technically superior drive of OLCF to install and maintain the most advanced and powerful compute platforms has made possible otherwise unachievable discoveries and findings. The attention to user needs, the breadth of science supported, and the responsible and caring administration are major factor in the success of this special asset in the service of the best science."

1.1.2 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 (3.5/5.0) in the previous year's survey. All questions scored above 3.5 on both the 2020 and 2021 surveys. Although we had no results that scored below satisfactory on the 2021 survey, we did a thorough review of the survey and identified areas in which we could improve or add new services to enhance the user experience at the OLCF. In some cases, the issue has already been addressed or a solution is in the works and forthcoming in short order.

- User cgroup Query Tool

- Summit's login nodes are resources shared by all users of the system. Runaway scripts or large workflows can impact all users of the resource. Following testing to find limits that would not impact most workflows, the OLCF implemented cgroups on Summit's login nodes. Each user is now limited to a set number of hardware threads and GPUs, amount of available memory, and total time spent using the CPU. The use of cgroups has helped to ensure availability of resources to all users, but for users who reach the cgroup limits, it was difficult to view which limit impacted their workflow. The OLCF created a tool that gathers log files of cgroup impact and a front-end tool that users can run to view impact. Users can now run the command "check_cgroup_user" on the Summit login nodes to view whether they have been impacted by cgroups on any login node.

- Improve data analytics and visualization support
  - The OLCF improved documentation in the "ML/DL & Data Analytics" section of the User Guide and created a new "Visualization Tools" area that includes documentation on how to use specific data analytics and visualization tools on Andes and Summit. The pages in the new section outline how to install, connect, and run the visualization tools, as well as provide example scripts for users to run. The new section can be found here: https://docs.olcf.ornl.gov/software/viz_tools/index.html
  - The OLCF also developed a "Data Visualization and Analytics" training series that started in December 2021. The "Visualization Tools Overview" event in December was attended by approximately 70 people and provided an overview of popular visualization tools we offer on our systems. In 2022, this series has four hands-on training events planned from July through October and two additional "user calls" connected to the series will occur in September and August. The hands-on events cover how to efficiently use VisIt and ParaView on our systems, along with how to analyze data using Jupyter and NVIDIA Rapids. The user calls will cover the smaller community (not as widespread) visualization tools available on our systems that have been requested by users.

- Viewing Support Tickets in MyOLCF
  - The OLCF utilizes the Jira ticketing software to help gather, track, and respond to user inquiries. Historically, communication through the ticket system has occurred through email. Recently, the OLCF added the ability to see support tickets and communication within the tickets on the MyOLCF site. The feature is live but still in beta testing. We expect the feature to improve users' ability to track the state of their submitted tickets as well as see the full action history for each issue. The feature will continue to be improved and will enter full production in 2022.

- SecurID Video Verifications
  - OLCF systems in the moderate enclave require two-factor authentication using RSA SecurID tokens. Each remote OLCF user is shipped a physical RSA token. The tokens are shipped disabled. Enabling the token requires identity proofing to authenticate a user's identity and possession of the token. Historically, most validations were performed by a notary since most of the OLCF users are offsite. For some, locating a notary could be time-consuming, and the process could be further complicated by country or state policy. Periods of restrictions over the past two years also complicated access to a notary. The OLCF began using verification over video calls—historically used for cases where a notary was not an option—more often over the past two years. Because video verifications allow approved staff to complete the verification over a video call in a matter of minutes, they provide a quick and easy option for users. During the past year, the OLCF has added the ability for users to schedule video verifications within MyOLCF. The new feature allows users to schedule calls in 10 minute blocks throughout the business day at a time most convenient to the user. The option has

become popular, and video verifications have become the primary verification process, thus saving the user effort and decreasing the amount of time it takes to provide a user with access to the OLCF resources.

- Improvements to software efforts
  - The survey feedback identified software documentation as one of the areas where users would like to see improvements. The OLCF engaged the OLCF User Group Executive Board to solicit feedback on the current "Software" page and has begun updating the various sections at https://docs.olcf.ornl.gov/software/index.html. The "Profiling Tools" section was updated to include information on using the latest release of the Vampir software on OLCF systems. In addition, the OLCF removed outdated software pages and plans to continue expanding the content provided in the "Software" section.
  - In response to user requests, the OLCF began providing Spack environments for Spock and Summit to allow users to build software on top of the facility-provided Spack instance. We also developed a new section on the OLCF documentation site to detail the steps needed to leverage the facility-provided software using Spack. The documentation is available at: https://docs.olcf.ornl.gov/software/spack_env/index.html.
  - In 2021, the OLCF deployed the Extreme-scale Scientific Software Stack (E4S) on Summit and Spock. The E4S stack provides a wide range of scientific libraries that are often requested by the OLCF and ECP user communities. To provide details on how to leverage the E4S stack, the OLCF also added a new section to the OLCF documentation: https://docs.olcf.ornl.gov/software/e4s.html.
  - The OLCF created a new "Python on OLCF Systems" page on the OLCF documentation website that covers how to access and run Python on both Andes and Summit. We also provide a best practices guide with a short tutorial of how to use Anaconda on our systems. The Python page also includes step-by-step guides on how to install specific Python packages that users commonly request via tickets (mpi4py, CuPy, and h5py). The new page is located here: https://docs.olcf.ornl.gov/software/python/index.html.

1.2 PROBLEM RESOLUTION METRICS

The following operational assessment review 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.

- Average satisfaction ratings for questions on the user survey related to problem resolution are satisfactory or better.

1.2.1 Problem Resolution Metric Summary

In most instances, the OLCF resolves reported problems directly, including 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 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.

During 2021, the OLCF changed its problem tracking software from Request Tracker (RT) to Jira. Like RT, Jira is used to track problem reports and to ensure response goals are met or exceeded. Users can submit tickets in a variety of ways, including email, telephone, and an online request form. Email remains the most common method. Ticket metrics are gathered from both RT and Jira data with logic in place to handle the small number of tickets that were transitioned from RT to Jira. During CY 2021, the OLCF issued 3,130 tickets in response to user inquiries. The OLCF resolved 96% of issues within 3 business days, as shown in Table 1.4, which is an increase of 6% over the previous year. The increase can be attributed in part to a maturing Summit system.

More than three-quarters (77%) of the survey respondents submitted between one and five queries to the OLCF (via phone or email) in 2021. 95% of respondents were satisfied or very satisfied with the OLCF's response to reported issues, with similarly high ratings for the timeliness of responses to reported issues (96%) and the quality of technical advice given (94%).

Table 1.4. Problem resolution metric summary

Survey Area	CY 2020		CY 2021	
	Target	Actual	Target	Actual
Percentage of problems addressed in 3 business days	80%	90%	80%	96%
Average of problem resolution ratings	3.5/5.0	4.6	3.5/5.0	4.6

1.3 USER SUPPORT AND ENGAGEMENT

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

- a user support staff made up of account management liaisons, User Assistance and Outreach (UAO) analysts, Scientific Computing Group (SciComp) liaisons, data liaisons, and visualization liaisons;
- multiple vehicles to communicate with users, sponsors, and vendors;
- developing and delivering training to current and future users; and
- strong outreach to engage the next generation of HPC users, the external media, and the public.

1.3.1 User Support

The OLCF recognizes that users of HPC facilities have a wide range of needs requiring diverse solutions, from immediate, short-term, trouble ticket–oriented support—such as assistance with debugging and optimizing code—to more in-depth support requiring total immersion in and collaboration on projects. The facility provides complementary user support vehicles that include user assistance and outreach staff; liaisons in respective scientific, data, and visualization areas; and computer scientists who assist with issues surrounding the programming environments and tools. The following sections detail some of the high-level support activities during CY 2021 and the specific OLCF staff resources available to assist users.

1.3.2 User Support Roles

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 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 2021 that helped improve the overall user experience are provided below, although some of them are very much behind the scenes.

1.3.2.1 MyOLCF

The OLCF provides users with a recently launched web-based self-service application called myOLCF (https://my.olcf.ornl.gov) that offers PIs and project members timely, accurate data to empower decision making with respect to OLCF projects and self-service tools for project administration. To accomplish this, 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 center-mandated 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 monthly on average.

This was the first year that we collected information in the survey on MyOLCF. Since the application was new this year, users were asked to rate their agreement with three statements about the application, rather than rate their satisfaction or dissatisfaction with the service directly (See Table 1.5). Respondents agreed that the application has provided them with more information about their project's utilization and that it has made the account process easier.

Table 1.5. OLCF user rating for MyOLCF for 2021.

MyOLCF Statements	% Agree
MyOLCF has made getting and/or renewing my account easier	85%
MyOLCF has provided me with improved information about my project's utilization	87%
MyOLCF is easy to navigate	86%

Twenty-two versions of myOLCF were deployed to production in 2021, adding many new features and improvements, including, but not limited to the following:

- Ability to check status of project membership applications
- Simple PI approvals for project membership applicants
- User-centric help ticket visibility
- Streamlined account and project application forms

- Streamlined account renewal approach with auto-populated data, making the account renewal process much less burdensome

- RSA video verification scheduling system, saving users time and effort to get their token enabled

1.3.2.2 RATS CRM

The center's customer relationship management software, called RATS CRM, is under continuous development. Thirty-six versions of the software were deployed in 2021, adding many new features and improvements.

- Security enhancements to support International Traffic in Arms Regulations (ITAR) and Health Insurance Portability and Accountability Act of 1996 (HIPAA) projects

- Redesign of scheduler configuration datasets for Frontier, the OLCF's exascale computer

- New APIs for help ticket integrations

- Support for user-managed software (UMS) projects (see Section 4.1.4 for more information about this new initiative)

1.3.2.3 Preparing Users for Frontier

In 2021, the OLCF installed and deployed two systems to help users prepare their applications and workflows for the upcoming Frontier supercomputer. The first system, Spock, is the early access system with 36 nodes, each with 1 AMD EPYC CPU and 4 AMD Instinct MI100 GPUs. The second system, Crusher, is a test and development system (TDS) with 192 nodes, each with 1 AMD EPYC CPU and 4 AMD Instinct MI250X GPUs.

Spock's hardware is one generation removed from what will be available for users on Frontier and will allow users to port their applications to an AMD GPU-based ecosystem. Spock was opened to Exascale Computing Project (ECP) and the Center for Accelerated Application Readiness (CAAR) users on May 17, 2021. To help users transition from Summit to Spock, the OLCF developed the Spock quick-start guide (https://docs.olcf.ornl.gov/systems/spock_quick_start_guide.html) and held the Spock Training session on May 20, 2021. The session provided users with an introduction to the HPE/Cray environment and the AMD GPU ecosystem.

Crusher's compute nodes have the same architecture as what will be available on Frontier. Crusher will allow users to gain experience using the AMD Instinct MI250X GPU and optimize their applications in preparation for Frontier. The OLCF enabled the first wave of early access users from ECP and CAAR on November 29, 2021 and developed the Crusher quick-start guide, which was made available to early users in 2021 and has since been published to the OLCF documentation site (https://docs.olcf.ornl.gov/systems/crusher_quick_start_guide.html). In addition, in collaboration with partner vendors, the OLCF scheduled a workshop and two hackathons in 2022 to give users the opportunity to learn about the system and work alongside OLCF and vendor mentors to optimize their applications for Crusher.

Due to the bleeding-edge nature of both Spock and Crusher, it is expected that these systems will potentially exhibit more issues than what we expect to see from a production system. To help manage the evolving state of the software on these systems, the OLCF developed a mechanism that allows us to integrate JIRA issues with our public OLCF documentation site. As a result, we can publish issues identified on the system during staff testing, or received via user reports, automatically. These issues are tracked via JIRA and get automatically published and updated to the quick-start guide corresponding to the system in which the issue was identified. This in-house–developed mechanism has reduced the overhead of maintaining a separate list of issues and has simplified the process of communicating issues with our users.

During CY 2021, the OLCF issued 196 tickets in response to Spock user queries and 23 tickets in response to Crusher user queries.

1.3.2.4 OLCF Quantum Computing User Program

In 2021, we continued to develop the OLCF's strategic capability to provide OLCF users with access to state-of-the-art quantum computing resources for purposes of discovery and innovation in scientific computing applications, given the ever-increasing demand. The program provides users with the opportunity to become familiar with the unique aspects and challenges of quantum computing, as well as to implement and test quantum algorithms on the available systems. Users can explore prospective computational research applications and potentially accelerate existing scientific applications using quantum processors and architectures. Research projects supported include work from every program of DOE's Office of Science—including Advanced Scientific Computing, Basic Energy Science, Biological Environmental Research, High-Energy Physics, Fusion Energy Science, and Nuclear Physics, among others. This continuing effort to support and expand the OLCF Quantum Computing User Program (QCUP) required a team effort consisting of significant contributions from staff in many different areas inside and outside of the OLCF.

The team renewed contracts with two quantum computing hardware vendors (IBM, Rigetti Computing), negotiated a new contract with a new hardware vendor (Honeywell/Quantinuum), and is currently evaluating three new hardware vendors (ColdQuanta, QuEra, and IonQ) to provide the quantum computing resources necessary to meet the needs of the user program. Each contract was a significant effort: the negotiations addressed user agreements, resource availability, pricing, and policies. In addition, the team worked with each of these vendors to maintain and establish the multi-step processes required for issuing user and project accounts, monitoring allocations, enforcing user agreements, providing user support, tracking, and reporting, among other tasks. An important point is that the team also worked to integrate the QCUP proposal and merit review processes into the standard OLCF procedures by which researchers submit proposals for access to computing resources. The team also worked with the vendors to establish and maintain the user support model, whereby QCUP users request help and report issues with the resources. As each vendor has very specific requirements for both access and support; there was no easy one-size-fits-all solution when establishing these processes. The team maintained and expanded communication mechanisms with users to inform them of system downtimes, changes, training events, etc. To accommodate our increasing user base, and to provide the most effective user experience possible, the QCUP program expanded the operations teams by onboarding new User Assistance members to strengthen our quantum computing user support capability and to initiate a quantum science engagement objective, as well as by adding a subject-matter expert to that effort. The QCUP website, (https://www.olcf.ornl.gov/olcf-resources/compute-systems/quantum-computing-user-program/quantum-computing-user-support-documentation/) was maintained and updated throughout the year, containing background on the priorities of the QCUP program, FAQs, QCUP event info, and highlighted publications. Additionally, the website seeks to provide user documentation and training materials to help approved users access and utilize the quantum computing resources supplied by our current vendors. The effort to provide a quick-start guide has expanded into a full-fledged quantum computing user guide, a quantum computing crash-course training resource (in development), as well the quantum science engagement content (in development)—all of which reflect the unique aspects and services of each hardware vendor. Additionally, the website was updated regularly to reflect the changing capabilities and offerings of our vendor partners. Through the partnership between QCUP and the vendor partners, users were invited to participate in several quantum training events. For example, in January, users were invited to participate in Xanadu's Quantum Machine Learning Hackathon, QHack 2021 (Qhack.ai). In April, OLCF hosted the 2021 OLCF User Forum, featuring a tabletop session "Quantum at the OLCF." In July, users were invited to attend IBM Quantum's Qiskit Global Summer School 2021, as well as numerous other IBM Quantum webinars and workshops throughout the year.

The OLCF supported 79 total QCUP projects and 278 QCUP users in 2021 and tracked 63 publications resulting from the use of these resources.

1.3.2.5 User Engagement: Primary ExaSky OLCF Development and Contribution

The ExaSky project aims to produce science to solve challenging cosmological problems by producing a suite of very large cosmological simulations run with the HACC cosmological simulation code. ExaSky will ultimately return a set of large-scale cosmological simulations supported by exascale computing, producing theory for future wide-field sky surveys to better understand the large-scale structure of the universe.

At the moment, we are approaching this project on multiple levels. HACC has two branches that need to perform well: the gravity-only code and a code that includes hydrodynamics.

During 2021, the HACC code was in preparation to produce cosmological simulations of different implementations: (1) large-volume, high-mass, gravity-only simulations; (2) large-volume, high-mass hydrodynamic simulations; (3) and small-volume, very high–mass hydrodynamic simulations. The HACC code is hybrid MPI–OpenMP and does not have major external library dependencies.

Accomplished Tasks:

- N-body gravity-only solver for cosmological simulations at scale on CPU/GPU systems
- Optimized performance of the Lagrangian hydrodynamics method, CRK-SPH, in HACC for GPU systems
- Data reduction capability for cosmological simulations that reduces storage and IO requirements by factors ranging from ~5 to on the order of ~100

OLCF Science Engagement staff helped to create a mini-app to run tests for memory allocation optimization of the HACC code. The mini-app was created to address the memory allocation usage per node. We wanted to achieve maximum usage of memory to up to 95% on the Spock system. We worked on issues that seemed to appear in cases when we used ~10–30 GB. For such allocations in memory, the test runs became slow and eventually crashed. The mini-app used 8 ranks per node. The issue was persisting and was recently fixed in the Crusher system.

The last pre-exascale HACC/Nyx code releases include the efforts of the team to incorporate subgrid modeling into the main codes, add CosmoTools capabilities, and a final report on the verification activity across the two codes. The OLCF's efforts were focused on running tests for performance profiling and optimization across the Center of Excellence (COE) test machines and the OLCF system Spock. The HACC and Nyx codes seemed to perform successfully and with improved performance on these systems.

1.3.3 OLCF User Group and Executive Board

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. Conference Calls are held each month, with the exception of June and November. These webinars are attended by representatives of several groups within the OLCF and provide users with the opportunity to interact with OLCF staff. During the webinars, OLCF staff also provide updated information on upcoming system events and training events. A central focus of each User Conference Call is a training topic. Training topics in 2021 are listed in Appendix B. The 10 User Conference Calls in 2021 had a total of 945 attendees, nearly doubling 2020's total of 502 attendees.

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 such that three new members are elected each year. Additionally, an outgoing chair remains on the board for 1 year as an *ex officio* member if the term as chair is their third year on the board. In 2021, Eric Nielsen of NASA replaced Sarat Sreepathi of ORNL as chair. André Walker-Loud was elected as vice chair and will become chair for the 2022–23 term. Evan Schneider, Sara Isbill, and Steven Gottlieb began new 3-year terms, which will conclude in 2024. The current Executive Board is listed at https://www.olcf.ornl.gov/about-olcf/oug/.

The OLCF hosted its annual User Meeting on June 22–24, 2021; it was an online/virtual event. The first day's morning session featured talks delivered by OLCF staff members, including updates on Summit, an overview of the OLCF Training program, updates on support for Jupyter Notebooks, and a virtual facility tour. The afternoon session featured talks from three users highlighting their work at the OLCF. As with Day 1, Day 2's morning session began with talks from OLCF staff, including a review of the 2020 User Survey in the morning and four additional user-led talks in the afternoon. Day 3 looked forward to Frontier, with reviews of the timeline, tips for developing for Frontier, user tools, and a virtual tour highlighting facility improvements to support Frontier. The afternoon contained user talks related to preparing for Frontier.

Each day of the meeting also included virtual "tabletop" breakout discussions between the morning and afternoon meeting sessions. These breakout sessions provided users with the opportunity to meet with center experts to discuss a variety of topics. Several breakout sessions focused on storage resources, including node-local storage, the center-wide filesystem, and HPSS as well as data transfer. Several sessions had a focus on optimizing use of the systems, including discussions of CPU/GPU/memory affinity on Summit, Jupyter Notebooks, performance measurement, ML/DL at scale, containers, and visualization/data analytics. A session gave the opportunity to learn about/discuss OLCF training opportunities. Others gave users the opportunity to further discuss a forthcoming Frontier system, the Quantum User program. Finally, a session was held to give advice on preparing proposals for INCITE and Director's Discretion projects.

A total of 193 people attended the 2021 User Meeting, down very slightly from last year's 199. The agenda for the user meeting, including links to slide decks and recording of various sessions, is available at https://www.olcf.ornl.gov/calendar/2021-olcf-user-meeting/.

1.3.4 Training

The OLCF training program serves to educate our user community in general HPC training as well as special topics needed to fully leverage the facility's cutting-edge HPC resources. The OLCF offers training in the form of workshops, webinars, tutorials, seminars, and hackathons. In most cases, the training events are recorded, and the slides, recordings, and hands-on materials are made available to our users through the OLCF training archive: https://docs.olcf.ornl.gov/training/training_archive.html.

Survey respondents rated training 4.6/5.0, and 97% were highly satisfied with training overall. The highest rated aspect of training was the usefulness of the online training archive (98%), whereas the lowest rated aspect was the breadth of training events offered (92%).

Again, in 2021, the global COVID-19 pandemic required the OLCF adapt their training program to a fully virtual format. Message board tools such as Slack were used for each event to maintain communication between users and trainers throughout the event. One of the advantages of purely virtual training is that it was logistically easier for multiple laboratories and vendors to co-organize and contribute to training events. Following the trend from last year, we continued to increase collaboration with staff from other DOE laboratories and our vendors to organize and present training events.

In 2021, the OLCF facilitated or collaborated in several hackathons, including seven multi-day virtual NVIDIA GPU hackathons, two virtual OpenMP ECP SOLLVE hackathons, and a virtual NCI/DOE Pilot 3 Hackathon. The facility also hosted or contributed to several training series, including a three-part CUDA training series in collaboration with NVIDIA and NERSC, ten OLCF user conference calls/webinars, twelve HPC Best Practices-ECP (IDEAS-ECP) webinars, and four Hands-On Summit outreach events. In addition, the OLCF delivered a number of individual virtual trainings and workshops, including a two-day virtual HPCToolKit for GPU Tutorial in collaboration with NERSC; a three-day Heterogenous-computer Interface for Portability (HIP) Training workshop in collaboration with AMD; a two-day Kokkos training in collaboration with Kokkos developers; a four-day CMAKE training in collaboration with Kitware, ALCF, and NERSC; a two-day OpenMP Offloading workshop in collaboration with NERSC; a two-day Deep Learning at Scale Training; a two-day Score-P and Vampir Performance Analysis workshop in collaboration with Vampir developers; and an Introduction to Leadership Computing workshop aimed at US government agencies.

See Appendix B for a complete summary of these events.

Some notable events are highlighted in Sections 1.3.4.1–1.3.4.6.

1.3.4.1 The OLCF Slate Hackathon and Workshop

The OLCF hosted its first ever Slate Hackathon for users of the OLCF's Slate Platform, a resource that provides container orchestration services and gives users the opportunity to run specialized tools and workflows that support computational campaigns.

More than 30 people attended the event, which targeted users who are currently or intend to take advantage of Slate's cluster compute systems.

The OLCF developed training materials for the hackathon, offering self-paced, hands-on, "run-this-command" modules on Github. Users could jump in using existing allocations on Slate or request an allocation specifically for the hackathon.

Gaurab KC and Sarat Sreepathi from the Computational Earth Sciences Group at ORNL attended the hackathon to learn how they might augment the capabilities of their monitoring framework, Performance Analytics for Computational Experiments (PACE) (https://climatemodeling.science.energy.gov/technical-highlights/performance-analytics-computational-experiments-pace), which summarizes performance data from computational runs. The two are members of the project that develops the Energy Exascale Earth System Model (E3SM), a high-resolution coupled-Earth system model designed to address energy-related science challenges of national interest while effectively using DOE supercomputers. Sreepathi and KC developed the PACE framework to aggregate performance data collected from E3SM experiments from various supercomputers to derive insights and identify bottlenecks and targets for performance engineering and optimization.

“We have now implemented an automatic upload capability using the Jenkins server on Slate that tracks and exports the computational climate experiments on Summit into the PACE database,” KC said. “With this capability, domain scientists can access the PACE web portal to view completed experiments, examine model configurations, and look at the performance data.”

Event page: https://www.olcf.ornl.gov/calendar/olcf-slate-hackathon/

1.3.4.2 2021 GPU Hackathon Series

In 2021, the OLCF and its partners continued delivering their annual (virtual) GPU hackathon series. As with the in-person format, these hackathons are multi-day coding events during which teams of developers prepare their own applications to run on GPUs or focus on optimizing their applications that already run on GPUs. Each team typically consists of three or more developers of an application and two mentors (identified by the organizers) with GPU programming expertise. These hackathons offer a unique opportunity for application teams to set aside time for development, surround themselves with experts in the field, and push toward their development goals. During the events, teams have access to GPU-enabled compute systems ranging from workstations and local clusters to world-class HPC systems such as OLCF’s Summit.

The OLCF and its partners delivered seven GPU hackathons with the following organizations:

- San Diego Supercomputer Center (May 4, 11–13)
- Princeton University (June 2, 8–10)
- NERSC (July 19, 26–28)
- NOAA (August 23, August 30–September 1)
- NASA (September 20, 27–29)
- ORNL (October 18, 25–27)
- NERSC (December 2, 8–10)

These (virtual) sites hosted a total of 65 application teams, bringing dedicated GPU programming expertise to 433 unique developers from 89 unique organizations. During the events, teams developed applications spanning a wide range of scientific domains, including (but not limited to) astrophysics, climate modeling, combustion, computational fluid dynamics (CFD), ML, molecular dynamics, plasma physics, and quantum mechanics.

During the ORNL event, a team from Georgia Tech—with an OLCF director’s discretionary allocation on Summit—worked alongside mentors from OLCF and NVIDIA to accelerate their CFD application using the OpenACC GPU programming model. Using a simple test case, they were able to obtain a ~3× speedup with 1 GPU as compared to 7 CPU cores on Summit, and they later demonstrated near-ideal weak scaling on up to 128 Summit compute nodes. Another team of developers at the ORNL hackathon brought their radiation-hydrodynamics code as part of an ECP project. During the event, they profiled their code, performed a roofline analysis, ported two kernels to GPUs, and optimized some of their existing GPU code to obtain a ~5× speedup in their test application when running on 1 GPU relative to 7 CPU cores on Summit.

Of particular note in 2021 was our successful collaboration with NASA and NVIDIA during the hackathon hosted by NASA’s Langley Research Center. During the event, 50 developers from 10 application teams located at NASA's Ames, Glenn, Goddard, and Langley field centers, as well as the US Department of Defense, participated in the event. Eric Nielsen, who was the on-site coordinator of the 2021 NASA event and who has participated in multiple hackathons with his own application team, said:

"With GPU technology comprising the backbone of the nation's exascale roadmap and also serving as an enabler for the machine learning community, the event offered a unique opportunity for ten application teams to receive two weeks of intensive, hands-on mentoring. Feedback provided by all of the teams during the final session was overwhelmingly positive and each achieved substantial acceleration of their target application. I am confident that the experience will factor heavily into their future successes in the field of computational science."

In addition to providing application teams (many of which are OLCF users) with development support, these events benefit the OLCF by introducing participants to our facility and resources, nurturing relationships with our industry partners and leaders of the host organizations, generating articles in peer-reviewed journals, recruiting new OLCF staff members, and producing new OLCF users and projects.

1.3.4.3 NCI/DOE Pilot 3 Hackathon

In conjunction with the National Institutes of Cancer (NIC) and DOE, the OLCF organized a hackathon on January 21 and 22, which was aimed at helping researchers further understand how to develop and deploy robust and scalable AI solutions for automated information extraction from free text pathology reports. Researchers from the cancer informatics community, particularly from academic cancer centers and the federal public health community, were invited to the event. In addition to a coding hackathon, the event featured a workshop on deep natural language processing (NLP) for information, novel data analytic techniques for patient information integration, and data-driven integrated modeling and simulation for precision oncology. The 46 participants ran 256 jobs on the Ascent training computer.

1.3.4.4 Early-Frontier Training

In addition to "production" training efforts, the OLCF also delivered "pre-production" training events for CAAR, ECP, and OLCF staff members with access to the OLCF's early-Frontier systems. In 2021, the OLCF and our COE vendor partners, HPE and AMD, delivered two technical sharing webinars, a Spock Training Workshop, and a Spock Application Readiness Workshop.

During the technical sharing webinars, AMD discussed their Tensile tool (for building backend libraries for GEMM-like problems), and three ECP teams gave talks on their experiences on early HPE + AMD systems. The Spock Training Workshop was given in May 2021, soon after the CAAR and ECP teams gained access to the Spock system (Rome CPU + MI100 GPUs). This event delivered new Spock users the information they needed to get up and running on the system. Talks covered topics such as system architecture, AMD MI100 GPUs, the programming environment, compilers, GPU-aware Cray-MPICH, process/thread/GPU mapping, and performance tools. The event was open to all CAAR and ECP teams, totaling about 800 invited participants. The Spock Application Readiness Workshop (August 2021) targeted the 8 CAAR and 12 OLCF-supported ECP engagement teams, at which the teams gave status updates, reported existing issues impeding their progress, and identified the correct channels to resolve their issues and drive their progress forward.

Additional events have been scheduled for 2022 on the latest Frontier hardware (Optimized 3rd Gen AMD EPYC CPUs + MI250X GPUs) to which the teams currently have access. These include a Crusher Training Workshop (delivered January 2022) and two Crusher Hackathons (delivered February 2022).

1.3.4.5 HIP Training Workshop

In May, OLCF staff collaborated with AMD to present a three-day hands-on workshop using HIP to accelerate algorithms that are common in HPC workflows. HIP is a C++ runtime API that allows developers to write portable code to run on AMD and NVIDIA GPUs. This workshop focused on teaching the OLCF user base how to prepare codes with HIP for Frontier. Since not all OLCF users have access to AMD GPUs, AMD provided access to the AMD accelerator cloud for the hands-on exercises. Each day, presenters from AMD began with a series of lectures that covered topics such as how to accelerate CPU code with HIP, how to convert CUDA code to HIP code, and how to measure code performance. AMD and OLCF staff prepared hands-on exercises based on common CUDA examples and the Cholla astrophysics code for users to practice converting to HIP or accelerating with HIP. AMD and OLCF staff helped participants with the exercises over Slack through the duration of the event. Each day, OLCF staff collected lessons learned from the users and shared them during the closing discussions. The training was attended by 127 users.

Event page: https://www.olcf.ornl.gov/calendar/2021hip/

1.3.4.6 CUDA Training Series

The OLCF, NERSC, and NVIDIA again co-organized and delivered a CUDA training series as a follow-on to the 2020 series. Each module consisted of a 1-hour presentation followed by a hands-on session where participants worked through example exercises meant to reinforce the material covered in the presentation. The following topics were covered:

- CUDA Multithreading with Streams: https://www.olcf.ornl.gov/calendar/cuda-multithreading/
- CUDA Multi Process Service: https://www.olcf.ornl.gov/calendar/cuda-multi-process-service/
- CUDA Graph: https://www.olcf.ornl.gov/calendar/cuda-graphs/

The series was attended by 185 unique users.

1.3.5 Community Engagement

OLCF staff design seminars, workshops, and tutorials to inform current and future scientists and engineers about the methods and techniques used for HPC and how HPC can be used to advance scientific research. Aspects of these activities are also targeted at the general public to convey the benefits and advances rendered by HPC-enabled science and engineering.

1.3.5.1 Virtual Hands-On Summit Series

A series of hands-on training sessions offered through the OLCF is opening the doors of HPC to a wider audience. The organizers behind the “Hands-on with Summit” training worked to expand the scope of the course to reach audiences that are often underrepresented in the HPC community, both within and outside ORNL.

During the training, students are introduced to the overall user facility, to the IBM AC922 Summit supercomputer, and to the OLCF’s Ascent training system. Ascent is essentially a single cabinet’s worth of Summit nodes and is allocated for training new users—including students—on Summit’s node architecture, albeit at a smaller scale. In recent years, the organizers have provided an overview of the science that HPC enables: advanced earthquake simulations, studying an x-ray burst moving across a neutron star, and modeling an aerosolized SARS-CoV-2 virus for the first time, to name a few. Students are free to work through a set of self-paced challenges as they progress through increasingly complex tasks designed to teach them the basics of using a supercomputer. OLCF staff are on hand via Slack to answer questions and help anyone who gets stuck.

The "Hands-on with Summit" training was first offered to students in the Students@SC program at SC19. Thomas Papatheodore and Jack Morrison (now at Rescale) designed this course for students who had little to no background in HPC. In 2020 and 2021, several OLCF staff members contributed new exercises to cover a broader range of HPC topics, including Python coding as well as additional C coding exercises to teach different methods for offloading work to the GPU. The modular design of the training and its large Git repository of lesson slides and coding exercises with self-guided instructions make the training easy to prepare and deliver to large numbers of students, either live or virtually.

In 2021, the Hands-On Summit training was given four times and the organizers are now working to expand the training into a multi-day *HPC Crash Course* that includes a *day zero module* to teach the foundational skills for a true HPC beginner, including remote login and programming basics. This expanded course will be offered to universities and university programs that might not have a computing track or a computing focus, with the goal of exposing HPC to students who might not otherwise have the chance to explore scientific computing.

GitHub Repositories: https://github.com/olcf/hands-on-with-summit, https://github.com/olcf/foundational_hpc_skills

1.3.5.2 HOLA Outreach

The Hispanic and Latino Organization for Leadership and Awareness (HOLA) is one of ten Employee Resource Groups at ORNL, which are a key part of the lab's Diversity, Equity, and Inclusion efforts to "enlist and empower our employees to help create a community where everyone feels respected and valued…."

In addition to a virtual tour of Summit led in Spanish, HOLA offered a virtual "Hands-on with Summit" training session that was open to everyone in the community, regardless of computing experience or background. The 55 attendees included ORNL staff, university students from the United States and Mexico, and even a high school student.

1.3.5.3 Introduction to Leadership Computing

On November 12, OLCF staff gave an "Introduction to Leadership Computing," a workshop targeted at potential users from US government agencies. The workshop described the largest-scale capabilities for HPC and provided examples of the breadth of scientific discoveries enabled by them. An overview was given of how research teams get access to, write programs for, and run workloads on Summit. The event was attended by current and former American Association for the Advancement of Science Policy Fellows and participants from eight different government agencies.

Event page: https://www.olcf.ornl.gov/introduction-to-leadership-computing/

2. OPERATIONAL PERFORMANCE

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

OLCF RESPONSE: Yes. The OLCF provides a series of highly capable and reliable systems for the user community. The 2021 reporting period includes full CY production periods for the following HPC resources: the IBM AC922 Summit, the GPFS file system (Alpine), and the archival storage system (HPSS). In 2021, 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 2021 for the Summit computational resources (Table 2.1), HPSS data resources (Table 2.2) and the Alpine resources (Table 2.3). Supporting systems such as EVEREST, Andes, and data transfer nodes were also offered. Metrics for these supporting systems are not provided. See Appendix D for more information about each of these systems. The following tables describe the resource availability for OLCF resources. More details on the definitions and formulae describing scheduled availability (SA), overall availability (OA), mean time to interrupt (MTTI), and mean time to failure (MTTF) are provided in Appendix E.

Operational performance metrics are provided for the OLCF computational resource Summit, the HPSS archive system, and the external GPFS file system Alpine (Tables 2.3–2.5).

Table 2.1. OLCF Operational performance summary for Summit.

	Measurement	2020 target	2020 actual	2021 target	2021 actual
IBM AC922 Summit	Scheduled availability	90%	99.59%	95%	99.71%
	Overall availability	85%	98.73%	90%	97.88%
	MTTI[a] (hours)	NAM[c]	788	NAM	612
	MTTF[b] (hours)	NAM	2,187	NAM	2,184

[a] MTTI = Mean time to interrupt.
[b] MTTF = Mean time to failure.
[c] 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 HPSS.

	Measurement	2020 target	2020 actual	2021 target	2021 actual
HPSS	Scheduled availability	95%	99.92%	95%	99.96%
	Overall availability	90%	98.33%	90%	97.56%
	MTTI[a] (hours)	NAM[c]	432	NAM	342
	MTTF[b] (hours)	NAM	1,463	NAM	1,459

[a] MTTI = Mean time to interrupt
[b] MTTF = Mean time to failure
[c] 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 Alpine, the external GPFS file system.

	Measurement	2020 target	2020 actual	2021 target	2021 actual
Alpine	Scheduled availability	95%	99.43%	95%	99.88%
	Overall availability	90%	98.93%	90%	99.07%
	MTTI[a] (hours)	NAM[c]	621	NAM	789
	MTTF[b] (hours)	NAM	873	NAM	2,187

[a] MTTI = Mean time to interrupt
[b] MTTF = Mean time to failure
[c] NAM = Not a metric. No defined metric or target exists for this system. Data provided as reference only.

Alpine saw dramatic increases in both MTTI and MTTF in 2021. An increase in drive failures resulted in a lower MTTI/MTTF in the previous year, but continued partnerships with IBM, OLCF development of new monitoring tools, and enhanced debugging sessions have led to more proactive efforts in resolving issues before they result in a problem.

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 2, the SA target for an HPC compute resource increases to at least 90%, and the OA target increases to at least 85%. For year 3 through the end of life for the associated compute resource, the SA target for an HPC compute resource increases to 95%, and the OA target increases to 90%. Consequently, 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 2 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%. SA targets are thus described as 90%/95%. OA targets are thus described as 85%/90%.

2.1.1 Scheduled Availability

SA is the percentage of time 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 prior to the event and preferably as much as 7 calendar days prior. If that regularly scheduled maintenance is not needed, 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 2020 and 2021 (Table 2.4).

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

Table 2.4 OLCF operational performance summary: Scheduled availability.

	System	2020 target	2020 actual	2021 target	2021 actual
Scheduled availability	IBM AC922	90%	99.59%	95%	99.71%
	HPSS	95%	99.92%	95%	99.96%
	Alpine	90%	99.43%	95%	99.88%

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

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

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

As shown in Table 2.5, the OLCF exceeded the OA targets of the facility's resources for 2020 and 2021.

Table 2.5 OLCF operational performance summary: Overall availability.

	System	2020 target	2020 actual	2021 target	2021 actual
Overall availability	IBM AC922	85%	98.73%	90%	97.88%
	HPSS	90%	98.33%	90%	97.56%
	Alpine	90%	98.93%	90%	99.07%

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 2.6.

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

Table 2.6. OLCF operational performance summary: Mean time to interrupt.

	System	2020 actual	2021 actual
MTTI[a] (hours)	IBM AC922	788	612
	HPSS	432	342
	Alpine	621	789

[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 2.7.

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

Table 2.7. OLCF operational performance summary: Mean time to failure.

	System	2020 actual	2021 actual
MTTF[a] (hours)	IBM AC922	2,187	2,184
	HPSS	1,463	1,459
	Alpine	873	2,187

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

2.2 TOTAL SYSTEM UTIIZATION IN 2021

2.2.1 Resource Utilization Snapshot

For the IBM AC922 Summit during the operational assessment period January 1–December 31, 2021, 36,402,388 Summit node hours were used outside of outage periods from an available 40,590,815 Summit node hours. The total system utilization for the IBM AC922 Summit was 89.68%.

2.2.2 Total System Utilization

2.2.2.1 2021 Operational Assessment Guidance

System utilization (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.

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

The measurement period is for 2021, irrespective 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 node hours by job. This method is extended to track the consumption of Summit node hours by program, project, user, and system with high fidelity. Figure 2.1 summarizes the IBM AC922 Summit utilization by month and by program for CY 2021. 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 third production year of Summit, utilization was 89.68%—remaining very much in line with year two at 90.11%.

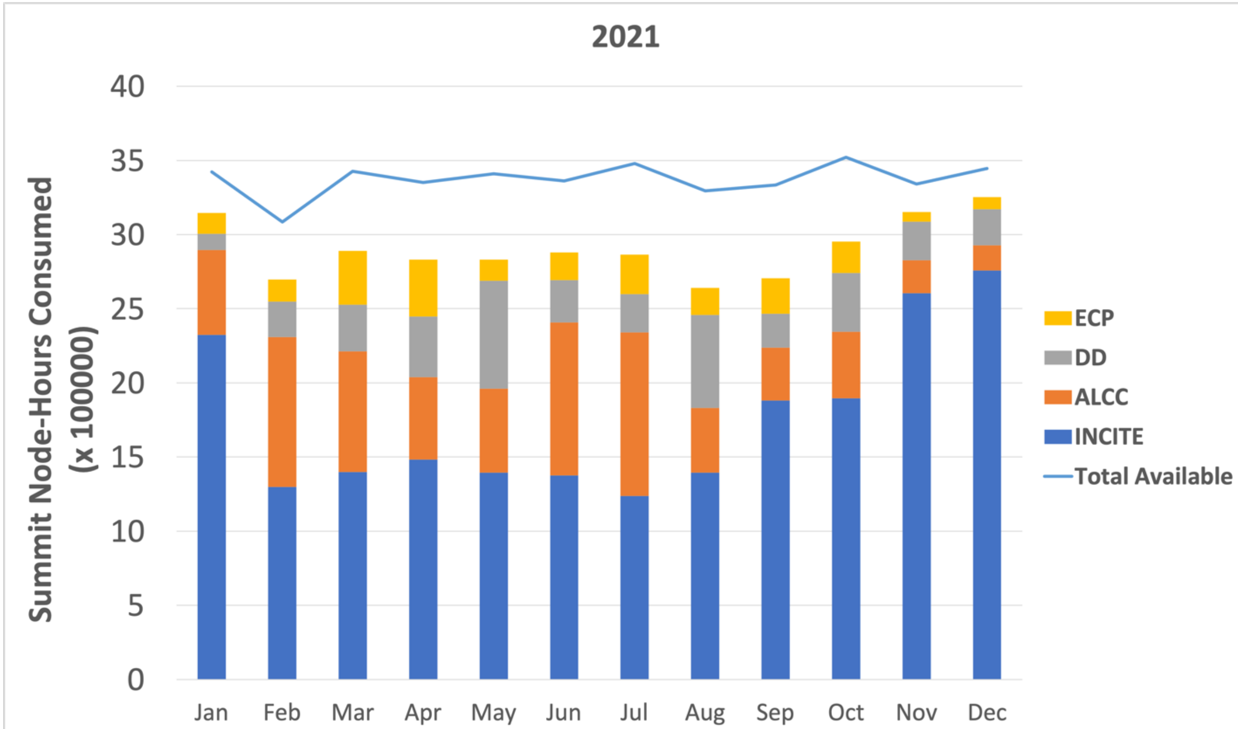


Figure 2.1. 2021 IBM AC922 resource utilization—Summit node hours by program.

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 2021 INCITE allocation program was the largest program in 2021, with a commitment of 19.3 million Summit node hours. The consumption of these allocation programs is shown in Table 2.8. As shown, all commitments were exceeded for each allocation program on Summit for 2021. This programmatic overachievement is due in part to the high uptime and diligent work of the OLCF staff.

Table 2.8. The 2021 allocated program performance on Summit.

Program	Allocation	Hours consumed	Percent of total
INCITE[a]	19,325,000	19,659,283	59%
ALCC[b]	Allocation spans multiple CY	7,289,852	22%
DD[c]	–	6,505,656	19%
Total[d]		33,454,791	100%

[a] Includes all 2021 INCITE program usage (including the 13th bonus month usage in Jan 22).
[b] Includes all ALCC program usage for CY 2021.
[c] Includes ECP.
[d] Does not include usage outside of the three primary allocation programs.

Non-renewed INCITE projects from 2021 continued running through January 2022 under the OLCF's 13th month policy. This policy is in place to permit an additional, final month for completion and was recognized as a best practice during a previous Operational Assessment Review (OAR). It also serves to maintain high utilization while new projects establish a more predictable consumption routine. ALCC projects from the 2021 allocation period (ending June 30, 2021) were also granted extensions where appropriate.

CY 2021 saw a continuation of Summit allocations to the COVID-19 High Performance Computing Consortium. Over 1 million Summit hours were provided to the consortium spanning both the ALCC and DD programs. The allocations declined as the year came to an end.

2.3 CAPABILITY UTILIZATION

To be classified as a capability job, any single job must use at least 20% of the available nodes of the leadership system. 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 node hours will be from jobs requiring 20% or more of nodes available to the users. The metric for capability utilization describes the aggregate number of node hours delivered by capability jobs. The metric for CY 2021 was 35% for Summit based on years of service, as described above. The OLCF continues to exceed expectations for capability usage of its HPC resources (Table 2.9 and Figure 2.2). Keys to successful demonstration of capability usage include the liaison role 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), which actively engage with code developers to promote application portability, suitability to hybrid systems, and performance. The OLCF also aggressively prioritizes capability jobs in the scheduling system.

Table 2.9. OLCF capability usage on the IBM AC922 Summit system.

Leadership usage	CY 2020 target	CY 2020 actual	CY 2021 target	CY 2021 actual
INCITE	NAM[a]	49.79%	NAM	50.83%
ALCC	NAM	32.26%	NAM	36.76%
All projects	35%	43.42%	35%	50.68%

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

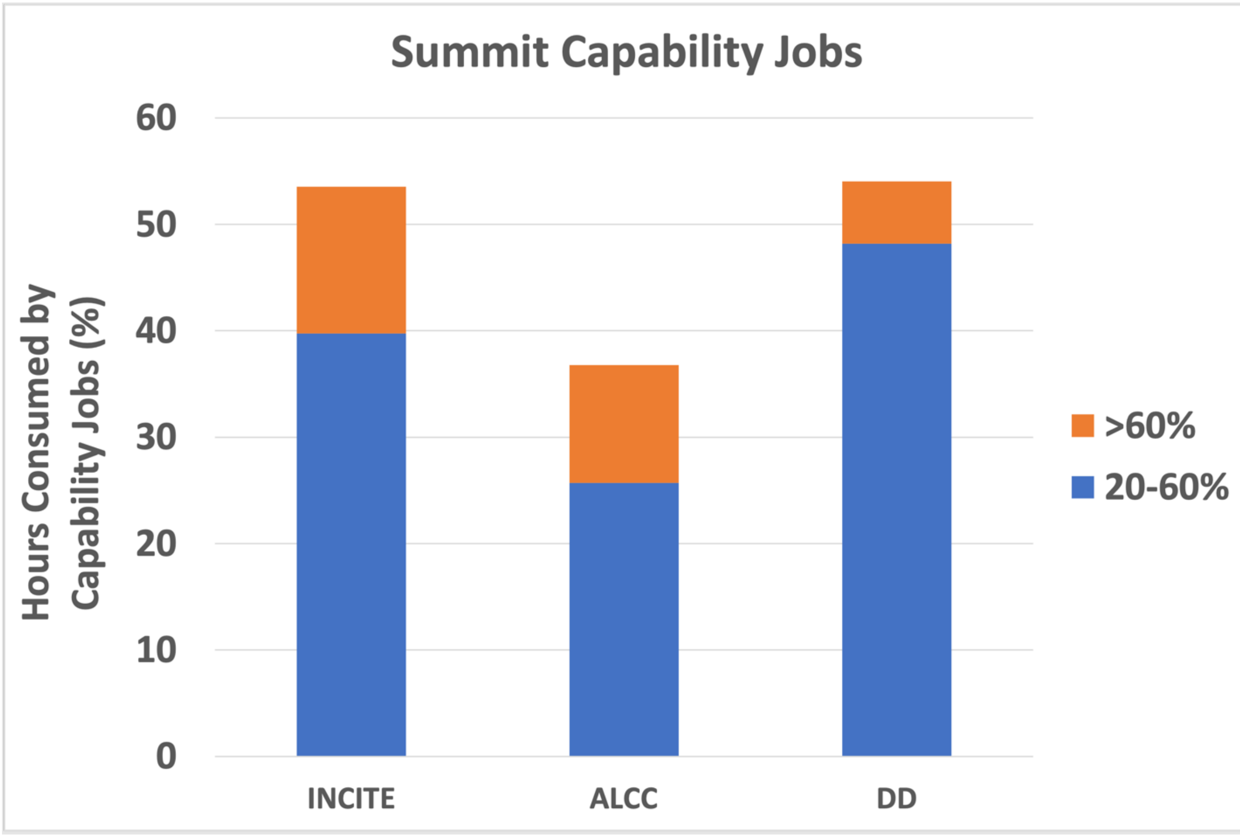


Figure 2.2. Summit capability usage by job size bins and project type.

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, Director's Discretionary, ECP)? (b) Was the allocation of Director's Discretionary 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. The OLCF successfully modified job priorities for INCITE jobs relative to our usual policies to ensure that all of the allocated hours were delivered to the program in CY 2021.

3.1 SUMMARY OF ALLOCATION PROGRAMS

The primary allocation programs providing 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 the INCITE program, 20% to ALCC, and 20% to DD, with up to half of the DD allocation (i.e., 10%) dedicated to ECP Application Development and Software Technology projects during each quarter. We report these ECP hours as a separate category in the following discussion. In CY 2021, the delivered fraction of available time to these individual programs was as follows: INCITE, 59%; ALCC, 22%; DD, 12%; ECP, 7%. The ECP resources were generally under-allocated in each quarter of CY 2021, and the resulting time available in each quarter was often used by DD projects. In particular, those DD projects that were allocated via the COVID-19 HPC Consortium and not moved to the ALCC program after July 2021, represent a significant portion of this minimal over-utilization.

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 in 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 DOE and ORNL strategic agendas

These goals are aligned particularly well with three of the OLCF's four 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. They may be entirely new areas with respect to HPC, or they may be areas in need of nurturing. Examples of projects are those associated with the ORNL Laboratory Directed Research and Development program; programmatic science areas (fusion, materials, chemistry, climate, nuclear physics, nuclear engineering, and 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 carry out research that would not otherwise be possible. See Section 8.2.7 for more information about ACCEL.

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 is roughly 22,500 Summit node hours, but awards can range from thousands to 200,000 node hours, or more. In 2021, the OLCF DD program participants used approximately 19% of total user resources (on Summit), consuming more than 6.5 million Summit node hours. Of that total, 7% of the hours were consumed by ECP Application Development (AD) and Software Technology (ST) projects. Note that several COVID-19–related projects were initially allocated under the DD program but were subsequently moved to the 2020–2021 ALCC program at ASCR direction. See Appendix F for a full list of DD projects for CY 2021.

3.3 ALLOCATION PROGRAM CHALLENGES

The INCITE program suffered from significant under-utilization throughout much of CY 2021. The OLCF modified job priorities for INCITE jobs relative to our usual policies to ensure that all of the allocated hours were delivered to the program in CY 2021. In particular, starting in September 2021, we increased the relative priority of INCITE jobs by increasing the effective number of days they appeared to be queued by 5. 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 enabled INCITE projects to use the remainder of their time. In fact, the 2021 INCITE projects consumed 334,283 node hours over the allocated time by the end of the allocation year. It is important to note that this priority boost was essential, but not sufficient, to achieve this result. The projects also significantly ramped up their usage during this time, effectively driving up the demand for this increase in “supply.”

Innovation

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2022

4. OPERATIONAL INNOVATION

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: Yes. The OLCF actively pursues innovations to enhance facility operations, improve the user experience, support integrated research infrastructure, and grow our workforce pipeline. Through collaborations with users, other facilities, vendors, and the broader digital infrastructure community, many of these innovations are disseminated and/or adopted across the country. In addition, the OLCF is training the future computational science and HPC workforce through the Computational Scientists for Energy, the Environment, and National Security (CSEEN) postdoctoral program and other engagement activities. Since the facility's inception in 2004, OLCF staff have provided leadership in the HPC community, spearheading the creation and development of tools and policies necessary for computing and computational science and disseminating that knowledge throughout the community through a variety of outlets. It is not possible to highlight all the innovative work carried out by the OLCF. Instead, this section will highlight certain specific areas of innovation in operations and research in 2021. We also highlight the postdoctoral fellows who were trained in the OLCF in CY 2021, the contributions they made to OLCF applications projects, and, when appropriate, their career paths after their OLCF postdoctoral appointment.

4.1 OPERATIONAL INNOVATION – TECHNICAL

4.1.1 Software Deployment, CI, and E4S

In collaboration with ALCF and NERSC, the OLCF has shaped the ECP-led scientific software ecosystem to ensure that facility requirements and needs are met both now and in the future. ECP software teams have standardized their software packaging, testing, and deployment framework through a series of community guidelines developed by the E4S project (See Figure 4.1). By packaging their software into Spack, the OLCF software integration team can build, test, and vet the broad scientific software stack at facilities when E4S releases a curated set of packages and versions once a quarter. The OLCF deployed several E4S releases on Summit, Spock, and Crusher in 2021, resulting in a general increase in the number of installed and supported scientific software packages available to end users. Additionally, the OLCF has led the development of an "L2 software support model" through E4S. User tickets created as a result of a software bug or performance regression caFedn be routed to the E4S team, which can in turn be routed to the appropriate ECP Software Technology development team to be addressed. This support model was finalized in 2021 and will serve as a prototype for a sustainable scientific software ecosystem at least through the end of ECP, leading to broad efficiencies within the DOE system.

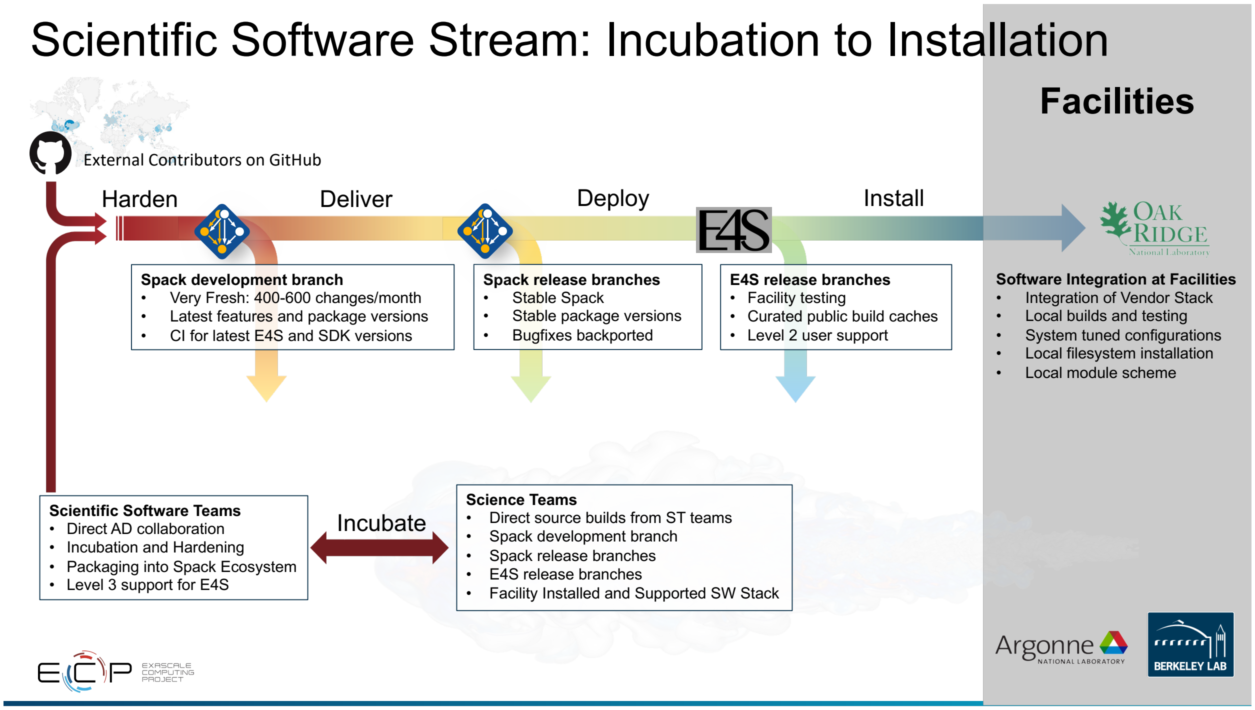


Figure 4.1. Scientific software stream framework from incubation to installation. Credit: Adamson, et al./ORNL.

GitLab CI platforms are also available on OLCF resources and continue to be maintained by OLCF systems administrators. In the last half of 2021, over 17,000 CI jobs were executed, resulting in nearly 2,000 hours of unit, smoke, and regression tests by application developers and software integrators. The OLCF CI platform was migrated to Spock from Ascent in 2022 and will serve CI needs on Frontier and Frontier TDS in the coming years.

4.1.2 Revealing Power, Energy, and Thermal Dynamics of the Summit Supercomputer through Analysis of High-Volume Operational Data

To address the power and energy challenges in the exascale era, the OLCF is continuing its effort in innovative HPC data center cooling that uses energy-efficient cooling methods and extensive operational data.

After successfully deploying a warm-water (70°F) direct liquid cooling system for its major systems such as Summit and Alpine, the center is pushing the limits by employing higher temperature water for its post-exascale systems such as Frontier and Orion.

Compared to traditional usage of air cooling and chilled water (45°F), higher temperature enabled the usage of energy efficient cooling methods such as evaporative cooling over extensive periods of the year and gave significant energy savings in data center cooling operations. The cost performance of liquid cooling has been maximized by the extensive usage of per-cabinet warm/high water-based rear-door heat exchangers (RDHX) that reject more than approximately 10% to 15% of the ambient compute drawer heat generated by the lowest heat density components. This heat rejection method now extends to all of the infrastructure and supporting systems.

For continuous improvement of such energy-efficient cooling operations, the OLCF has been extensively utilizing high-volume operational data that cover the entire HPC data center, reaching not only the compute system but also the building infrastructure. Although the OLCF benefits from an in-house, near–real-time monitoring system that gives a full end-to-end high-resolution power and thermal view of the data center, the center also benefits from long-term insights from large scale-data analytics on this data. In 2021, a data science team at the OLCF (led by Dr. Woong Shin) performed a comprehensive analysis on the accumulated data from the "Cooling Intelligence for Summit" project to study power, energy, and thermal behavior, as well as the impact of the Summit system and its cooling and electrical infrastructure. This study was a foundational work in understanding Summit's power and energy operational data and its use for advanced analytics and AI/ML-based methods for OLCF systems, especially for energy-efficient and sustainable HPC in the post-exascale era. The result of this analysis was submitted, accepted, and won the best paper award at the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC 2021).

Related paper: Woong Shin, Vladyslav Oles, Ahmad Maroof Karimi, J. Austin Ellis, and Feiyi Wang. 2021. "Revealing power, energy and thermal dynamics of a 200PF pre-exascale supercomputer." In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC21). Association for Computing Machinery, New York, NY, USA, Article 12, 1–14. DOI: https://doi.org/10.1145/3458817.3476188

4.1.3 Supporting Integrated Data and Workflow Infrastructures

Several observational scientific instruments, especially commercial instruments, are disconnected from the laboratory network (air-gapped) because the instrumentation software may not be compatible with newer versions of operating systems, other software updates, or cannot meet our cyber policies. Consequently, researchers resort to the use of portable storage drives to manually and slowly move acquired observational data from the instrument computer to scalable storage and computational resources. Additionally, researchers often use the limited computational power of the instrument computer or their personal laptop computer to slowly process the acquired data. Both approaches lower the productivity of researchers, especially given the data explosion most domains are facing. Furthermore, valuable data go unused because they are not married to the contextual information that remains in physical laboratory notebooks.

DataFlow is a software stack deployed on a physical server in proximity to air-gapped instrument computers that selectively and securely provides access to remote data services directly to the instrument computer. Using DataFlow, researchers can use a web-based interface to securely authenticate themselves, enter crucial metadata associated with an experiment, and drag-and-drop data files into the web interface. DataFlow uploads both the metadata and data files to a designated, remote file system that is more resilient and scalable than the instrument computer's storage drive, in addition to being accessible to powerful cloud and HPC capabilities for data processing. The availability of metadata instantly makes the data more valuable for downstream applications, including AI. The data and workflow from observational facilities at ORNL and from collaborators are typically enabled by institutional computing and data capability in ORNL's Compute and Data Environment for Science (CADES) and bridged to OLCF when the campaigns request and gain access to OLCF resources. Draft Data Management Plans (DMPs) are being offered to facility scientists to help them manage their science-data lifecycle.

The web-based interface to DataFlow was made available in August 2021 at https://dataflow.ornl.gov. A facility-local instantiation of DataFlow was deployed to connect two atomic force microscopes at the Center for Nanophase Materials Science to the data and compute resources at the Compute and Data Environment for Science. Fifteen more instruments are expected to be connected to this DataFlow server in 2022.

4.1.4 User-Managed Software

User-Managed Software (UMS) provides a middle ground between software that is deployed, maintained, and supported by the facility and software installed by individual users or projects that is hard to discover. UMS was developed to enable software developers to make their packages widely available to facility users without the facility having to take on new commitments to handle its deployment, maintenance, or support. UMS packages are available through the well-known module system; thus, users do not need to discover an obscure user or project directory. The module system also supports versioning straightforwardly, so that providers can maintain multiple versions and users can change versions on timelines that suit them. This can also facilitate collaborations among projects where one provides a tool, which may be under active development, and one or more other projects make use of it. Users can apply for a UMS project through a simplified process (compared to other allocation programs). A set of policies has been developed for UMS that delineates the responsibilities of both the UMS project members and the facility, among the most important being that the UMS project team agrees to provide support for the software and will include as a part of the module banner a note that the software is not facility-supported and how to obtain assistance. The facility retains the right to disable or remove any software that is not appropriately maintained or supported or causes other problems. UMS projects are provided with a directory in the UMS software tree, can be used to deploy and maintain the software. In 2021, the OLCF received nine UMS project applications providing a range of software packages.

4.1.5 NCI/NIH Whole-Slide Imaging Service on Slate

The Slate infrastructure has a new use case that is improving access to OLCF's infrastructure as well as enhancing the OLCF user experience. Staff members in the Analytics and AI Methods at Scale Group and the Platforms Group deployed a whole-slide imaging service called caMicroscope on the Slate platform. Slate provides a container orchestration service to deploy user-customized software stacks at the OLCF. With this novel use case, pathologists from the National Cancer Institute (NCI) will be able to annotate gigapixel-scale cancer images interactively for later training on the Summit supercomputer. caMicroscope on Slate provides remote visualization, collaboration, and large-scale storage capabilities to NCI's pathologists across their annotation and labeling workflows. These capabilities provide operational efficiency in that pathologists—who are typically not well-versed HPC users—can host their data sets at the OLCF and access the caMicroscope service to effectively and transparently use the OLCF data infrastructure.

4.2 OPERATIONAL INNOVATION – MANAGEMENT/WORKFORCE

4.2.1 Crash Courses for HPC and ML/DL Ecosystem Community Training

The OLCF has developed a virtual crash course for HPC that covers the foundational skills needed to learn about HPC (e.g., UNIX, command-line text editors, introductory C programming) before moving on to the basics of HPC itself. The courses are hosted on Zoom and use Slack for hands-on help during the exercises. During the training, students are introduced to the overall user facility, the IBM AC922 Summit supercomputer, and the OLCF's training system, Ascent—essentially a single cabinet of Summit, with the same node architecture, allocated for students to use during the training. In recent years, the organizers have provided an overview of the science that HPC makes possible: advanced earthquake simulations, studying an x-ray burst moving across a neutron star, and modeling an aerosolized SARS-CoV-2 virus for the first time. Then, students are free to work through a set of self-paced challenges, progressing through increasingly complex tasks designed to teach them the basics of working on a supercomputer.

The inaugural "Hands-on with Summit" training was first offered to students in the Students@SC program at SC19, and we quickly realized that we could build upon what we did in 2019 to help bring a more diverse audience to HPC. In 2020 and 2021, several OLCF staff members contributed new exercises to cover a broader range of HPC topics that included Python coding as well as more C coding exercises to teach different methods for off-loading work to the GPU. The modular design of the training and its large Git repository of lesson slides and coding exercises with self-guided instructions make the training easy to prepare and deliver to a large number of students either live or virtually.

We piloted the course in 2021 with a number of different organizations. Each opportunity brought an opportunity to make improvements and continue refining the course.

We have now developed a multi-day "HPC Crash Course" that includes a set of modules to teach the foundational skills for a true HPC beginner, including remote login and programming basics. This expanded course can be offered to universities and university programs that might not have a computing track or a computing focus, with the goal to expose students to who might otherwise never have the chance to explore HPC. We are launching a webpage in 2022 to advertise this course and invite organizations and academic institutions to partner with us to offer this event to their students. The focus of this page will be on providing an introduction to HPC for under-represented groups, bringing HPC to a broader audience and potentially expanding the future workforce and/or user community.

GitHub Repositories: https://github.com/olcf/hands-on-with-summit, https://github.com/olcf/foundational_hpc_skills

4.2.1.1 ML/DL ecosystem training and support

ML/DL is rapidly and fundamentally transforming the way science uses data to solve problems. As these models and data grow, the need for scalable methods and software to train them grows accordingly. Given the rapid development cycle for DL frameworks, significant effort is required to maintain the software environment, especially on non-x86 platforms such as Summit. To facilitate OLCF users' ML/DL development, we deploy and support popular DL frameworks (TensorFlow and PyTorch) based on Open-CE, a community effort for maintaining DL frameworks. Furthermore, to provide our facility users, especially domain scientists, with a working knowledge of DL on HPC-class systems, we provide hands-on trainings at both the OLCF and SC, covering DL core concepts, scientific applications, performance optimization tips, and techniques for scaling.

Related training: Introduction to Open-CE, OLCF User Meeting, Junqi Yin; Deep Learning at Scale, SC21 tutorial/OLCF tutorial, Aristeidis Tsaris, Junqi Yin, in partnership with NERSC and NVIDIA teams.

4.2.2 Winter Classic Invitational Student Cluster Competition

The OLCF was asked in late 2021 by Intersect 360 to serve as a mentor institution for the Winter Classic Invitational Student Cluster Competition (SCC, https://www.winterclassicinvitational.com/) in which twelve teams from historically black colleges and universities (HBCUs) and Hispanic-serving institutions (HSIs) will receive hands-on experience in HPC and an opportunity to execute, analyze, and optimize scientific applications. The OLCF will be providing compute resources, mentorship, and a set of challenge problems for the students to complete. The main goal of the Winter Classic Invitational SCC is to attract young new talent from diverse communities to HPC by providing them with hands-on experience with real HPC systems and applications. Teams from the following HBCUs and HSIs will be participating in 2022:

- California State Fullerton (HSI)
- California State University Channel Islands (HSI)
- Florida Agricultural and Mechanical University (HBCU)
- Spelman College (HBCU)
- Morehouse College (HBCU)
- Tennessee State University (HBCU)
- Texas Tech University (HSI)
- Fayetteville State University (HBCU)
- Prairie View Agricultural and Mechanical University (HBCU)
- The University of Texas at El Paso (HSI)

4.2.3 Pathways to Supercomputing Initiative

The OLCF has launched a new initiative to assist researchers and institutions currently underrepresented in HPC, equipping them with the tools, mentorship, and resources needed to develop competitive allocation proposals and providing a pathway to the OLCF's programs, including DD, the Innovative and INCITE, and ALCC programs. The initiative is open but not limited to community colleges, HBCUs, HSIs, and liberal arts colleges. If selected for the Pathways to Supercomputing initiative, projects will receive a startup allocation on the center's IBM AC922 Summit supercomputer and other OLCF resources, as well as mentoring from OLCF staff to design an experimental campaign with the goal of submitting an allocation proposal at the end of the startup period. The first successful application for the Pathways to Supercomputing Initiative is from North Carolina Agricultural and Technical State University (HBCU).

4.2.4 OLCF Recruiting, Mentoring, and Workforce Development Efforts

Members of the OLCF participated in a number of recruiting and outreach events, conferences, and workshops throughout the year to help introduce HPC to a broader audience and increase awareness of the user programs and career opportunities. Below are just a few of examples of ways the OLCF plugged in to the community in 2021.

- Supercomputing
    - SC is among the largest venues to introduce students and early-career researchers to leadership-scale HPC. OLCF staff participate in this annual event as speakers, committee members, session chairs, and workshop and tutorial organizers. Examples of the tutorials delivered at SC21 are given below.

        - As part of the conference's Students@SC program, on October 29, OLCF staff led a half-day session titled "Hands-on with Summit." The session was attended virtually by 17 undergraduate and graduate students participating in the Students@SC program.

        - As part of the conference's general tutorial program on Sunday, November 14, the OLCF—in collaboration with Lawrence Berkeley National Laboratory (LBNL) and NVIDIA—gave "Deep Learning at Scale," a set of lessons about deep learning on HPC-class systems, including core concepts, scientific applications, performance optimization tips, and techniques for scaling. The tutorial was motivated by the fact that deep learning is rapidly and fundamentally transforming the way science and industry use data to solve problems. Participants were provided with training accounts, example code, and datasets to allow them to experiment hands-on with optimized, scalable, distributed training of deep neural network ML models.

        - As part of the conference's general tutorial program, on Monday, November 15, the OLCF—in conjunction with ANL and Los Alamos National Laboratory—presented "Better Scientific Software," a tutorial for the Computational Science and Engineering (CSE) community aimed to address the challenges presented by the confluence of disruptive changes in computing architectures, demand for greater scientific reproducibility, sustainability and quality, and new opportunities for greatly improved simulation capabilities. The topics included software processes for (small) teams, including agile processes, collaboration via version control workflows, reproducibility and scientific software design, refactoring, and testing. Its goals were to improve the productivity of those who develop CSE software and increase the sustainability of software artifacts.

- PEARC
    - Members of the OLCF took leadership roles in the PEARC Student Volunteer program and organized and led several events targeted at student attendees and once again served as mentors to students and other early-career attendees. The OLCF engaged our colleagues from both NERSC and ALCF, who also helped serve as mentors and gave their time to help with the other student centric activities, including the resume workshop.

- OLCF Virtual Career Fair:
    - The OLCF organized our first standalone career fair to feature HPC positions within our division. We held the virtual event using GatherTown. We invited participants by sharing links with groups including Latinas in Computing, Hispanics in Computing, Americas HPC Collaboration, SC21 and SC22 Students@SC program, PEARC Student Volunteers, and using our social media and regular communication channels. During the career fair, attendees had the opportunity to listen to four short talks at the top of the hour from OLCF staff and users, as well as take part in guided virtual tours of our facilities. Attendees were also invited to engage with OLCF staff at a number of recruiting tables to discuss the open positions within our teams. A total of 115 people participated in the Virtual Career Fair. The OLCF plans to offer another career fair in 2022.

- Tapia Diversity in Computing Conference
    - The ACM Richard Tapia Celebration of Diversity in Computing Conference (TAPIA 2021) is the premier venue to acknowledge, promote, and celebrate diversity in computing. As part of TAPIA, on September 16, OLCF staff members gave an OpenMP tutorial during the HPC in a Nutshell Session designed to give participants an introduction to important topics in HPC. During the tutorial, students learned the basics of how to utilize shared-memory parallelism with OpenMP. This session was a collaboration between multiple DOE laboratories.

- CARLA 2021 Americas HPC Collaboration Workshop
    - The Latin America High Performance Computing Conference (CARLA 2021) is "an international conference aimed at providing a forum to foster the growth and strength of the High Performance Computing community in Latin America through the exchange and dissemination of new ideas, techniques, and research in HPC and its applications areas." On October 5 at CARLA, OLCF staff member Verónica Melesse Vergara and ANL staff member Silvio Rizzi co-chaired the Americas High Performance Computing Collaboration Workshop, which sought to showcase collaborations that have resulted from the partnerships

formed between researchers and entities across the Americas, from Patagonia to Alaska, including Caribbean institutions. The event shared opportunities and experiences between HPC networks and laboratories from countries in North, Central, and South America, as well as the Caribbean. The goal of this workshop was to highlight the current state of the art in continental collaboration in HPC, and the latest developments of regional collaborative activities.

- ECP Broader Engagement
  - Members of the OLCF contributed to the new Broader Engagement efforts undertaken by ECP to help in the establishment of a sustainable, long-term approach to recruit and retain a diverse workforce in the DOE HPC community. The approach has three complementary thrusts, which leverage and augment existing workforce efforts in DOE national laboratories, computing facilities, and the HPC computational science community.
    - Establishing an HPC Workforce Development and Retention (HPC-WDR) Action Group, whose primary focus is fostering a supportive and inclusive culture in ECP and DOE labs and communities. Its goal is to promote the workforce pipeline for, and the retention of, a diverse DOE lab HPC workforce. In addition to engaging underrepresented communities and strengthening our workforce pipelines, we must also focus on HPC culture within DOE laboratories. Although the HPC community, including DOE labs, have undertaken a variety of initiatives to improve culture with respect to diversity, equity, and inclusion, sharing of information across institutions is limited. The HPC-WDR Action Group will provide a mechanism for better sharing and communicating about these initiatives, best practices, lessons learned, and other DEI experiences, thereby benefiting the HPC community across the DOE lab complex.
    - Building an "Introduction to HPC" Training and Workforce Pipeline Program (HPC-Intro) to provide accessible introductory material on HPC, scalable AI and analytics, thereby addressing gaps for—and expanding the pipeline of—people with foundational HPC skills. Basic HPC is not typically taught at early stages of students' careers, and the capacity and knowledge of HPC at many institutions are limited. Even so, such topics are prerequisites for advanced opportunities such as the Argonne Training Program for Extreme-Scale Computing (ATPESC) and the DOE Computational Science Graduate Fellowship (CSGF) Program, and ultimately for careers in HPC. The national lab computing complex has expertise, capabilities, and a long history of cross-lab coordination on joint efforts. The HPC intro training and pipeline program will target advanced undergraduates, students in gap years, and early graduate students.
    - Creating Sustainable Research Pathways for HPC (SRP-HPC); based on the Sustainable Research Pathways Program (a partnership with Sustainable Horizons Institute and Lawrence Berkeley Laboratory) and the Broader Engagement Program at SIAM-CSE Conferences to enable a multilab cohort of students from underrepresented groups (and faculty working with them) to work side-by-side with ECP teams and facilities staff on world-class HPC projects.

- 2023 AIAA SciTech Conference
  - Stephen Nichols is serving as the Technical Discipline Chair of the Meshing, Visualization, and Computational Environments technical sessions for the 2023

AIAA SciTech conference (https://www.aiaa.org/SciTech/utility/about/future-scitech-dates). To prepare for this role, he served as the Deputy TDC for the 2022 AIAA SciTech conference that was held Jan 3 (https://aiaascitech2022.azurewebsites.net/). The yearly AIAA SciTech conference is the world's largest event for Aerospace R&D.

4.2.5 Development of a Cross-Facility ASCR Cybersecurity Working Group

The OLCF security team has worked with representatives from the ALCF, NERSC, and ESNet to form a cybersecurity-focused working group. This team meets regularly to discuss a number of topics ranging from somewhat tactical incident response and threat intelligence sharing to more strategic policy planning and cybersecurity workforce development issues. Work in CY 2021 culminated in a draft working group charter, which is expected to be fine-tuned and then socialized with facility and ASCR leadership in early 2022. Early successes from the working group already include collaboration on responses to zero-day vulnerabilities, including technical deep dives of Log4j and dbus privilege escalation exploits. Other notable outcomes have been consistent messaging from the ASCR facilities to their broader laboratory cyber security teams regarding the viability of implementing technical controls found in recent executive orders, including the Executive Order on Improving the Nation's Cybersecurity. (https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity/)

4.3 POSTDOCTORAL FELLOWS

In 2021, ORNL instituted a Postdoctoral Program Engagement Committee to provide input on the postdoctoral program and to ensure its quality in both the hiring process of postdoctoral fellows and in their experience during their tenure at ORNL. The OLCF is represented on this committee by Markus Eisenbach, a staff member in the Scientific Engagement (SE) section. Within SE, the postdoctoral fellows engage in a variety of research and development projects across the four groups. We are providing opportunities for their scientific and technical progress and development by engaging them in research and code development in their specific scientific domains, as well as giving them opportunities for development in technical trainings, such as HIP, CUDA, and OpenMP training for current and future HPC architectures, and workshops on career developments, such as proposal writing workshops.

4.3.1 Computational Scientists for Energy, the Environment, and National Security Postdoctoral 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 mission needs to maintain the US competitive advantage in high-performance and data-intensive scientific computing. 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. In recognition of this need, DOE proposed that ASCR establish a postdoctoral training program at its user facilities, including the OLCF, ALCF, and NERSC, for future Computational Scientists for Energy, the Environment, and National Security (CSEEN). The objectives of this program are to (1) help 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) to 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 develop 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 scientific 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 2021, a total of 14 postdocs were members of the OLCF workforce. Of those 14, 10 were fully supported by OLCF funds (including six postdoctoral fellows at least partially supported by ECP Application Integration). One of the postdocs was supported by sources outside the OLCF, including individual ECP AD projects. The remaining three postdocs were supported partially by the OLCF and partially by individual ECP projects. The background and current work of these postdocs in the Science Engagement section is described below.

- David Eberius joined the Algorithms and Performance Analysis group in October 2020 after receiving his PhD in computer science from The University of Tennessee, Knoxville (UTK), where he researched techniques for performance measurement and analysis of HPC software. During his graduate studies, he focused on message-passing communications software and collaborated as part of the team that develops the OpenMPI implementation of MPI. Since joining the OLCF, Eberius has been evaluating programming models (implementation and performance optimization) to enable performance portability for the ECP ExaBiome project—in particular, the tools and techniques for understanding the observed performance of code running on GPUs and which architectural feature(s) bottleneck that performance.

- Paul Eller joined the Algorithms and Performance Analysis Group in January 2020 after earning his PhD in computer science from the University of Illinois at Urbana-Champaign. During his doctoral studies, he focused on techniques for measuring, modeling, and improving the scalability and performance of blocking and non-blocking Krylov solvers. To support this research, he developed the Scalable Algorithm Testbed, a collection of tools for measuring and analyzing the performance (and performance variation) of computational kernels and network communication. While at the OLCF, Eller collaborated with the CoMet comparative genomics project team to study the performance benefit of using reduced-precision numerical methods in computationally intensive GPU kernels. Paul left ORNL in January 2022 to work for General Dynamics at the DOD Army Research Lab in Aberdeen proving Ground, MD, where he will be working to optimize codes on DOD supercomputers.

- Antigoni Georgiadou joined the Advanced Computing for Nuclear, Particle, and Astrophysics Group in September 2019 after earning her PhD in mathematics from Florida State University, where she worked on optimization in stellar evolution applications. During her graduate studies, she was also a visiting scholar with the Theoretical Astrophysics Group and the Machine Intelligence and Reconstruction Group at Fermilab, where she worked on an analysis to develop a statistical framework with Gaussian processes and ML techniques to optimize the input parameter space of cosmological simulations. Georgiadou collaborated with the ECP ExaStar project, which builds on the current capabilities of astrophysics codes such as FLASH for multi-physics astrophysics simulations run on exascale machines. The target science is to explain the origin of the elements via stellar explosion simulations and to study the conditions for nucleosynthesis in stars for nuclear data experiments. Georgiadou also collaborated with the ExaSky project, which aims to maximize the computing power for cosmological simulations. In August 2021, Georgiadou transitioned to staff in the Algorithms and Performance Analysis Group as an Applied Mathematician.

Harris, J. A., Chu, R., Couch, S. M., Dubey, A., Endeve, E., Georgiadou, A., . . . Weide, K. (2022). Exascale models of stellar explosions: Quintessential multi-physics simulation. *International Journal of High Performance Computing Applications*, 36(1), 59–77. doi:10.1177/10943420211027937

- Swarnava Ghosh joined the Advanced Computing for Chemistry and Materials Group in August 2020. Previously, he had a postdoctoral appointment from October 2016 to August 2020 in the Department of Mechanical and Civil Engineering at the California Institute of Technology. Ghosh earned his PhD from the Georgia Institute of Technology in August 2016 for his work on efficient large-scale real-space electronic structure calculations. His research at ORNL investigates the magnetism in disordered systems using first principles real-space multiple scattering calculations, such as the dependence of magnetism on disorder in Mn–Sb–Te quantum material systems and the influence of spin on the mechanical behavior of magnetic structural alloys. Ghosh contributes to the development of the ORNL locally self-consistent multiple scattering code (LSMS). In particular, he is extending the capability of this code to the efficient calculation of interatomic forces and non-spherical atomic potentials. Ghosh is interviewing for a Computational Scientist, Materials position in the Advanced Computing in Chemistry and Materials Group.

  Ghosh, S. (2021). Precipitation during creep in magnesium–aluminum alloys. *Continuum Mechanics and Thermodynamics*, 33(6), 2363–2374. doi:10.1007/s00161-021-01047-7

  Teh, Y. S., Ghosh, S., & Bhattacharya, K. (2021). Machine-learned prediction of the electronic fields in a crystal. *Mechanics of Materials*, 163. doi:10.1016/j.mechmat.2021.104070

- Darren Hsu joined the Advanced Computing for Chemistry and Materials Group in May 2021. Previously, he had a postdoctoral appointment at AstraZeneca from October 2020 to April 2021. Hsu earned his PhD from Northwestern University in 2020 in chemistry. His research focused on utilizing ultrafast time-resolved x-ray scattering and molecular dynamics to track structural changes of small molecules and proteins. Since joining the OLFC, Hsu has been working on developing machine-learned Boltzmann generators to sample particle configurations efficiently. In addition, he has been contributing to the molecular modeling effort by modifying mdgx.cuda code for high-throughput simulations, as well as a transformer-based AI model for protein-ligand structure prediction.

- Mariia Karabin joined the Advanced Computing for Chemistry and Materials Group in July 2020 after earning her PhD from Clemson University for her work on the application of statistical mechanics methods to the construction of interatomic potentials and sampling algorithms. Her research at ORNL focuses on the first principles investigation of the physics of high-entropy alloys based on first principles density functional calculations. Developing multiple scattering-based methods and first-principles derived Hamiltonians to Monte Carlo simulations. Application of these methods to the phase transitions and atomic behavior of alloys. Working in the center of accelerated application readiness (CAAR) to prepare new capabilities (Lloyd's formula) and accelerate the LSMS first principles multiple scattering code for efficient use on Frontier and its combination with a model training and Monte Carlo workflow on this platform. Karabin has applied for a Year Three extension in 2022.

- Nicholson Koukpaizan joined the Algorithms and Performance Analysis Group in June 2020 after receiving his PhD in aerospace, aeronautical, and astronautical engineering from the Georgia Institute of Technology. As a member of Georgia Tech's Nonlinear Computational Aeroelasticity Laboratory, he researched techniques for modeling fluidic actuation for aerodynamic flow control with a special focus on jet interaction fluidic oscillators sited on helicopter bodies. Since joining the OLCF, Koukpaizan has worked on evaluating programming models (implementation and performance optimization) to enable performance portability for ECP applications, particularly for the ExaSGD (Power Flow Optimization at Exascale) project. ExaGO and HiOp depend on RAJA and MAGMA to target GPUs. Koukpaizan is developing a GPU-accelerated CFD code for aerodynamic flow control applications; the code now supports OpenACC and OpenMP offloading to target the GPUs. In February 2022, Koukpaizan will be transitioning to staff in the Advanced Computing for Life Sciences and Engineering Group as a Computational Scientist, Fundamental Turbulence.

  Nicholson K. Koukpaizan, Ari Glezer, and Marilyn J. Smith. (2021). Computational Characterization and Boundary Condition Models of a Jet Interaction Fluidic Oscillator. *AIAA Journal*, 59(9).

  Koukpaizan, N. K., Glezer, A., & Smith, M. J. (2020). Towards full-scale fuselage drag reduction computations using fluidic oscillators. Paper presented at the Vertical Flight Society's 76th Annual Forum and Technology Display.

- Justin Lietz joined the Advanced Computing for Nuclear, Particle, and Astrophysics Group in September 2019 after completing his PhD in nuclear physics from Michigan State University (MSU). His research focuses on quantum many-body physics calculations of nuclear matter and electron gases. These calculations can be used to provide inputs for nuclear astrophysics simulations and as a theoretical study of extreme states of matter. Since joining the OLCF, Lietz has been working at CAAR to prepare the nuclear many-body physics code NUCCOR for the exascale supercomputer Frontier. Extending coupled-cluster calculations for nuclear physics, quantum chemistry, and solid-state physics to new levels of precision and system size. Developing massively parallel data structures and algorithms so that large-scale tensor operations can run at the scale of thousands of GPUs. In September 2021, Lietz transitioned to staff in the Algorithms and Performance Analysis Group as a Performance Engineer.

- Isaac Lyngaas joined the Advanced Computing for Life Sciences and Engineering Group in September of 2019 after receiving his PhD in computational science from Florida State University, where he worked on the development of new numerical approaches for solving nonlocal continuum models. During his time at ORNL, he has worked on refactored cloud resolving model Fortran code base to C++ using a performance portability library for single-source portability. Implemented OpenMP, SYCL and CUDA backends in C++ portability library for accelerated hardware. The development of Clang-based precompiler tools to eliminate GPU run time errors in C++ Portability Library. Collaborating with the CANDLE group working on natural language processing applications. Lyngaas is interviewing for a Computational Biomedical Natural Language Processing Scientist position in the Advanced Computing for Life Sciences and Engineering Group.

Blanchard, A. E., Shekar, M. C., Gao, S., Gounley, J., Lyngaas, I., Glaser, J., & Bhowmik, D. (2022). Automating genetic algorithm mutations for molecules using a masked language model. *IEEE Transactions on Evolutionary Computation*. doi:10.1109/TEVC.2022.3144045.

Norman, M. R., Bader, D. A., Eldred, C., Hannah, W. M., Hillman, B. R., Jones, C. R., …, Yuan, X. (2022). Unprecedented cloud resolution in a GPU-enabled full-physics atmospheric climate simulation on OLCF's summit supercomputer. *International Journal of High Performance Computing Applications*, 36(1), 93–105. doi:10.1177/10943420211027539.

Lyngaas, I., & Peterson, J. S. (2021). Using radial basis function-generated quadrature rules to solve nonlocal continuum models. *Numerical Methods for Partial Differential Equations*, doi:10.1002/num.22825.

Lyngaas, I., Norman, M., & Kim, Y. (2021). SAM++: Porting the E3SM-MMF cloud resolving model using a C++ portability library. *International Journal of High Performance Computing Applications*, doi:10.1177/10943420211044495.

- Xingze Mao joined the Advanced Computing for Nuclear, Particle, and Astrophysics Group in August 2020 after earning a dual PhD in nuclear physics and scientific computing from MSU on the complex-energy description of molecular and nuclear open quantum systems. While at the OLCF, Mao worked on the NUCLEI SciDAC-4 project, where his research focused on coupled-cluster theory for atomic nuclei. Specifically, he worked on a new approach to compute properties of deformed but axially symmetric atomic nuclei. These nuclei will be studied experimentally at the new DOE Facility for Radioactive Ion Beams under construction at MSU, but a precise and accurate theoretical description of these nuclei will require exa-scale computing resources using tightly coupled distributed algorithms. Mao left ORNL in August 2021 to work in industry.

- Elvis Maradzike joined the Scientific Computing group in December 2019 after earning his PhD in chemistry from Florida State University. During his PhD, Maradzike focused on developing approaches for ground and excited electronic states based on the variational two-electron reduced density matrix complete active space self-consistent field method. His work at the OLCF will be in collaboration with the NWChemEx ECP project in developing computational tools to accelerate electronic structure computations as well as modeling the physics of strongly correlated electrons. Maradzike received a Year Three extension in 2021.

- Muralikrishnan (Murali) Gopalakrishnan Meena joined the Advanced Computing for Life Sciences and Engineering Group in July 2020 after earning his PhD from the University of California, Los Angeles. His thesis work was on using network (graph) theory to characterize, model, and control vortical interactions in turbulence and wake flows. In particular, Meena introduced a network community–based framework to formulate reduced-order models and perform flow modification of complex laminar and turbulent vortical flows. Since joining the OLCF, his main research focuses on using ML to formulate subgrid-scale turbulence closures for cloud-resolving models. Additionally, Meena is also using ML, graph theory, and other data-driven techniques to characterize and model complex systems in physical and biological sciences. These are targeted at applications such as identifying extreme climate events, geophysical flow modeling using sensor measurements (flood forecasting), and identifying key interactions in various fungal communities. In March 2022, Meena will be transitioning to staff in the Advanced Computing for Life Sciences and Engineering Group as a Computational Scientist, Combustion.

Rush, T. A., Shrestha, H. K., Gopalakrishnan Meena, M., Spangler, M. K., Ellis, J. C., Labbé, J. L., & Abraham, P. E. (2021). Bioprospecting Trichoderma: A Systematic Roadmap to Screen Genomes and Natural Products for Biocontrol Applications. *Frontiers in Fungal Biology*, 41.

- Vassili Mewes joined the Advanced Computing for Nuclear, Particle, and Astrophysics Group in September of 2019 after a postdoctoral stint at the Rochester Institute of Technology. He obtained his PhD in numerical relativity at the University of Valencia in July 2016 after performing the first general relativistic hydrodynamics simulations of tilted accretion tori with a fully dynamical spacetime evolution using the Einstein Toolkit. Since joining the OLCF, Mewes has been expanding his expertise at ORNL, working on neutrino interaction physics with the ExaStar collaboration as part of the ECP. Developing hardware-accelerated code to include nucleons–nucleon Bremsstrahlung opacities in the thornado code for neutrino radiation hydrodynamics using OpenMP / OpenACC offloading, as well as code development for the weaklib code as part of the ExaStar project in the ECP. Developing a double fast Fourier transform (FFT) filter algorithm to alleviate the severe Courant–Friedrichs–Lewy restriction when solving hyperbolic PDEs in spherical coordinates in the SphpericalNR framework. Mewes performed the first binary neutron star simulation with prompt collapse to a black hole in spherical coordinates using filtering. In May 2021, Mewes transitioned to staff in the Nuclear, Particle, and Astrophysics group as a Computational Scientist.

  Armengol Lopez, F. G., Combi, L., Campanelli, M., Noble, S. C., Krolik, J. H., Bowen, D. B., … Nakano, H. (2021). Circumbinary disk accretion into spinning black hole binaries. *Astrophysical Journal*, 913(1) doi:10.3847/1538-4357/abf0af

  Mahlmann, J. F., Aloy, M. A., Mewes, V., & Cerdá-Durán, P. (2021). Computational general relativistic force-free electrodynamics: II. characterization of numerical diffusivity. *Astronomy and Astrophysics*, 647. doi:10.1051/0004-6361/202038908

  Mahlmann, J. F., Aloy, M. A., Mewes, V., & Cerdá-Durán, P. (2021). Computational general relativistic force-free electrodynamics: I. multi-coordinate implementation and testing. *Astronomy and Astrophysics*, 647. doi:10.1051/0004-6361/202038907

  Del Rio, A., Sanchis-Gual, N., Mewes, V., Agullo, I., Font, J. A., & Navarro-Salas, J. (2020). Spontaneous creation of circularly polarized photons in chiral astrophysical systems. *Physical Review Letters*, 124(21). doi:10.1103/PhysRevLett.124.211301

  Mewes, V., Zlochower, Y., Campanelli, M., Baumgarte, T. W., Etienne, Z. B., Armengol, F. G. L., & Cipolletta, F. (2020). Numerical relativity in spherical coordinates: A new dynamical spacetime and general relativistic MHD evolution framework for the einstein toolkit. *Physical Review D*, 101(10). doi:10.1103/PhysRevD.101.104007

- Paul Mott joined the Advanced Computing in Chemistry and Materials Group in December 2019 after earning his PhD in theoretical chemistry from the University of Tennessee, Knoxville. His research focused on implementing path integral quantum Monte Carlo methods to explore the role of zero-point motion and three-body interactions in the lattice dynamics of solid He-4 systems. While at ORNL, Mott collaborated with the GAMMESS ECP project to optimize their software for the Frontier exascale system. Mott left ORNL in December 2021 to work at The University of Tennessee, Knoxville.

- Charles (CJ) Stapleford joined the Nuclear, Particle, and Astrophysics group in October 2020 after obtaining his PhD from North Carolina (NC) State University. His thesis work was on neutrino oscillation physics in core-collapse supernova environments, including the use of Boltzmann transport solvers coupled to approximate oscillation physics. Stapleford is continuing and expanding this work with the ECP ExaStar collaboration and has worked primarily on establishing a verification suite for new neutrino interaction models used in the FLASH-X code and the inclusion of muonic degrees of freedom in both the matter equation of state and the neutrino transport.

Risk Management

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2022

5. RISK MANAGEMENT

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

OLCF RESPONSE: Yes, the OLCF has a very successful history of 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, when appropriate, retired, reclassified, or mitigated. A change history is maintained for historical reference.

The major operational risks for the OLCF in CY 2021 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 as the risk management approach is to continuously review and assess for new risks, that could change.

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 (OLCF-5 during CY 2021). Each project execution plan refers to the main Risk Management Plan but may 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, 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, 24 active entries are in the OLCF operations risk register that fall into two categories: risks for the entire facility and risks for a specific portion of the facility. Facility-wide risks are concerned with such issues as safety, funding, expenditures, and staffing. The specific, 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 2021

Table 5.1 contains the major risks tracked for OLCF operations in 2021. The selected risks are all rated medium or high in impact.

Table 5.1. 2021 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, daily toolbox safety meetings, periodic surveillances using independent safety professionals, joint walk-downs by management and work supervisors, and encouraging 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.
1006: Inability to acquire sufficient staff	Medium/low	Accept	The OLCF reduced the probability of encountering this risk to medium in 2015 and has maintained that rating through aggressive hiring and extensive succession planning. The number of open positions has been lower than the threshold determined to trigger this risk (10%). Succession candidates have been identified for key positions.
1245: System unavailability due to mechanical/electrical system failure	Low/high	Mitigating	System was designed with leak detection from the start; sensors have been triggered during preventative maintenance activities and also indicated issues when a new cabinet was brought in and the cooling water connection was out of alignment. Perform all preventative maintenance activities; perform inspections and monitoring where possible.

5.3 NEW OR RECHARACTERIZED RISKS SINCE LAST REVIEW

5.3.1 Recharacterized Risks

In 2021, there were no recharacterized risks in the OLCF Operations Risk Register.

5.3.2 New Risks in This Reporting Period

In calendar year 2021, there were no new risks entered into the OLCF Operations Risk Register.

5.4 RISKS RETIRED DURING THE CURRENT YEAR

In calendar year 2021, there were no risks retired from the OLCF Operations Risk Register.

5.5 MAJOR RISKS FOR NEXT YEAR

Summit's operations are critical to the success of the OLCF in 2022, and one risk directly impacts Summit's ability to operate: Risk ID 1245—System unavailability due to mechanical/electrical system failure.

Construction activities are winding down in and around Building 5600. The remaining activities may require closure of some areas and ingress/egress paths for staff and visitors. These closures require increased focus on safety in and around the work areas by all staff (Risk ID 723).

Finally, the programming environment for the OLCF-5 system will be different enough that experiences on Summit may be insufficient to prepare the user community for Frontier (Risk ID 1063). The OLCF is working closely with IBM, the HPC community and standards bodies, and Cray/HPE to deploy tools and versions of the programming environment and compilers to make the transition as smooth as possible.

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

The following risks in Table 5.2 were encountered and effectively mitigated in 2021. A short summary of the status and impact of the risk on the operations of the OLCF is included.

Table 5.2. Risks encountered and effectively mitigated in CY 2021.

Risk No. 1328	OLCF Operations impacted by COVID-19 outbreak		
Risk owner	Georgia Tourassi		
Status	Mitigate		
Probability	Low		
Impact	Cost: Low	Schedule: Low	Scope/Tech: Low
Trigger Event	On March 16, 2020, the laboratory instructed all staff who could effectively work from home to do so until further notice.		
Mitigations	ORNL has added capacity to remote access tools like VPN and Citrix. Additionally, direct connect for Windows systems reduces consumption on VPN and Citrix.		
Triggers	Feedback from users or increased unscheduled downtime during the work-from-home period.		

Table 5.2. Risks encountered and effectively mitigated in CY 2021 (Continued).

Risk No. 1079	OLCF-4 post deployment issues
Risk owner	Don E. Maxwell
Status	Mitigate
Probability	Medium
Impact	Cost: Low Schedule: Low Scope/Tech: Low
Trigger Event	In 2019, a set of users experienced issues with Summit nodes when their codes were running. After a full root cause analysis with IBM, it was determined that a defect exists in the power supply for the AC922 nodes that is only exacerbated by fast/high frequency transient current. IBM had developed a strategy to replace all the power supplies in Summit in CY 2020 but was impacted by the COVID-19 pandemic, and that activity was completed in CY 2021.
Mitigations	Working closely with the vendor to track and identify root cause of every failure as quickly as possible.
Triggers	Problems encountered during early science and ongoing operations.

Environment Safety and Health

HIGH PERFORMANCE COMPUTING FACILITY
2020 OPERATIONAL ASSESSMENT
OAK RIDGE LEADERSHIP COMPUTING FACILITY
April 2022

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. ORNL is committed to operating under the DOE safety regulations specified in 10 CFR 851, which outlines the requirements for a worker safety/health program to ensure that DOE contractors and their workers operate a safe workplace. Additionally, 10 CFR 851 establishes procedures for investigating whether a violation of a requirement has occurred, for determining the nature and extent of any such violation, and for imposing an appropriate remedy. These safety requirements are incorporated into ORNL contracts as required compliance documents. To implement these safety requirements in a consistent manner across ORNL, UT-Battelle LLC deploys an online procedure management system referred to as the Standards-Based Management System (SBMS). Within SBMS, work control requirements describe the processes to be used in ORNL operations and R&D activities to implement integrated safety management functions and principles.

A key feature of the Integrated Safety Management (ISM) process is the development and implementation of specific work control. Research work in the OLCF is controlled by research safety summaries (RSSs), which define the scope of work, identify and analyze hazards, and establish safety procedures. Each RSS is reviewed and approved by line managers, qualified safety and health professionals, and research staff. An RSS provides the means by which ORNL management and staff plan and conduct research in a safe manner. It is used to control work, train participants, and provide information about operations and emergency services if needed. In addition to RSSs, ISM also requires work control for maintenance and non-employees. Maintenance work in the OLCF is conducted under a work plan developed by Facilities and Operations Directorate (F&O) line management and reviewed by subject-matter experts as required. Work plans are also written before maintenance work can proceed to ensure that work is conducted safely. Work by non-employees or subcontractor/vendors is performed in accordance with a hazard analysis. The subcontractor/vendor hazard analysis is a requirement included in the contract-specific language. The following highlights provide additional information regarding the subcontractor hazard analysis process.

Safety assessments are conducted for RSSs, work plans, and subcontracts, as well as inspections of job sites throughout each year. Lessons learned, safety snapshots, safety talks, and management assessments are conducted and 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 is reflected in these processes, which seeks to reduce and prevent injuries to personnel and potential exposure to hazards associated with operation of the facility.

Normal day-to-day operations of the OLCF related to small system installations, daily maintenance, and system upkeep remained safe, efficient, and effective: there were zero total recordable cases, zero Days

Away Restricted or Transferred (DARTs), and zero first aid cases in FY 2021. However, the Frontier Installation Project did experience one recordable case in 2021. That case and the center's actions related to the case are explained later in this section.

The following activities are ES&H highlights from CY 2021 for the non-Frontier work.

- The OLCF continued to operate 24/7/365 support for center operations and had several manpower challenges in 2021 due to COVID-19 cases and contact tracing quarantines. COVID protocol as established by ORNL was implemented throughout the year with the mindset to protect the operating staff. At times during the review period, the center was short staffed with operators and facility support craft personnel. During these manpower shortages, support organizations made additional manpower available and schedule/shift changes were implemented and overtime was utilized to staff the operations team 24/7/365. Managing each unique situation effectively was critical for ensuring appropriate coverage.

- Annual occupational exposure monitoring was conducted for noise in all three OLCF data centers. The results of these surveys are documented in the Research Safety Summary for the centers.

- The center safely received and deployed several new systems in all three spaces. Considerable communication about COVID-19 requirements was relayed to the various vendors/suppliers. In addition, center support personnel were not permitted to work "shoulder to shoulder" with vendors/suppliers that had traveled to ORNL.

- The Authorized Access to ORNL Computing Centers access training was revised and re-issued. Several new slides were added to communicate different access controls for specially controlled spaces. In addition, clarifications and other reported deficiencies were corrected.

- The two Research Safety Summaries (RSSs) for the spaces were revised, reviewed, approved, and re-issued for the spaces. In addition, all the required reading was completed by staff. Late in the CY, the multiple RSSs were combined into one RSS and issued for review. Most of the hazards and controls are consistent in all three spaces. The RSS format allows unique hazards for a certain space to be identified as occurring only in that space. This goal of this effort provides one consistent document, removes an additional review of like materials, and hopefully creates more efficient work control.

6.1 FRONTIER INSTALLATION

The installation of the Frontier project and the effective safety and health management of the subcontractor and their sub-tiered subcontractors was a primary focus on the OLCF this CY. Some of those highlights include:

- Worked with the subcontractor to develop an Installation Plan that highlighted key personnel, morning stand-up meetings, and other relevant process for the safe delivery of Frontier.

- Developed an installation Hazard Analysis, reviewed by the subcontractor, to be the backbone for the work control and the ES&H process.

- Briefed all incoming personnel on the hazard analysis. The briefing covered the process from responding to plant announcements, COVID-19, receiving, installing, troubleshooting, and environmental compliance. Each briefing was conducted by the ORNL Installation Manager and included expectations for questioning attitude, performing work safely, correct behaviors, and the requirement to suspend/stop work when adverse conditions or unsafe acts are noted. The briefing duration took ~45 minutes, and 99 people were briefed.

- The delivery schedule was coordinated with ORNL support personnel, including the Laboratory Shift Superintendent, Security, Fire Department, and ORNL Roads and Grounds. The coordination of receiving up to five truck loads per day eliminated campus impact to normal operations and enabled an efficient turnaround time at the facility dock.

- A process to control the facility dock and facility corridors was developed with the facility manager. In addition to the data centers, building 5600 houses over 400 people in office spaces. Many of these individuals utilize the same entry doors and hallway that were used during the off-loading of Frontier. An effective way to block those corridors and doors as well as how to communicate the frequent closures was developed and implemented. This process protected building occupants and provided the subcontractor a work area free of unwanted personnel.

6.2 FRONTIER INJURY

- As noted above, the Frontier Project did incur one recordable injury during the installation process. In late September, a sub-tiered subcontractor for the Frontier vendor fell from a step ladder while descending. The employee was on a six-foot ladder and was descending from the fourth step. A review of the video from the area yielded that the employee was descending correctly, did not have anything in his hands, and was maintaining correct contact. The root cause of the fall was that the employee’s left foot missed the third step.

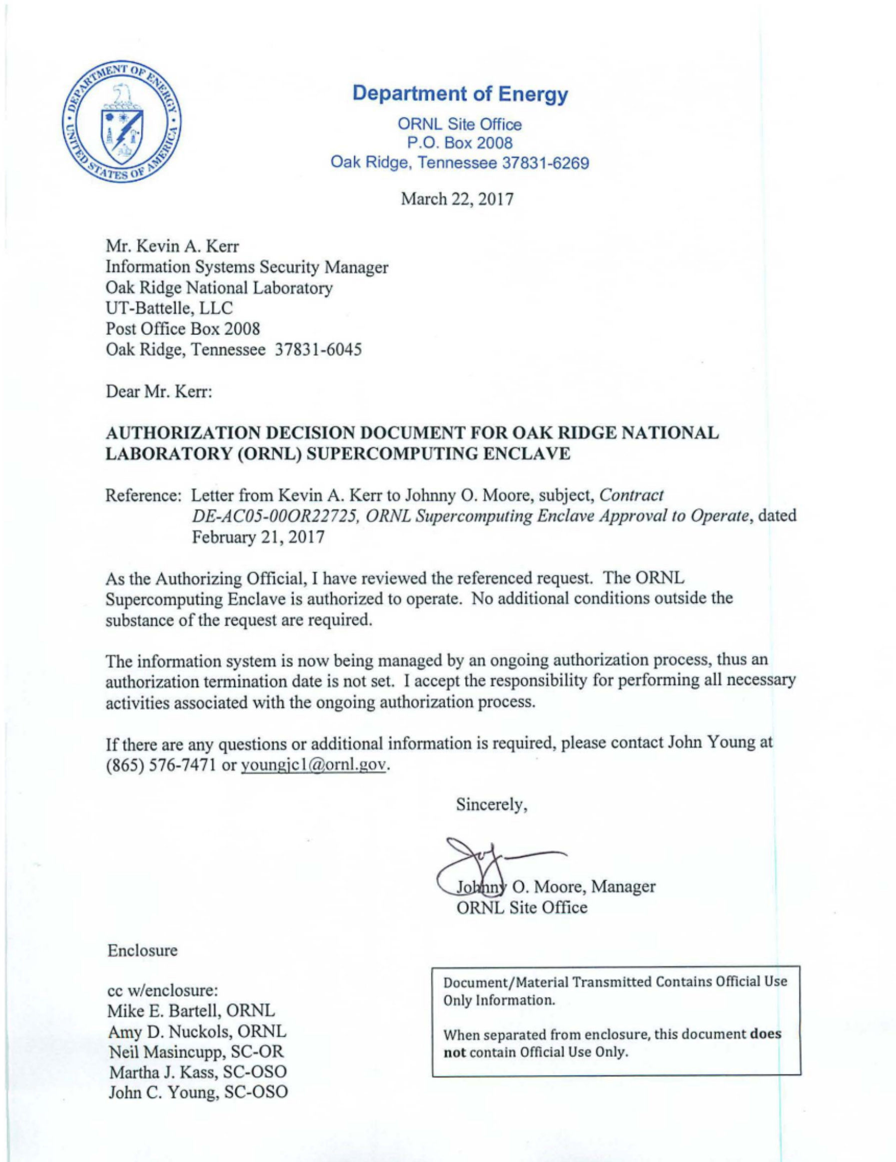
- Actions after the event
    - Employee was tended to and was immobilized until emergency medical personnel arrived for evaluation.
    - Work was ceased for the remainder of the afternoon, and all employees were dismissed except potential witnesses.
    - A safety stand-down was issued by the center manager.
    - An all-hands meeting was conducted the following morning. Expectations for attention to detail, safe work, and questioning attitude were again relayed to the team. All vendor personnel, as well as the National Science for Computational Sciences (NCCS) Division Director, attended and participated. NCCS is the ORNL division that is home to the OLCF.
    - A critique/fact-finding meeting was conducted.
    - The hazard analysis for the project was reviewed and the hazard and controls were defined.
    - The RSSs for the centers were reviewed and found to adequately address ladder safety.
    - The event and the message of “personal attention to detail” was used in multiple divisional meetings and during the remaining hazard analysis briefings for incoming project personnel.

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 February 21, 2017, and is managed through an ongoing authorization process; no authorization termination date is set (Figure 7.1). The technical staff members track and monitor for 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 operating 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 National Institute of Standards and Technology Special Publication 800-53, revision 4, 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, National Institute of Standards and Technology. The OLCF is accredited at the moderate level of controls for protecting the confidentiality and integrity of user and system information, which 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 as needed, is used to enumerate and deploy both security and operational configuration required by policy and best practices. HPC system images delivered from the vendor are augmented with Puppet configuration to bring all nodes of a system into compliance. Other important controls include enforcement of multi-factor authentication in the OLCF moderate enclave, lightweight and well-adopted configuration management procedures, adoption of DevSecOps principles such as tight inter-group coordination and use of CI, and strong incident response and triage capabilities such as operational and security dashboards and frequent practice.

In the future, cybersecurity planning will become more complex as the center continues its mission to produce great science. The center is very proactive, viewing its cybersecurity plans as dynamic documentation to which it will preemptively respond and modify as its needs change to provide an appropriately secure environment. The OLCF abides by the Health Insurance Portability and Accountability Act (HIPAA) Privacy and Security Rule to provide supercomputing resources to projects containing PHI, as well as the International Traffic and Arms Regulations (ITAR) for projects containing that type of sensitive information.

Over the past year, the focus has been on implementing a new, complex processing ability, specifically the Scalable Protected Infrastructure (SPI), also known as Citadel Version 2. The policy documentation and multiple third-party assessments have been completed, and operations started in the first quarter of the 2021 calendar year. The third-party assessments granted HIPAA certifications and Cybersecurity Maturity Model Certification (CMMC) readiness to the SPI, with the expectation that once the CMMC certifications are given, we would receive a level 3 certification.

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 where possible and 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 led the development of a software development lifecycle (SDLC) framework plan in CY 2021 in collaboration with the Software Services Development and Platforms groups. The purpose is to identify stakeholders early in the lifecycle of a project and engage other non-security operational teams to gather requirements and influence the design to harden the eventual software product. Additionally, various software scanning tools were evaluated and tested, which will result in a new software scanning capability for OLCF-developed custom software to help identify software weaknesses and potential vulnerabilities. Significant amounts of technical debt were also paid off this year: major upgrades to vulnerability scanning software versions and authenticator synchronization application containerization reduced the amount of break/fix work required to maintain those systems.

In addition to project work, the security team also tracks significant cyber security events that are above and beyond normal baseline threats in the OLCF ticket tracking system. In CY 2021, the OLCF tracked 13 above-baseline threats—none of which resulted in compromise of OLCF systems by malicious attackers. We treat each of these events as a live incident to practice the OLCF incident response plan several times throughout the year. Most of these issues are related to newly discovered vulnerabilities in the Linux kernel or other critical component of the OLCF technology stack. The rest are related to tracking interesting activity discovered by on-call security engineers or reported by admin staff or users.

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.

Security is, at its core, a data science problem. Network intrusion detection, host-based intrusion detection, system and firewall logging, vulnerability scanning, netflow collection, and databases such as hardware inventory, software inventory, and user CRM each may 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.

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 are defined in more detail 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 OLCF's cybersecurity team 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 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.	• Approved PAS is required for applicants born in, residing in, or citizens of ○ China, ○ Russia, ○ Iran, ○ Sudan, ○ Syria, ○ Crimea, ○ Cuba, and ○ North Korea. • All other applicants and their institutions go through RPS screening.	Must have valid user agreement.	N/A	Required
Category 2	PIs must approve all user account requests to access to their project.	• Approved PAS is required for all applicants that are not US citizens or lawful permanent residents unless they reside in one of the countries listed above.	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. The officer then works with the Counterintelligence Office to look at the applicant or institution in more detail and informs the OLCF whether it is acceptable to proceed.

[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 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 users' 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 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 high performance computing platforms? (c) Has the Facility demonstrated effective engagements with critical stakeholders (such as the SC Science Programs, DOE Programs, DOE National Laboratories, SC User Facilities, and/or other critical U.S. 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 US Department of Energy 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 our 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 resulting (at least in part) from 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. In addition, the Facility may report other publications where appropriate.

The Facility also regularly searches for and actively solicits information about possible scientific highlights produced by projects using 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 possibilities for features.

8.1.1 OLCF Publications Report

In 2021, 497 publications resulting from the use of OLCF resources were published, based on a data collection completed on April 5, 2022. In this document, "year" refers to the calendar year unless it carries the prefix FY. In the 2020 OLCF OAR, 355 publications were reported (this number has been significantly revised upward; see below). A list of 2014–2021 publications is available on the OLCF website (https://www.olcf.ornl.gov/leadership-science/publications/) or 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 2012–2021 period, showing consistent growth in both the total publications count and the number of publications in journals with high impact factors.

Table 8.1 Summary of unique OLCF publications for 2012–2020.

Year	Unique, confirmed OLCF publications	High-impact publications with JIF* >10
2021	497	17
2020	496	19
2019	442	20
2018	486	20
2017	472	27
2016	459	33
2015	359	21
2014	317	16
2013	368	9
2012	347	20

*JIF = Journal impact factor

8.1.2 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 outreach. The following subsections provide brief summaries of select scientific and engineering accomplishments, as well as resources for obtaining additional information. Though they cannot capture the full scope and scale of achievements enabled by the OLCF in 2021, these accomplishments advance the state of the art in science and engineering R&D across diverse disciplines and are advance DOE's science programs toward their targeted outcomes and mission goals. OLCF users published many breakthrough publications in high-impact journals in 2021, which is an additional indication of the breadth of these achievements. These publications are shown in Table 8.2.

Table 8.2. Publications in high-impact journals in 2021.

Journal	Number of publications
Nature	4
Nature Communications	2
Nature Physics	4
Advanced Materials	1
ACS Nano	3
Accounts of Chemical Research	1
Chemical Reviews	2

Altogether in 2021, OLCF users published 55 papers in journals with a journal impact factor (JIF) of greater than 7 and 17 papers in journals with a JIF greater than 10.

8.1.3 Science Highlights

8.1.3.1 Computer Simulations Shed Light on Nanomaterial Structures

A team of researchers at ORNL and ANL used supercomputers to uncover unique nanomaterial properties

PI: Paul Kent, Oak Ridge National Laboratory
Allocation Program: INCITE

The Science

2D nanomaterials, with a thickness of a single layer of atoms, have unique electrical and optical properties that make them good candidates for use in electronics and optical sensors. An increasing range of these ultrathin materials can now be synthesized, and the ways in which they emit and absorb light can change based on their structures and positions. Now, scientists have used the Summit supercomputer at the OLCF and the Theta supercomputer at the Argonne Leadership Computing Facility to more accurately simulate the optical properties of a 2D nanomaterial called germanium selenide (GeSe) and discovered that the properties of a single layer of GeSe are extremely dependent on the way its atoms are arranged. The team employed a new method that allowed the team members to model geometries of different GeSe structures.

The Impact

Attractive for applications such as solar cells and photodetectors—devices used to identify the presence of light—GeSe has been of interest in the applied physics community for some time. Understanding the properties of new and forthcoming materials such as GeSe can help scientists determine which ones warrant a deeper look. Simulations on Summit allow the team to make more confident predictions or analyze more complex materials that might have applications in photodetectors, gas sensors, and anode materials for lithium-ion batteries.

Summary

Scientists used Summit and the Theta supercomputer at the Argonne Leadership Computing Facility to more accurately simulate the optical properties of GeSe and discovered that the properties of a single layer of GeSe are extremely dependent on the way its atoms are arranged. 2D nanomaterials, with a thickness of a single layer of atoms, have unique electrical and optical properties that make them good candidates for use in electronics and optical sensors. Through an INCITE allocation, the team used Theta, Summit, and QMCPACK, an electronic structure code for quantum Monte Carlo algorithms, to accurately estimate the bandgap energies for different geometric arrangements of GeSe.

Funding

This work was supported by the DOE Office of Science, Basic Energy Sciences (BES), Materials Sciences and Engineering Division, as part of the Computational Materials Sciences Program and Center for Predictive Simulation of Functional Materials.

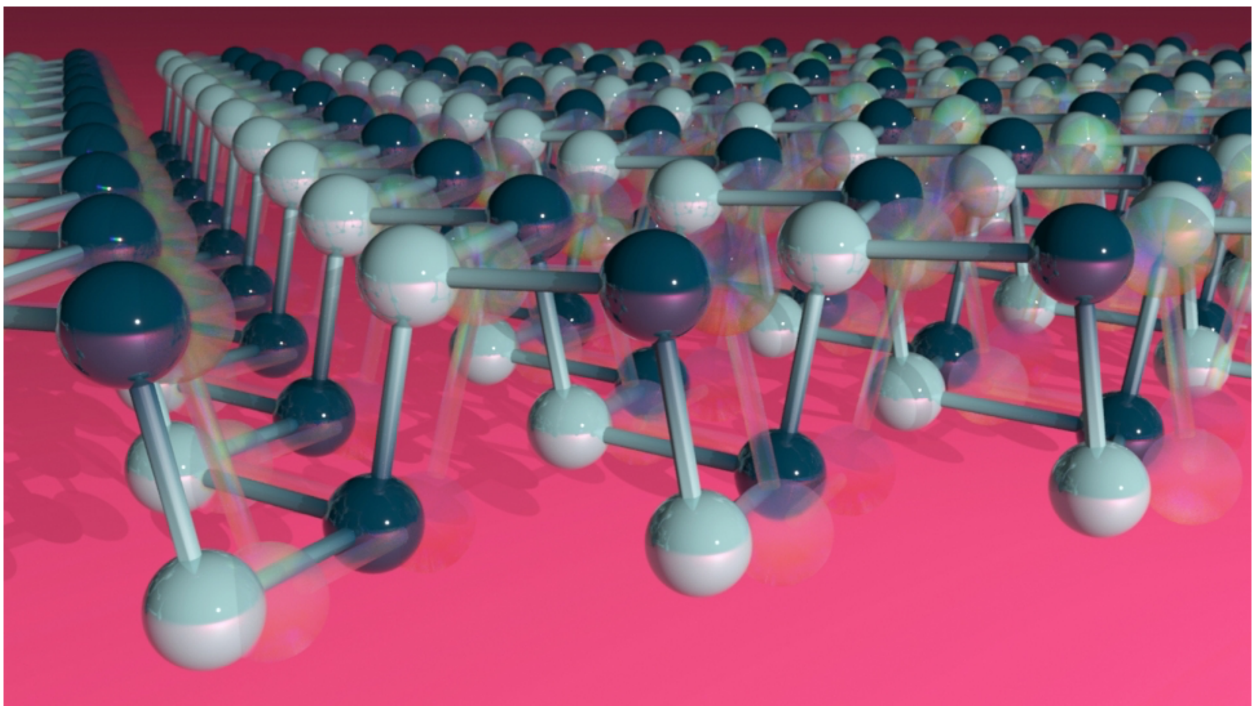


Figure 8.1. Optimized geometry of GeSe monolayer. Using a newly developed QMC structural optimization algorithm, the GeSe structure is fully optimized (colored structure) from the initial structure (clear structure). (Image: Janet Knowles, Joseph Insley, Silvio Rizzi, and Victor Mateevitsi, Argonne National Laboratory.)

Publication: Hyeondeok Shin, Jaron T. Krogel, Kevin Gasperich, Paul R. C. Kent, Anouar Benali, and Olle Heinonen, "Optimized Structure and Electronic Band Gap of Monolayer GeSe from Quantum Monte Carlo Methods," *Physical Review Materials* 5, no. 2 (2021): 024002, doi:10.1103/PhysRevMaterials.5.024002.

Related Links: "Computer Simulations Shed Light on Nanomaterial Structures," OLCF News (April 22, 2021).

8.1.3.2 Physicists Crack the Code to Signature Superconductor Kink Using Supercomputing

A team of condensed-matter physicists at LBNL uses the Summit supercomputer to understand electron behavior in superconductors.

PI: Jack Deslippe, Lawrence Berkeley National Laboratory
Allocation Program: INCITE

The Science

Over the past 35 years, scientists have investigated a special type of materials called superconductors. When cooled to the correct temperatures, these materials allow electricity to flow without resistance. One team is researching superconductors using Summit. The team found that negative particles in the superconductors interact strongly with the smallest units of light in the materials. This interaction leads to sudden changes in these materials' behavior. This interaction is at the root of understanding how a certain type of copper-based superconductor works.

The Impact

The team wanted to determine how the interactions between particles in the material change when they are in a crowded space with lots of other interacting particles. They hope that the results will help them better understand a unique class of superconducting materials based on copper. These materials will be more efficient than typical superconductors, thanks to their ability to work at relatively warm temperatures. This work could eventually lead to extremely efficient future electronic devices.

Summary

Researchers modeled the complicated interactions between negatively charged electron particles in a material and the interactions between electrons and phonons. Phonons are the smallest units of vibrational energy in a material. These models involved millions of particle states, each state comprising distinct characteristics. The result is one of the team's largest calculations to date of copper-based superconductors. The method gives the researchers a framework to study the so-called "self-energy" of electrons. The results could help the team get closer to understanding the mechanisms of a unique family of copper-based superconductors, which would be more efficient than typical copper-based superconductors.

Funding

The work was supported by the Department of Energy Office of Science through the Theory of Materials Program at LBNL and by the National Science Foundation. Advanced codes were provided by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM). The Oak Ridge Leadership Computing Facility provided computational resources in this study. The Texas Advanced Computing Center and National Energy Research Scientific Computing Center provided additional computational resources in this study.

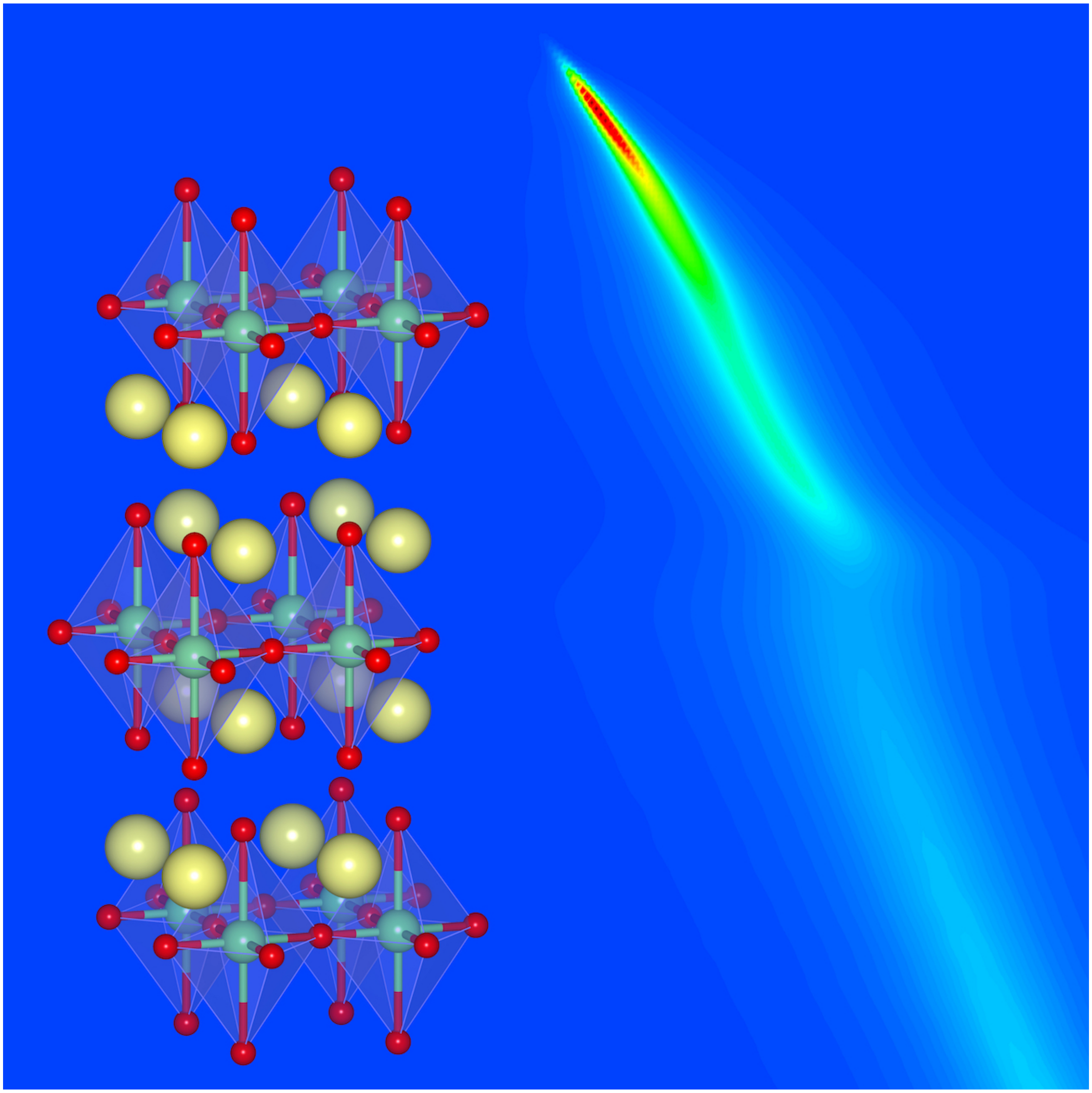


Figure 8.2. Crystal structure of a representative cuprate superconductor and simulated photoemission spectrum. (Image Credit: Zhenglu Li, Berkeley Lab.)

Publication: Zhenglu Li, Meng Wu, Yang-Hao Chan, and Steven G. Louie, "Unmasking the Origin of Kinks in the Photoemission Spectra of Cuprate Superconductors," *Physical Review Letters* 126 (2021): 146401, doi:10.1103/PhysRevLett.126.146401.

Related Link: "Physicists Crack the Code to Signature Superconductor Kink Using Supercomputing," OLCF News (May 19, 2021)

8.1.3.3 Team Earns Gordon Bell Prize Finalist Nomination for Simulating Carbon at Extreme Pressures and Temperatures

A team employed Summit, the nation's fastest supercomputer, to model the behavior of carbon under extreme temperatures and pressures.

PI: Ivan Oleynik, University of South Florida
Allocation Program: INCITE

The Science

A team used machine-learned descriptions of interatomic interactions on Summit to model more than 1 billion carbon atoms at quantum accuracy and observe how diamonds behave under some of the most extreme pressures and temperatures imaginable. The researchers found that under extreme conditions, a shock wave strongly compresses the diamond as it passes through and forces it to crack under the pressure. These cracks are healed through the formation of amorphous carbon, which is eventually converted into regions of hexagonal diamond, thus explaining the underlying mechanism of diamond's strength. The feat earned the team a finalist spot for the Association of Computing Machinery Gordon Bell Prize.

The Impact

The study will help scientists better understand how carbon behaves under extreme conditions. This understanding is crucial for inertial confinement fusion, in which hydrogen fuel is kept inside a diamond capsule and nuclear fusion reactions are initiated by compressing the collapsing diamond shell. It is also important for uncovering the internal structure of carbon-rich planets—like Uranus—and carbon-rich exoplanets. Exoplanets exist around stars outside of our solar system, and observations suggest that they may be rich in diamond and silica.

Summary

A research team led by scientists at the University of South Florida (USF), DOE's Sandia National Laboratories, DOE's National Energy Research Scientific Computing Center (NERSC), and the NVIDIA Corporation, used machine-learned descriptions of interatomic interactions on the 200 petaflop Summit supercomputer at Oak Ridge National Laboratory to model more than a billion carbon atoms at quantum accuracy and observe how diamonds behave under some of the most extreme pressures and temperatures imaginable. The study will help scientists better understand how carbon behaves under extreme conditions, which is crucial for inertial confinement fusion, in which hydrogen fuel is kept inside a diamond capsule and nuclear fusion reactions are initiated by compressing the collapsing diamond shell.

Funding

The research is supported by NNSA; the Exascale Computing Project, a collaborative effort of the DOE's Office of Science and NNSA; and DOE's Advanced Scientific Computing Research Leadership Computing Challenge and Innovative and Novel Computational Impact on Theory and Experiment awards. This research used resources of NERSC and the OLCF.

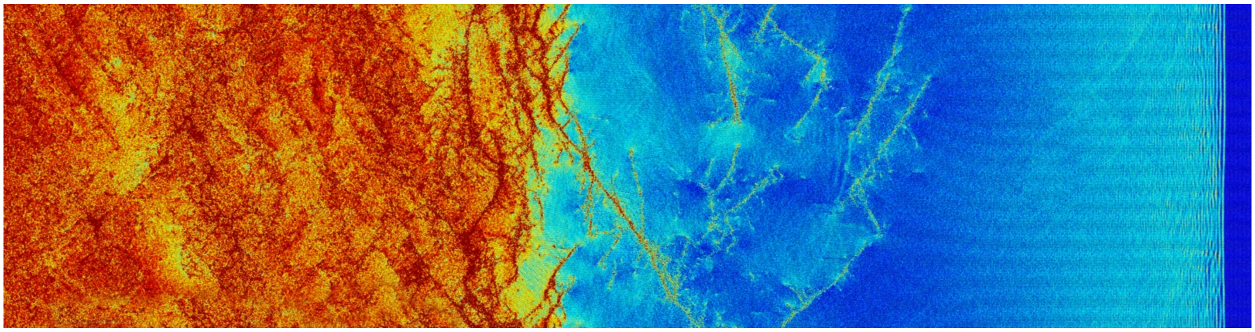


Figure 8.3. The Oleynik team simulated a split elastic-inelastic shock wave moving through a single crystal diamond. (Image Credit: Jonathan Willman, Materials Simulation Laboratory, Oleynik's Group at USF.)

Publication: Nguyen-Cong, Kien, Jonathan T. Willman, Stan G. Moore, Anatoly B. Belonoshko, Rahulkumar Gayatri, Evan Weinberg, Mitchell A. Wood, Aidan P. Thompson, and Ivan I. Oleynik. "Billion Atom Molecular Dynamics Simulations of Carbon at Extreme Conditions and Experimental Time and Length Scales." Paper presented at SC21: The International Conference for High Performance Computing, Networking, Storage and Analysis, St. Louis, MO, November 2021.

Related Link: "Team Earns Gordon Bell Prize Finalist Nomination for Simulating Carbon at Extreme Pressures and Temperatures," OLCF News (Nov. 16, 2021)

8.1.3.4 Darwin on Fast Forward: ORNL Study on COVID-19 Earns Gordon Bell Special Prize Nomination

Scientists used the nation's fastest supercomputer to teach an algorithm the language of molecules to search for COVID-19 treatments.

PI: John Gounley, Oak Ridge National Laboratory
Allocation Program: ECP and DD

The Science

A team led by John Gounley and Andrew Blanchard at ORNL used a DL language model known as Bidirectional Encoder Representations from Transformers, or BERT, to sort through billions of chemical sequences to find molecules that might block two of the primary protein components of the novel coronavirus, SARS-CoV-2. Blanchard and Gounley used Summit, ORNL's 200 petaflop supercomputer, to pretrain BERT on a dataset of 9.6 billion molecules. That training enabled BERT to perform work in a matter of hours that might have otherwise taken years. The huge amount of data—nearly 10 times the size of previous training sets for molecular language models—finished training the model in only 2 hours by running on 4,032 of Summit's nodes. That process taught the model to recognize the most minute details of the myriad potential molecular combinations.

The Impact

BERT relied on data recognition and NLP techniques honed on Summit, the nation's fastest supercomputer, to sift through the molecular combinations, each represented as a simple text sequence. This approach could work for any virus. The training aspects of the process are what's most compute-intensive, and that training won't have to be done every time. The team continues to refine the model, and biomedical researchers have begun experiments to test the most promising inhibitors suggested by the results. Next steps include tweaking the model to predict how molecules could bind and to provide detailed explanations behind the suggestions and rankings.

Summary

An ORNL team used Summit to streamline the search for potential COVID-19 treatments. They used a DL language model, BERT, to sort through billions of chemical sequences to find molecules that might block two of the primary protein components of the coronavirus. BERT relied on data recognition and natural language processing techniques honed on Summit to sift through the molecular combinations, each represented as a simple text sequence. That training enabled BERT to perform work in a matter of hours that might have otherwise taken years. The study earned the team a finalist nomination for the Association of Computing Machinery Gordon Bell Special Prize for High Performance Computing–Based COVID-19 Research.

Funding

This research was supported by the ECP and the DOE's National Virtual Biotechnology Laboratory, with funding provided by the Coronavirus CARES Act and the DOE Office of Science's Advanced Scientific Computing Research program. The OLCF is a DOE Office of Science user facility.

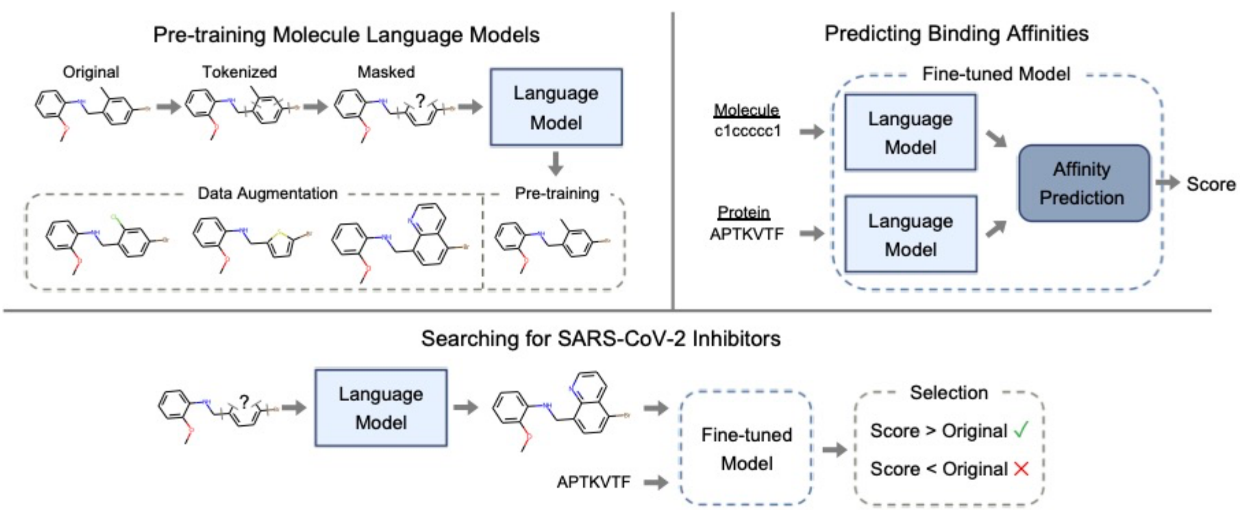


Figure 8.4. An ORNL study used a DL language model to sort through billions of chemical sequences to find molecules that might block two of the primary protein components of the coronavirus. (Credit: Blanchard, et al./ORNL.)

Publication: Andrew Blanchard, John Gounley, Debsindhu Bhowmik, Mayanka Chandra Shekar, Isaac Lyngaas, Shang Gao, Junqi Yin,Aristeidis Tsaris, Feiyi Wang, and Jens Glaser. "Language Models for the Prediction of SARS-CoV-2 Inhibitors." Paper Presented at the International Journal of High Performance Computing Applications, 2021.

Related Link: "Darwin on Fast Forward: ORNL Study on COVID-19 Earns Gordon Bell Special Prize Nomination," OLCF News (Nov. 16, 2021)

8.1.3.5 Closing In on Fusion

A team modeled plasma turbulence on the nation's fastest supercomputer to better understand plasma behavior.

PI: Emily Belli, General Atomics
Allocation Program: ALCC and INCITE

The Science

The same process that fuels stars could one day be used to generate massive amounts of power here on Earth. Nuclear fusion—in which atomic nuclei fuse to form heavier nuclei and release energy in the process—promises to be a long-term, sustainable, and safe form of energy. But scientists are still trying to fine-tune the process of creating net fusion power. A team led by computational physicist Emily Belli of General Atomics used Summit to simulate energy loss in fusion plasmas. The team used Summit to model plasma turbulence, the unsteady movement of plasma, in a nuclear fusion device called a tokamak. The team's simulations will help inform the design of next-generation tokamaks like ITER—the world's largest tokamak, which is being built in the south of France—with optimum confinement properties.

The Impact

Until now, almost all fusion simulations have included only deuterium or tritium isotopes, but Summit enabled the team to include both as two separate species, model the full dimensions of the problem, and resolve it at different time and spatial scales. The results provided estimates for the particle and heat losses to be expected in future tokamaks and will help scientists and engineers understand how to achieve the best operating scenarios in real-world tokamaks.

Summary

A team led by computational physicist Emily Belli of General Atomics used the 200 petaflop Summit supercomputer at the OLCF to simulate energy loss in fusion plasmas. The team used Summit to model plasma turbulence, the unsteady movement of plasma, in a nuclear fusion device called a tokamak. The team's simulations will help inform the design of next-generation fusion devices with optimum confinement properties. The insights can also inform operations of ITER, which will be the world's largest tokamak and is now under construction in France.

Funding

This work was supported by the US Department of Energy under the Edge Simulation Laboratory and AToM SciDAC-4 project. An award of computer time was provided by the INCITE program and the ALCC program. This research used resources of the Oak Ridge Leadership Computing Facility and the National Energy Research Scientific Computing Center, DOE Office of Science User Facilities.

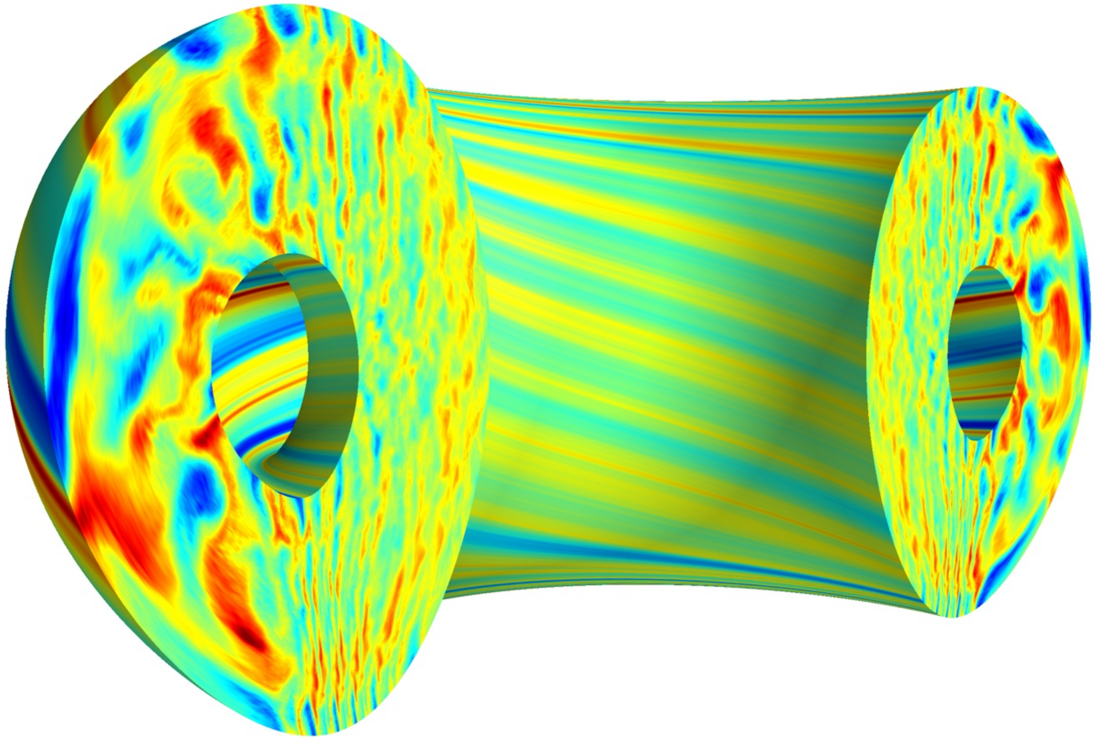


Figure 8.5. A visualization of deuterium-tritium density fluctuations in a tokamak driven by turbulence. Areas of red are representative of high density and areas of blue are representative of low density. (Image Credit: Emily Belli, General Atomics.)

Publication: Emily A. Belli and Jeff Candy, "Asymmetry between Deuterium and Tritium Turbulent Particle Flows," *Physics of Plasmas* 28, no. 6 (2021), doi:10.1063/5.0048620.

Related Link: "Closing In on Fusion," OLCF News (Dec. 13, 2021)

8.2 RESEARCH ACTIVITIES / VENDOR ENGAGEMENT FOR FUTURE OPERATIONS

8.2.1 Summit on Summit

A "Summit on Summit" is regularly organized by NVIDIA and OLCF to engage with the OLCF in highly technical discussions, including working groups focused on delivering solutions in support of current and future users of the OLCF facility. It has grown to include Lawrence Livermore National Laboratory (LLNL) and NERSC to support their accelerated computer systems Sierra and Perlmutter, respectively. The so-called Workstreams set short-term deliverables and report back every four months on progress made, with topical areas including compilers, math libraries, data center monitoring, development tools, portable memory, communication, data analytics, reduced precision, coding for performance, python, Monte Carlo algorithms, structured and unstructured mesh, particle methods, ADIOS, GPU-direct Storage, HPC and AI, HPC at the Edge, ARM on NVIDIA GPUs, and multi-tenancy. During 2021, virtual meetings were held every 4 months in lieu of the in-person workshops held at NVIDIA in Santa Clara before the pandemic. The Workstreams are led by NVIDIA staff, and the milestones and deliverables are tracked monthly by both NVIDIA and the OLCF.

8.2.2 Accelerated Data Analytics and Computing (ADAC) Institute

In 2016, a consortium of HPC computer centers was created with ORNL, Tokyo Tech, and ETH Zürich as the founding members. Since then, the ADAC institute has grown to eleven members representing computer centers in Australia, Japan, the United States, and Europe. The activities are organized in three topic areas: Applications, Resource Management, and Performance Portability. Twice a year the institute meets for a workshop to discuss progress on joint activities and planning for the next 6 month period. Vendors also participate in these workshops, including NVIDIA. Work in 2021 focused on heterogeneous computing, from using distinct computer components in the same system to leveraging computing resources across wide-area networks. In 2021, the bi-annual meetings were held in a virtual setting.

8.2.3 Enabling Novel Capability on OLCF Machines

In mid-2021, Google's Deepmind published a paper in *Nature* describing Alphafold2—a DL-based computational approach for predicting protein structures. This method far outperformed existing methods. A team at ORNL wanted to scale the use of Alphafold2 to much larger problem sizes with Summit. The OLCF experimented with building Alphafold2 natively on Summit, but this proved to be a problem due to lack of support for Summit's PowerPC architecture. Individual components had to be rebuilt from scratch because the existing publicly available builds would not work. Alphafold2's build system also had specific file system structure requirements that were incompatible with native Summit. The OLCF worked on building container images that compiled Alphafold2 and all its components that they could run at scale, with patches to support Summit's architecture and working around the file system requirements. Containerizing Alphafold2 was the critical step. It was deployed and orchestrated at scale on Summit with Dask.

Related paper: The documentation of this effort and the results were published in SC's 7th Workshop on Machine Learning in High Performance Environments. The Alphafold2 container image has also been made publicly available for other researchers to run their own experiments with the application, along with instructions for deployment at scale.

8.2.4 ML/DL Acceptance, Comparative analysis

With Frontier acceptance ramping up, the OLCF will soon be operating two leadership-class supercomputers. Given the architecture differences—one being accelerated by NVIDIA GPUs (Summit) and the other by AMD GPUs (Frontier)—this poses some challenges in both managing the systems and understanding their performance, especially for ML/DL workloads, which are becoming important factors to consider as we design, develop, and deploy next-generation HPC systems. To gain a holistic understanding from compute kernels, models, and frameworks of DL stacks, and to assess their impact on science-driven, mission-critical applications, we performed a comparative study between the NVIDIA V100–based Summit system with its CUDA stack and an AMD MI100–based tested system with its ROCm stack. The insight from this evaluation will help with both the operation of current systems and the design of future ones.

Related paper: "Comparative Evaluation of Deep Learning Workloads for Leadership-class Systems," Junqi Yin, Aristeidis Tsaris, Sajal Dash, Ross Miller, Feiyi Wang, and Mallikarjun (Arjun) Shankar, Bench Council Transactions on Benchmarks, Standards and Evaluations, 2021. (Best paper.)

8.2.5 Exploring Power, Energy, and Thermal Dynamics of the Summit Supercomputer through Analysis of High-Volume Operational Data

A data science team at the OLCF performed a comprehensive analysis on the accumulated data from the "Cooling Intelligence for Summit" project to study power, energy, and thermal behavior and impact of the Summit system and its cooling and electrical infrastructure (See Figure 8.6 and Figure 8.7). This study was a foundational work in understanding Summit's power and energy operational data and its use for advanced analytics and AI/ML based methods for OLCF systems, especially for energy-efficient and sustainable HPC in the post-exascale era. The work classified job classes and their power consumption on Summit. The result of this analysis was submitted, accepted, and won the best paper award at SC21.

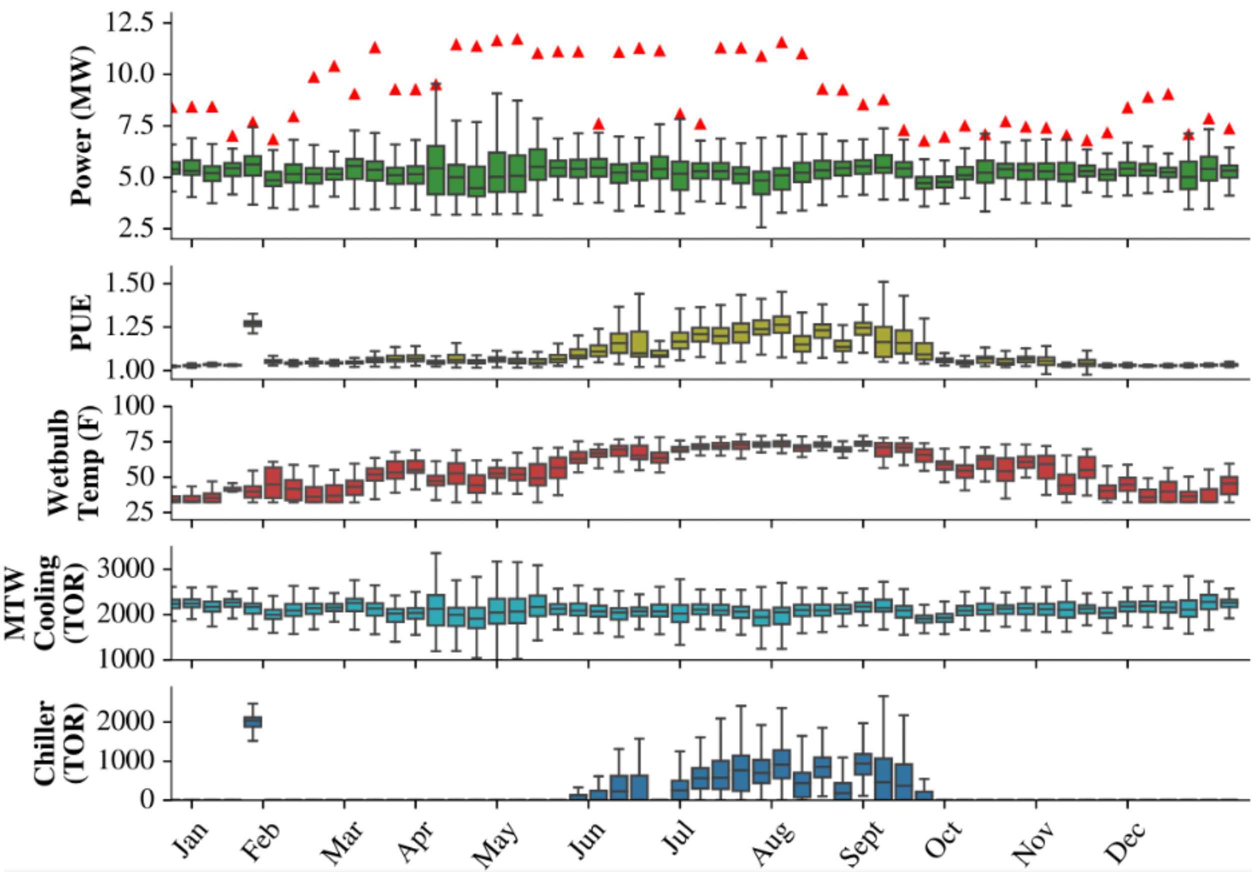


Figure 8.6. Summit power and energy trends (CY 2020).

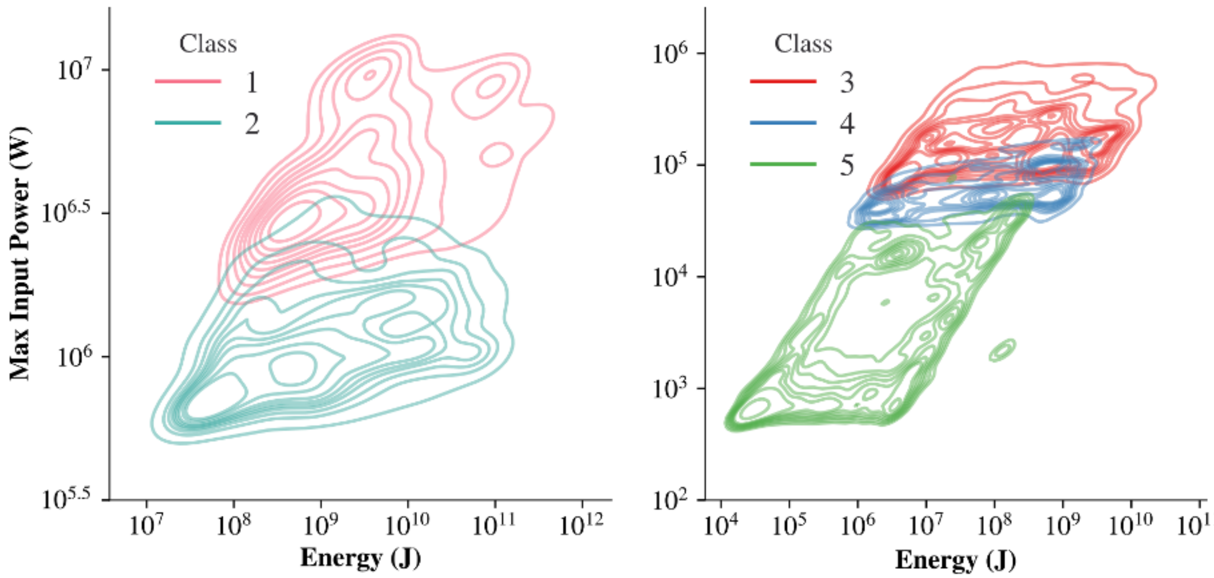


Figure 8.7. Distribution of the total energy consumed during a job run versus the maximum input power for each Summit scheduling class.

Related paper: Woong Shin, Vladyslav Oles, Ahmad Maroof Karimi, J. Austin Ellis, and Feiyi Wang. 2021. "Revealing power, energy and thermal dynamics of a 200PF pre-exascale supercomputer." In Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC21). Association for Computing Machinery, New York, NY, USA, Article 12, 1–14. DOI: https://doi.org/10.1145/3458817.3476188

8.2.6 Industry Engagement

ACCEL, the OLCF's industrial partnership program, "stayed the course" in 2021 as industry and the nation continued to manage through the many challenges brought on by the ongoing COVID-19 pandemic. Thirty-seven industry projects were underway on Summit during the year: 18 new projects launched, and 19 that began prior to 2021 continued throughout the year.

- The 37 industrial projects underway in 2021 represented 10% of the total number of projects provided to external user programs, INCITE, ALCC, and DD (including ECP).
- These projects used 3,913,375 Summit node hours, representing approximately 11% of the total Summit hours.
- 51% of the total industrial project hours on Summit were allocated through INCITE, 42% via ALCC, and 7% through the OLCF DD program.
- The 18 new projects received their awards via INCITE (6 projects), ALCC (4 projects), and DD (8 projects).

8.2.6.1 Observations about the Industrial Projects

- This was the strongest year yet for industry INCITE awards. Industry had six new awards, along with one INCITE project that began in 2020 and continued into 2021, for a total of seven INCITE projects.

- Industry partnered with universities in three of the new industry INCITE awards as companies and universities realize the benefits of collaborating to present compelling proposals and experienced teams to the INCITE review committee.

- With INCITE providing the largest computational allocations across the three external user programs, these seven projects accounted for over half of the hours allocated to the 37 industry projects underway in 2021.

- Numeca (now Cadence), the first commercial application software vendor (sometimes referred to as an Independent Software Vendor/ISV) to receive an INCITE award at the OLCF (in 2020), received another award in 2021.

- Industry continued to fare well in the ALCC program with four new awards. And GM returned to continue research into predicting engine knock with a DD award.

- GE received a DD allocation to test the ECP's E4S software stack. GE is the first (and only) OLCF industrial user to do this, reflecting the advancements the company has made through its projects at the OLCF in using leadership-scale systems to address larger and more complex problems that exceed its internal computing capabilities.

- Cascade Technologies, the only other commercial application software vendor who has invested the time and resources to port its application software to Summit, was a co-PI on an ALCC award with Stanford, MIT, and Penn State. The team is further exploring the use of high-fidelity large eddy simulation (LES) techniques to achieve higher efficiency turbomachinery designs for products such as jet engines and gas turbines for power generation. Summit's scale anticipates that engineers soon will be able to use these detailed and highly computationally intensive techniques to predict heat flows in compressors and heat transfer characteristics on turbine blades within engineering design cycle times—something that is not possible today.

8.2.6.2 Industrial Project Highlight

GE Spins up Supercomputer Models to Zero in on Energy Loss in Turbines
A team at GE Aviation and the University of Melbourne is studying turbulent flows on the Summit supercomputer for better engines.

PI: Richard Sandberg, Univ. Of Melbourne; Sriram Shankaran, GE Aviation
Allocation Program: INCITE

The Science

High-pressure turbines are vital components of gas turbines used to propel jet engines. The more efficient these jet engines are, the better they are for the aircraft industry and their customers. But these large, dynamic systems are difficult to study via experiments and physical testing. A team led by scientists at GE Aviation and the University of Melbourne used the Summit supercomputer to run for the first time real-engine cases capturing the largest eddies, or circular fluid movements, down to those that were tens of microns away from the turbine blade surface. From the simulations, the researchers determined which regions near a turbine blade experience a greater loss of energy. For the case with the highest Mach number, which describes the flow's velocity compared with the speed of sound, they discovered an extra loss of energy resulting from strong shock waves, or violent changes in pressure, that interact with the edge and wake of the flow to cause a massive amount of turbulence.

The Impact

More accurate prediction of real engine conditions will lead to more efficient engines that consume less fuel and other positive derivative effects. A 1% reduction in fuel consumption across a fleet of engines is equal to about 1 billion dollars a year in fuel cost savings. Reduced fuel consumption also translates into reduced emissions—a 1% reduction in fuel burn reduces CO_2 emissions by roughly 1.5%.

Summary

A team led by scientists at GE Aviation and the University of Melbourne has been using supercomputers to model turbulent flows, tumultuous mixtures of combusted fuel and air, for the last decade to better determine the effects of turbulence on performance (See Figure 8.8). The problem, though, is that the models often used for turbulence are not entirely accurate. Using the HiPSTAR code on Summit, the team ran for the first time real-engine cases—a total of five of them—capturing the largest eddies all the way down to those that were tens of microns away from the blade surface. Specifically, they performed direct numerical simulations (DNS), CFD simulations that directly capture the full range of scales of turbulence without using separate models that can estimate only turbulent effects.

Funding

This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project ID s977. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.

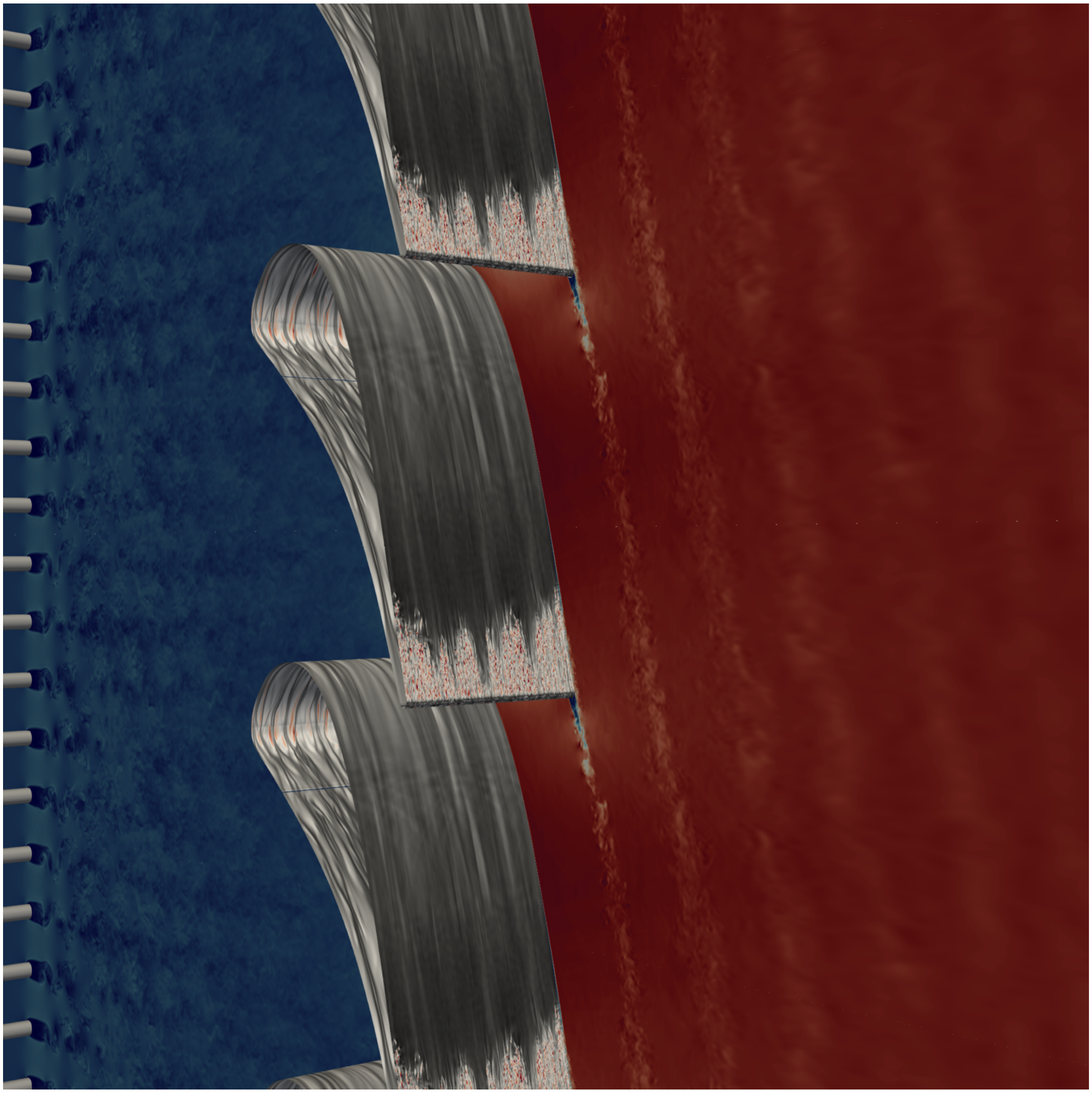


Figure 8.8. A row of upstream bars produces highly turbulent flow that gets accelerated through a high-pressure turbine blade row and interacts with the blade surface, causing significant temperature variations. (Image Credit: Richard Sandberg, University of Melbourne.)

Publication**:** Y. Zhao and R. D. Sandberg, "High-Fidelity Simulations of a High-Pressure Turbine Vane Subject to Large Disturbances: Effect of Exit Mach Number on Losses," Journal of Turbomachinery 143, no. 9 (2021), doi:10.1115/1.4050453.

Related Link**:** "GE Spins up Supercomputer Models to Zero in on Energy Loss in Turbines," OLCF News (June 9, 2021).

8.3 DOE PROGRAM ENGAGEMENTS / REQUIREMENTS GATHERING

Through ORNL and DOE partnerships, the OLCF cultivates a number of strategic engagements with other DOE Office of Science programs, DOE Applied Programs, the Exascale Computing Project, and other Federal agencies with the goals 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. 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 SC Observational Facilities Investigators

Although the use of OLCF by PIs funded by DOE SC program offices is commonplace, in recent years, many PIs have explored discretionary allocations to establish workflows from their observational facilities to OLCF resources. Office of Science Biological and Environmental Research (BER)-sponsored investigators in the Atmospheric Radiation Monitoring program use operational data workflows originating from their sensors through their data management facility using institutional resources and have set up exploratory efforts with OLCF data and analytics (JupyterHub) portals. BES-sponsored scientists who are users of facilities such as ORNL's Center for Nanophase Materials Science (CNMS) and the Spallation Neutron Source (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 data movement 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). The lessons learned from workflows for HEP 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 the deployment of workflows on Slate for Summit, and the design of the Frontier login nodes.

8.3.2 Ongoing Engagements with Atmospheric Radiation Measurement (ARM)

NCCS operates ARM's two computing clusters, Stratus and Cumulus, which process petabytes of atmospheric data that is collected from hundreds of instruments and observation facilities around the world. The OLCF offers data management support and expertise in archival storage, ensuring access to data required for critical climate research and ensuring safe long-term storage.

The hosting of ARM's HPC resources and data, coupled at OLCF, enables several benefits to both the OLCF and ARM. One of the primary ARM projects, actively using Stratus and Cumulus at OLCF, is the LES ARM Symbiotic Simulation and Observability (LASSO) project. The LASSO project has developed "data bundles" that combine ARM observations and high-resolution model output, these bundles are then made accessible to researchers through a searchable web interface. Supporting projects like this showcases the unique data and compute service capabilities that can be provided, targeting a global research community, using underlying infrastructure hosted within NCCS.

8.3.3 Engagement with the National Institutes of Health (NIH) and the National Cancer Institute (NCI)

The collaboration between ORNL and NIH/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.

The research collaboration between NIH/NCI and ORNL, known as Modeling Outcomes using Surveillance data and Scalable Artificial Intelligence for Cancer (MOSSAIC), 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 our models in real-world cancer registry settings; and (5) enabling large-scale Transformer training on leadership computing facility (LCF) systems. Collectively, these efforts aim at modernizing the national cancer surveillance program and enabling the US achieve near real-time cancer surveillance.

This research poses several significant computational challenges. Transformer models require pre-training on very large datasets with billions of elements, which requires access to large HPC resources, as these models need to be trained on hundreds of Summit nodes using data parallelism. Additionally, Transformer models need to be trained on the raw text of cancer pathology reports, which includes protected health information (PHI).

To better support these efforts, the OLCF launched a security framework called CITADEL, allowing researchers to use supercomputers for research using protected health data for the first time. ORNL's unmatched combination of secure PHI enclave and the CITADEL framework – a protected computing environment —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 be leveraged for the benefit of multiple programs, including other OLCF user communities.

The 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 enables 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 Application Program Interfaces (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.

8.3.4 Engagement with the US Department of Veterans Affairs (VA)

In 2016 the US Department of Veterans Affairs (VA) partnered with DOE and ORNL to revolutionize the health care of veterans, and by extension all Americans, via advanced data analytics and high-performance computing. The DOE-VA collaboration advances the mission of both the VA and DOE and leverages resources unique to each entity to create a symbiotic relationship allowing for efforts that would not otherwise exist. The VA's Office of Research and Development (ORD) uses DOE's supercomputing facilities, deep expertise in big data, and high-performance computing 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 Summit and computationally efficient transformer models to capture interactions within and between genes, a capability not viable without supercomputers.

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

8.3.5 Engagement with Air Force Weather (AFW) and National Oceanic and Atmospheric Administration (NOAA)

ORNL and AFW launched Miller and Fawbush, two Hewlett Packard Enterprise (HPE) Cray EX supercomputers, reaching Certification of Operational Readiness in February 2021, to support the AFW's numerical weather modeling at much higher speed and fidelity. Thanks to the powerful systems, AFW is now 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 our 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-learned paid very significant benefits to the OLCF-5 Frontier project, supporting key decisions early in that project, in advance of the delivery of the initial test and development systems.

ORNL and NOAA established the National Climate-Computing Research Center through a strategic partnership in 2009 with a goal to develop, test, and apply state-of-the science computer-based global climate simulation models, based upon a strong scientific foundation while leveraging leading edge high-performance computing and information technologies. In 2021, based on success achieved over the program's lifespan, the program was extended another five years. The NCCS hosts the program's primary computation resource named Gaea which consists of a pair of HPE XC40 HPC systems providing 5.29PF peak computation. The NCRC supports and complements NOAA's climate mission to understand climate variability and change to enhance 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, nurturing 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. A collaborative project, which began in 2021 and includes members from the NCCS and NOAA, aims to develop the ability to run portable workloads within containers on NOAA's HPC resources. The NCCS collaboration members are also undertaking similar container efforts for the OLCF Summit resource. Through this method, lessons learned on one resource benefits both NOAA and the OLCF. The broadened perspective also helps to harden and improve the end result for all programs within the NCCS.

8.3.5.1 Other Engagement Activities

On an ongoing basis, the OLCF seeks to better understand current and future mission needs related to the Facility. Below are two examples of how the OLCF plugged into the community to better assess future needs for LCF resources.

National Science and Technology Council (NSTC)'s Biodefense R&D Interagency Working Group

A mini-workshop with the National Science and Technology Council (NSTC)'s Biodefense R&D Interagency Working Group was held on July 12 and July 22, 2021, to discuss how the LCFs can support and advance national biodefense goals and identify areas of opportunity arising from the most recent advances in biosciences with leading-edge discoveries in other areas of physical, computational, and life sciences research. The role of HPC in biodefense, anticipated computational challenges in the next 10–15 years, and related computational infrastructure needs were explored.

ASCR Workshop with the DOE Energy Efficiency and Renewable Energy Program Offices

On June 28–29, 2021, the OLCF and ALCF facilitated an ASCR workshop with the DOE Energy Efficiency and Renewable Energy program offices to assess the future need for advanced computing resources in the areas of clean energy and advanced manufacturing. In part, this discussion served as an update to earlier workshops and town halls. Participants were drawn from the communities of leading domain scientists, experts in computer science and applied mathematics, ASCR facility staff, and DOE program managers in ASCR and the respective program offices. The goal of the workshop was to identify challenges at the intersection of applied office mission priorities and the mission of the LCFs. DOE program office stakeholders identified present and future use cases, strategic objectives, and anticipated needs. Speakers were asked to consider the following questions:

- What existing efforts do you have that utilize HPC? AI and ML? Analysis of large data sets? Are current efforts limited by lack of computational resources?

- What are some challenges in your domain that will be dominant in 5–15 years and that HPC and data science can help solve?

- What computational resources (including storage and networking) are needed to achieve your 5–15 year goals?

- How will achieving these goals impact DOE mission priorities?

- If you could design the ideal computational infrastructure/ecosystem to address the challenges in your mission space, what would it look like? How would it differ from today's HPC systems, if at all?

Application domain sessions included overviews from relevant program office representatives and subject-matter experts followed by brief technical presentations from across the Lab complex, and covered the following areas:

- Transportation: Vehicles, Mobility Systems, and Biofuels
- Renewable Power: Wind Energy, Solar Energy
- Energy Efficiency: Advanced Manufacturing and Building Technologies
- Crosscuts: Electric Grid, Energy Storage, and Carbon Management

8.3.6 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. The ECP continues to fund efforts at national labs, academia, and industry with the goal of producing usable software and applications and influencing the development of hardware technology for the exascale systems in the 2021–2022 time frame. These investments are very timely and will significantly aid the OLCF in delivering capable exascale systems with robust system software and application software that can address the science gaps immediately upon delivery and acceptance of the systems. The OLCF's engagement with ECP includes three primary thrust areas, described below, as well as many pairwise and other interactions (e.g., staff involvement in ECP AD and ST projects).

OLCF-ECP Application Development/Software Technology engagements

The OLCF has a long history of readying applications for its forthcoming architectures, dating back 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 has chosen to partner with the ECP to augment the OLCF CAAR portfolio with an additional 12 ECP AD teams. These teams were selected to diversify OLCF applications readiness efforts funded through the OLCF-5 CAAR such that the OLCF will have 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 are currently partnering with the OLCF and their OLCF Scientific Computing liaison 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 (FNAL)	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 (ANL)	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

[a] FNAL = Fermi National Accelerator Laboratory, ORNL = Oak Ridge National Laboratory, Sandia = Sandia National Laboratories, PPPL = Princeton Plasma Physics Laboratory, LBNL = Lawrence Berkeley National Laboratory, ANL = Argonne National Laboratory, PNNL = Pacific Northwest National Laboratory.

With ECP funding, these ECP AD teams have dedicated 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 potential support from postdoctoral researchers through the CSEEN program as availability allows. In addition, because of the dependence of the AD project on the ECP ST, the ECP ST projects also have access to the CoE resources and access to early testbed hardware.

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

The OLCF continued to contribute to ECP project in the Hardware Integration (HI) area as well, leading or participating in efforts in application support, software integration, facility resource utilization, and training and productivity.

APPENDIX A. RESPONSES TO RECOMMENDATIONS FROM THE 2020 OPERATIONAL ASSESSMENT REVIEW

In April 2021, the operational activities of the OLCF for CY 2020 were presented to the DOE sponsor. The review committee of that report identified no recommendations.

APPENDIX B. TRAINING, WORKSHOPS, AND SEMINARS

Table B.1 lists the 2021 OLCF training and outreach events.

Table B.1. 2021 OLCF training and outreach events.

Event type	Event title	Date	Participants
Monthly User Con Call	2021 OLCF User Conference Call: JupyterHUB@OLCF	January 27	128
Monthly User Con Call	2021 OLCF User Conference Call: New User Training/Best Practices@OLCF	February 24	97
Monthly User Con Call	2021 OLCF User Conference Call: NVIDIA Rapids	March 31	100
Monthly User Con Call	2021 OLCF User Conference Call: myOLCF	April 28	70
Monthly User Con Call	2021 OLCF User Conference Call: GPU Concurrency	May 26	105
Monthly User Con Call	2021 OLCF User Conference Call: NVIDIA HPC SDK	July 28	122
Monthly User Con Call	2021 OLCF User Conference Call: NVIDIA Rapids Update@OLCF	August 25	68
Monthly User Con Call	2021 OLCF User Conference Call: OLCF's User Managed Software Program	September 29	85
Monthly User Con Call	2021 OLCF User Conference Call: Node-Local Storage on Summit	October 27	92
Monthly User Con Call	2021 OLCF User Conference Call: Analysis Tools@OLCF	December 8	78
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Extreme-scale Scientific Software Stack (E4S)	January 13	223
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Good Practices for Research Software Documentation	February 10	300
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: An Overview of the RAJA Portability Suite	March 10	168
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: A Workflow for Increasing the Quality of Scientific Software	April 7	174
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Automated Fortran-C++ Bindings for Large-Scale Scientific Applications	May 12	186
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Using the PSIP Toolkit to Achieve your Goals - A Case study at The HDF Group	June 9	75
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Mining Development Data to Understand and Improve Software Engineering Processes in HPC Projects	July 7	97
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Software Engineering Challenges and Best Practices for Multi-Institutional Scientific Software Development	August 4	221
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: What I Learned from 20 Years of Leading Open Source Projects	September 15	218
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Migrating to Heterogeneous Computing: Lessons Learned in the Sierra and El Capitan Centers of Excellence	October 13	222
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: 55+ years in High-Performance Computing: One Woman's Experiences and Perspectives	November 10	253

Event type	Event title	Date	Participants
IDEAS-ECP/Facility Webinar Series	HPC Best Practices Series: Scientific software ecosystems and communities: Why we need them and how each of us can help them thrive	December 8	148
Workshop/Training Meeting	NCI/DOE Pilot 3 Hackathon	January 21–22	46
Workshop/Training Meeting	SIAM CSE21: Hands-On Learning with the Summit Supercomputer	March 4	38
Workshop/Training Meeting	HPCToolKit for GPU Tutorial	March 29, April 2	286
Workshop/Training Meeting	OLCF Spock Training Webinar	May 20	298
Workshop/Training Meeting	HIP Training Workshop	May 24–26	127
Workshop/Training Meeting	Students' Hands-On with Summit	June 21	173
Workshop/Training Meeting	2021 OLCF User Meeting	June 22–24	193
Workshop/Training Meeting	Computational Mission Needs in Clean Energy and Manufacturing	June 28–29	117
Workshop/Training Meeting	CUDA Multithreading with Streams	July 16	103
Workshop/Training Meeting	Kokkos Training	July 16, July 19	34
Workshop/Training Meeting	Computational and Autonomous Workflows Workshop: CAW 2021	7/20–21/21	60
Workshop/Training Meeting	CUDA Multi Process Service	August 17	89
Workshop/Training Meeting	Cmake Training	August 23–26	125
Workshop/Training Meeting	Slateathon	August 25–26	34
Workshop/Training Meeting	HOLA Hands-On Summit	September 13	55
Workshop/Training Meeting	CUDA Debugging	September 14	81
Workshop/Training Meeting	Introduction to OpenMP GPU Offloading	September 22–23	72
Workshop/Training Meeting	E4S at DOE Facilities with Deep Dive at NERSC Confirmation	October 6	65
Workshop/Training Meeting	CUDA Graphs Workshop	October 13	68
Workshop/Training Meeting	Deep Learning at Scale	October 28	53
Workshop/Training Meeting	SC Student's Hands on Summit	October 31	17
Workshop/Training Meeting	OLCF Virtual Career Fair	November 11	115

Event type	Event title	Date	Participants
Workshop/Training Meeting	Introduction to Leadership Computing	November 12	34
Hackathons	Argonne National Laboratory Hackathon	April 20–29	
Hackathons	San Diego Supercomputing Center Hackathon	May 4, 11–13	
Hackathons	Princeton Hackathon	June 2, 8	55
Hackathons	NERSC Hackathon	July 19–28	
Hackathons	OLCF Hackathon	October 11, 18–27	64
Seminar Series	I/O Considerations for Artificial Intelligence/Machine Learning	January 8	
Seminar Series	Computational Approaches to Understanding Core Collapse Supernovae: An Overview of the CHIMERA and SLIP Codes	January 15	
Seminar Series	Scaling Transformers for Clinical Text Classification	January 29	
Seminar Series	Developing High Performance Computing in Astrophysics Simulations	February 4	
Seminar Series	Reactive and Non-Reactive Atomistic Simulations of Materials under Non-Equilibrium and/or Extreme Conditions	February 4	
Seminar Series	Incorporating X-ray Scattering-derived Force using GPU for Molecular Dynamics	February 8	
Seminar Series	Modeling Phase Transitions in Liquid Crystals and Polymers: Selected Topics	February 9	
Seminar Series	Healthcare AI Applications and Transcriptomics Data Pipeline Using High-Performance Computing	February 12	
Seminar Series	Towards "Extreme-Scale" Scientific Discoveries with Astrophysics Simulations	February 17	
Seminar Series	Supramolecular Assembly: How Does Local Packing of Identical Molecules Select Equilibrium Complex Crystals	February 18	
Seminar Series	Lattice Quantum Chromodynamics and the Calculation Nuclear Matrix Elements Calculations	March 5	
Seminar Series	Parallelization and Scalability Analysis of the 3D Spatially Variant Lattice Algorithm	March 8	
Seminar Series	Adventures in Engaging HPC to Solve Scientific Challenges over Multiple Scales	April 8	
Seminar Series	Evaluation of a Distributed Computing Application	April 9	
Seminar Series	Experiences Transitioning a Molecular Modeling Group to GPU-accelerated Computing	April 15	
Seminar Series	Large-Scale Computing of Multi-Physics/Multi-Scale Problems	April 22	
Seminar Series	GPU-I-Tasser: a GPU Accelerated I-Tasser Protein Structure Prediction Tool	April 22	
Seminar Series	Using Machine Learning and HER Data to Change Familial Hypercholesterolemia Health Outcomes	April 26	
Seminar Series	Core-Collapse Supernova Simulations – The Power of Summit and Rhea	April 30	

Event type	Event title	Date	Participants
Seminar Series	Explainable and Network-based Machine Learning for Critical Infrastructures	June 3	
Seminar Series	Stance Detection with Weakly-supervised NLP	June 7	
Seminar Series	A Retrospective Look at the Hobbes Exascale Project and Directions for the Future	June 8	
Seminar Series	Insights into Cellular Signaling Mechanisms of Kras Monomer and Dimer on Lipid Membranes	June 8	
Seminar Series	Dispersion, Mixing, and Combustion in Dense Flows	June 10	
Seminar Series	Numerical Methods and Simulations of Multiphase and Multiphysics Flows: From Fundamental Physics to Engineering Applications	June 11	
Seminar Series	History of the Exascale Project and How Frontier Solves the Four Exascale Challenges	June 14	
Seminar Series	Explorations in Machine Learning based Fault-Tolerance and Quantum Circuit Simulations in High-Performance Computing Settings and Graph Sampling for Complex Networks	June 25	
Seminar Series	ATS Seminar Series: Machine Learning for Weather and Climate Predictions	June 28	
Seminar Series	How to Leverage Multi-tiered Storage to Accelerate I/O	July 12	
Seminar Series	Getting Ahead of the Virus with Machine Learning	July 26	
Seminar Series	Accelerated Data Science on National Lab Resources an Introduction to RAPIDS	August 9	
Seminar Series	Design Structure and Algorithms of the Nuclear Tensor Contraction Library (NTCL)	August 19	
Seminar Series	Performance Portability of Computational Applications: Challenges and Solutions in Pharmaceutical and Biological Science Use Cases	August 20	
Seminar Series	An Overview of ORNL's AI Initiative	August 23	
Seminar Series	Taking Risks to Push the Boundaries	August 23	
Seminar Series	How Projects Drive Solutions	August 30	
Seminar Series	Building on Frontier – What's Next for OLCF?	September 1	
Seminar Series	Scalability Engineering for Parallel Programs Using Empirical Performance Models	September 2	
Seminar Series	Integrated Cyberinfrastructure for Ensemble Cryo-Em Applications & Modeling (ICE-CREAM)	September 13	
Seminar Series	Partial Differential Equation Solvers on Vector Architecture Supercomputers	September 13	
Seminar Series	Using Applications to Guide Data Management for Emerging Memory Technologies	September 15	
Seminar Series	Efficient Scaling of Scientific Computing Applications in Cloud	September 27	
Seminar Series	OpenStack and the Software-Defined Supercomputer	September 27	
Seminar Series	Data-Driven Modeling of Turbulent Flows: Approaches Using Graph Theory and AI	September 29	
Seminar Series	Computational Tools for the Pandemic	October 6	

Event type	Event title	Date	Participants
Seminar Series	Towards Aerodynamic Flow Control Simulations on Exascale Supercomputers	October 8	
Seminar Series	Supercomputing and the Scientist: How HPC and Large-scale Data Analytics are Transforming Experimental Science	October 11	
Seminar Series	Fighting the Curse of Dimensionality with Tensor Network Manifolds and High Performance Computing	October 20	
Seminar Series	Accelerating Multi-dimensional Sparse Computations on Heterogeneous Parallel Architectures	October 20	
Seminar Series	NASA's Earthdata Cloud Migration: One Approach for Serving and Integrating Large Volumes of Earth Observation Data	October 25	
Seminar Series	Fine-Grain Parallel Linear and Nonlinear Solvers in Science and Engineering	October 26	
Seminar Series	The Advantage of Selection of Configurations and Electron Propagator Concepts for Describing the Electronic Correlations in Atoms and Molecules	October 28	
Seminar Series	Designing Metal Organic Frameworks for Carbon Dioxide Capture: Computational Discovery Guiding Experimental Design	October 29	
Seminar Series	Embracing Heterogeneity in Cloud Platforms through Adaptable Resource Management Framework	November 1	
Seminar Series	GronOR, a Massively Parallel and GPU Accelerated Program for Non-Orthogonal Configuration Interaction with Fragments	November 3	
Seminar Series	Efficient Genomic Collection Analysis with Randomized Algorithms	November 4	
Seminar Series	Ecosystem Computing with Sensitive Data	November 8	
Seminar Series	Journey from the Universe to Quarks with High Performance Computing and Quantum Computing	November 9	
Seminar Series	Fighting the Curse of Dimensionality with Tensor Network Manifolds and High Performance Computing	December 1	
Seminar Series	Production Deployment of Machine-Learned Surrogate Models on HPC	December 13	
Seminar Series	JupiterNCSM: A Pantheon of Nuclear Physics – An Implementation of Three-Nucleon Forces in the No-Core Shell Model	December 14	
Seminar Series	Zettascale Computing on Exascale Platforms	December 15	
Seminar Series	Simulating the Evolution of Nuclear Systems with Custom Gates	December 15	

APPENDIX C. OUTREACH PRODUCTS

Date	Science Highlights Submitted to DOE Title	Writer	URL
1/5/21	Former National Center for Computational Sciences Director James Hack Retires from 40-Year Computing Career	Rachel McDowell	https://www.olcf.ornl.gov/2021/01/05/former-national-center-for-computational-sciences-director-james-hack-retires-from-40-year-computing-career/
1/12/21	DOE and NOAA Extend Strategic Partnership	Coury Turczyn	https://www.olcf.ornl.gov/2021/01/12/doe-and-noaa-extend-strategic-partnership/
1/12/21	University of Delaware Team Tightens up Code for Exascale Computing on Frontier	Rachel McDowell	https://www.olcf.ornl.gov/2021/01/12/university-of-delaware-team-tightens-up-code-for-exascale-computing-on-frontier/
1/12/21	Rethinking GPU Hackathons for a Pandemic and Beyond	Rachel McDowell and Coury Turczyn	https://www.olcf.ornl.gov/2021/01/12/rethinking-gpu-hackathons-for-a-pandemic-and-beyond/
1/28/21	A New Approach for Bigger and Better Earthquake Modeling	Rachel McDowell	https://www.olcf.ornl.gov/2021/01/28/a-new-approach-for-bigger-and-better-earthquake-modeling/
1/28/21	Supercomputers Aid Scientists Studying the Smallest Particles in the Universe	Rachel McDowell	https://www.olcf.ornl.gov/2021/01/28/supercomputers-aid-scientists-studying-the-smallest-particles-in-the-universe/
2/1/21	OLCF, UTK Researchers Explore Potential of Quantum Annealing to Solve Optimization Problems	Andrea Schneibel	https://www.olcf.ornl.gov/2021/02/01/olcf-utk-researchers-explore-potential-of-quantum-annealing-to-solve-optimization-problems/
2/9/21	US Air Force, ORNL Launch Next-Generation Global Weather Forecasting System	Coury Turczyn	https://www.olcf.ornl.gov/2021/02/09/us-air-force-ornl-launch-next-generation-global-weather-forecasting-system/
2/18/21	Computational Staff at ORNL among Those Recognized with DOE Secretary's Honor Awards	Rachel McDowell	https://www.olcf.ornl.gov/2021/02/18/computational-staff-at-ornl-among-those-recognized-with-doe-secretarys-honor-awards/
2/18/21	Pioneering Frontier: Planning Ahead	Coury Turczyn	https://www.olcf.ornl.gov/2021/02/18/pioneering-frontier-planning-ahead/
2/18/21	Scientists Use Supercomputers to Study Reliable Fusion Reactor Design, Operation	Rachel McDowell	https://www.olcf.ornl.gov/2021/02/18/scientists-use-supercomputers-to-study-reliable-fusion-reactor-design-operation/

Date	Science Highlights Submitted to DOE Title	Writer	URL
2/19/21	UT–ORNL Data Center Course Takes Students on a Virtual Ride through ORNL's Supercomputing Facilities	Rachel McDowell	https://www.olcf.ornl.gov/2021/02/19/utornl-data-center-course-takes-students-on-a-virtual-ride-through-ornls-supercomputing-facilities/
2/26/21	Bronson Messer Named 2020 ORPA Postdoc Mentor of the Year	Laurie Varma	https://www.olcf.ornl.gov/2021/02/26/bronson-messer-named-2020-orpa-postdoc-mentor-of-the-year/
2/26/21	OLCF's Buddy Bland Earns DOE Secretary of Energy Exceptional Service Award	Rachel McDowell	https://www.olcf.ornl.gov/2021/02/26/olcfs-buddy-bland-earns-doe-secretary-of-energy-exceptional-service-award/
2/26/21	Arthur "Buddy" Bland, World-Leading Supercomputing Project Director, Retires after 40 Years Dedicated to HPC	Rachel McDowell	https://www.olcf.ornl.gov/2021/02/26/arthur-buddy-bland-world-leading-supercomputing-project-director-retires-after-40-years-dedicated-to-hpc/
3/31/21	Oak Ridge National Laboratory joins ATOM Consortium to Accelerate Drug Discovery	Katie Bethea	https://www.olcf.ornl.gov/2021/03/31/oak-ridge-national-laboratory-joins-atom-consortium-to-accelerate-drug-discovery/
3/31/21	ORNL's Supercomputer Utility Squad	Coury Turczyn	https://www.olcf.ornl.gov/2021/03/31/ornls-supercomputer-utility-squad/
3/31/21	Pioneering Frontier: Keeping the Power On	Coury Turczyn	https://www.olcf.ornl.gov/2021/03/31/pioneering-frontier-keeping-the-power-on/
3/31/21	The Magic is Gone for Certain Atomic Nuclei	Coury Turczyn	https://www.olcf.ornl.gov/2021/03/31/the-magic-is-gone-for-certain-atomic-nuclei/
4/12/21	U.S. Department of Energy's INCITE Program Seeks Proposals for 2022	Coury Turczyn	https://www.olcf.ornl.gov/2021/04/12/us-department-of-energys-incite-program-seeks-proposals-for-2022/
4/22/21	Computer Simulations Shed Light on Nanomaterial Structures	Rachel McDowell	https://www.olcf.ornl.gov/2021/04/22/computer-simulations-shed-light-on-nanomaterial-structures/
4/22/21	Pioneering Frontier: Staying on Track	Rachel McDowell	https://www.olcf.ornl.gov/2021/04/22/pioneering-frontier-staying-on-track/
4/27/21	ORNL Licenses Revolutionary AI System to General Motors for Automotive Use	Coury Turczyn	https://www.olcf.ornl.gov/2021/04/27/ornl-licenses-revolutionary-ai-system-to-general-motors-for-automotive-use/
5/5/21	Virtual Workshop on HPCToolkit Profiler Gives Computational Codes a Boost	Rachel McDowell	https://www.olcf.ornl.gov/2021/05/05/virtual-workshop-on-hpctoolkit-profiler-gives-computational-codes-a-boost/

Date	Science Highlights Submitted to DOE Title	Writer	URL
5/5/21	Pioneering Frontier: Managing Milestones	Coury Turczyn	https://www.olcf.ornl.gov/2021/05/05/pioneering-frontier-managing-milestones/
5/5/21	Scientists Tap Supercomputing to Study Exotic Matter in Stars	Rachel McDowell	https://www.olcf.ornl.gov/2021/05/05/scientists-tap-supercomputing-to-study-exotic-matter-in-stars/
5/5/21	NCCS Introduces CITADEL Security Framework	Coury Turczyn	https://www.olcf.ornl.gov/2021/05/05/nccs-introduces-citadel-security-framework/
5/5/21	Developing a Data Ecosystem	Coury Turczyn	https://www.olcf.ornl.gov/2021/05/05/developing-a-data-ecosystem/
5/10/21	Register Now for Lustre User Group Conference 2021	Matt Lakin	https://www.olcf.ornl.gov/2021/05/10/register-now-for-lustre-user-group-conference-2021/
5/19/21	Physicists Crack the Code to Signature Superconductor Kink Using Supercomputing	Rachel McDowell	https://www.olcf.ornl.gov/2021/05/19/physicists-crack-the-code-to-signature-superconductor-kink-using-supercomputing/
5/19/21	Pioneering Frontier: Planning Against Pitfalls	Matt Lakin	https://www.olcf.ornl.gov/2021/05/19/pioneering-frontier-planning-against-pitfalls/
5/20/21	OLCF Announces Storage Specifications for Frontier Exascale System	Matt Lakin	https://www.olcf.ornl.gov/2021/05/20/olcf-announces-storage-specifications-for-frontier-exascale-system/
6/9/21	GE Spins up Supercomputer Models to Zero in on Energy Loss in Turbines	Rachel McDowell	https://www.olcf.ornl.gov/2021/06/09/ge-spins-up-supercomputer-models-to-zero-in-on-energy-loss-in-turbines/
6/11/21	Building a Better Compiler	Coury Turczyn	https://www.olcf.ornl.gov/2021/06/11/building-a-better-compiler/
6/11/21	Pioneering Frontier: Forging a New Compiler	Coury Turczyn	https://www.olcf.ornl.gov/2021/06/11/pioneering-frontier-forging-a-new-compiler/
6/14/21	Spock and HIP Seminars Look Forward to Supercomputing's Near Future	Coury Turczyn	https://www.olcf.ornl.gov/2021/06/14/spock-and-hip-seminars-look-forward-to-supercomputings-near-future/
6/17/21	Argonne and Oak Ridge National Laboratories Award Codeplay Software to Further Strengthen SYCL™ Support	Katie Bethea	https://www.olcf.ornl.gov/2021/06/17/argonne-and-oak-ridge-national-laboratories-award-codeplay-software-to-further-strengthen-syclx2122-support-extending-the-open-standard-software-for-amd-gpus/
6/28/21	Benchmarking Mixed-Precision Performance	Coury Turczyn	https://www.olcf.ornl.gov/2021/06/28/benchmarking-mixed-precision-performance/

Date	Science Highlights Submitted to DOE Title	Writer	URL
7/2/21	ALCC Program Announces 2021–2022 OLCF Research Grants	Coury Turczyn	https://www.olcf.ornl.gov/2021/07/02/alcc-program-announces-2021-22-olcf-research-grants/
7/9/21	Unstoppable Science	Matt Lakin	https://www.olcf.ornl.gov/2021/07/09/unstoppable-science/
7/21/21	Pioneering Frontier: Fine-Tuning "Serial No. 1"	Matt Lakin	https://www.olcf.ornl.gov/2021/07/21/pioneering-frontier-fine-tuning-serial-no-1/
7/21/21	Introduction to Artificial Intelligence Inspires Next Generation	Elizabeth Rosenthal	https://www.olcf.ornl.gov/2021/07/21/introduction-to-artificial-intelligence-inspires-next-generation/
7/22/21	Hands-On Supercomputing	Matt Lakin	https://www.olcf.ornl.gov/2021/07/22/hands-on-supercomputing/
7/22/21	Titan Study Takes Jet Turbine Design to New Heights	Matt Lakin	https://www.olcf.ornl.gov/2021/07/22/titan-study-takes-jet-turbine-design-to-new-heights/
8/9/21	Shape-Based Model Sheds Light on Simplified Protein Binding	Rachel McDowell	https://www.olcf.ornl.gov/2021/08/09/shape-based-model-sheds-light-on-simplified-protein-binding/
8/31/21	Pioneering Frontier: Test-Driving with "User Zero"	Matt Lakin	https://www.olcf.ornl.gov/2021/08/31/pioneering-frontier-test-driving-with-user-zero/
8/31/21	OLCF's Advanced Technologies Section Hosts Seminar Series	Rachel McDowell	https://www.olcf.ornl.gov/2021/08/31/olcfs-advanced-technologies-section-hosts-seminar-series/
9/9/21	Reaching Remarkable States of Matter with Laser Simulations	Rachel McDowell	https://www.olcf.ornl.gov/2021/09/09/reaching-remarkable-states-of-matter-with-laser-simulations/
9/29/21	Stunning Specs: What's inside the Nation's First Exascale Supercomputer Facility?	Rachel McDowell	https://www.olcf.ornl.gov/2021/09/29/stunning-specs-whats-inside-the-nations-first-exascale-supercomputer-facility/
9/29/21	OLCF Organizes First Hackathon on Slate Service	Rachel McDowell	https://www.olcf.ornl.gov/2021/09/29/olcf-organizes-first-hackathon-on-slate-service/
9/29/21	Pioneering Frontier: Packaging a User-Friendly Supercomputer Environment	Coury Turczyn	https://www.olcf.ornl.gov/2021/09/29/pioneering-frontier-packaging-a-user-friendly-supercomputer-environment/
9/29/21	Pioneering Frontier: Meeting Industry at Scale	Rachel McDowell	https://www.olcf.ornl.gov/2021/09/29/pioneering-frontier-meeting-industry-at-scale/
10/18/21	Exascale Day 2021	Katie Bethea	https://www.olcf.ornl.gov/2021/10/18/exascale-day-2021/

Date	Science Highlights Submitted to DOE Title	Writer	URL
10/18/21	Exascale Computing's Four Biggest Challenges and How They Were Overcome	Coury Turczyn	https://www.olcf.ornl.gov/2021/10/18/exascale-computings-four-biggest-challenges-and-how-they-were-overcome/
10/18/21	The Road to Exascale	Coury Turczyn	https://www.olcf.ornl.gov/2021/10/18/the-road-to-exascale/
10/25/21	Team Publishes Massive Dataset Unravelling Details of the Expanding Universe	Rachel McDowell	https://www.olcf.ornl.gov/2021/10/25/team-publishes-massive-dataset-unravelling-details-of-the-expanding-universe/
11/11/21	Unraveling the "Big Boom"	Matt Lakin	https://www.olcf.ornl.gov/2021/11/11/unraveling-the-big-boom/
11/15/21	Digging Deep into HPC Power and Cooling	Coury Turczyn	https://www.olcf.ornl.gov/2021/11/15/digging-deep-into-hpc-power-and-cooling/
11/15/21	INCITE Program Awards Supercomputing Time to 51 Projects to Advance Open Science	Katie Bethea	https://www.olcf.ornl.gov/2021/11/15/incite-program-awards-supercomputing-time-to-51-projects-to-advance-open-science/
11/16/21	Darwin on Fast Forward: ORNL Study on COVID-19 Earns Gordon Bell Special Prize Nomination	Matt Lakin	https://www.olcf.ornl.gov/2021/11/16/darwin-on-fast-forward-ornl-study-on-covid-19-earns-gordon-bell-special-prize-nomination/
11/16/21	Waltzing the Virus: Study on COVID-19 Reproduction Earns Gordon Bell Special Prize Nomination	Matt Lakin	https://www.olcf.ornl.gov/2021/11/16/waltzing-the-virus-study-on-covid-19-reproduction-earns-gordon-bell-special-prize-nomination/
11/16/21	We Know #COVIDisAirborne—Now We Have the First Ever Model of an Aerosolized Viral Particle	Rachel McDowell	https://www.olcf.ornl.gov/2021/11/16/we-know-covidisairbornenow-we-have-the-first-ever-model-of-an-aerosolized-viral-particle/
11/16/21	Team Earns Gordon Bell Prize Finalist Nomination for Simulating Carbon at Extreme Pressures and Temperatures	Rachel McDowell	https://www.olcf.ornl.gov/2021/11/16/team-earns-gordon-bell-prize-finalist-nomination-for-simulating-carbon-at-extreme-pressures-and-temperatures/
11/17/21	OLCF's Bing Xie Receives Early Career Award	Matt Lakin	https://www.olcf.ornl.gov/2021/11/17/olcfs-bing-xie-receives-early-career-award/
11/22/21	Meet Gina Tourassi, Director of the National Center for Computational Sciences	Rachel McDowell	https://www.olcf.ornl.gov/2021/11/22/meet-gina-tourassi-director-of-the-national-center-for-computational-sciences/

Date	Science Highlights Submitted to DOE Title	Writer	URL
12/2/21	"Unfreezing" Molecular Proteins to Discover Their Functions in Action	Coury Turczyn	https://www.olcf.ornl.gov/2021/12/02/unfreezing-molecular-proteins-to-discover-their-functions-in-action/
12/3/21	ORNL's AAIMS Group Scores Two Awards	Coury Turczyn	https://www.olcf.ornl.gov/2021/12/03/ornls-aaims-group-scores-two-awards/
12/13/21	A New Design for Quantum Computer–Monitored Electrical Grids	Coury Turczyn	https://www.olcf.ornl.gov/2021/12/13/a-new-design-for-quantum-computermonitored-electrical-grids/
12/13/21	Oak Ridge National Laboratory Supercomputers Support Nobel Prize–Winning Research	Rachel McDowell	https://www.olcf.ornl.gov/2021/12/13/oak-ridge-national-laboratory-supercomputers-support-nobel-prizewinning-research/
12/13/21	Closing In on Fusion	Rachel McDowell	https://www.olcf.ornl.gov/2021/12/13/closing-in-on-fusion/
12/29/21	2021 at the OLCF: Year in Review	Rachel McDowell	https://www.olcf.ornl.gov/2021/12/29/2021-at-the-olcf-year-in-review/

APPENDIX D. OPERATIONAL SYSTEMS SUMMARY

The OLCF provided the Summit computational resource (Table D.1) and the Alpine and HPSS data resources for production use in 2021. 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.

IBM AC922 (SUMMIT) RESOURCE SUMMARY

The OLCF installed and deployed an IBM AC922 system named Summit, which became available for full production on January 1, 2019. Summit is composed of 4,662 high-density compute nodes, each equipped with two IBM POWER9 CPUs and six NVIDIA Volta GPUs. 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 (HPL) and Conjugate Gradient (HPCG) 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 a hardware spare parts cache for Summit if an optional sixth year is exercised in the contract.

Table D.1. Summit 2021.

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)	1530 MHz NVIDIA V100 (Volta)	4,662	IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node)	(4608) 512 GB DDR4 and 96 GB HBM2 per node; (54) 2TB DDR4 and 192 GB HBM2 per node; >10PB DDR4 + HBM + Non-volatile aggregate	Mellanox EDR 100G InfiniBand (Non-blocking Fat Tree)

[a] SMP = symmetric multiprocessing
[b] SM = streaming multiprocessor

GPFS FILE SYSTEMS (ALPINE AND WOLF) RESOURCE SUMMARY

In January 2019, the OLCF released Alpine, its next-generation global file system to support the computational resources in the OLCF. Alpine is a single GPFS namespace named Alpine, with a usable capacity of 250 PB and a filesystem-level performance of 2.5 TB/s. The Alpine file system is the default high-performance parallel file system for all the OLCF's moderate compute resources.

In March 2017, the OLCF procured, installed, and deployed the Wolf General Parallel File System (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.

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 INCITE- or ALCC-supported projects are automatically given accounts on the data analytics cluster. DD projects may also request access to this cluster. Andes is a 704-node cluster, each node containing two 16-core 3.0 GHz AMD EPYC processors and 256GB 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, LANL, Sandia National Laboratories (SNL), LLNL, LBNL, and ORNL. The ORNL HPSS instance is currently over 116 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 housing 81 IBM TS1155 tape drives and over 17,000 tape media slots—giving ORNL a current capacity of 180 PB, expandable well into hundreds of petabytes. The archive has ingested over 1 PB in a single day within the last year with the average daily ingestion ranging between 100–150 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 ft × 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, providing 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.

PROTECTED DATA INFRASTRUCTURE SUMMARY

The OLCF now provides a production enclave to support the processing of HIPAA and ITAR data. 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 filesystem named 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 in Table D.2. This 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 D.2. OLCF HPC system production dates, 2008–present.

System	Type	Production date[a]	Performance end date[b]	Notes
Summit	IBM AC922	March 19, 2021	Present	Production with 4,680 hybrid CPU-GPU nodes [IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node)]. Three additional cabinets added as a spare parts cache for an optional sixth year.
Summit	IBM AC922	July 1, 2020	Present	Production with 4,626 hybrid CPU-GPU nodes [IBM POWER9 CPUs (2/node) + NVIDIA Volta V100 GPUs (6/node)]. 3 additional cabinets added for COVID-19 research.
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 parallel file system	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.
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.

System	Type	Production date[a]	Performance end date[b]	Notes
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.

GPU USAGE

The leadership computing systems at OLCF provide heterogeneous architectures that allow users to exploit a hybrid compute node containing both CPUs and GPUs. Hybrid nodes provide researchers with a diverse architecture that is well suited for certain operations. As such, the use of this diverse architecture is optional and is exercised in different ways by research teams.

In 2021, the OLCF continued tracking GPU usage on Summit through NVIDIA's system management interface tool. Figure D.1 shows the percentage of GPU-enabled hours compared with CPU-only hours of each of the three primary allocation programs at the OLCF (INCITE, ALCC, and DD). The INCITE program on Summit exceeds all programs with an incredible 95% of hours utilized on Summit using the GPUs. Certainly, the overwhelming majority of the capability of the Summit nodes resides in the GPUs, and this number as well as the percentages from the ALCC and DD programs reflects the recognition and use of that capability from the OLCF user programs. Overall, Summit boasted a hefty 86% in 2021 for GPU-enabled applications matching the 2020 value.

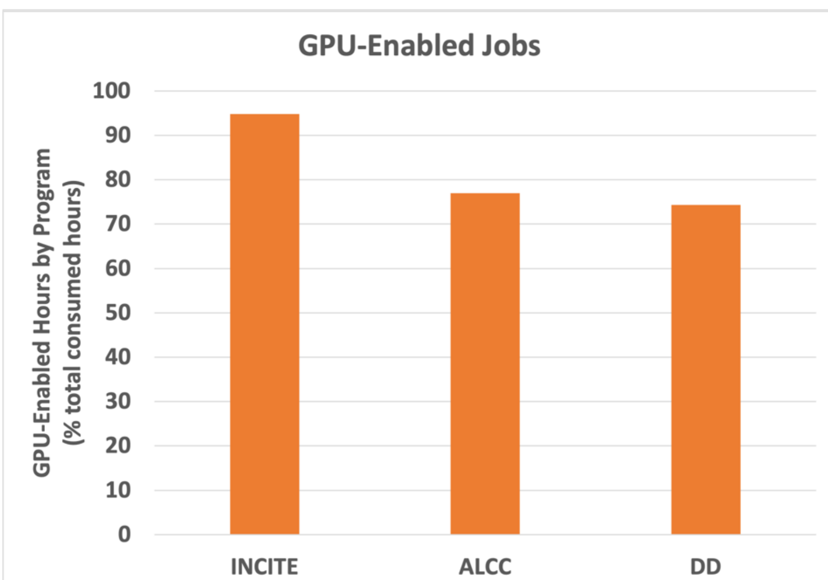


Figure D.1. 2021 GPU-enabled usage by program.

APPENDIX E. BUSINESS RESULTS FORMULAS

2021 OPERATIONAL ASSESSMENT GUIDANCE

Scheduled Availability

Scheduled availability (Eq. [E.1]) in HPC facilities is the percentage of time 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 h before the event and preferably as much 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.

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

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

Overall Availability

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

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

Mean Time to Interrupt

Mean time to interrupt is the time, on average, to any outage of the full system, whether unscheduled or scheduled. It is also known as *mean time between interrupt (MTBI)*, as shown in Eq. (E.3).

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

Mean Time to Failure

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

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

System Utilization

SU is the percent 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. [E.5]).

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

APPENDIX F. DD PROJECTS ENABLED AT ANY POINT IN CY 2021

Table F.1 shows the DD projects enabled at any point in CY 2021.

Table F.1. DD projects enabled at any point in CY 2021.

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
ARD137	Eric Nielsen	NASA Langley Research Center	20,000	14,795	FUN3D Scaling for CFD Simulations Requiring General Gas Chemistry
ARD138	Ngoc-Cuong Nguyen	MIT	20,000	10,730	Direct Numerical Simulation of Transitional Turbulent Flows Past Hypersonic Vehicles
ARD139	Zhi Wang	University of Kansas	15,000	1,086	Large eddy simulation of a high-lift configuration on GPUs
ARD143	Wesley Brewer	Computational Solutions	10,000	87	iBench Extensions for ML Surrogate Model Inferencing at Petascale
ARD145	Parviz Moin	Stanford	20,000	19,920	Large Eddy Simulation of High-Lift Aircraft
ARD147	Wesley Harris	MIT	20,000	2,576	Direct numerical simulations of hypersonic boundary layer transition on a flat plate
AST031	Pierre Ocvirk	Universite de Strasbourg, Strasbourg Astronomical Observatory	20,000	730,346	Reionization And Its Impact on The Local Universe: Witnessing Our Own Cosmic Dawn
AST146	Brian O'Shea	Michigan State University	20,000	3,863	Measuring Performance and Scaling of Kokkos-accelerated Athena++ on Summit
AST154	Philipp Moesta	University of Amsterdam	20,000	0	Dynamical Space-time GRMHD Simulations of Neutron-star Mergers and Remnants
AST158	Robert Fisher	UMass Dartmouth	20,000	3,639	Explorations of the D6 Scenario of Type Ia Supernovae on Summit
AST162	Bronson Messer	ORNL	50,000	0	Three-dimensional Stellar Explosions with Detailed Nuclear Kinetics
AST163	Christian Cardall	ORNL	20,000	310	INCITE Preparation for 3D+1D core-collapse supernova simulations with GenASiS
AST165	Francis Halzen	U. of Wisconsin	15,000	0	IceCube Simulation at Summit
AST166	Gaurav Khanna	UMass Dartmouth	20,000	682	Mixed-Precision WENO Method for Hyperbolic PDE Solutions
AST167	Richard Sarmento	United States Naval Academy	20,000	2	Chemical Evolution and Reionization In The Early Universe

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
AST169	Brant Robertson	UC Santa Cruz	70,000	34,100	Preparing Cholla for Extreme-Scale Cosmological Simulations using Summit
AST170	Yuran Chen	U. of Colorado	20,000	944	Simulations of the Magnetospheres of Neutron Stars and Black Holes
AST171	Charles Gammie	UIUC	10,000	2,364	Next-Generation Modeling for the EHT
AST172	Ulrich Steinwandel	Flatiron Institute	20,000	1,067	Multiphysics galaxy cluster simulations on the GPU with Gadget3
ATM112	Wei Zhang	ORNL	60,000	4,928	Aim High: Air Force R&D Collaboration
ATM114	Richard Loft	UCAR	12,000	4,292	Increasing MPAS-A's GPU Production Readiness at Convection Resolving Scales
ATM115	Stan Posey	NVIDIA	10,000	12,614	WRFg
ATM116	Todd Hutchinson	IBM	15,000	0	IBM GRAF
ATM117	Branko Kosovic	NCAR, UCAR	7,000	7,018	GPU-resident Real-time Large-eddy Simulations of Atmospheric Boundary Layer
ATM121	Chao Sun	U. of Maryland	15,000	0	Dashboard for Agricultural Water use and Nutrient management (DAWN)
ATM122	Branko Kosovic	NCAR, UCAR	10,000	0	GPU-resident Real-time Large-eddy Simulations of Stably-stratified Atmospheric Boundary Layers
BIE104	Daniel Jacobson	ORNL	100,000	0	CBI Contribution in Kind Allocation
BIE108	Peter St. John	NREL	10,000	0	Enzyme Engineering Directly from Protein Primary Sequence Via Natural Language Processing
BIF113	Kjiersten Fagnan	LBNL	0	0	JGI Data Archive
BIF117	Yanling Liu	FNLCR, Frederick National Laboratory	20,000	5,928	Systematic Annotation Creation on H&E WSIs for Automated Digital Pathology Informatics
BIF119	Ali Torkamani	Scripps Research	20,000	20,079	Genetic Imputation with Deep Learning
BIF121	Chongle Pan	OU	10,000	946	Acceleration And Parallelization of Bioinformatics Software For Processing Large-scale Proteogenomics Data

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
BIF122	Ramu Anandakrishnan	VCOM	15,000	0	Identifying Multi-hit Combinations of Genetic Mutations in Cancer
BIF132	Jianlin Cheng	U. of Missouri	20,000	21,092	Improving Deep Learning Prediction of Protein Inter-residue Distance by High Performance Computing
BIF133	Jeffrey Skolnick	Georgia Tech	20,000	19,958	Application Of a Deep-learning Based Sequence Alignment Algorithm To Annotate The Proteome of Cable Bacteria
BIF136	Andrew Blanchard	ORNL	20,000	24,559	Scalable Transformer language models for drug discovery
BIP167	Philip Kurian	Howard University	20,000	0	Computing Many-body Van Der Waals Dispersion Effects in Biomacromolecules
BIP169	Martin Karplus	Harvard University	20,000	252	Computational Design of HIV Vaccination Schedule
BIP196	Kirk Jordan	IBM	20,000	0	Molecular Mechanisms of Resistance to Last Defence Antibiotics
BIP198	Ada Sedova	ORNL	10,000	3,192	Genes to Interactomes
BIP207	Darrin York	Rutgers	20,000	0	Computational Tools for High-throughput Lead Optimization
BIP208	James Gumbart	Georgia Tech	80,000	46,899	Determining The Contribution of Glycosylation To SARS-CoV-2 S-protein Conformational Dynamics
BIP211	Rommie Amaro	UC San Diego	50,000	124,202	SARS-CoV-2 in Respiratory Aerosols
BIP212	Abhishek Singharoy	Arizona State University	21,000	55,955	Molecular Simulations of Energy Transfer in Supercomplexes
BIP213	Juan Perilla	UIUC, U. of Delaware	20,000	14,887	Molecular Basis of The Native SARS-CoV-2 Liposome Physical Properties from All-atom Simulations
BIP214	Jens Glaser	ORNL	15,000	15,352	Data Centric Approach to Drug Discovery for COVID-19
BIP215	Julie Mitchell	ORNL	50,000	16,233	Training An Advanced Model for Structure-based Proteomics
BIP216	Arvind Ramanathan	ANL	40,000	1,529	Understanding SARS-CoV-2 Replication-Transcription Complex by Integrating Experiments, Simulations and ML Techniques

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
BIP217	Alan Hicks	ORNL,	20,000	7	Sampling Closed-Open Transition of SARS-CoV-2 Nsp15 in Real Time
CFD124	Bamin Khomami	UTK	120,000	77,958	Elucidating the Molecular Rheology of Entangled Polymeric Fluids via Direct Comparison of NEMD Simulations and Model Predictions
CFD129	Ramesh Balakrishnan	ANL	10,000	1	Performance Modeling of CFD Solvers on Heterogeneous Computing Platforms
CFD131	Sumanta Acharya	Illinois Tech	0	0	Demonstration of OpenFOAM software usage on Summit
CFD132	Azardokht Hajiloo	GE Power	20,000	7,665	Combustion Dynamics Predictions in The “Cyber Space”, From “Singing Flames” To A “Roaring Fire”
CFD135	Stephen de Bruyn Kops	University of Massachusetts, UMass, UMass Amherst	0	0	Stratified Turbulence at Very Low Froude Number
CFD136	John Gounley	ORNL	15,000	1,430	Performant High-Order Lattice Boltzmann for Exascale Applications
CFD138	Duane Rosenberg	Colorado State University	20,000	0	GPU Performance Analysis for Fluid Turbulence Codes
CFD139	Marc-Olivier Delchini	ORNL	15,000	0	High-fidelity solution of turbulent flow in a HFIR cooling channel with NekRS
CFD140	Stephan Priebe	GE, GE	20,000	9,989	Large Eddy Simulations of SBLI using High-Order Solver Genesis
CFD141	Kirk Jordan	IBM	20,000	20,543	Investigating CFD Models Involving Complex Geometries for ML-driven Searches at Scale
CFD144	Saumil Patel	ANL	18,000	4	Scaling Internal Combustion Engine Simulations Using nekRS
CFD146	Ramanan Sankaran	ORNL	5,000	7,508	Performance Studies of Quilt for Direct Numerical Simulation of Multiphase Flows
CFD147	Eric Johnsen	U. of Michigan	15,000	7	Cavitation Dynamics in The Spallation Neutron Source Target
CFD149	Davide Modesti	Delft University of Technology	10,000	19,624	Direct Numerical Simulation of Turbulent Boundary Layers Over Acoustic Liners

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CFD150	Fazle Hussain	Texas Tech University	20,000	3,264	Direct Numerical Simulation of Turbulent Pipe Flow Using Nek-OpenMP/OpenACC
CFD154	Spencer Bryngelson	Georgia Tech	10,000	739	Accelerated Sub-grid Multi-component Flow Physics
CHM154	Coen de Graaf	ICREA, Rovira i Virgili University, URV	70,000	101,554	Intermolecular energy and electron transfer by non-orthogonal configuration interaction
CHM156	Remco Havenith	U. of Groningen	10,000	1,015	TURTLE
CHM160	Andre Severo Pereira Gomes	CNRS	25,500	59,357	PRECISE: Predictive Electronic Structure Modeling of Heavy Elements
CHM174	Monojoy Goswami	UTK, ORNL	20,000	2,202	Multi-scale/multi-physics molecular simulations at the Chemical Sciences Division
CHM175	Thomas Miller	Entos	12,000	2,946	Accelerating COVID-19 Drug Discovery Through Leadership-scale Graph Neural Net Machine Learning
CHM177	Edward Hohenstein	SLAC National Accelerator Laboratory	20,000	4	Development Of Heterogeneous Parallel Electronic Structure Methods
CHM179	Eugen Hruska	Rice U., Emory University	12,000	27	Extensible and Scalable AdaptiveAIMD
CHM180	Aditya Savara	ORNL	25,000	1,889	Simulating Mesoscale Structural Dynamics of Vitrimers
CHP108	Timmy Ramirez-Cuesta	ORNL	20,000	9,907	Development of SNS-Spectroscopy Leadership-Computational Methods
CHP110	Neeraj Rai	Mississippi State University	0	0	BioMass-Rai
CLI120	Giri Prakash	ORNL	0	0	Symbiotic Simulation and Observation (LASSO) - Initialization, Forcing, and Multiscale Data Assimilation
CLI137	Forrest Hoffman	ORNL	10,000	0	Land Model Testbed (LMT) for Rapid Assessment of Multiscale Complex Biogeochemistry in Earth System Models
CLI138	Moetasim Ashfaq	ORNL	0	0	Analytical Frameworks for Sub-Seasonal to Multi-Decadal Climate Predictions and Impact Assessments

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CLI143	Nathaniel Collier	ORNL	15,000	0	Identifying Ecosystems Vulnerable to Climate Change
CLI144	Peter Thornton	ORNL	20,000	4,209	Ultra-high Resolution Land Process Simulation Using GPUs And OpenACC on Summit
CLI900	Valentine Anantharaj	ORNL	5,000	0	Provisioning of Climate Data
CMB103	Jacqueline Chen	Sandia	23,000	10,845	DNS of Turbulent Combustion Towards Efficient Engines with In Situ Analytics
CMB124	Ronald Grover	GM	20,000	3,554	Prediction of Engine Knock in a Gasoline Direct Injection Engine
CMB147	Joseph Oefelein	Georgia Tech	20,000	526	Analysis of Combustion and Wave Dynamics in Rotating Detonation Engines
CMB148	Aditya Konduri	IISc, IISc Bangalore	10,000	5,715	Scalable Mathematically Asynchronous Algorithms for Flow Solvers
CPH005	Dario Alfe	University College London	80,000	164	New Frontiers for Material Modeling via Machine Learning Techniques with Quantum
CPH111	Andreas Glatz	ANL	18,000	279	Genetic Algorithms for Tailored Superconducting Materials
CPH122	Cristian Batista	UTK	20,000	14,437	Transport Properties and Dynamics of Quantum Materials
CPH123	Panchapakesan Ganesh	ORNL	20,000	0	Defects, Interfaces and Disorder in Correlated Quantum Materials
CPH127	Mark Oxley	ORNL	20,000	0	The Application of Machine Learning to 4D-STEM data
CPH128	Corey Melnick	U. of Michigan, BNL	25,000	31,564	Application of GPU-accelerated Quantum Monte-Carlo Impurity Solver to Plutonium Compounds
CPH130	Jihong Ma	UVM	20,000	72	Conductive Polymer
CPH132	Ashley Shields	ORNL	20,000	25,275	Exascale Lattice Dynamics of Uranium Materials
CPH136	Marco Govoni	ANL	20,000	14,266	benchMBPT
CPH137	Emanuel Gull	U. of Michigan	10,000	1,930	Exploration of Time-dependent Second Order Perturbation Theory
CPH138	Sumit Sharma	OU	20,000	26,512	Computational Studies of Interactions of Small Molecules with Biological and Polar Interfaces
CPH139	Zhanghui Chen	LBNL	20,000	0	Real-time time-dependent density functional theory method for spin dynamics simulations

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC040	Neena Imam	ORNL	30,000	2,019	Durmstrang
CSC143	Norbert Podhorszki	ORNL	40,000	46,401	ADIOS - The Adaptable IO System
CSC343	Sergey Panitkin	University of Montreal	5,000	429	Porting the ATLAS Experiment Software and Workload Management System
CSC345	Ramakrishnan Kannan	ORNL	20,000	44	Parallel Low-rank Approximation with Nonnegative Constraints (PLANC)
CSC349	Markus Rampp	MPCDF	10,000	21	Porting, Benchmarking, and Optimization of MPCDF Codes
CSC356	Edmon Begoli	ORNL	1,000	0	CITADEL Pilot
CSC357	Laxmikant Kale	UIUC	15,000	2,307	CharmRTS
CSC362	Wen-mei Hwu	UIUC	20,000	27,459	Petascale 3D Image Reconstruction with Ptycho-Tomography
CSC369	Dmitry Pekurovsky	UC San Diego	5,000	1,291	Scalable Software Framework for Multidimensional Fourier Transforms
CSC377	Allan Grosvenor	Microsurgeonbot	20,000	49	Expertise-as-a-Service via Scalable Hybrid Learning: R&D Supporting Improvements to Hybrid Intelligence Agent Tooling
CSC380	Bronson Messer	ORNL	150,000	40,569	CAAR for Frontier
CSC381	Ramakrishnan Kannan	ORNL	20,000	0	Graph500 Algorithms for Summit
CSC382	Catherine Schuman	UTK	5,000	52	Scalable Neuromorphic Simulation and Training
CSC387	Una-May O'Reilly	MIT	25,000	4,925	Towards A Robust and Scalable Adversarial Learning Framework
CSC388	Thomas Uram	ANL	1,000	1	Balsam – An HPC Job Campaign Management Framework
CSC392	Thomas Deakin	University of Bristol	15,000	1,469	Scalability of unstructured mesh SN particle transport on GPUs
CSC395	Dali Wang	ORNL	15,000	41	An Extreme Scale Deep Reinforcement Learning Framework for Solving Big Graph Problems
CSC397	Michael McCarty	NVIDIA	15,000	0	Testing Legate Deployment
CSC398	Thomas Naughton	ORNL	5,000	7	Federated Scientific Instruments

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC399	Douglas Doerfler	LBNL	20,000	0	NESAP for Perlmutter
CSC400	Hari Sundar	University of Utah	10,000	0	Massively Parallel Simulations of Binary Black Hole Intermediate-Mass-Ratio Inspirals
CSC401	Barbara Chapman	SUNY Stony Brook	10,000	3,940	Support for the Deployment of OpenMP in Large-Scale Scientific Applications
CSC416	David Keyes	KAUST	20,000	0	Geostatistical Modeling and Prediction in Three Precisions
CSC418	Tjerk Straatsma	ORNL	60,000	1	ADAC Applications Readiness
CSC422	Luke Olson	UIUC	10,000	95	Node-Aware Communication
CSC424	Theodore Papamarkou	ORNL	11,200	0	Uncertainty Quantification by Combining Data Augmentation and Ensemble Learning
CSC425	Terry Jones	ORNL	20,000	16,070	WEAVABLE
CSC426	Athirai Aravazhi Irissappane	U. of Washington	20,000	0	Multi-Modal Brain Tumor Segmentation
CSC427	Michela Taufer	UTK	15,000	6,667	A4NN: Accelerating Scientific AI Leveraging Open Data and Open Models
CSC434	Bronson Messer	ORNL	20,000	29,093	Portable Performance on Exascale Hybrid Architectures
CSC436	John Paul Walters	USC	15,000	3,695	CASPER: Compiler Abstractions Supporting High Performance on Extreme-scale Resources
CSC443	Seung-Hwan Lim	ORNL	20,000	1,418	Generalized Scalable Parallel Training of Deep Graph Neural Networks
CSC445	Rama Vasudevan	ORNL	8,000	2,191	Variational Scalable Multi-Agent Reinforcement Learning
CSC452	Abhinav Bhatele	LLNL, U. of Maryland	20,000	50	Performance Analysis and Tuning of HPC and AI Applications
CSC453	Abhinav Bhatele	LLNL, U. of Maryland	8,000	135	Analyzing and Optimizing Parallel I/O and Performance Tools
CSC454	John Michalakes	NREL, UCAR	100	10	U.S. Navy NEPTUNE Weather Model Development for GPU
CSC455	Jacob Hinkle	ORNL	20,000	7,103	Design of Intrinsically Scalable Deep Learning Algorithms
CSC456	Benjamin Nebgen	LANL	15,000	20,561	Distributed Non-negative Tensor Factorization for Global Seismic Analysis

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
CSC457	Massimiliano Lupo Pasini	ORNL	20,000	24,641	Scalable stable numerical optimization for artificial intelligence applied to computed tomography
CSC462	Antonios Kougkas	Illinois Tech	15,000	0	Hermes: Extending the HDF Library to Support Intelligent I/O Buffering for Deep Memory and Storage Hierarchy System
CSC467	Andrew Gallo	GE	5,000	221	Genesis CI/CD
CSC470	Christopher Stanley	ORNL	20,000	0	Flexible Privacy-enabled Platform for Sensitive Applications
CSC471	Rishi Khan	Extreme Scale Solutions	10,000	238	Scaling Distributed GPU FFT to Exascale
ENG107	Andrew McGough	Newcastle University	20,000	169	NUFEB: Newcastle University Frontiers in Engineering Biology
ENG120	George Karniadakis	Brown	10,000	125	PINNs on Multi-GPUs
ENG130	Zhiting Tian	Cornell	18,000	0	Understanding Heat Transport in Complex, Ultra-low Conductivity Materials
FUS131	Randy Churchill	Princeton Plasma Physics Laboratory	20,000	401	Advanced Sequence Models and Self-supervised Learning for Fusion Energy Diagnostics
FUS134	Xianzhu Tang	LANL	12,000	0	Tokamak Disruption Simulations
FUS135	Chang Liu	Princeton Plasma Physics Laboratory	20,000	1,133	Kinetic-MHD Hybrid Simulation of Alfvén Instabilities Interacting with Energetic Particles
FUS137	William Fox	Princeton Plasma Physics Laboratory	20,000	361	Energetics of Collisionless Plasmas in the Laboratory and Space
FUS138	Axel Huebl	LBNL	0	0	Elucidating Acceleration Mechanisms in Laser-Plasma Ion Accelerators
FUS144	Predrag Krstic	SUNY Stony Brook	20,000	0	Classical and quantum-classical molecular dynamics for the fusion plasma-material interface
GEN042	Jason Kincl	ORNL	1,000	0	Platforms - Managed Work Activities
GEN156	Dale Stansberry	ORNL	0	0	DataFed
GEO132	Ethan Coon	LANL, ORNL	20,000	82	ExaSheds
GEO137	Dalton Lunga	ORNL	20,000	13,651	Accelerated Infrastructure Mapping from Trillion Pixels
GEO139	Sangkeun Lee	ORNL	5,000	357	URBAN-NET

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
HEP121	Corey Adams	ANL	20,000	0	Sparse Convolutional Neural Networks for Neutrinos
LGT100	Andre Walker-Loud	LBNL	80,000	105,091	The Structure and Interactions of Nucleons from the Standard Model
LGT107	Rajan Gupta	LANL	20,000	32,677	Nucleon Matrix Elements: Probes of New Physics
LGT117	Rajan Gupta	LANL	20,000	0	Nucleon Matrix Elements: Probes of New Physics
LRN009	George Siopsis	UTK	5,000	15	Hybrid Quantum-Classical Reinforcement Learning in Controlled Quantum Networks
LSC115	Vipin Sachdeva	Silicon Therapeutics	20,000	15,371	Weighted Ensemble Simulations for In Silico Modeling of Targeted Protein Degradation
MAT193	Trung Nguyen	Northwestern University	15,000	0	Development of New Models and Sampling Techniques for LAMMPS
MAT197	Yangyang Wang	ORNL	20,000	3,174	Exploring New Paradigms for Understanding Ionic Transport in Polymer Electrolytes
MAT199	Lou Kondic	NJIT	5,000	199	Exploiting Non-equilibrium Processes, Free Surface and Substrate Effects to tailor Phase Separation in Metal Alloys: a Molecular Dynamics Study
MAT201	Panchapakesan Ganesh	ORNL	100,000	41,202	Center for Nanophase Materials Sciences
MAT203	Zhiting Tian	Cornell	15,000	4,687	Thermal Transport Properties of Covalent Organic Frameworks (COFs)
MAT204	Ram Devanathan	PNNL	20,000	0	Simulated Adhesive Response to Aging
MAT208	Volker Blum	Duke	20,000	0	Distributed-Parallel Dense Eigenvalue Solutions for Electronic Structure Calculations in ELSI and ELPA
MAT223	Neil Gershenfeld	MIT	10,000	0	Performance Scaling of Particle Systems for Discovery of Multiphysics Models
MAT224	Dongwon Shin	ORNL	50,000	40,428	Computational Design of High-temperature Alloys

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
MAT226	Jan Michael Carrillo	UTK, ORNL	20,000	12,822	Molecular Dynamic Simulations of Amphiphilic Oligomer Membranes: Design Rules Towards Stable Membranes Capable of Learning and Memory
MAT230	Vikram Gavini	U. of Michigan	20,000	59,649	DFT-FE
MAT231	Zachary Ulissi	Carnegie Mellon University	35,000	10,373	Deep learning for electrocatalyst applications with the Open Catalyst 2020 Dataset
MAT233	Ying Wai Li	LANL	15,000	409	Data Driven Modeling of Non-Equilibrium Dynamics in Chemical and Materials Systems
MAT235	Guannan Zhang	ORNL	20,000	9,217	AI-based Scalable Feature Extraction for Real-time Processing And Analytics Of Neutron Scattering Data
MAT240	Yi Yao	Duke	20,000	33	GPU Acceleration of an All-Electron Full-Potential G0W0 Method
MAT242	Yangyang Wang	ORNL	10,000	1,594	Resolving Flow-induced Mesoscopic Structures In Polymeric Materials
MED107	John Gounley	ORNL	10,000	4,564	AI-Enabled Computational Cancer Phenotyping for Precision Oncology
MED112	Sumitra Muralidhar	US Dept. of Veteran Affairs	200,000	72,606	Genome-Wide Phenome-Wide Association Study in the Million Veteran Program
MED114	James Christian	ORNL	2,000	0	Construction of Environmental Determinants of Health Data for VA and NIH
MED115	Hong-Jun Yoon	ORNL	15,000	74	Studies of Scalable Machine Learning Algorithms for Histopathological Image Analysis
MPH113	Balakrishnan Naduvalath	UNLV	30,000	37,209	Ultracold Controlled Chemistry
NFI120	Justin Watson	University of Florida	20,000	8,535	Neural Networks for Exa-Scale Reactor Parameter Prediction
NPH133	Eric Church	PNNL	15,000	8,311	NEXT – Neutrino Experiment in a Xenon Time Projection Chamber
NPH136	Rick Archibald	ORNL	20,000	0	Development of Design Optimization Code for the Transformational Challenge Reactor

Project ID	PI	Institution	Most recent Summit allocation	Summit usage	Project name
NPH143	Kenneth McElvain	UC Berkeley	30,000	80	Neutron Star Interiors: Nuclear-matter Calculations from Chiral Effective Field Theory
NPH145	Kehfei Liu	University of Kentucky	20,000	3,005	Beyond Standard Model Physics from Lattice QCD
NRO106	Kristofer Bouchard	LBNL	15,000	3,398	DL4Neurons
NTI009	Lin-Wang Wang	LBNL	20,000	2,334	First Principle Heat Generation Simulations of On-die Interconnects
NTI114	Michael Shirts	U. of Colorado	19,000	0	Nanoscale phase behavior of and transport in lyotropic liquid crystal membranes
PHY130	Zhihong Lin	UC Irvine	20,000	1,210,689	Integrated Simulation of Energetic Particles in Burning Plasmas
PHY149	Dmitry Liakh	ORNL	20,000	885	Pushing the Limits of Classical Simulation of Hard Quantum Circuits via Novel Tensor Network Algorithms and Accelerated High Performance Computing
PHY151	Suzanne Parete-Koon	ORNL	5,000	2,494	2020 SMC Data Challenge
PHY155	Edwin Pednault	IBM	24,500	0	Leveraging Secondary Storage in Quantum Circuit Simulation
PHY158	Phillip Lotshaw	ORNL	25,000	11	QAOA simulations
PHY159	Suzanne Parete-Koon	ORNL	5,000	0	SMC Data Challenge 2021
PSS101	Jonathan Jara-Almonte	Princeton Plasma Physics Laboratory	20,000	0	Reconnection Turbulence in the Magnetotail
SYB105	Daniel Jacobson	ORNL	100,000	36,194	CBI Contribution in Kind Allocation
TUR120	Pui-kuen Yeung	Georgia Tech	60,000	17,305	High Resolution Study of Intermittency in Turbulence and Turbulent Mixing
TUR135	Sanjiva Lele	Stanford	20,000	17,772	Compressibility and Curvature Effects on Turbulent Mixing
TUR137	Justin Sirignano	Oxford	20,000	8,154	Deep Learning Closure Models for Large-Eddy Simulation of Unsteady External Aerodynamics
TUR138	Ramesh Balakrishnan	ANL	73,173	55,792	Simulating the Transmission and Dispersion of Aerosol in Turbulent Flows